LONG-TERM CHANGES IN SEASONAL TEMPERATURE EXTREMES WORLDWIDE

by

Allison Grace Van Ormer

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Earth Sciences

Charlotte

2023

Approved by:

____
Dr. Brian Magi

____
Dr. Jacob Scheff

____
Dr. Matthew Eastin

ABSTRACT

ALLISON GRACE VAN ORMER. Long-Term Changes in Seasonal Temperature Extremes Worldwide. (Under the direction of DR.BRIAN MAGI)

The temperatures a country, state, or any particular area experience over many decades tend towards a bell curve probability distribution, such that the mean value is the center of the distribution, and the range of values experienced can be quantified by the standard deviation, or sigma, around the mean value. This is how a temperature climatology is typically defined. Global warming is the documented increase in the temperature averaged over the entire planet, due primarily to the increase in greenhouse gas concentration, but global warming also suggests that the climatology of any given region has changed over time. In order to assess the change in climatology across any region in time, I defined a climatological base period of 1951-1980 and calculated the frequency of observed local monthly standardized temperature anomalies above +2 and +3, or 2-sigma and 3-sigma events, using the NOAAv5 Global Land and Ocean Temperature Anomaly dataset from 1880 to 2020. From 1980 to 2020, the frequency of 2-sigma and 3-sigma events increased exponentially across all seasons for both ocean and land. In the most recent decade (2010-2020), 2-sigma events occurred 12-20 times statistical expectation, while 3-sigma events occurred 83-222 times expectation. The increase was largest in Summer (JJA), as well as over the ocean. By definition, a climatology suggests that hot extremes rarely occur in any particular region, but my results show that hot extremes are now far more common in every season and across most regions of the world.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my graduate advisor, Dr. Brian Magi, for guidance, support and patience throughout my Master's thesis. I would also like to thank committee members, Dr. Matthew Eastin and Dr. Jacob Scheff, for contributing their knowledge, technical advice, and suggestions to this project. Additionally, this endeavor would not have been possible without the grants I received from the Graduate School at The University of North Carolina at Charlotte. I also want to thank the outstanding faculty and staff of the Department of Geography and Earth Sciences at the University of North Carolina at Charlotte, who provided assistance throughout the course of my academic endeavors. Lastly, I want to thank my friends and colleagues, Roger Riggin, Andrew Robinson and Lauren Decker for assistance and support throughout the course of this project.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS	x
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: METHODOLOGY	9
2.1. Dataset	9
2.2. Data Processing	10
CHAPTER 3: RESULTS	20
3.1. Kolmogorov-Smirnov Test of Distribution Normality	20
3.2. Interquartile Range of the Distributions	31
3.3. Frequency of 2-sigma Events	41
3.4. Frequency of 3-sigma Events	55
CHAPTER 4: SUMMARY	65
REFERENCES	70
APPENDIX A: ANALYSIS CODE	73

LIST OF TABLES

TABLE 2.1: NOAAv5 Global average temperature anomalies from 1880 to 2020 11

LIST OF FIGURES

FIGURE 1.1: Summer (JJA) climatology in the SE United States from 1951-1980 2

FIGURE 1.2: A plot of normal distribution (or bell-shaped curve) 3

FIGURE 1.3: NOAAv5 Global average temperature anomalies from 1880 to 2020 4

FIGURE 1.4: Shifting distributions of temperature anomalies in the Northern Hemisphere (Land) in Winter (DJF) from Hansen et al. (2012) 6

FIGURE 1.5: Frequency of occurrence of local temperature anomalies from Hansen et al. (2012) 7

FIGURE 2.1: NOAAv4 and NOAAv5 surface air temperature trends in Winter months over land areas from 1988 to 2017 9

FIGURE 2.2: Global maps of monthly standardized temperature anomalies from 1880 to 2020 10

FIGURE 2.3: Fall Histogram of a monthly standardized anomaly frequency distribution from 1880-1909 and 2000-2020 13

FIGURE 2.4: Example plot of a Kolmogorov–Smirnov (KS) test 17

FIGURE 2.5: Normalized frequency of Summer (JJA) monthly standardized temperature anomalies 18

FIGURE 3.1: Maps of Summer (JJA) p-values from the KS test from 2000-2020 22

FIGURE 3.2: Maps of Fall (SON) p-values from the KS test from 1880-1979 23

FIGURE 3.3: Maps of Fall (SON) p-values from the KS test from 1980-2020 24

FIGURE 3.4: Maps of Winter (DJF) p-values from the KS test from 1880-1979 25

FIGURE 3.5: Maps of Winter (DJF) p-values from the KS test from 1980-2020 26

FIGURE 3.6: Maps of Spring (MAM) p-values from the KS test from 1880-1979 27

FIGURE 3.7: Maps of Spring (MAM) p-values from the KS test from 1980-2020 28

FIGURE 3.8: Maps of Summer (JJA) p-values from the KS test from 1880-1979 29

FIGURE 3.9: Maps of Summer (JJA) p-values from the KS test from 1980-2020 30

FIGURE 3.10: Seasonal trend in average KS p-values from 1880-2020 31

FIGURE 3.11: Maps of Fall (SON) IQR values from 1880-1979 33

FIGURE 3.12: Maps of Fall (SON) IQR values from 1980-2020 34

FIGURE 3.13: Maps of Winter (DJF) IQR values from 1880-1979 35

FIGURE 3.14: Maps of Winter (DJF) IQR values from 1980-2020 36

FIGURE 3.15: Maps of Spring (MAM) IQR values from 1880-1979 37

FIGURE 3.16: Maps of Spring (MAM) IQR values from 1980-2020 38

FIGURE 3.17: Maps of Summer (JJA) IQR values from 1880-1979 39

FIGURE 3.18: Maps of Summer (JJA) IQR values from 1980-2020 40

FIGURE 3.19: Seasonal trend in average IQR values from 1880-2020 41

FIGURE 3.20: Maps of Summer (JJA) -2-sigma frequency departure events from 1880-1979 43

FIGURE 3.21: Maps of Summer (JJA) -2-sigma frequency departure events from 1980-2020 44

FIGURE 3.22: Seasonal trend in spatially averaged -2-sigma frequency departure events from 1880-2020 45

FIGURE 3.23: Maps of Fall (SON) 2-sigma frequency departure events from 1880-1979 46

FIGURE 3.24: Maps of Fall (SON) 2-sigma frequency departure events from 1980-2020 47

FIGURE 3.25: Maps of Winter (DJF) 2-sigma frequency departure events from 1880-1979 48

FIGURE 3.26: Maps of Winter (DJF) 2-sigma frequency departure events from 1980-2020 49

FIGURE 3.27: Maps of Spring (MAM) 2-sigma frequency departure events from 1880-1979 50

FIGURE 3.28: Maps of Spring (MAM) 2-sigma frequency departure events from 1980-2020 51

FIGURE 3.29: Maps of Summer (JJA) 2-sigma frequency departure events from 1880-1979 52

FIGURE 3.30: Maps of Summer (JJA) 2-sigma frequency departure events from 1980-2020 53

FIGURE 3.31: Seasonal trend in spatially averaged 2-sigma frequency departure events from 1880-2020 54

FIGURE 3.32: Maps of Fall (SON) 3-sigma frequency departure events from 1880-1979 56

FIGURE 3.33: Maps of Fall (SON) 3-sigma frequency departure events from 1980-2020 57

FIGURE 3.34: Maps of Winter (DJF) 3-sigma frequency departure events from 1880-1979 58

FIGURE 3.35: Maps of Winter (DJF) 3-sigma frequency departure events from 1980-2020 59

FIGURE 3.36: Maps of Spring (MAM) 3-sigma frequency departure events from 1880-1979 60

FIGURE 3.37: Maps of Spring (MAM) 3-sigma frequency departure events from 1980-2020 61

FIGURE 3.38: Maps of Summer (JJA) 3-sigma frequency departure events from 1880-1979 62

FIGURE 3.39: Maps of Summer (JJA) 3-sigma frequency departure events from 1980-2020 63

FIGURE 3.40: Seasonal trend in spatially averaged 3-sigma frequency departure events from 1880-2020 64

LIST OF ABBREVIATIONS

DJF	December-January-February
ENSO	El Niño-Southern Oscillation
IQR	Interquartile Range
JJA	June-July-August
KS	Kolmogrov-Smirnov
MAM	March-April-May
SON	September-October-November

CHAPTER 1: INTRODUCTION

The temperatures a particular area (such as the globe or a country) experience over many decades tend towards a Gaussian (or normal or bell curve) probability distribution, such that the mean value is the center of the distribution, and the range of values experienced can be quantified by the standard deviation around the mean value. This is how a temperature climatology is typically defined. The statistical expectation is that a particular area will experience a temperature that is within one standard deviation of the mean value about two-thirds of the time. For example, if the average temperature in July is 75 F, and the standard deviation is 5 F, then the statistical expectation is that you would have a 68% chance of experiencing a temperature between 70 F and 80 F.

Figure 1.1 shows distributions of temperature, temperature anomalies, and standardized temperature anomalies for all summer months (JJA) averaged together from 1951-1980 for the southeastern United States "Climate Region" (Virginia, North Carolina, South Carolina, Georgia, Alabama, Florida; NOAA NCEI, accessed March 2023), which could be said to represent a summertime climatology for the region. Although each distribution looks similar, they tell different stories. Looking at the temperature distribution, the raw temperature data for the region is distributed across a range from 76-81 degrees F. The temperature anomaly distribution shows the departure from the average, where the average for 1951-1980 was subtracted from the raw data, creating a distribution that has the anomaly average centered on 0. Temperature anomalies are in units of temperature, such that a positive value is warmer than the average temperature (warm anomaly) and a negative value is cooler than average (cold anomaly). Lastly, the distribution of standardized anomalies of temperature is calculated as the temperature anomaly divided by the standard deviation of the temperatures from 1951-1980, which compares the anomaly to the observed variability in the temperature. A standardized anomaly is a unitless metric but is often stated to be in units of standard deviation. For example, the standardized temperature anomaly quantifies how many standard deviations an observed temperature is from the mean temperature, which allows for a more complete way to quantify how statistically unusual a warm or cold anomaly is. Using a standardized anomaly distribution helps understand a region's climatology in that it reflects the magnitude of each temperature anomaly relative to local variability, removing the influence of geographic location (such as elevation, or proximity to large bodies of water) on both the mean and standard deviation of temperatures. This allows for comparisons of regions that have different mean temperature and/or different temperature variability.

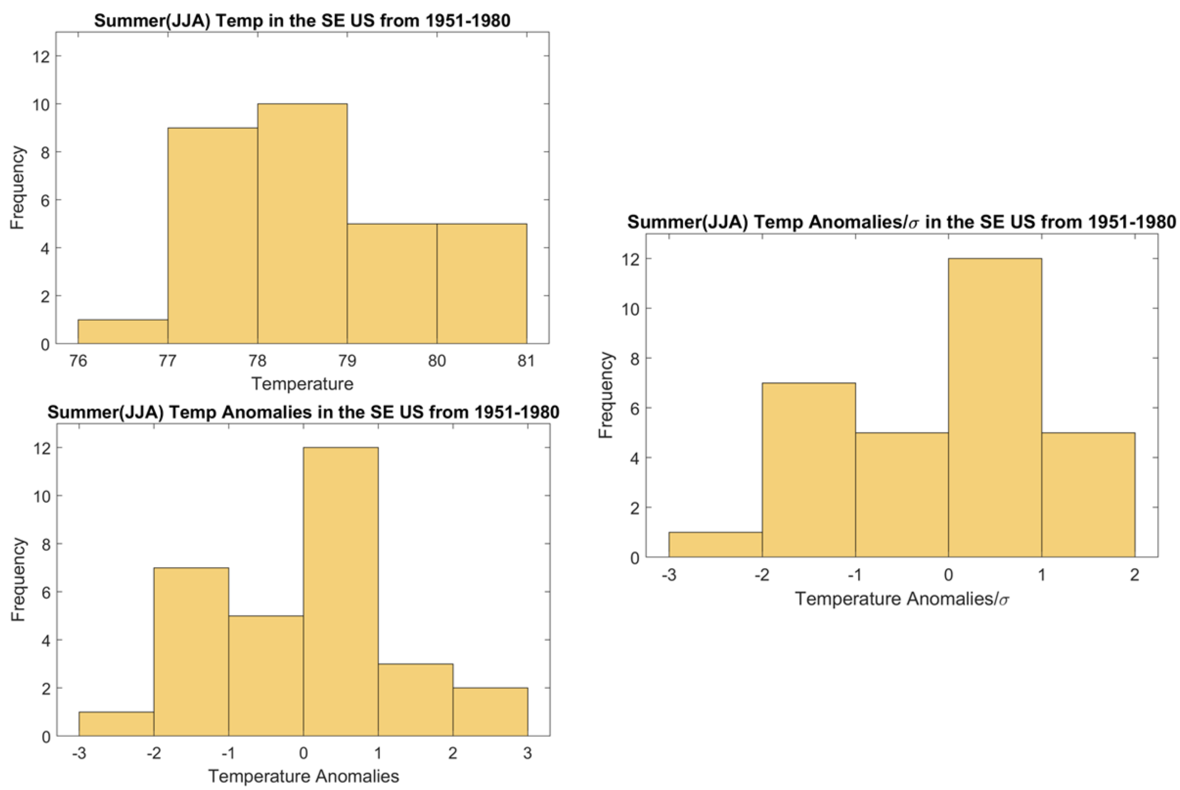

Figure 1.1. Summer (JJA) climatology in the SE United States "Climate Region" from 1951-1980 (i.e. N = 30) shown by (a) temperature (degrees Celsius) distribution, (b) temperature anomaly (degrees Celsius) distribution, and (c) temperature anomaly/sigma distribution. Data retrieved from NOAA NCEI database (accessed March 2023).

Figure 1.2 shows a Gaussian distribution, with 6 bands representing the probability of experiencing a value within different ranges of equal increments of standard deviations from the mean (Bland et al., 1996), where a standard deviation is denoted by the Greek letter “sigma”. Thus, the chance of a value randomly selected from a perfect Gaussian distribution of anomalies being greater than 0 would be 50%. Similarly, the probability of a randomly selected value being within 1 sigma (1 standard deviation) of the mean (i.e., anomaly of 0) is about 68.2%, while the probability for a randomly selected value being greater than +1 standard deviations, or a so-called “1-sigma” event, would be 15.8%. By extension, 2-sigma and 3-sigma events within a given Gaussian distribution are exceedingly rare occurrences, happening only 2.2% and 0.1% of the time, respectively (Figure 1.2). If the probability is inverted, then a 2-sigma temperature event occurs about once every 50 years, while a 3-sigma event occurs about once every 740 years.

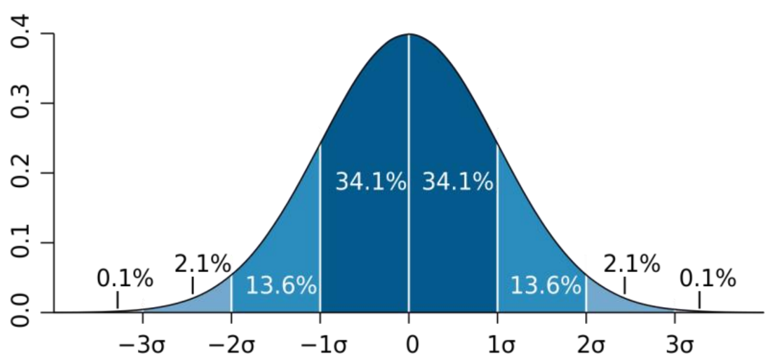

Figure 1.2. A plot of normal distribution (or bell-shaped curve) where each band has a width of 1 standard deviation. Taken from Wikipedia (Accessed March 2023).

With global warming, the climatology of regions across the globe is changing over time. Figure 1.3 shows the average global surface temperature from 1880 to 2020 as calculated from NOAAv5, and illustrates how mean surface temperature changes within different decadal time periods. From 1880 to 2010, the global average temperature anomaly steadily increased throughout the decades, where the global average temperature anomaly increased from -0.23 to 0.82 degrees Celsius from 1880 to 2010. Around 1980, the global average temperature ramps up, increasing from 0.24 to 0.82 degrees Celsius in the last 40 years, suggesting that the climatology of any given region has also shifted more dramatically in the last 40 years.

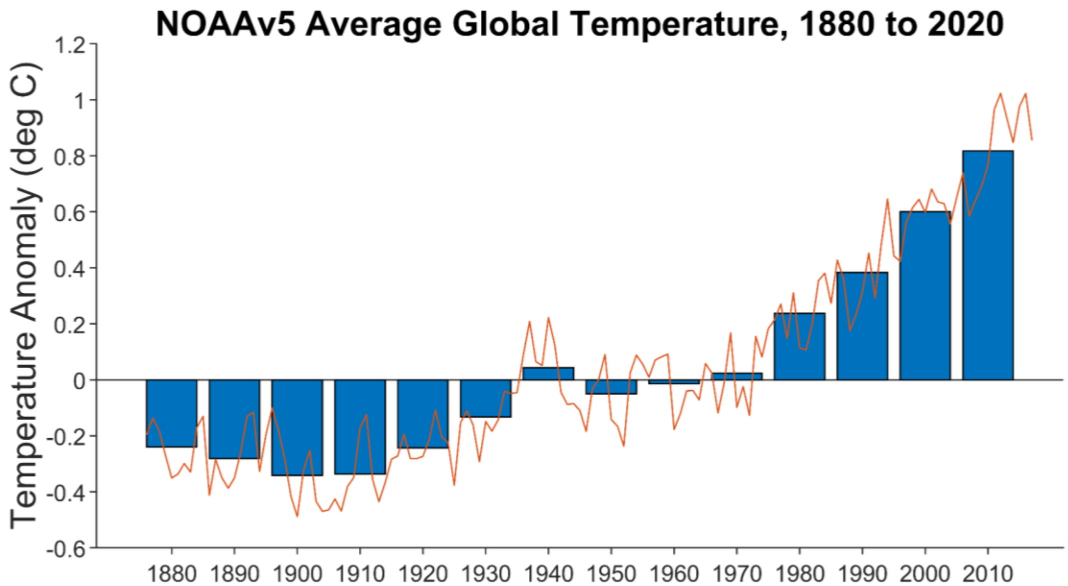

Figure 1.3 NOAAv5 Global average temperature anomalies (deg C) from 1880 to 2020.

In addition to temperature increase, the impacts of global warming are growing more apparent at continental (Gulev et al., 2021), regional (Jay et al., 2018), and even sub-regional and state scales (Kunkel et al., 2020), but it is often hard to grasp how a global scale problem translates to local scales (Howe et al., 2015; Mumenthaler, 2021; Gärtner et al., 2021). Yet, regional and local impacts of global warming are important for a number of societally-relevant reasons that include disruptions to physical, social, and economic structures in place within communities, where an increase in intensity and frequency of adverse weather events puts the health and safety of individuals at risk (USGCRP, 2018).

Although there is a well-documented shift in weather events today due to global warming (Sisco et al., 2017; Cohen et al., 2018; McBean, 2004; Painter et al., 2020), the changes so far have not always been obviously detrimental. Egan and Mullin (2016) calculated a "Weather Preference Index" using population and demographic models to show that from 1974 to 2013, shifts in where Americans lived largely coincided with movement to areas with "better" weather, such as warmer, milder winters (Egan and Mullin, 2016). Changes in the average weather are not always an immediately negative experience and can, for example, even improve individual experiences related to weather, at least in the short term. The research using the Weather Preference Index (Egan, 2016) points out, however, that projections of warming into the near and medium-term future implies that 88% of Americans will experience weather that is less favorable by the end of the century.

Long-term climate trends are more difficult to keep track of than the more immediate fluctuations in daily and seasonal temperatures, but the frequency of heat waves, or hot extremes, seem to hold more weight in the public memory (e.g., Perkins-Kirkpatrick and Lewis, 2020; Robinson et al., 2021; Simolo and Corti, 2022; Zhang et al., 2022). Global warming alters regional climatology which, in turn, increases the frequency of hot extremes.

Hansen et al. (2012) highlights changes in distributions of local seasonal temperatures at global and hemispheric scales from 1900-2020 using a "loaded dice" analogy to describe shifts in the distributions of local seasonal standardized temperature anomalies by comparing sides of a die to each represent a range of values in a Gaussian distribution that have equal chances of occurring. Specifically, Hansen et al. (2012) starts with the fact that two sides of the die have a 33% chance, and then divides the approximately Gaussian distribution of local seasonal standardized temperature anomalies (within the spatial domain in question) into sigma ranges that themselves each account for 33% of the occurrences. A seasonal standardized temperature anomaly less than -0.43 sigma has a 33% chance of occurring, as do values between -0.43 and +0.43 sigma, and values greater than +0.43 sigma. Correspondingly, Hansen et al. (2012) labels seasonal standardized temperature anomalies less than -0.43 sigma as two blue faces of a die ("colder than average"), values between -0.43 and +0.43 sigma as white faces of the die ("about average"), and values greater than +0.43 sigma as red ("warmer than average"). The "loaded dice" analogy then describes how the dice are essentially gaining red faces at the cost of the blue ones due to the increasing mean temperature that shifts the location of the seasonal climatology (Hansen et al., 1988).

Figure 1.4 shows the distributions of local standardized temperature anomalies in the Northern hemisphere in winter (DJF, averaged together such that there is 1 data point per year) from 1951-1980, 1990-2000, and 2010-2020 from Hansen et al. (2012). The 1951-1980 distribution is shaded blue, white, and red to correspond to a roll of the dice with approximately equal odds of landing in each shaded area, but in the 1990-2000 and 2010-2020 time periods, the odds are weighted to the red area and away from the blue area. Thus, the dice are loaded and there is a far greater chance of rolling red now compared to the relatively recent past.

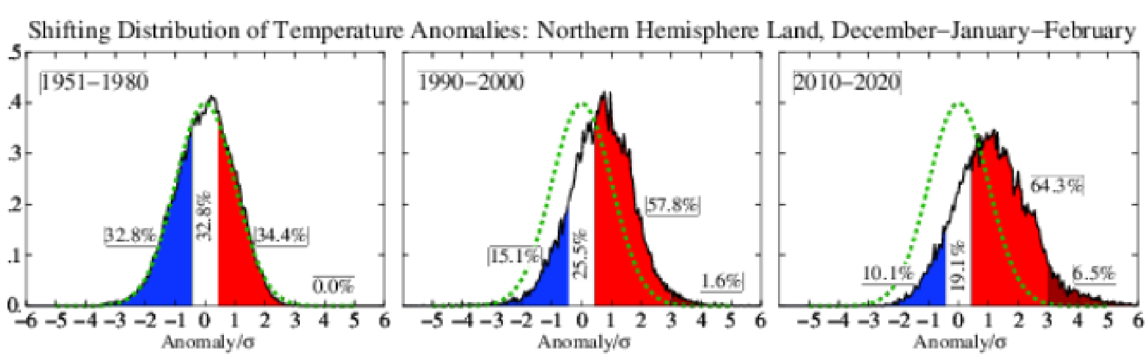

Figure 1.4. Series of plots depict shifting distributions of local temperature anomalies in the Northern Hemisphere (Land) in winter (DJF, averaged together such that there is 1 data point per year) from 1951-1980 to 2010-2020. Plots provided from Hansen et al. (2012).

Results from Hansen et al. (2012) also show that the frequency of large, positive standardized temperature anomalies (high sigma events) are increasing over time. From 1951-1980 to present day, the occurrence of 3-sigma temperature anomalies events has increased by a factor of 10 (Figure 1.5). All things being equal, the chance of experiencing a 3-sigma event is 0.1%, but Hansen et al. (2012) showed that the category of “extremely hot” 3-sigma summers increased from about 33% to about 75% between 1900 to 2020. This shift creates a new 6-sided die, with four sides of the die being red or “hot”, increasing the likelihood of a season that is much warmer than average on any “throw” of the dice (i.e., in any given year).

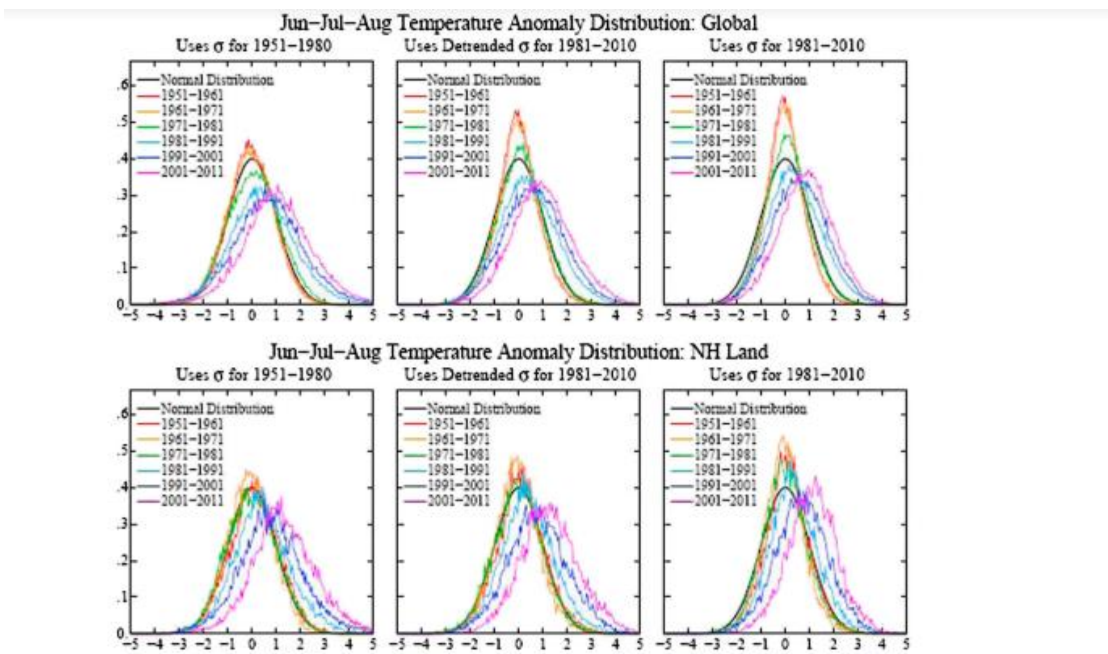

Figure 1.5. Frequency of occurrence (y axis) of local temperature anomalies (relative to 1951–1980 mean) divided by local standard deviation (x axis) obtained by counting grid boxes with anomalies in each 0.05 interval. Area under each curve is unity.

Hot extremes are memorable, and I will explore changes in seasonal climatologies using a methodology similar to Hansen et al. (2012). I aim to quantify how the frequency of large sigma events in distributions of monthly local standardized temperature anomalies have changed between 1880 and 2020. I compile the statistics related to those distributions for the months within each season of the year (i.e., up to three data points per year) at the spatial resolution of the input dataset. I use a gridded temperature dataset that extends from the late 1800s to the present day, and the methodology I describe could be applied to any gridded temperature dataset.

CHAPTER 2: METHODOLOGY

2.1: Dataset

I use the NOAAv5 Global Land and Ocean monthly temperature anomaly dataset (Huang et al., 2020; Zhang et al., 2019) to assess seasonal temperature trends and extremes. The dataset is constructed from “sea-surface temperature anomalies from the Extended Reconstructed SST version 5 (ERSSTv5) and land surface air temperature (LSAT) from the Global Historical Climatology Network monthly version 4 (GHCNm v4)” (Huang et al., 2020). The dataset provides temperature at a spatial resolution of 5 x 5 degrees from 1880 to present day, calculated as anomalies relative to the base period of 1971-2000. The goal in the creation of this dataset was to increase land and ocean temperature coverage relative to the previous version (NOAAv4), improve historical changes, and to provide an enhanced view of climate trends. Figure 2.1 shows temperature trends for winter months from 1988 to 2017 for land areas, highlighting an example of the differences between the NOAAv4 and NOAAv5 Global Land and Ocean Temperature Anomaly datasets in terms of spatial coverage.

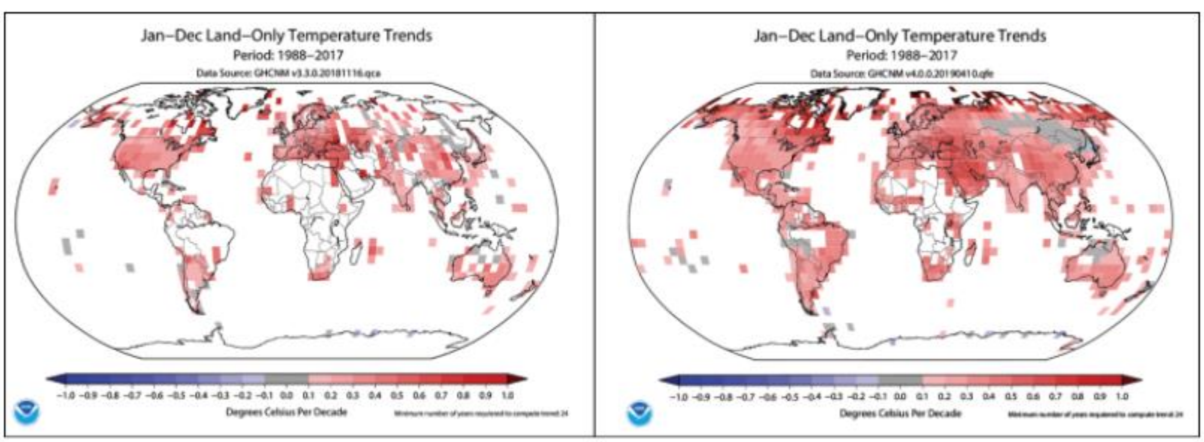

Figure 2.1. Surface air temperature trends in winter months over land areas from 1988 to 2017 for NOAAv4 (left) and NOAAv5 (right).

Figure 2.2 shows the overall evolution in spatial coverage for NOAAv5 from 1880 to 2020 in January. Throughout time, the coverage for both land and ocean improve due to an increase in temperature data availability (Huang et al., 2020). By 1930, large pockets in the Atlantic Ocean, Africa, Europe, and polar regions had become resolved. In contrast, uncertainty is more pronounced from 1880-1910 due to a lack of direct observations (Huang et al., 2020).

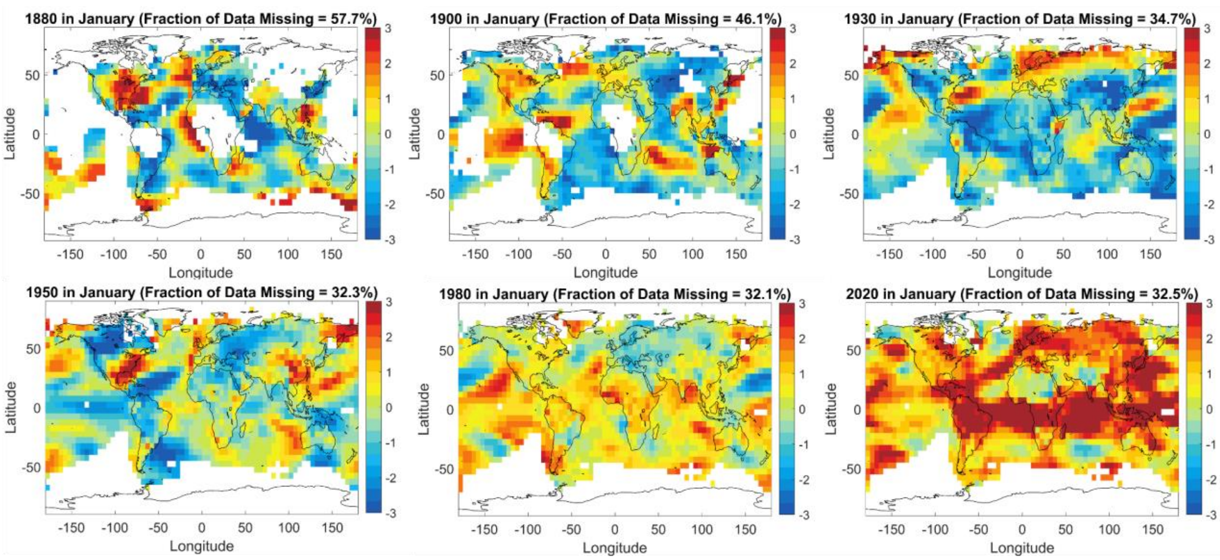

Figure 2.2. Global maps of monthly standardized temperature anomalies from 1880 to 2020. In addition to the anomalies, the maps also depict change in ocean and land data coverage over the time period, where white grid boxes indicate that not enough data was available for a monthly average.

2.2: Data Processing

The NOAAv5 dataset uses a 1971-2000 climatology period to calculate temperature anomalies. Table 2.1 expresses the global average surface temperature from 1880 to 2020 as calculated from NOAAv5, and lists the values displayed in Figure 1.3.

Table 2.1. NOAAv5 Global average temperature anomalies (degrees Celsius) from 1880 to 2020 using a base period of 1951-1980.

Time Period	**Global Average Temperature Anomaly (Decadal) in degrees Celsius**
1880-1889	-0.24
1890-1899	-0.28
1900-1909	-0.34
1910-1919	-0.34
1920-1929	-0.24
1930-1939	-0.13
1940-1949	0.04
1950-1959	-0.05
1960-1969	-0.01
1970-1979	0.02
1980-1989	0.24
1990-1999	0.38
2000-2009	0.60
2010-2019	0.82

Looking at the progression of global average temperatures, the change that occurs from the 1880s through the 1970s is an overall increase of 0.26 deg Celsius. From the 1970s to the 1980s, however, the global average temperature anomaly increased from 0.02 to 0.24 degrees Celsius, roughly the same as the increase from the 1880s to the 1970s. From the 1980s to nearly the present, the decadally-averaged anomaly increased from 0.24 to 0.82 degrees Celsius. Because of this feature, I changed the climatology period for the NOAAv5 data from 1971-2000 to 1951-1980, which is a period with minimal temperature change (Hansen et al., 2012) and that is consistent with what other global temperature datasets use (NASA GISS, Berkeley Earth, for example). The NOAA choice for the base period of 1971-2000 is useful since the time period was more recent (e.g. Hansen et al., 2012), however, the global average temperature anomaly changes by roughly 0.64 degrees Celsius during that time. This large change in mean temperature using a climatology period of 1971-2000 reduced its usefulness in my research because the non-stationary mean during those 30 years produces climatologies that are not Gaussian in shape, which affects the calculations used when determining the expected occurrence of a given sigma event.

To change the climatology period, I calculated the average monthly temperature anomaly for 1951-1980 for each grid cell, then subtracted that value from the monthly anomaly for all years in the corresponding grid cell (Huang, 2020; personal communication). This created temperature anomalies where each grid cell has a mean value of zero and standard deviation of 1 for the 1951-1980 period. To convert the anomalies to another base period would require the same process be applied to the NOAAv5 dataset using a different range of years.

Using the 1951-1980 NOAAv5 monthly temperature anomalies, I calculated monthly standardized temperature anomalies using monthly standard deviations from the 1951-1980 reference period in each grid cell. We used the following equation:

$$(1)\quad z = \frac{x - \underline{x}}{\sigma_x} = \frac{x'}{\sigma_x}$$

where z is the monthly standardized temperature anomaly, x is temperature, $\underline{x}$ is mean temperature for the reference period, σ_x is the monthly standard deviation over the reference period, and x' is the temperature anomaly, all at the monthly time scale. The calculation is done for each grid box of the NOAAv5 data, providing gridded monthly standardized temperature anomalies. The anomaly itself accounts for the local mean, while standardized anomalies account for both the local mean and the local variance (Wilks, 2011), noting that the Hansen et al. (2012) loaded dice analogy is also based on standardized temperature anomalies.

As an initial test of the methods, I created histograms of frequency distributions of monthly standardized temperature anomalies for different 3-month seasons in spatial domains in the USA. Figure 2.3 shows an example for Fall months (September, October, November) of 1880-1909 and 2000-2020 for a subset of the USA (red box in Figure 2.3), where the histogram shows local monthly standardized temperature anomalies from each of the 3 different months for each grid cell, noting that in this preliminary exploration, the base period was 1971-2000.

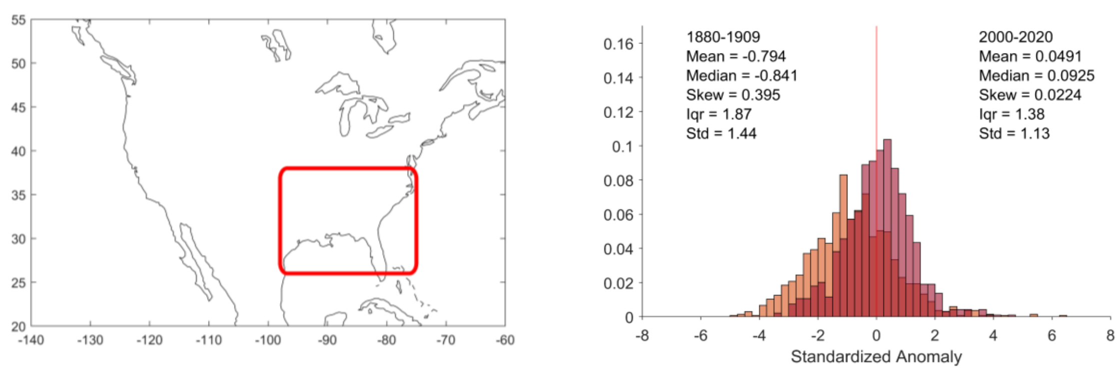

Figure 2.3. Histogram shows local monthly standardized anomaly frequency distributions from 1880-1909 and 2000-2020 during Fall (SON) in a subset of the United States (red box). Map highlights the region being assessed.

In this case, there is a positive shift in the mean value from -0.79 to +0.05, and an increase in the number of grid cell months with 2-sigma warm events (i.e., months with standardized anomaly greater than +2). I produced results for all seasons and four USA domains, and the same general finding was true in all cases: the distributions were generally Gaussian and the mean for every season shifted to larger sigma values. This initial analysis helped me confirm that the broader global and hemispheric scale results of Hansen et al. (2012) applied in smaller spatial domains. In other words, the climate dice are now loaded at regional scales.

I applied the same methods to every grid box on a global scale, working to quantify how the distributions of monthly standardized temperature anomalies were changing across time (from 1880 to 2020) for each season. The main metrics for quantifying this change were calculations of the location and spread of the distribution via the mean, standard deviation, and interquartile range, all calculated per grid box for monthly data and then aggregated using 3-month seasons in either 10- or 30-year time periods. In my analysis, I made calculations for the following time periods: 1880-1909, 1910-1939, 1940-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2009, 2010-2020, and for seasons defined as DJF, MAM, JJA, and SON. Note that the 2010-2020 time period is 11 years, but I account for that and the different time period lengths in general as described next.

I focused my calculations on quantifying how the extremes of the distributions were changing with global warming. Specifically, I calculated the frequency of -2 sigma events, +2 sigma events, and +3 sigma events, as well as large-scale grid-area weighted averages of the frequency changes for ocean and land grid boxes. I calculated the magnitude of the different sigma events (from the distributions of local monthly standardized anomalies) relative to the expected frequency for each sigma event.

To calculate the frequency of -2, +2 and +3 sigma events, I compiled all the monthly standardized anomalies that fit within the metric bounds, where each lat/lon grid cell contained a single number to characterize how many of those sigma events occurred within that season and time frame. I normalized the frequency by the variable number of years within each analysis period. For example, 2010-2020 is a 11-year period, while other analysis periods are 10 years, and others are 30 years. The choice of a 10-year time period was decided after we tested for how Gaussian each distribution was. Prior to deciding on 10-year chunks, I explored the possibility of looking at 30-year trends, however, the mean value of distributions changed on time scales less than 30 years. The distributions are more Gaussian in 10-year time frames over the last 40 years (1980 to 2020), during which the mean temperature anomaly increased from 0.24 to 0.82 degrees Celsius (Table 2.1). With this in mind, I looked at the last 40 years as decadal time steps to minimize how much the mean temperature itself was changing. The tradeoff with using a shorter analysis period is that the amount of possible data decreases, which reduces the statistical stability of the calculations. However, I describe a statistical test below to evaluate the symmetry and skewness of the distributions of monthly standardized temperature anomalies to assess whether the distributions remain similar to a traditional bell-curve shaped climatology.

Another processing step was that I had to account for gaps in the data within any analysis period. Some time periods, specifically during earlier years in the record, and some locations (such as the open ocean) simply did not have data points for some of the years in the analysis period. In order to account for this, I excluded any grid cells with less than 80% data coverage within the given time frame.

With those processing steps (variable time frame lengths, and variable data gaps) applied, I created global maps that plotted the frequency of different sigma events in each grid cell for each season and time frame, and displayed the frequencies as a ratio relative to the statistical expectation frequency for a perfect Gaussian centered on a zero mean with a standard deviation of 1. That final step varies as a function of the sigma level, such that 2-sigma events would be expected to occur (in a perfect Gaussian distribution) 2.2% of the time, and 3-sigma event frequency is 0.13% of the time (Figure 1.2). On the global maps, I included the spatially averaged frequencies relative to expectation (calculated as a grid area weighted average of ocean and land grid boxes) and the corresponding upper (95th) and lower (5th) percentiles of the ocean and land data.

I calculated the interquartile range (IQR) to assess how the spread of each distribution was changing over time using:

$$(2)\ IQR = Q_3 - Q_1$$

where IQR represents the interquartile range, Q_3 represents the third, or 75th, quartile of the distribution and Q_1 represents the first, or 25th, quartile of the distribution. I created global maps that display the IQR for each grid cell (in each season and time frame). Similar to the frequency maps, I calculated the spatially averaged IQR (land and ocean), and the corresponding upper (95th) and lower (5th) percentiles of the land and ocean IQR values as well. Lower values of the IQR imply a narrower distribution, while larger values imply a wider distribution. Thus, when evaluated over time, the IQR provides a way to see how the spread of the distribution is changing. I expected to see an increasing spread in the last several decades similar to what Hansen et al. (2012) found.

Since my calculations are implicitly dependent on the assumption of a Gaussian distribution for any time period and any season (i.e., a climatology), I assessed how Gaussian the distributions were using the non-parametric Kolmogorov–Smirnov (KS) test for each grid box. The KS test evaluates the null hypothesis that a given distribution is Gaussian by comparing the distribution determined from a given data set to values drawn from a perfect Gaussian distribution.

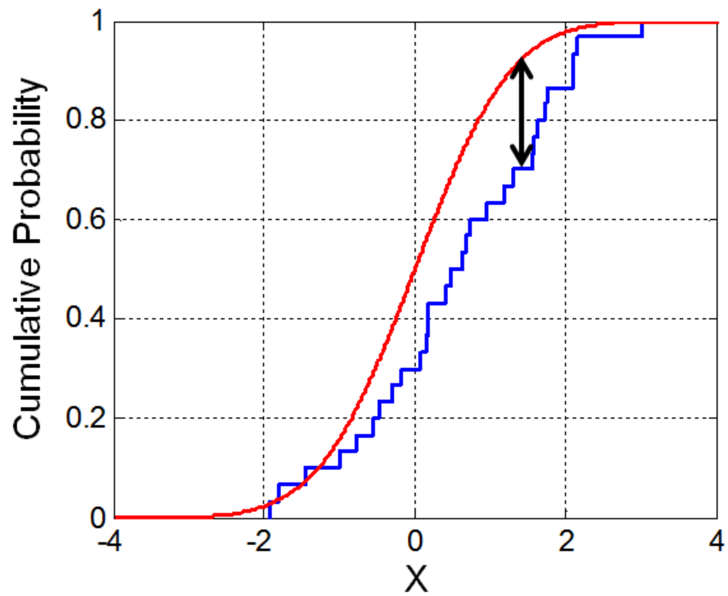

Figure 2.4. Example plot of a Kolmogorov–Smirnov (KS) test. The red line represents a Gaussian dataset, the blue line represents the sample data set and the black line represents the KS statistic (Figure source, Wikipedia: https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test accessed March 2023)

For example, Figure 2.4 shows the KS test on a sample dataset. The red line depicts a Gaussian distribution, the blue line is the sample dataset, and the black arrow reflects the KS statistic. The p-value of the KS test is used to determine if the null hypothesis that a dataset is Gaussian is accepted or rejected. A p-value of 0 indicates that the null hypothesis is rejected; thus, the dataset is not drawing from a Gaussian distribution. However, a p-value greater than 0 indicates that the null hypothesis could be accepted; for example, if the KS p-value is 0.10, then we could say there's a 10% chance that the data draws from a Gaussian distribution. For this project, I assumed a p-value greater than 0.50 was approximately Gaussian, but I did not exclude any results with KS p-values less than 0.50. I will show the KS p-value maps (and their corresponding ocean and land averages) in Chapter 3.

My main goal is to quantify the change in the frequency of extremes in the distributions. I assess these metrics across the entire globe, mapping the values at the spatial resolution of the dataset, which in my project is a 5x5 latitude by longitude spatial resolution of the NOAAv5 data. Each map quantifies how the distributions of monthly standardized anomalies of temperature are changing for each 3-month season across the globe and across the time range of the dataset (1880-2020, in my project).

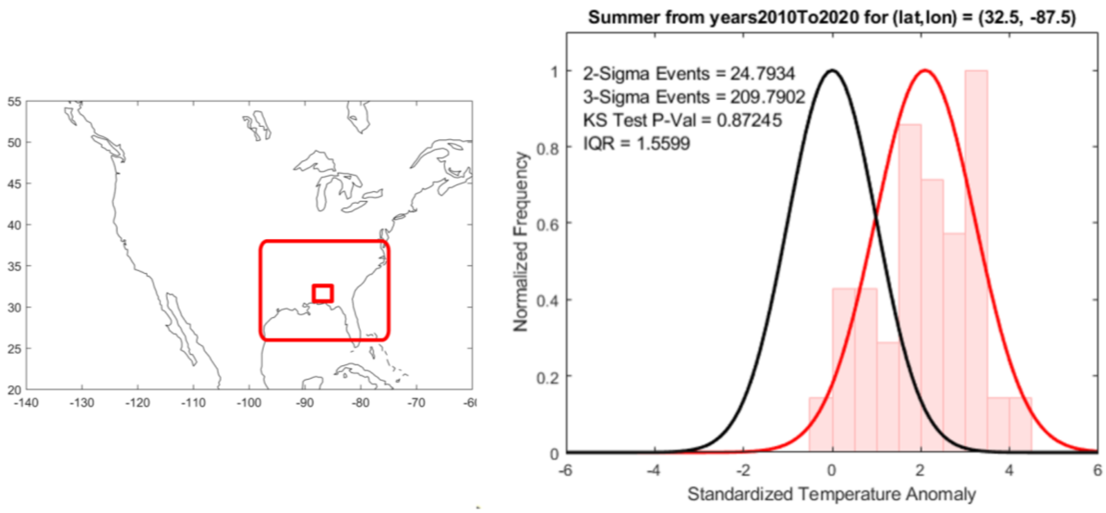

Figure 2.5. Normalized frequency of Summer (JJA) monthly standardized temperature anomalies at a single 5x5 NOAAv5 grid box (32.5, -87.5) denoted as a small red box within the Southeastern USA from the climatology period of 1951-1980 (black line) and 2010-2020 (red line).

As an example of the methodology outlined above, Figure 2.5 shows two distributions of monthly standardized temperature anomalies from NOAAv5 from one grid box during the Summer months; the first distribution (black) is from the climatology period, 1951-1980, while the second (red, red bars) is from 2010-2020. The left map shows the region and approximate grid cell in the southeastern United States that the distributions were calculated in for that specific season and time frame. The (red) distribution not only becomes more positively shifted, but there is an increase and appearance of large, positive sigmas above +2 and +3; thus, a change to the location of the distribution. From the climatology period of 1951-1980 to 2010-2020, the number of 2-sigma events occurred roughly 25 times greater than statistical expectation and 3-sigma events occurred roughly 210 times greater than statistical expectation. The KS test p-value of 0.88 suggests that there is an 88% chance that the data (red bars) draw from a Gaussian distribution, and the IQR of 1.6 suggests that the spread is slightly greater than the IQR of the 1951-1980 Gaussian curve of 1.34.

Figure 2.5 is an example of the analysis steps and corresponding results for a single grid cell, season (Summer), and time frame (2010-2020). My goal is to study how Summer 2010-2020 is changing across all time frames in this grid cell and every other grid cell in the NOAAv5 dataset, and then do this for other seasons as well. All the calculations are relative to the 1951-1980 base period (e.g., Hansen et al., 2012), which is a 30-year time frame with a semi-constant global mean temperature (Table 2.1). I study the changes via maps of the change in +2 and +3 sigma events, KS test p-values, and IQR values to better summarize the changes for every time frame and season. Then, I further summarize the map results in a time series of the changes in each metric to assess the local influence of global warming on different seasons.

CHAPTER 3: RESULTS

Mapping the distributions of monthly standardized temperature anomalies from 1880-2020 and for each season reveals patterns in temperature extremes across temporal and spatial scales. In order to fully understand the changes to the temperature extremes, it is necessary to check how Gaussian the distributions are so that the expected frequency of 2-sigma and 3-sigma events align with the expectation of a perfect Gaussian distribution, as well as evaluate the location and spread of the distributions. A Gaussian distribution signifies that the temperature data close to the mean would be experienced more frequently, where the extremes would be experienced less frequently. Noting back to Figure 2.5, temperature extremes happened less than 2.2% of the time in any month of a given season, however, my results help document the change in the frequency of extremes with global warming. This indicates that the mean temperature of the temperature distribution that represents our current climatology has increased, making it seem like what once was a temperature extreme, is now more common.

3.1: Kolmogorov–Smirnov Test of Distribution Normality

To objectively test how close to Gaussian the distributions of monthly standardized temperature anomalies are, we use the Kolmogorov–Smirnov (KS) test and calculate the p-value for each season and time frame for each grid cell (see Chapter 2.2). According to the KS test, a p-value less than 5% indicates that the null hypothesis can be rejected with greater than 95% confidence, but the p-values increasing beyond 5% imply that the data is more and more likely to draw from a Gaussian distribution. Thus, a KS p-value of 1 means there is a 100% chance that the data are from a Gaussian distribution. For this project, I consider that a distribution with a KS p-value greater than 50% is approximately Gaussian, such that the interpretation of 2-sigma and

3-sigma frequencies (Sections 3.3 and 3.4) in a multi-year climatology are well-grounded in the implicit assumption that a climatology behaves as a Gaussian distribution. Specifically, climatology is defined generally with the expectation that observed temperatures are warmer or cooler than average 50% of the time. Factors that can reduce the KS p-value include a non-stationary mean throughout the given time period (10 years or 30 years, in this project), as well as a positively or negatively skewed distribution.

Originally, we looked across 1880-2020 in approximately 30-year periods but found that looking at the data in 10-year periods improved how Gaussian the data was when plotted on a map. Referring to Figure 1.3 and Table 2.1, the globally averaged surface temperature increased at an increasing rate starting in about 1970-1980. The change in mean global temperature translated to smaller scales, and the distributions become less Gaussian (lower KS p-value). For example, Figure 3.1 shows the p-values during Summer from 2000-2020, where the top plots are broken into 10-year periods and the bottom plot shows the entire 2000-2020 time frame. The KS p-values averaged over ocean and land are 0.55 and 0.66 for 2000-2009 and 0.53 and 0.62 for 2010-2020, but decrease to 0.49 and 0.50 for 2000-2020. A longer time period in an era with rapidly changing mean temperature (Figure 1.3) generally results in distributions that are less Gaussian. Thus, I use 30-year time frames from 1880-1969 when the mean temperature changes less rapidly, and 10-year time frames from 1970-2020 when the mean changes more rapidly.

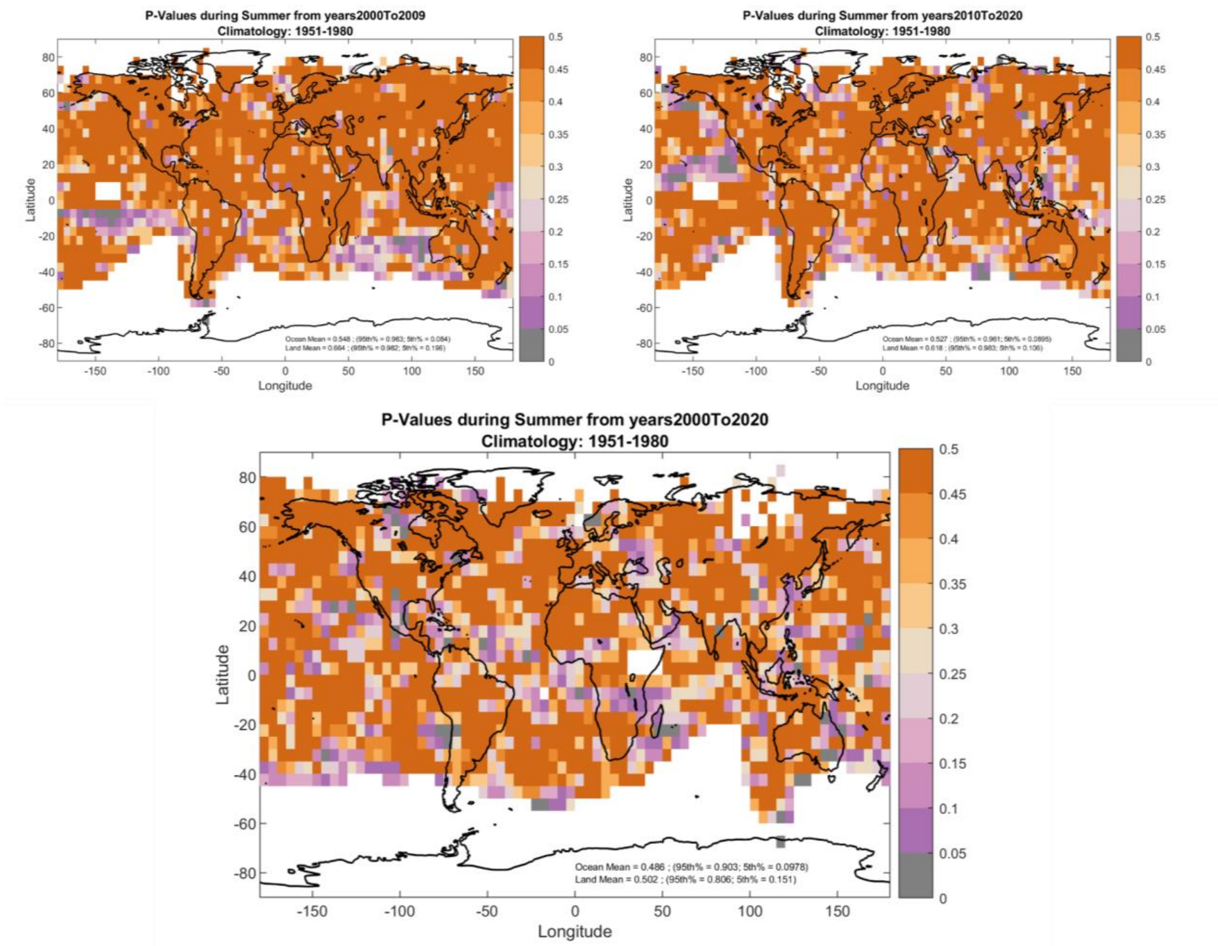

Figure 3.1. Maps depict summer (JJA) p-values from the KS test applied to distributions of monthly standardized temperature anomalies from (a) 2000-2009, (B) 2010-2020 and (C) 2000-2020, where the value is the likelihood the observations draw from a Gaussian distribution.

The following figures are compiled such that there are eight maps per season, and each of the eight maps show the different time periods: 1880-1909, 1910-1939, 1940-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2009, and 2010-2020. The maps in this section show the KS p-values for the distributions of monthly standardized temperature anomalies in each grid cell, and highlight when the p-values are less than 0.50 to emphasize where the Gaussian assumption weakens. Values greater than 0.50 are a single color, but I also show the large-scale spatial averages (and the range via the 5th and 95th percentiles) of the p-values for ocean and land to better convey the full range of the values on the map that are not shown by a color scale with a maximum of 0.50.

Figures 3.2 and 3.3 display Fall (SON) p-values from the KS test from 1880 to 2020, showing that the distributions remain mostly Gaussian ($p > 0.50$) throughout the Fall, where ocean and land grid-area weighted averages of the p-values range from 0.41 to 0.71. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.02 and 0.94 in 1880-1909 and 0.13 and 0.97 in 2010-2020. The value of the 5th percentile increases from 0.01-0.02 in 1880-1909 to 0.11-0.16 in 2010-2020, suggesting that the distributions are becoming more Gaussian on average. Regions in the Pacific, Atlantic and Indian ocean, as well as across Africa show clusters of low KS p-values, suggesting that those regions temperature distributions are not approximately Gaussian.

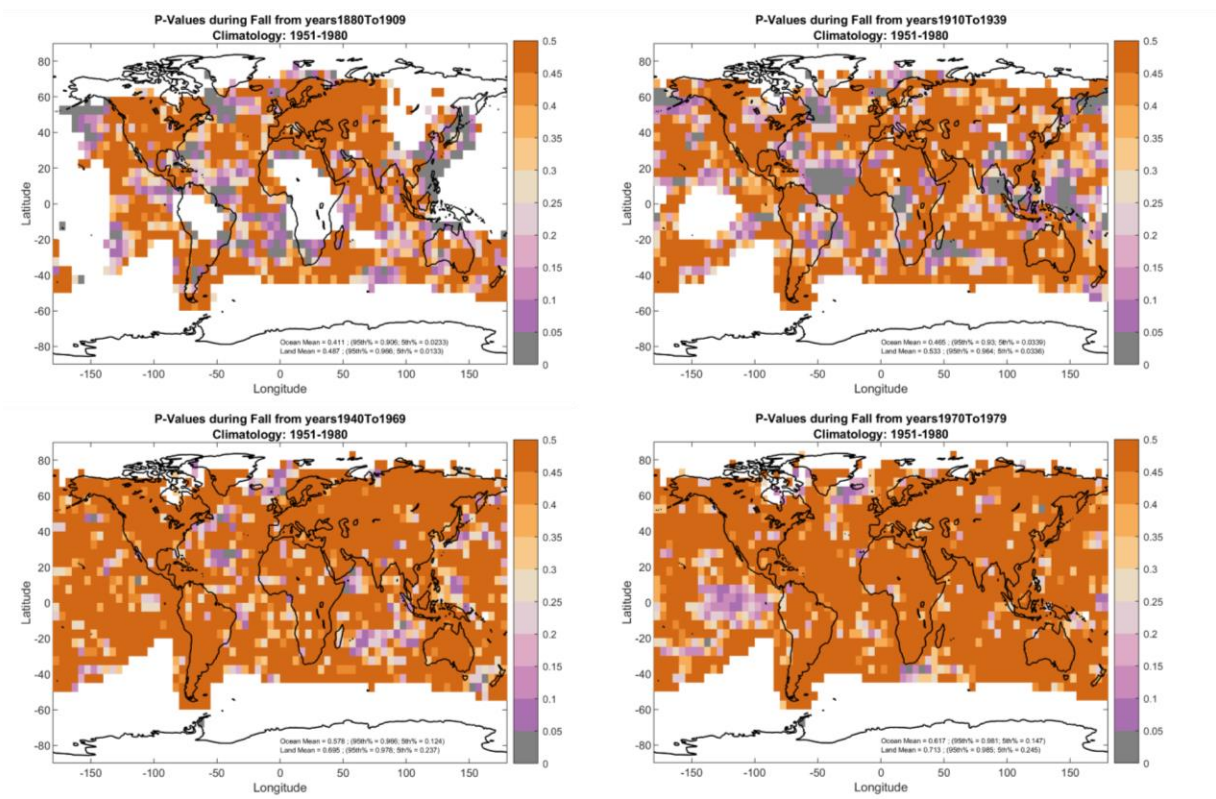

Figure 3.2. Maps depict Fall (SON) p-values from the KS test applied to distributions of monthly standardized temperature anomalies from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

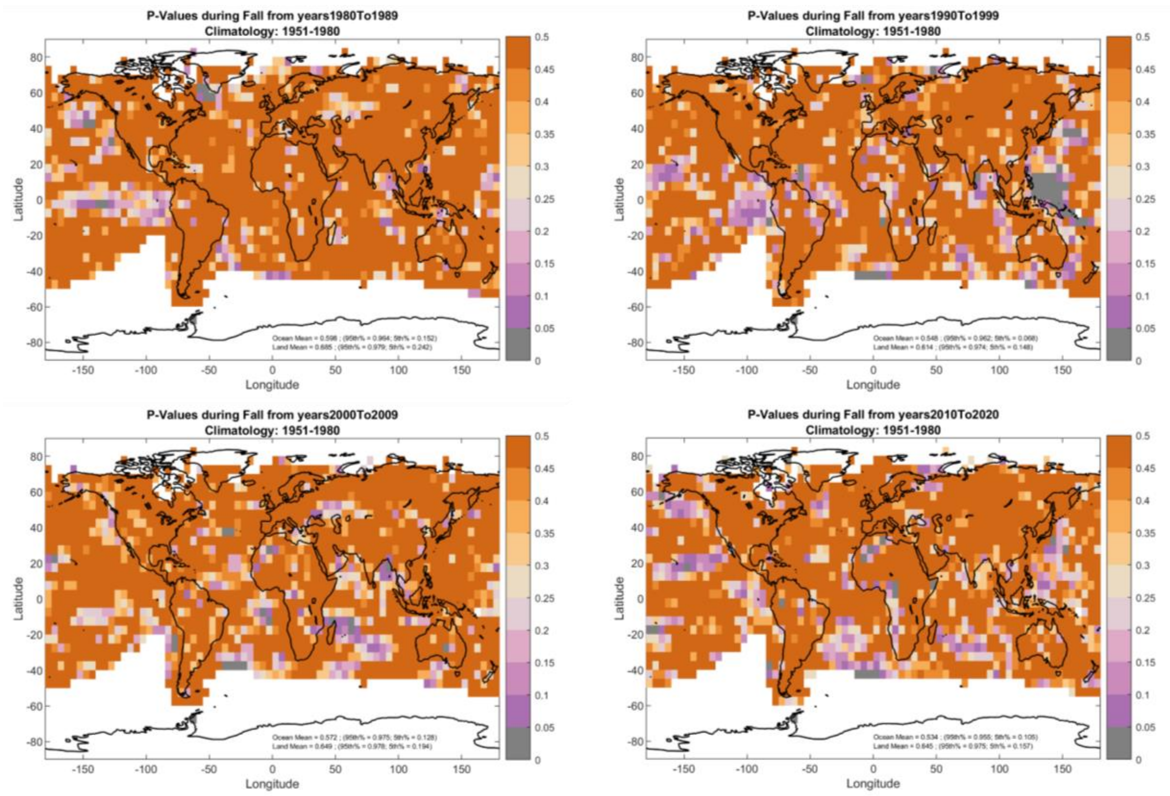

Figure 3.3. As per Figure 3.2, but for Fall (SON) p-values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.4 and 3.5 display Winter (DJF) p-values from the KS test from 1880 to 2020, showing that the distributions remain mostly Gaussian ($p > 0.50$) throughout the Winter, where ocean and land grid-area weighted averages of the p-values range from 0.49 to 0.67. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.02 and 0.96 in 1880-1909 and 0.14 and 0.97 in 2010-2020. Similar to the Fall, the value of the 5th percentile increases from 0.01-0.02 in 1880-1909 to 0.12-0.17 in 2010-2020, suggesting that the distributions are becoming more Gaussian on average. Similar to Fall, regions in the Pacific, Atlantic and Indian ocean, as well as across Africa show low KS p-values, suggesting that those regions temperature distributions are not approximately Gaussian.

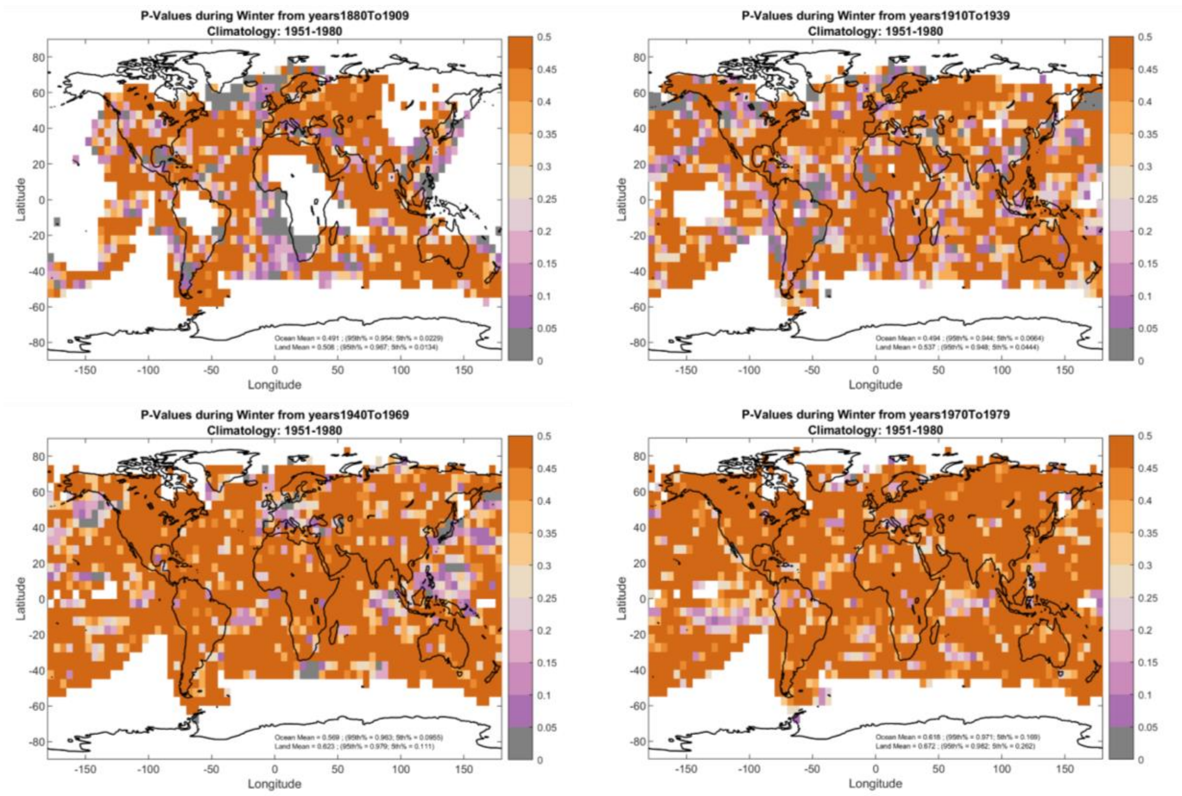

Figure 3.4. As per Figure 3.2, but for Winter (DJF) p-values from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

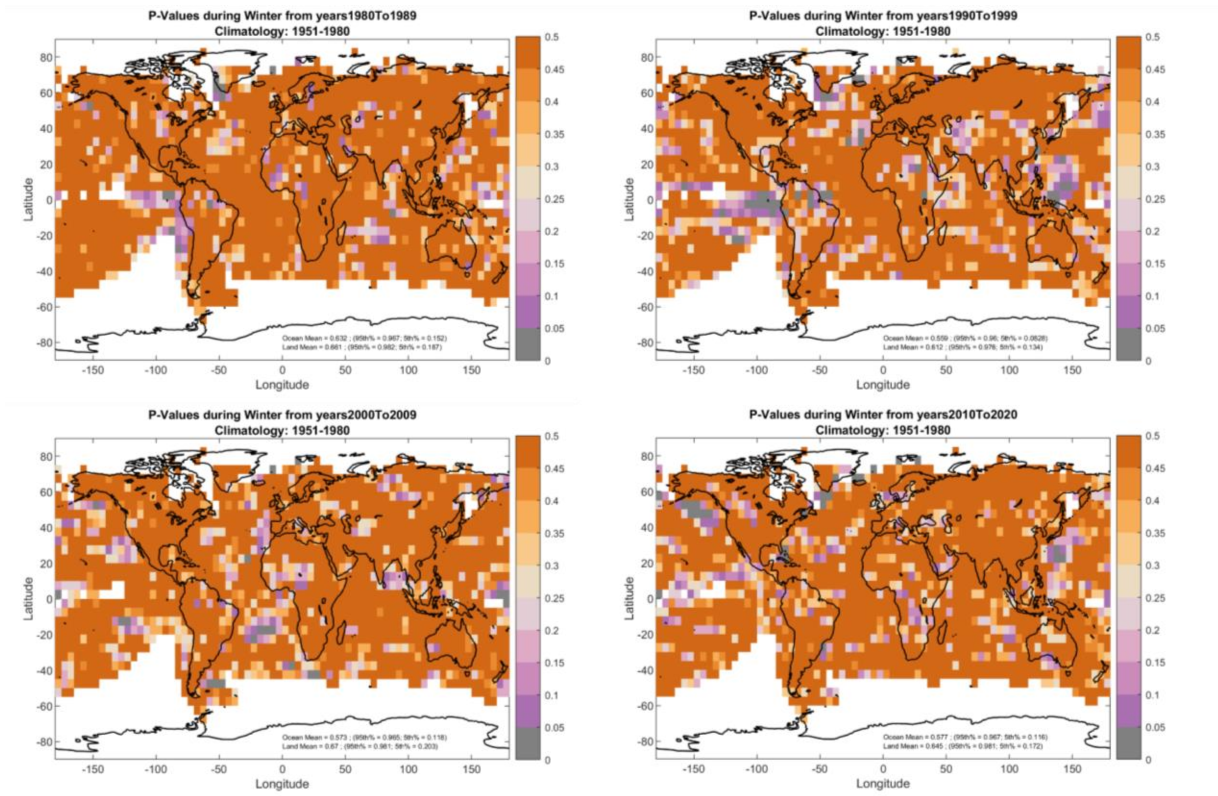

Figure 3.5. As per Figure 3.2, but for Winter (DJF) p-values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.6 and 3.7 display Spring (MAM) p-values from the KS test from 1880 to 2020, showing that the distributions remain mostly Gaussian ($p > 0.50$) throughout the Spring, where ocean and land grid-area weighted averages of the p-values range from 0.48 to 0.71. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.01 and 0.95 in 1880-1909 and 0.13 and 0.97 in 2010-2020. Similar to the Fall, the value of the 5th percentile increases from 0.01-0.02 in 1880-1909 to 0.10-0.20 in 2010-2020, suggesting that the distributions are becoming more Gaussian on average. Similar to Fall, regions in the Pacific, Atlantic and Indian ocean, as well as across Africa and Southeast Asia show low KS p-values, suggesting that those regions temperature distributions are not approximately Gaussian.

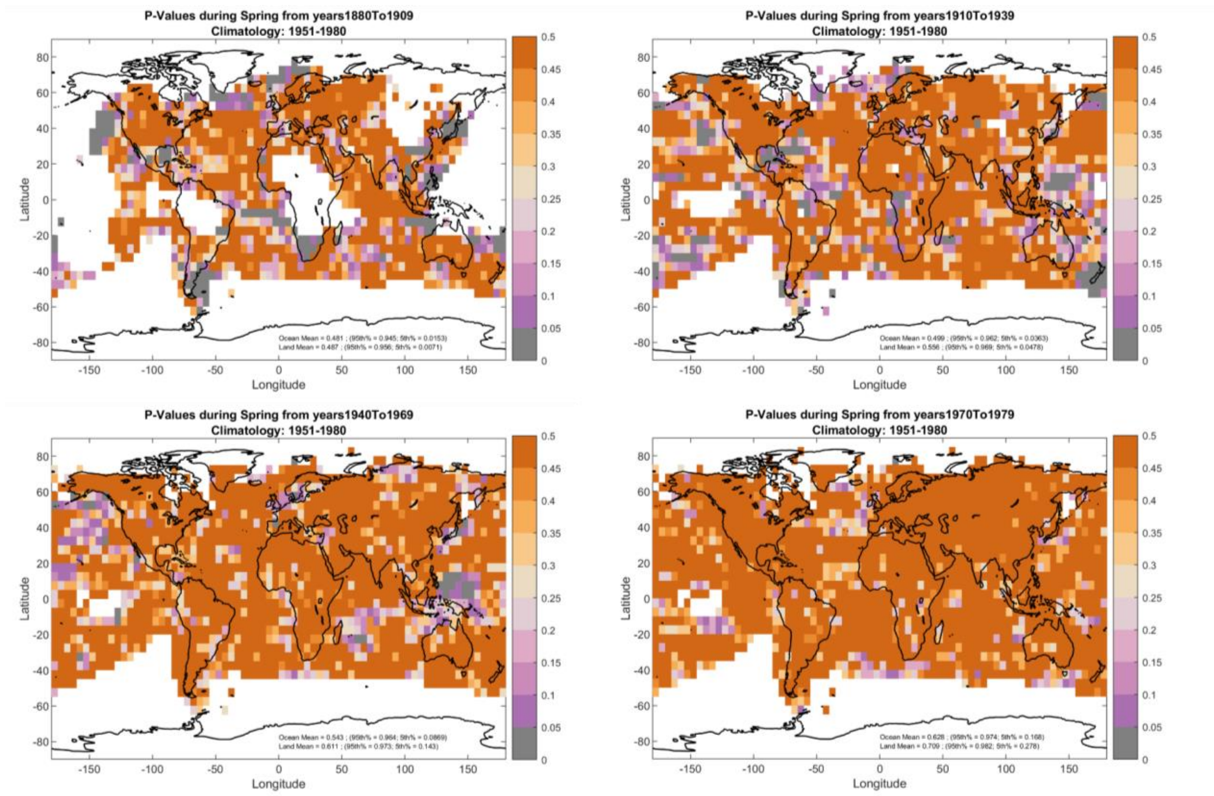

Figure 3.6. As per Figure 3.2, but for Spring (MAM) p-values from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

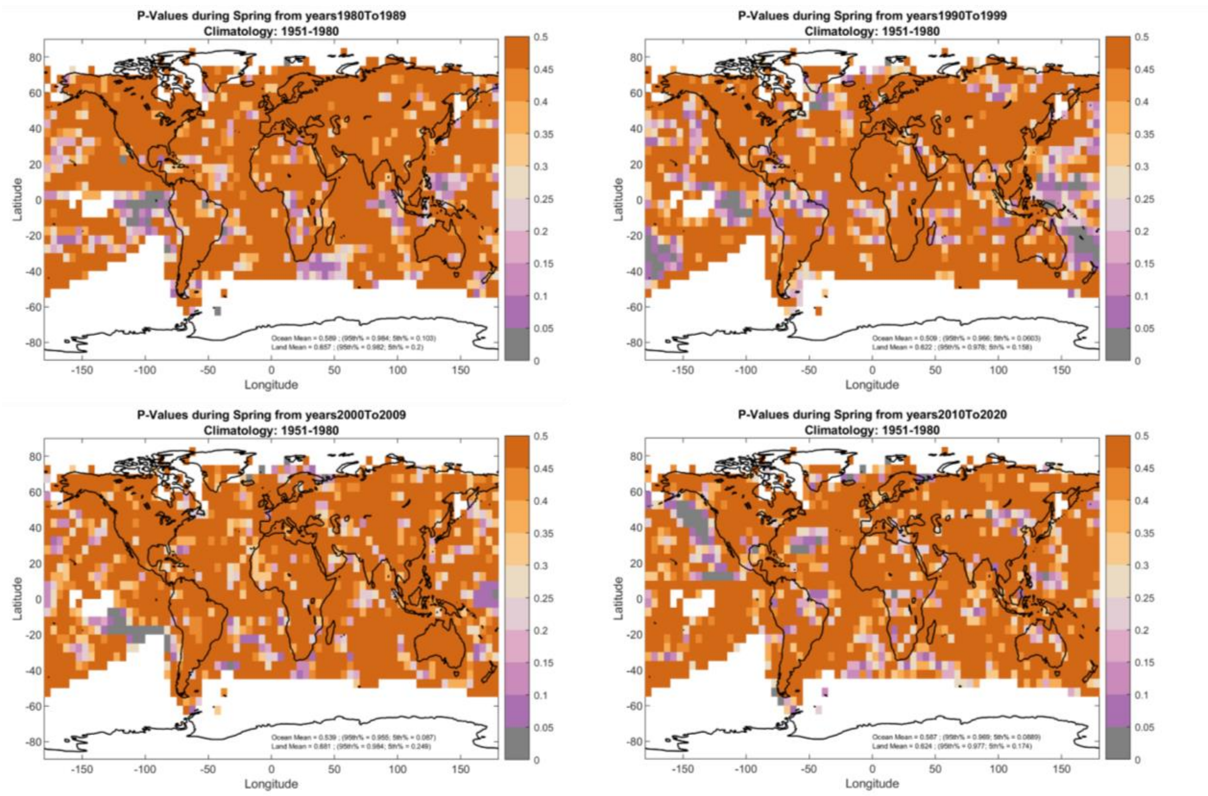

Figure 3.7. As per Figure 3.2, but for Spring (MAM) p-values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.8 and 3.9 display Summer (JJA) p-values from the KS test from 1880 to 2020, showing that the distributions remain mostly Gaussian ($p > 0.50$) throughout the Summer, where ocean and land grid-area weighted averages of the p-values range from 0.42 to 0.71. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.02 and 0.95 in 1880-1909 and 0.10 and 0.97 in 2010-2020. Similar to the Fall, the value of the 5th percentile increases from 0.01-0.02 in 1880-1909 to 0.10-0.11 in 2010-2020, suggesting that the distributions are becoming more Gaussian on average. Similar to Fall, regions in the Pacific, Atlantic and Indian ocean, as well as across northern Africa and eastern/central Europe, northern/southern South America and southern North America show low KS p-values, suggesting that those regions temperature distributions are not approximately Gaussian.

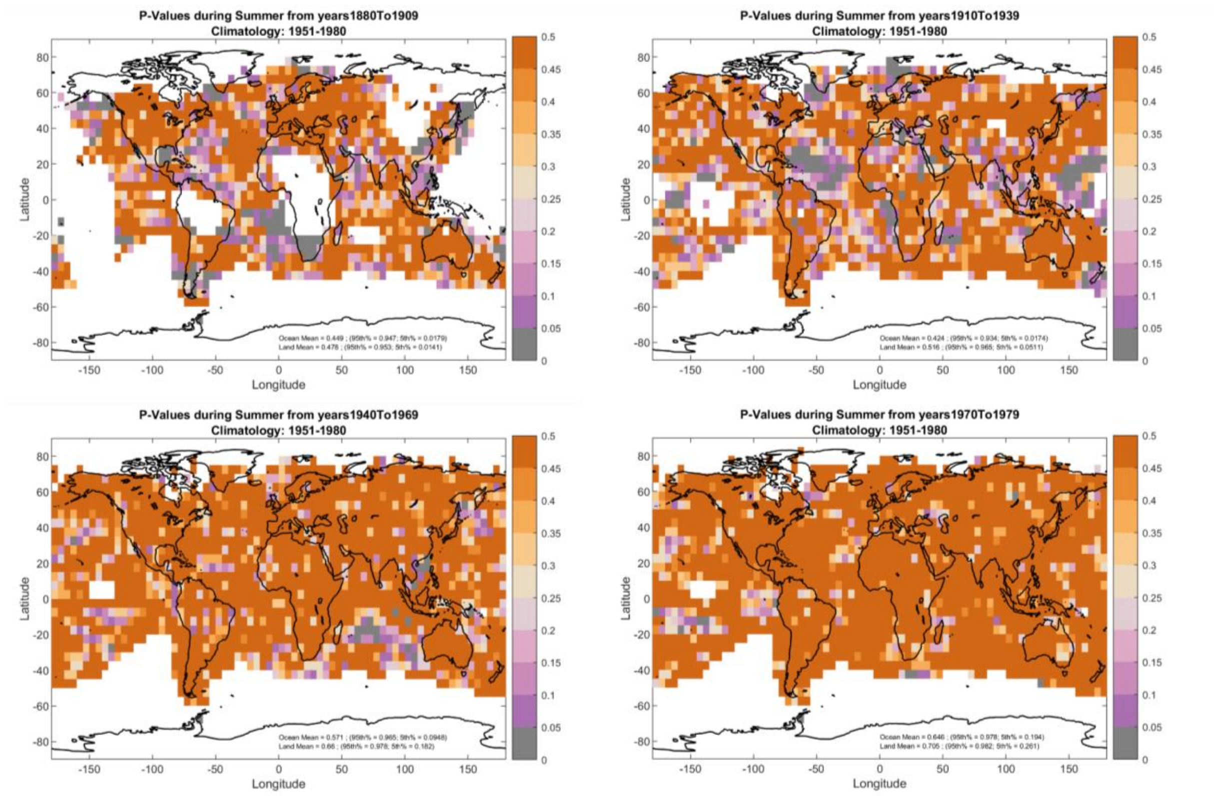

Figure 3.8. As per Figure 3.2, but for Summer (JJA) p-values from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

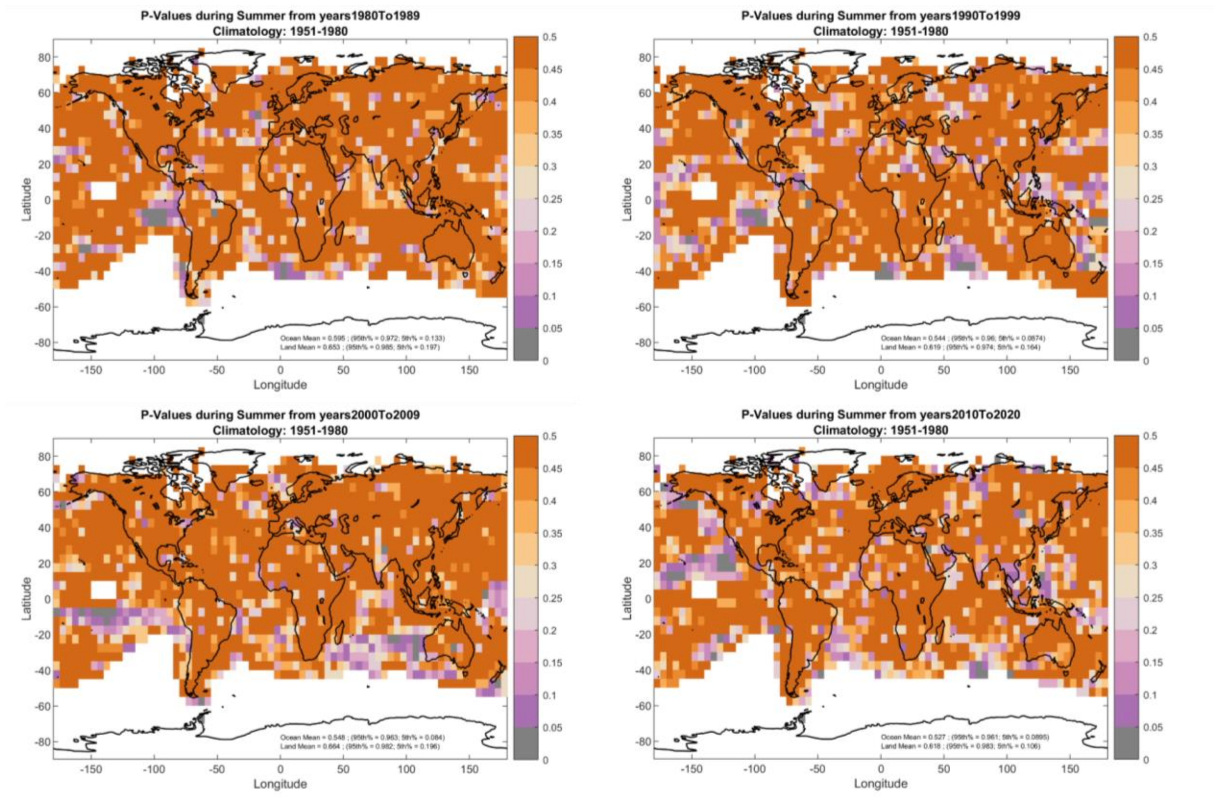

Figure 3.9. As per Figure 3.2, but for Summer (JJA) p-values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

In summary, p-values from the KS test largely suggest that the calculations of 2-sigma and 3-sigma frequency are reliable to use across all seasons given that most regions across the world have approximately Gaussian distributions, with the exception of the regions described above with low KS p-values. Figure 3.10 shows average p-values for ocean and land across all seasons and time frames. Across all seasons, every time period appears to have a p-value of 0.45 and greater, indicating that the data is mostly Gaussian, however, within the decadal time periods, the average p-value increases to roughly 0.6-0.7, suggesting that the decadal analysis is more Gaussian and more useful to use when assessing the frequency of 2-sigma and 3-sigma events.

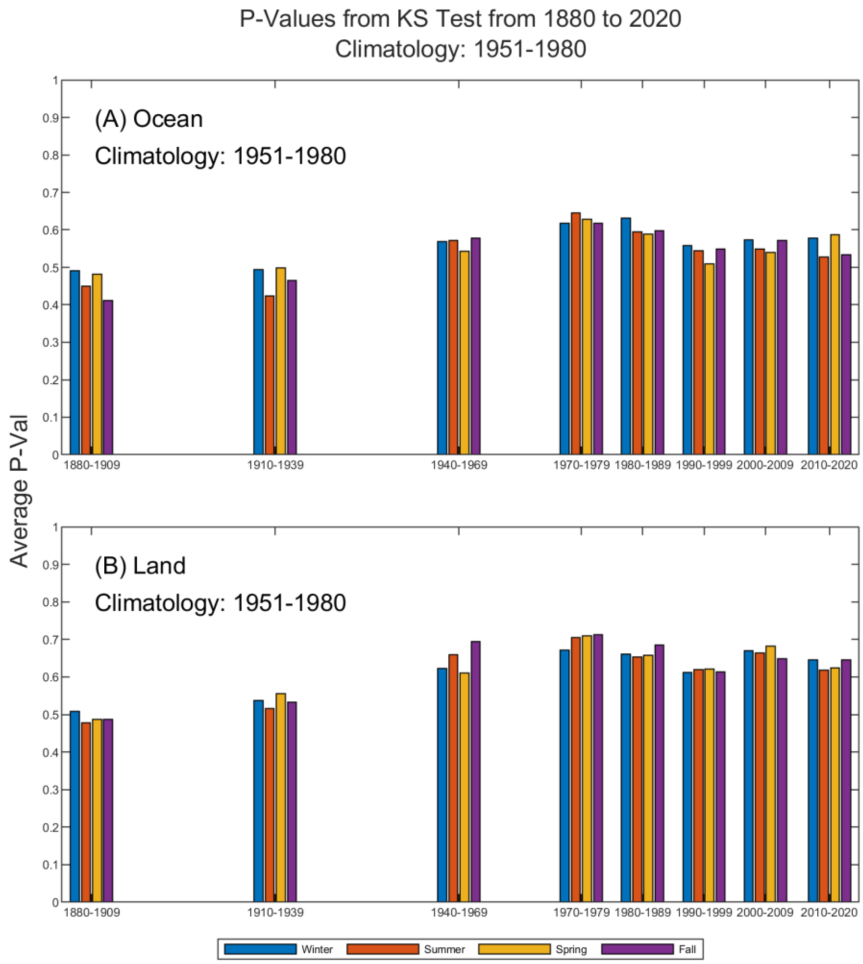

Figure 3.10. 1880 to 2020 seasonal trend in p-values from the KS test applied to distributions of monthly standardized temperature anomalies and spatially averaged over (a.) Ocean and (b.) Land grid boxes.

3.2: Interquartile Range of the Distributions

To assess how the spread of the distributions of monthly standardized temperature anomalies has changed from 1880 to 2020 within each time period and season, we created maps of the interquartile range. IQR values mainly range from 1-3, but I include the 5th and 95th percentile values for large-scale averages on the maps. Smaller values of IQR indicate a narrow distribution with less variability about the mean value, while larger values indicate a wider distribution with more variability about the mean value. Hansen et al. (2012) reported a distribution that widened from past to present.

Figures 3.11 and 3.12 display Fall (SON) IQR values from 1880 to 2020, showing that the spread of the distributions do not change much over time, where ocean and land grid-area weighted averages range from 1.21 and 1.58. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.98 and 2.21 in 1880-1909 and 0.85 and 2.18 in 2010-2020. The value of the 5th percentile changes from 0.92-1.04 in 1880-1909 to 0.75-0.94 in 2010-2020, suggesting that the distributions do not become wider with time, but stay relatively the same. Regions in the Pacific, Atlantic and Indian ocean, as well as across Africa, South America, Europe and Southeast Asia show high IQR values, suggesting that those regions' temperature distributions are wider. Based on this, I would expect a higher frequency of 2-sigma and 3-sigma events, given a wider distribution of temperature extremes.

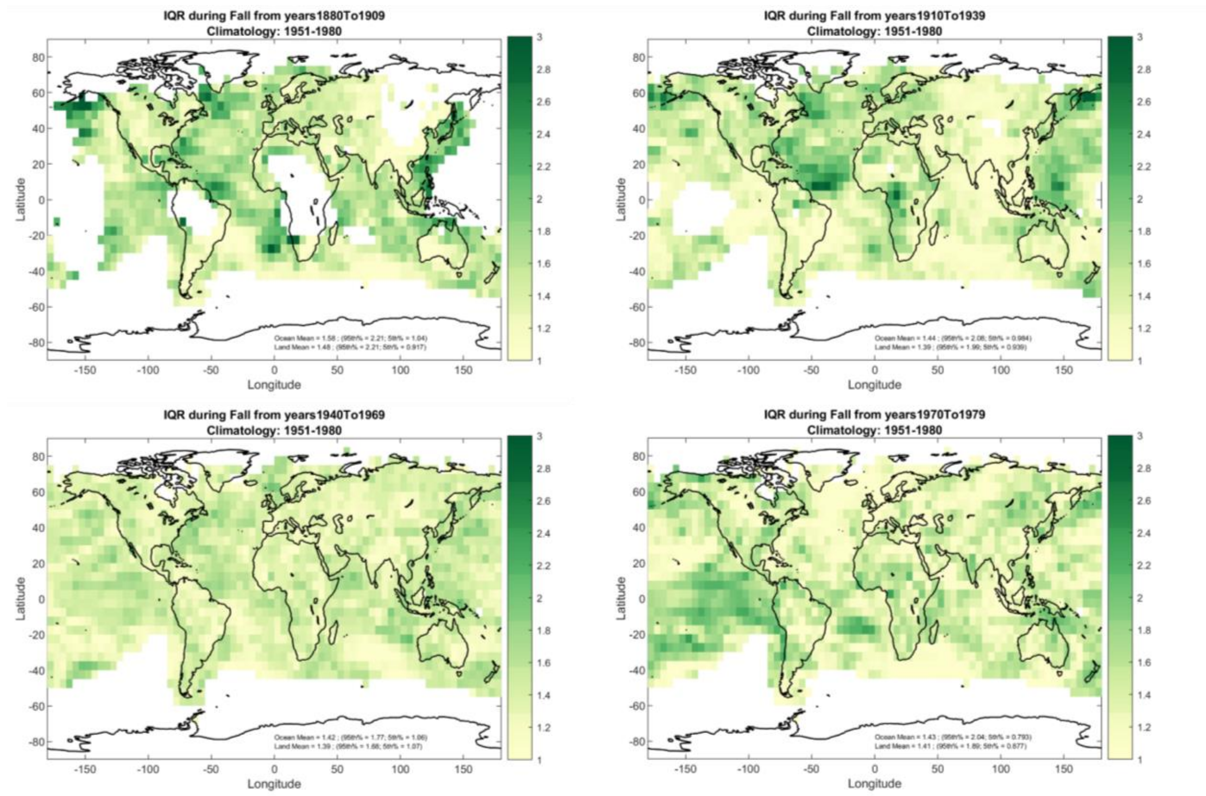

Figure 3.11. Maps depict Fall (SON) IQR values from the distributions of monthly standardized temperature anomalies from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

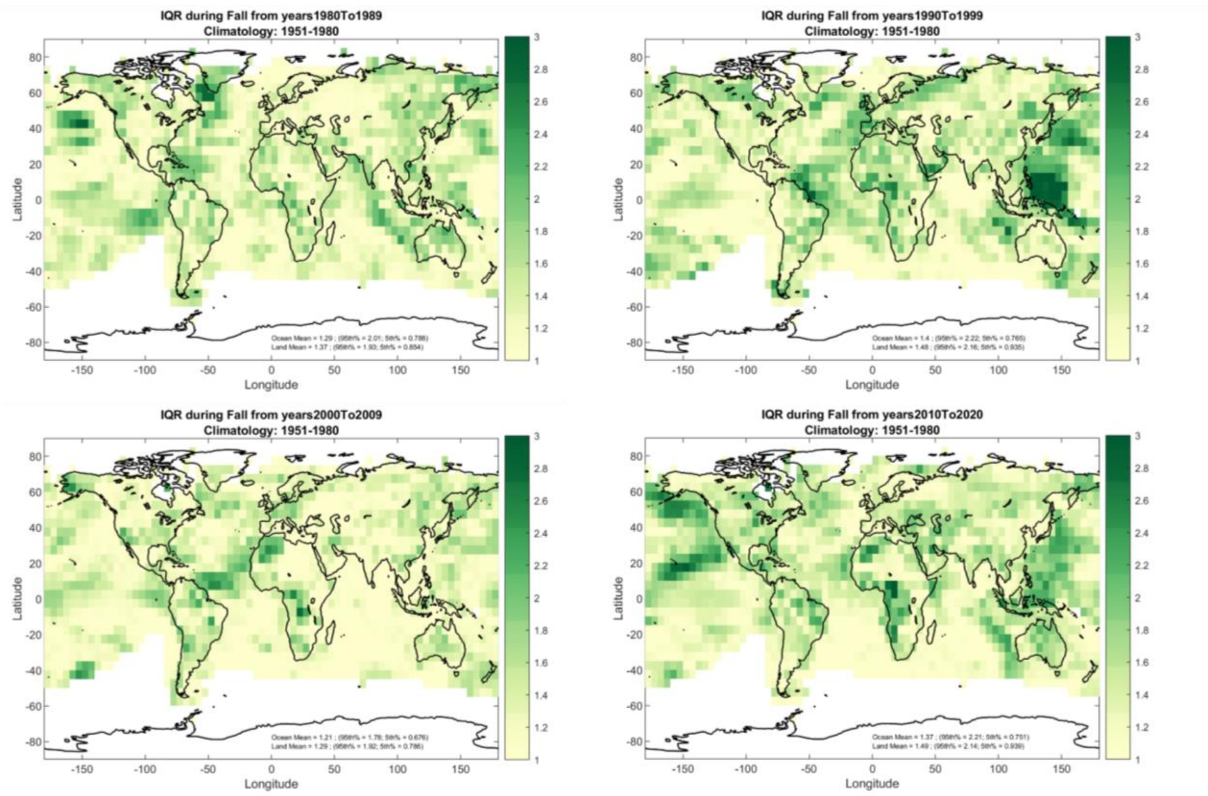

Figure 3.12. As per Figure 3.11, but for Fall (SON) IQR values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.13 and 3.14 display Winter (DJF) IQR from 1880 to 2020, showing that the spread of the distributions do not change much over time, where ocean and land grid-area weighted averages range from 1.31 and 1.50. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.94 and 2.12 in 1880-1909 and 0.92 and 2.12 in 2010-2020. The value of the 5th percentile changes from 0.89-1.00 in 1880-1909 to 0.91-0.93 in 2010-2020, suggesting that the distributions do not become wider with time, but stay relatively the same. Similar to Fall, regions in the Pacific, Atlantic and Indian ocean, as well as across Africa, South America, Europe and Southeast Asia show high IQR values, suggesting that those regions' temperature distributions are wider.

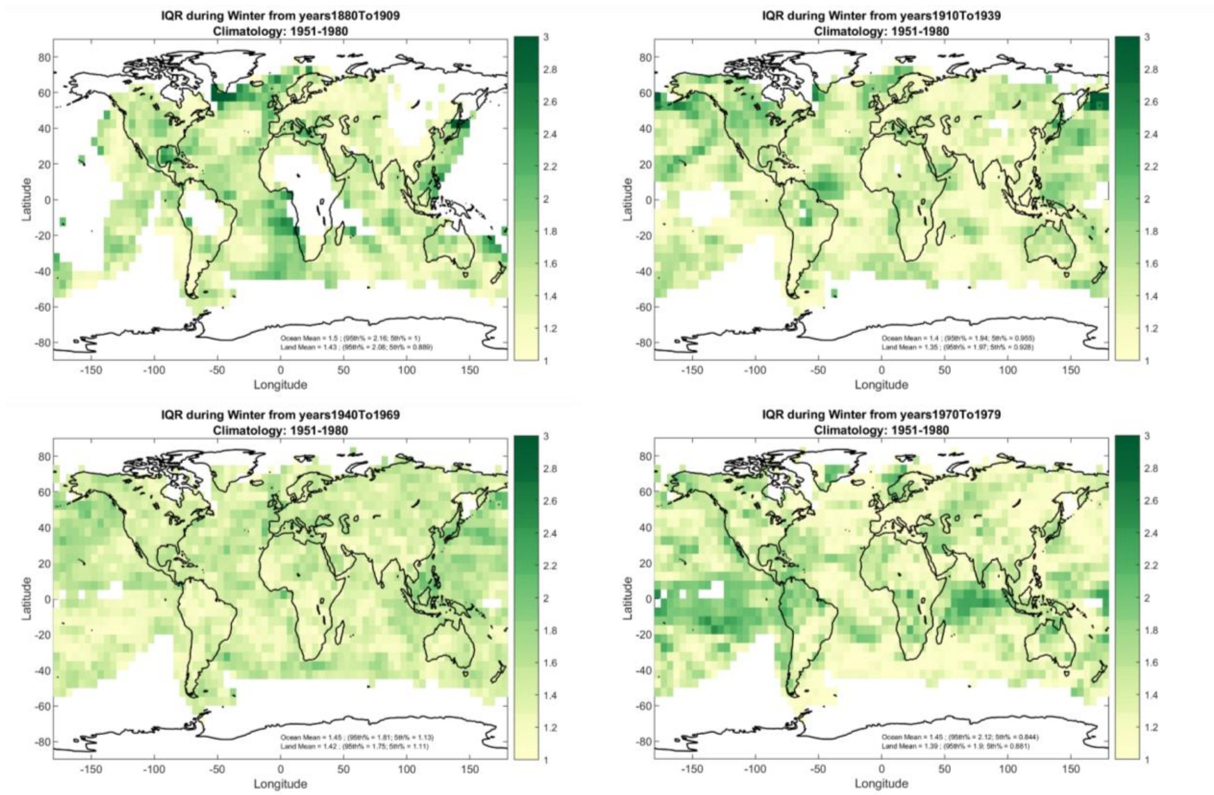

Figure 3.13. As per Figure 3.11, but for Winter (DJF) IQR values from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

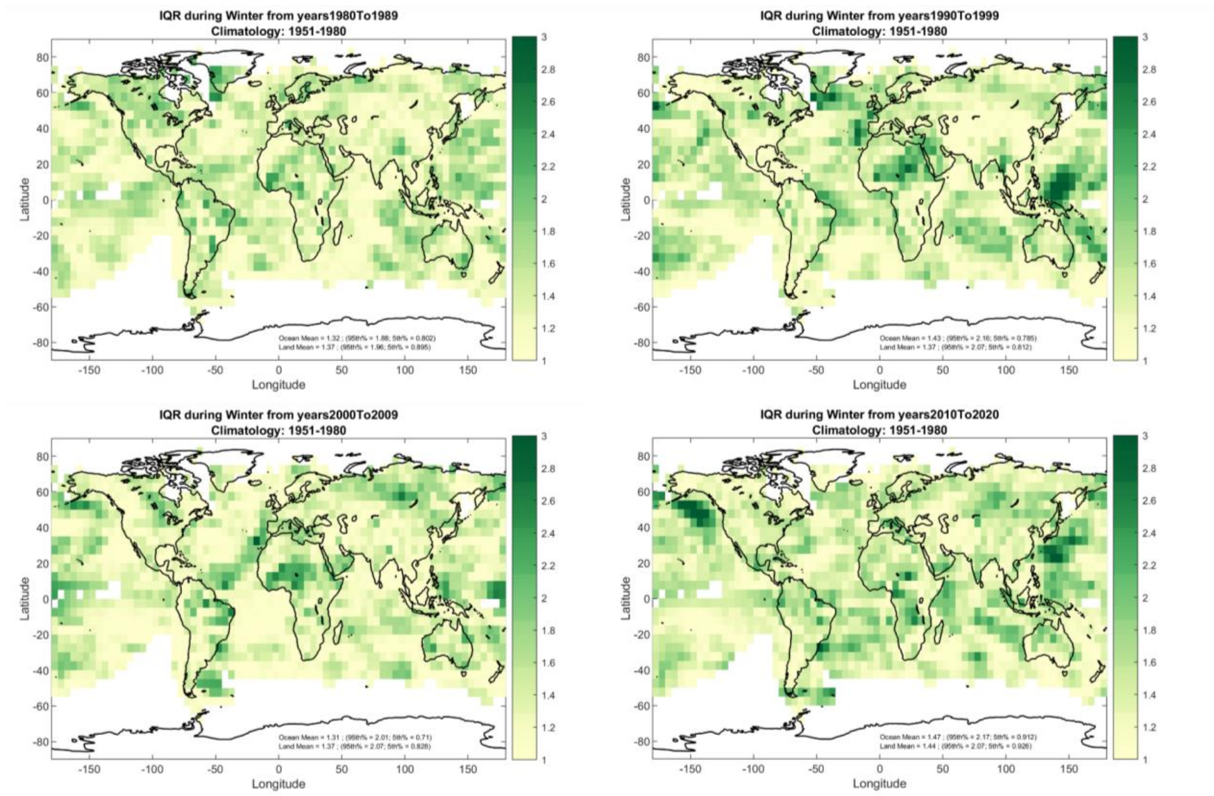

Figure 3.14. As per Figure 3.11, but for Winter (DJF) IQR values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.15 and 3.16 display the Spring (MAM) IQR from 1880-2020, showing that the spread of the distributions do not change much over time, where ocean and land grid-area weighted averages range from 1.27 and 1.53. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.96 and 2.24 in 1880-1909 and 0.89 and 2.24 in 2010-2020. The value of the 5th percentile changes from 0.88-1.03 in 1880-1909 to 0.82-0.96 in 2010-2020, suggesting that the distributions do not become wider with time, but stay relatively the same. Regions in the Pacific, Atlantic and Indian ocean, as well as across Africa, eastern/central Europe, northern South America and southeast Asia show high IQR values, suggesting that those regions' temperature distributions are wider.

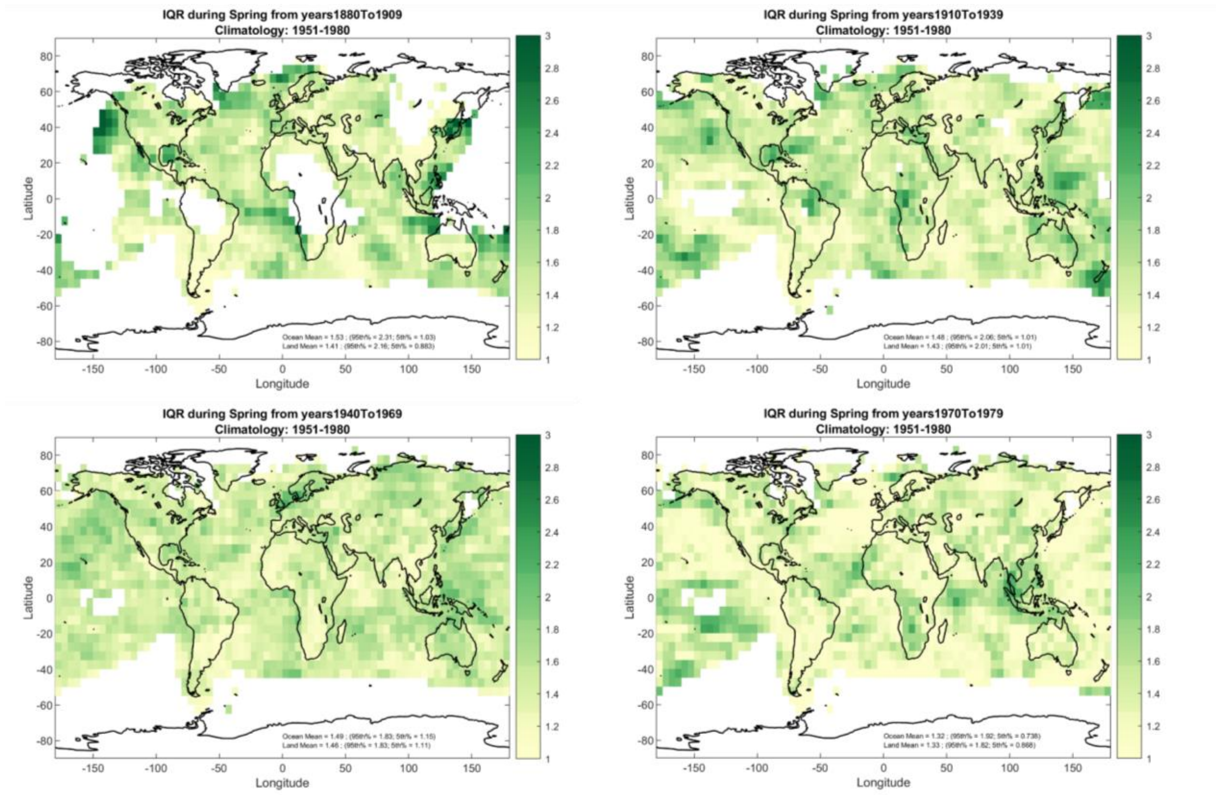

Figure 3.15. As per Figure 3.11, but for Spring (MAM) IQR values from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

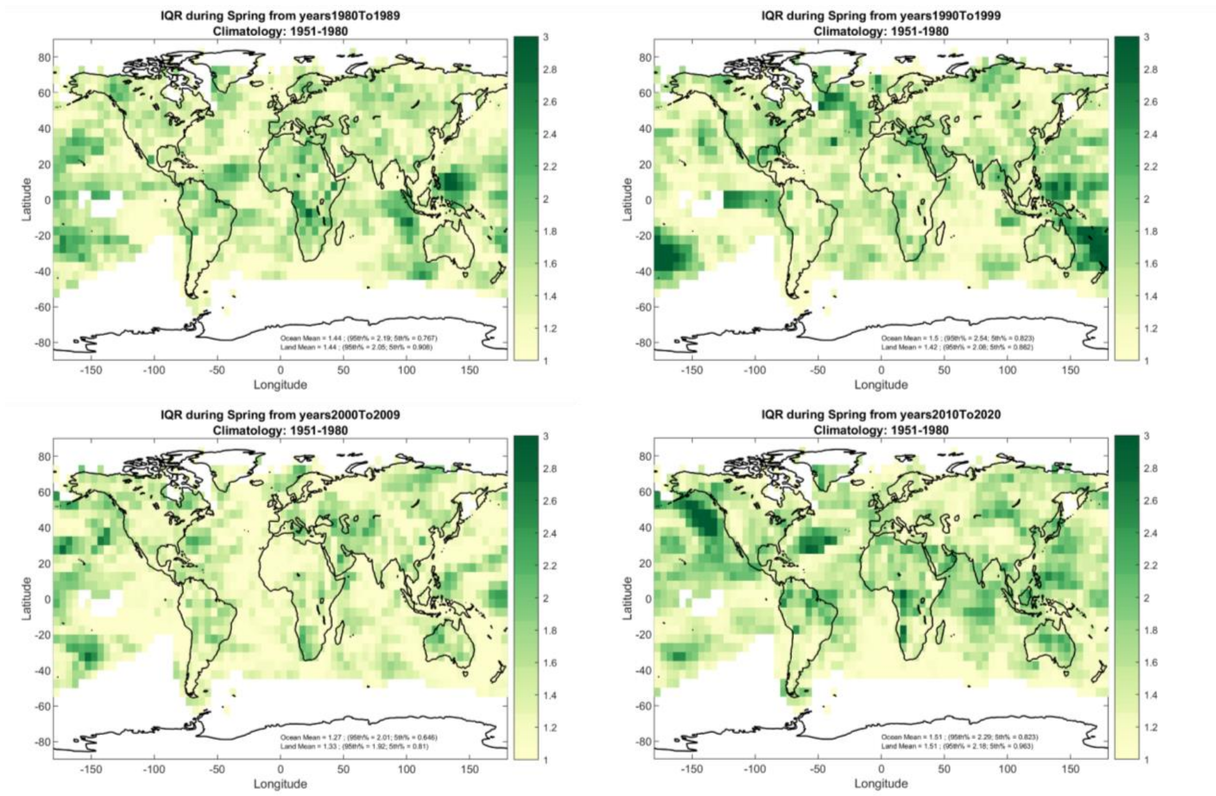

Figure 3.16. As per Figure 3.11, but for Spring (MAM) IQR values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.17 and 3.18 display the Summer (JJA) IQR from 1880-2020, showing that the spread of the distributions do not change much over time, where ocean and land grid-area weighted averages range from 1.16 and 1.56. The grid-area weighted averages of the 5th and 95th percentiles for land and ocean grid boxes are 0.96 and 2.17 in 1880-1909 and 0.81 and 2.29 in 2010-2020. The value of the 5th percentile changes from 0.91-1.00 in 1880-1909 to 0.76-0.86 in 2010-2020, suggesting that the summer distributions become slightly narrower over time. Regions in the Pacific, Atlantic and Indian ocean, as well as across Africa, eastern/central Europe, northern South America and southeast Asia show high IQR values, suggesting that those regions' temperature distributions are wider.

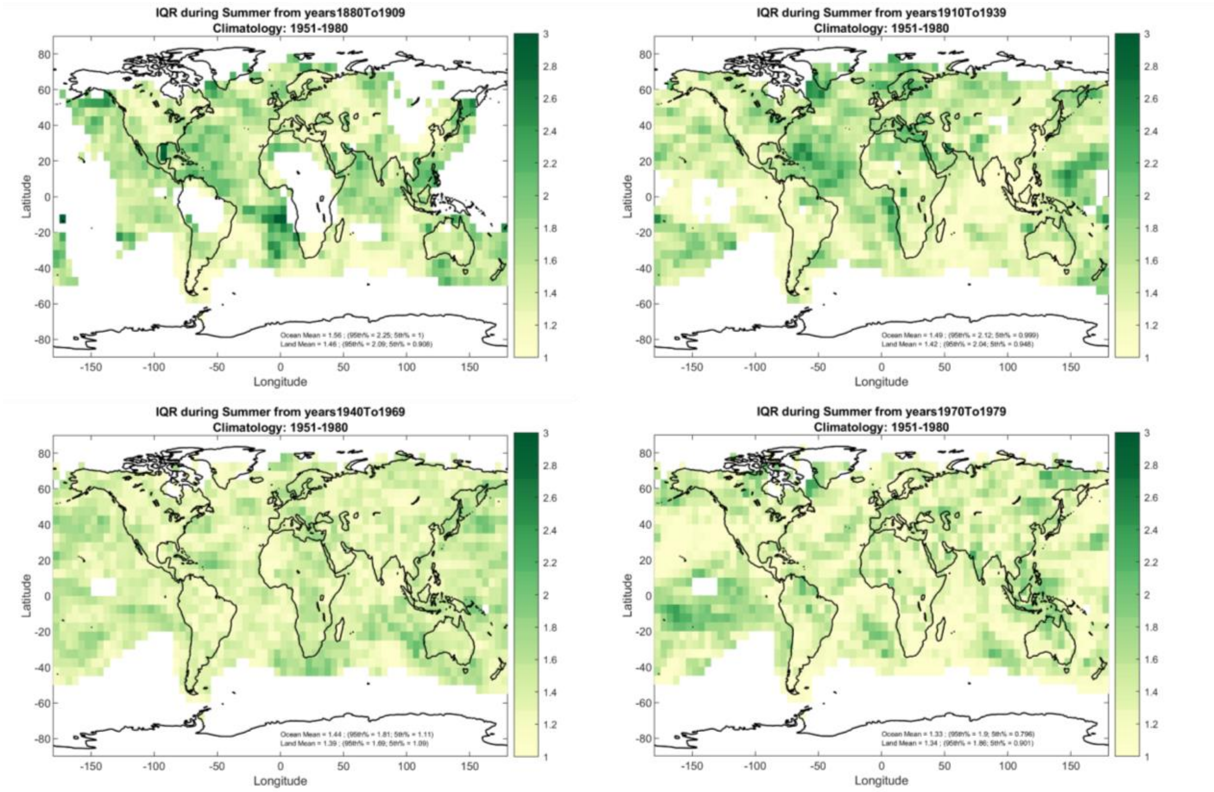

Figure 3.17. As per Figure 3.11, but for Summer (JJA) IQR values from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

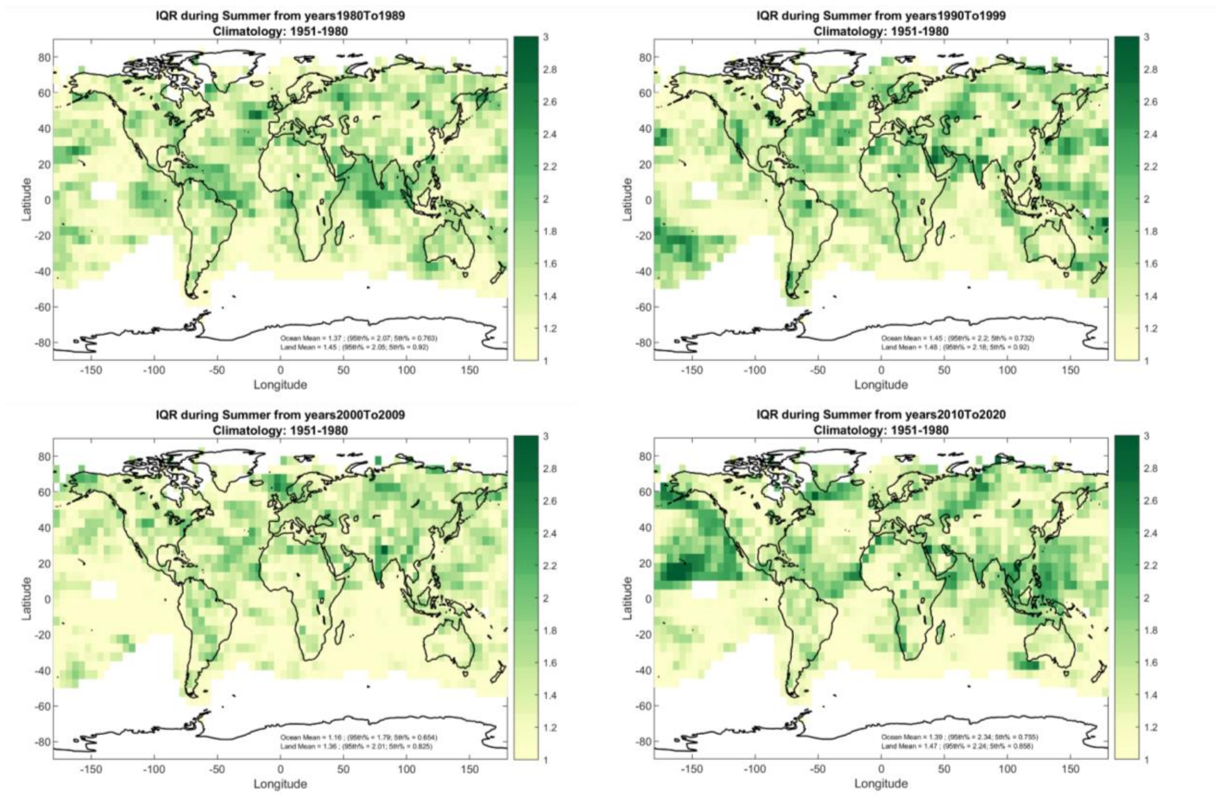

Figure 3.18. As per Figure 3.11, but for Summer (JJA) IQR values from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

In summary, IQR values largely suggest that there is a lack of consistency, both spatially and temporally, on the change in the spread of the distributions across all seasons and time frames. Because a wider distribution entails a higher frequency of larger, positive sigma values, I would expect to see higher frequencies of 2-sigma and 3-sigma events where the distributions are wider. Figure 3.19 shows average IQR values for ocean and land across all seasons and time frames. Based on this figure, the IQR does not change much throughout time and for each season, suggesting that the distributions do not get wider or narrower with increasing time, contrasting the results of Hansen et al (2012).

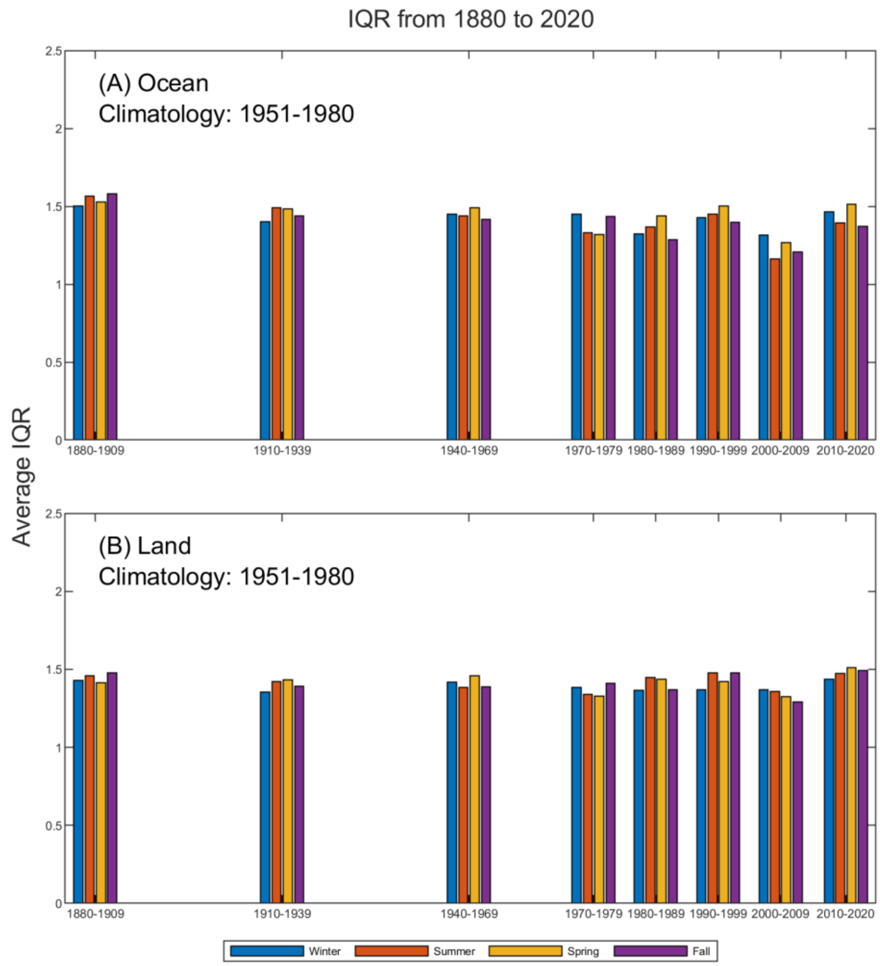

Figure 3.19. Average (a) Ocean and (b) Land IQR values applied to distributions of monthly standardized temperature anomalies for each season in decadal or multidecadal scale time frames from 1880 to 2020.

3.3: Frequency of 2-sigma Events

To assess how the changing location (i.e., the mean) of the monthly standardized temperature anomaly distributions are changing the way monthly temperature extremes are experienced per season, I calculated the frequency of 2-sigma events (where monthly standardized temperature anomalies are greater than +2 standard deviations) for each season in each time frame. This value was divided by the expected frequency of 2-sigma events of 2.2% and plotted on global maps, in order to show the magnitude of frequency relative to the expected frequency from a Gaussian distribution with a mean of zero and standard deviation of 1 (i.e., the base period). Gray values on the maps indicate that the frequency is roughly within expectation of the climatology period, where values above that is a departure of the climatological norm. I replicated the process for -2-sigma events, with the expectation that these events would decline across all seasons and time frames, but I only show examples of the Summer season maps for the -2-sigma events since the main focus of my project is on the changes in statistically extreme heat events. Thus, I show maps from all seasons for 2-sigma and 3-sigma events.

Figures 3.20 and 3.21 show the Summer (JJA) frequency of -2-sigma events from 1880 to 2020, showing a rapid decline of -2-sigma events over time for each season. In 2010-2020, Summer months essentially never experience -2-sigma events, and there is a corresponding decrease in the values of the 95th percentile from as high as 19.8 in the 1880-1909 time frame to 0 in the 2010-2020 time period. This means that while extreme cold (relative to the base period) was experienced regularly somewhere on the globe in the late 1800s, it was never experienced in the 2010s. I do not show maps for other seasons for -2-sigma frequency departures, but the results are similar as summarized in Figure 3.22.

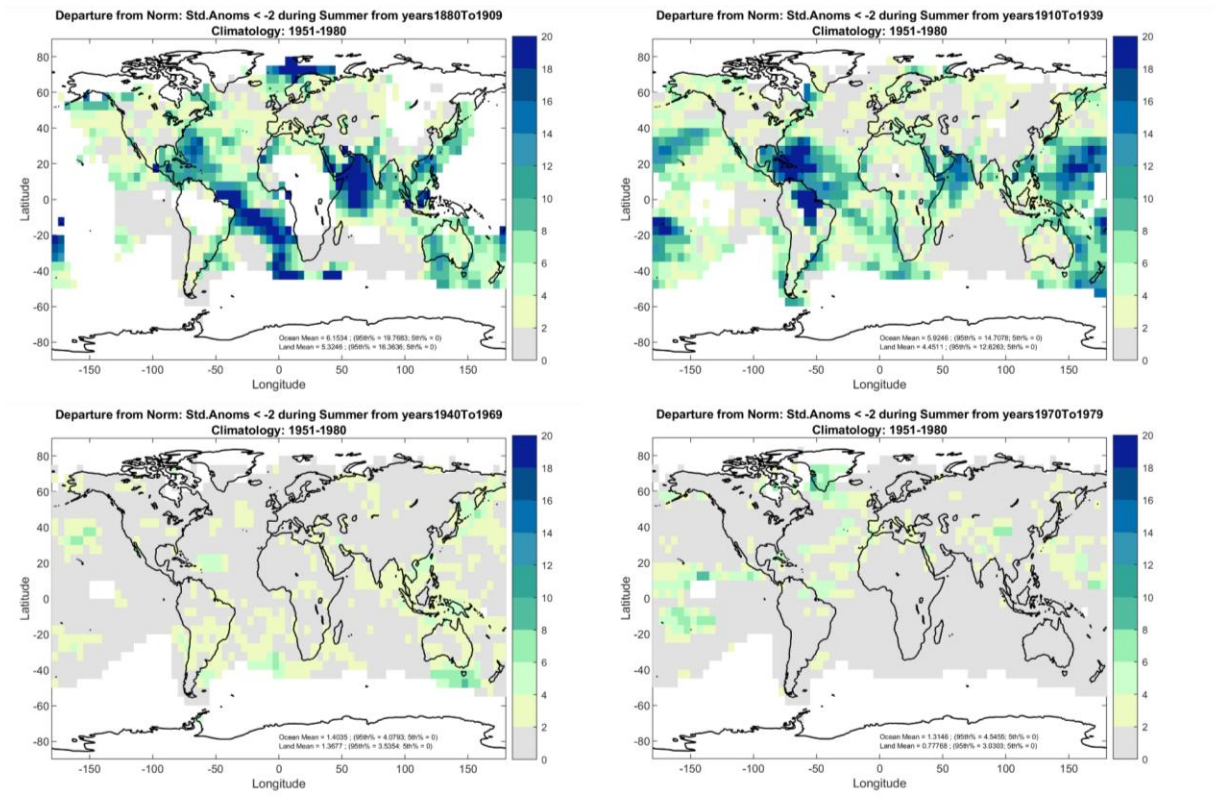

Figure 3.20. Maps depict the departure of the frequency of -2-sigma events in Summer (JJA) from statistical expectation of 2.2% as the ratio (observed frequency divided by expectation frequency) for (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979..

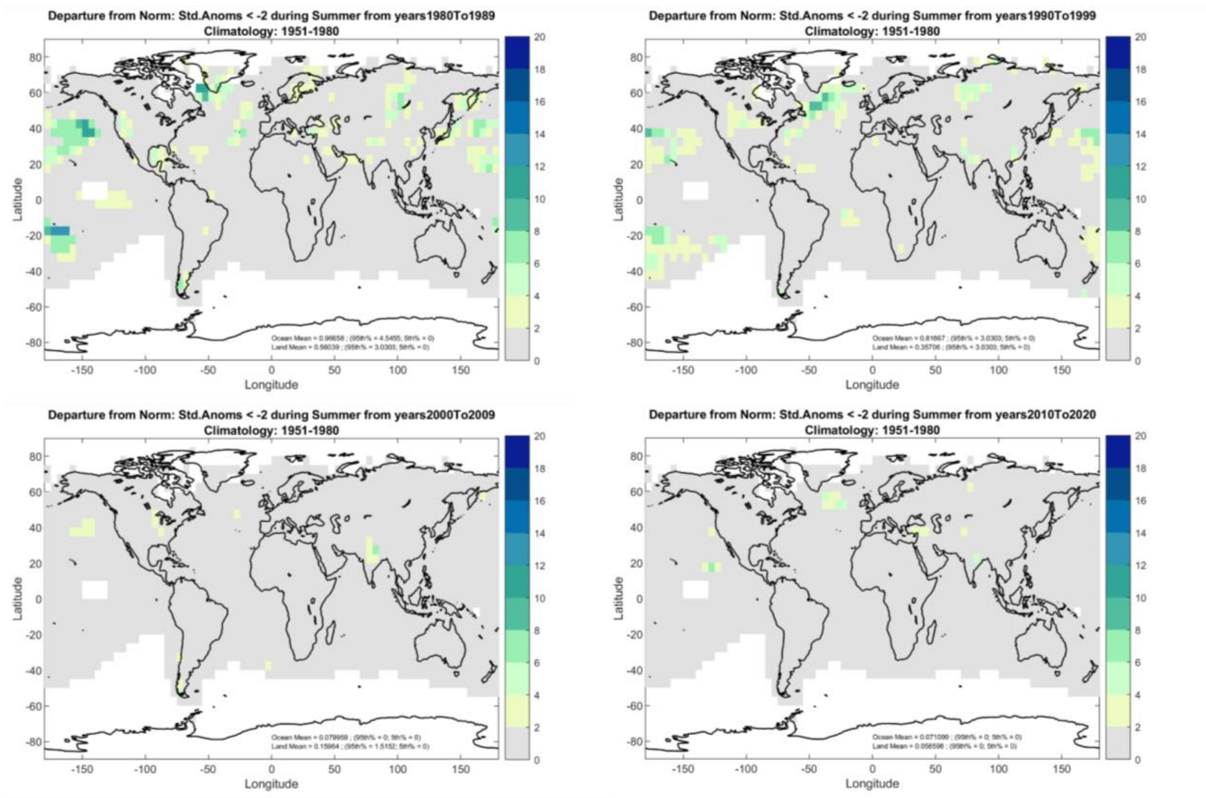

Figure 3.21. As per Figure 3.20, but for time periods of (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figure 3.22 shows spatially averaged departure frequencies relative to statistical expectation for -2-sigma events for ocean and land across all seasons and time frames. Results suggest that there is a rapid decline of -2-sigma events for each season over time, which is what I anticipated. Given that there is a decline in -2-sigma events, we can assume that there will be a rapid increase in 2-sigma events over time because the mean of the distributions of monthly standardized temperature anomalies are shifting to larger positive values, and the distributions are largely Gaussian (Section 3.1).

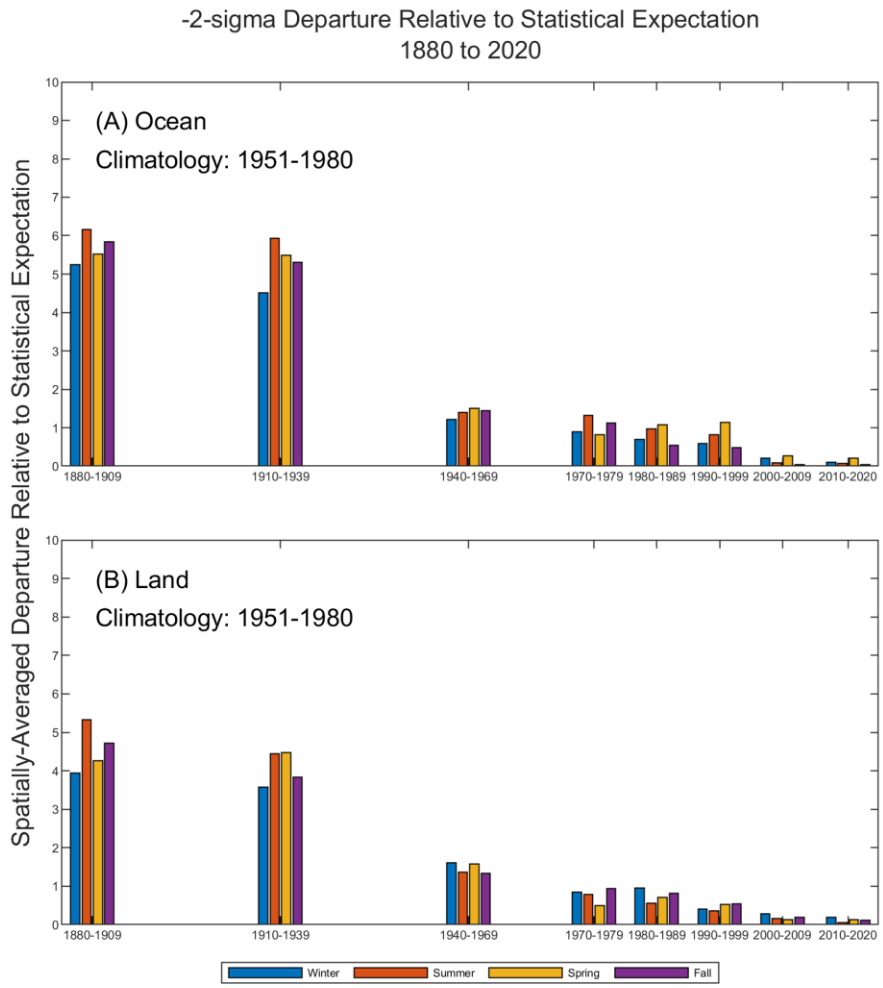

Figure 3.22. Average (a) Ocean and (b) Land spatially-averaged departure frequencies relative to statistical expectation for -2-sigma events applied to distributions of monthly standardized temperature anomalies for each season in decadal or multdecadal scale time frames from 1880 to 2020.

Figures 3.23 and 3.24 display Fall (SON) frequency departure values of 2-sigma events from 1880 to 2020. 2-sigma land and ocean departure values average 1.1 from 1880 to 1979 and 8 from 1980 to 2020. What this means is that from 1880 to 1979, 2-sigma events occurred at about the same rate as statistical expectation, while occurring around 8 times more than expected from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.24) suggests 5% of land and ocean areas experienced 2-sigma events about 9-12 times more frequently than expected, but from 2010 to 2020, that number increased to 36-43 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was about 1.4 (Figure 3.24), implying that greater than 95% of land and ocean area experienced extreme (2-sigma) warmth beyond statistical expectation.

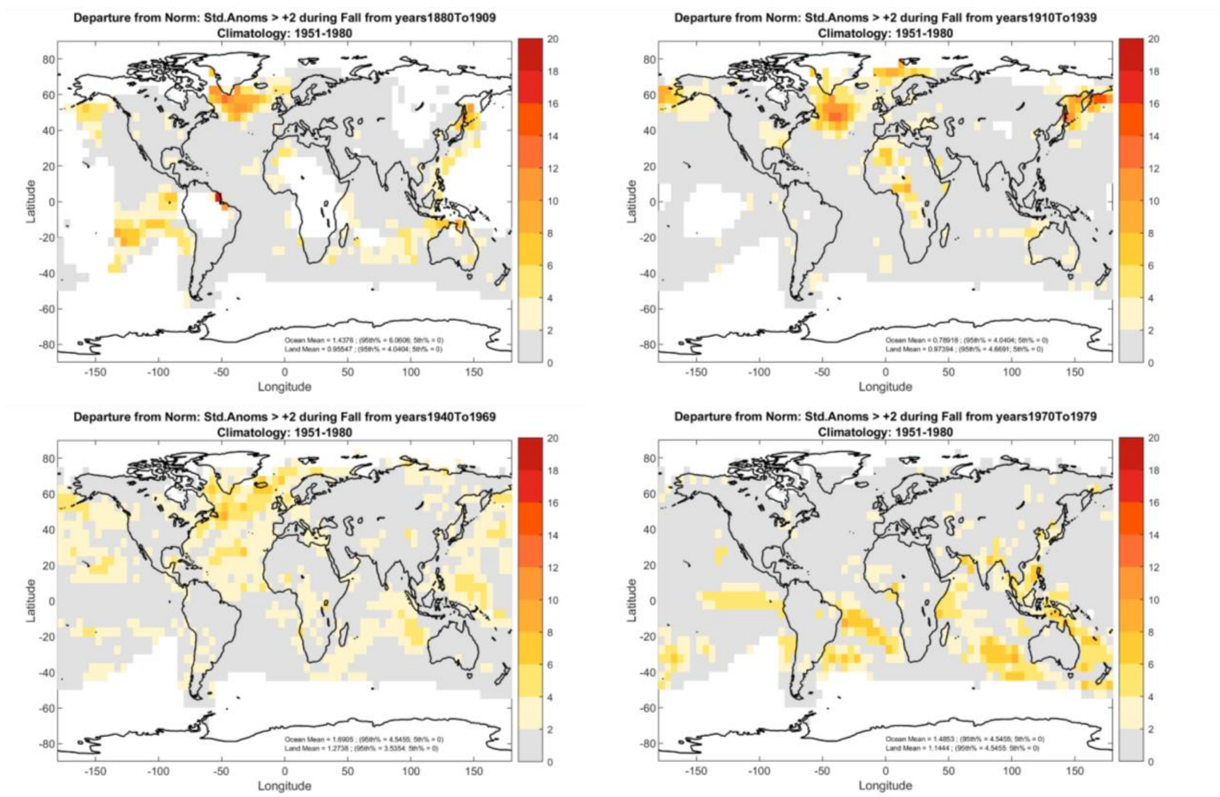

Figure 3.23. Maps depict the departure of the frequency of 2-sigma events in Fall (SON) from statistical expectation of 2.2% as the ratio (observed frequency divided by expectation frequency) for (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

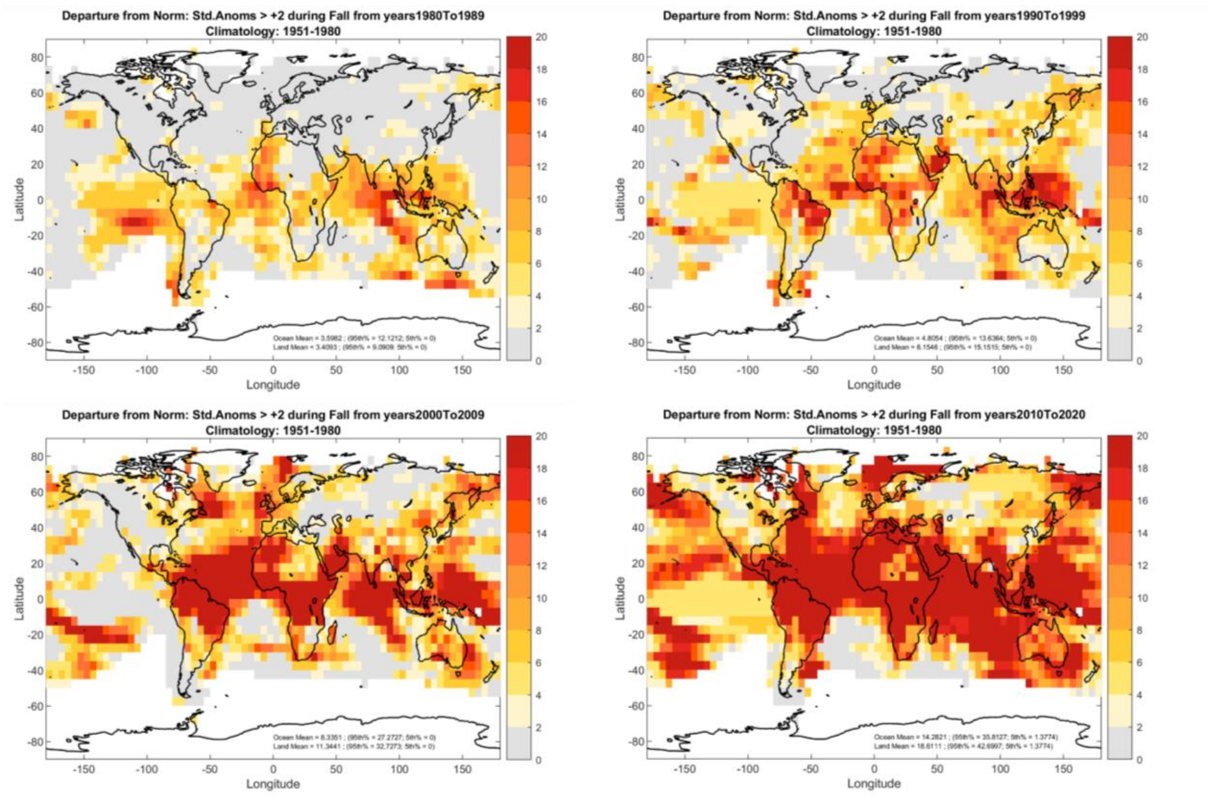

Figure 3.24. As per Figure 3.23, but for Fall (SON) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.25 and 3.26 display Winter (DJF) frequency departure values of 2-sigma events from 1880 to 2020. 2-sigma land and ocean departure values average 1.26 from 1880 to 1979 and average 7.18 from 1980 to 2020. What this means is that from 1880 to 1979, 2-sigma events occurred at about the same rate as statistical expectation, and occurred 7 times more frequently than expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.26) suggests 5% of land and ocean areas experienced 2-sigma events about 8-11 times more frequently than expected, but from 2010 to 2020, that number increased to 29-36 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 1.37 (Figure 3.26), implying that greater than 95% of land and ocean area experienced extreme (2-sigma) warmth beyond statistical expectation.

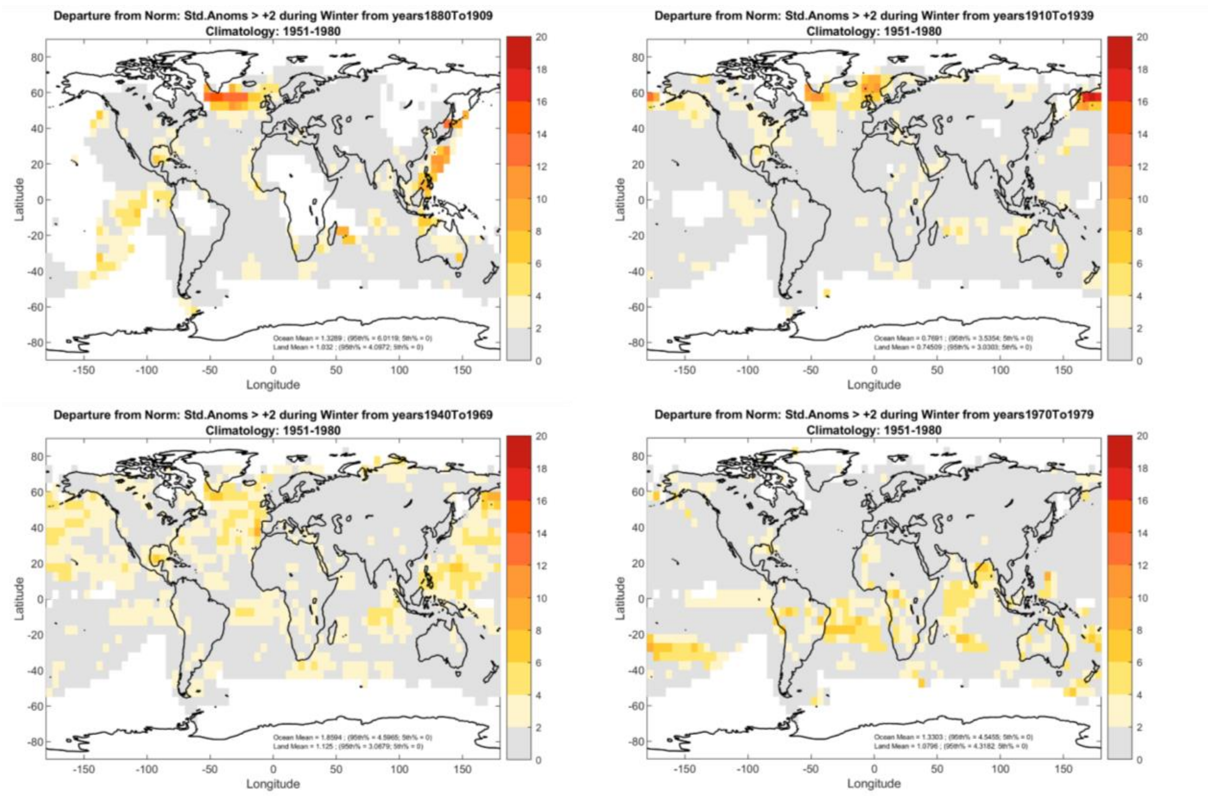

Figure 3.25. As per Figure 3.23, but for Winter (DJF) from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

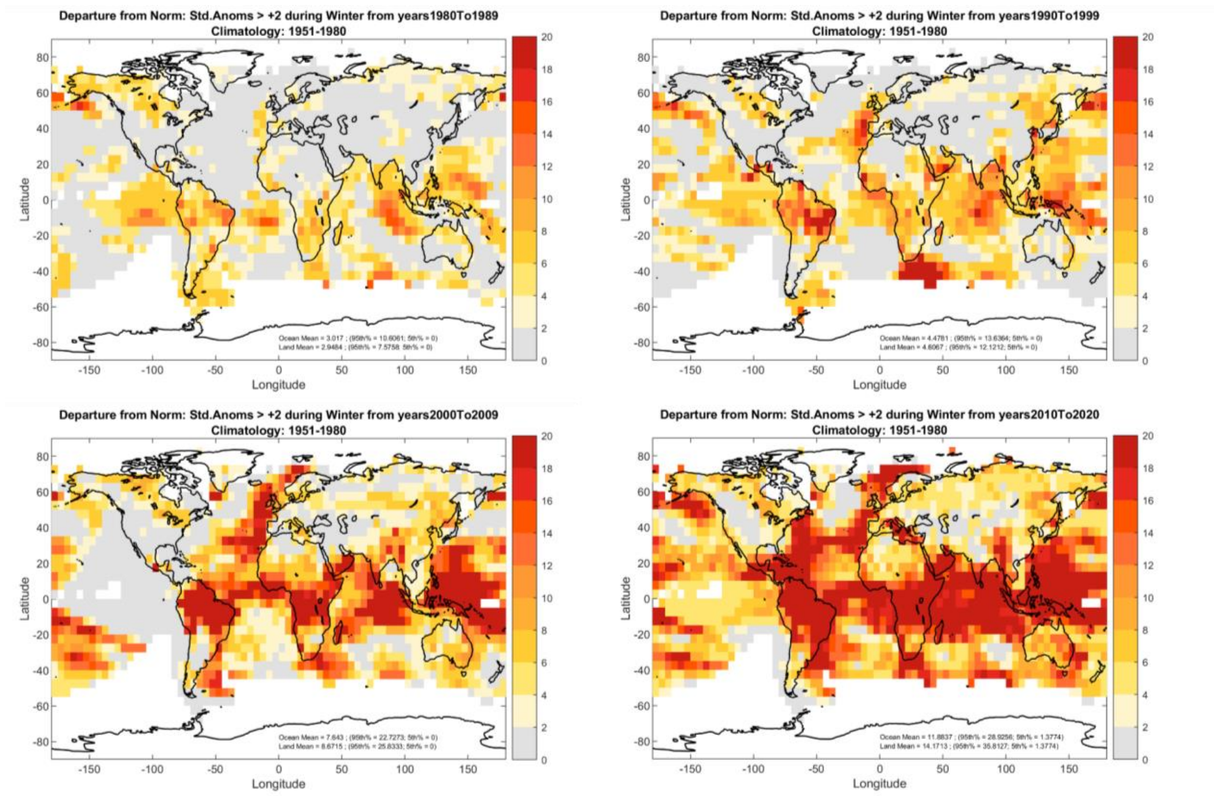

Figure 3.26. As per Figure 3.23, but for Winter (DJF) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.27 and 3.28 display Spring (MAM) frequency departure values of 2-sigma events from 1880 to 2020. 2-sigma land and ocean departure values average 1.10 from 1880 to 1979 and average 9.0 from 1980 to 2020. What this means is that from 1880 to 1979, 2-sigma events occurred at about the same rate as statistical expectation, and occurred 9 times more frequently than expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.28) suggests 5% of land and ocean areas experienced 2-sigma events about 12 times more frequently than expected, but from 2010 to 2020, that number increased to 34-40 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 1.35 (Figure 3.28), implying that greater than 95% of land and ocean area experienced extreme (2-sigma) warmth beyond statistical expectation.

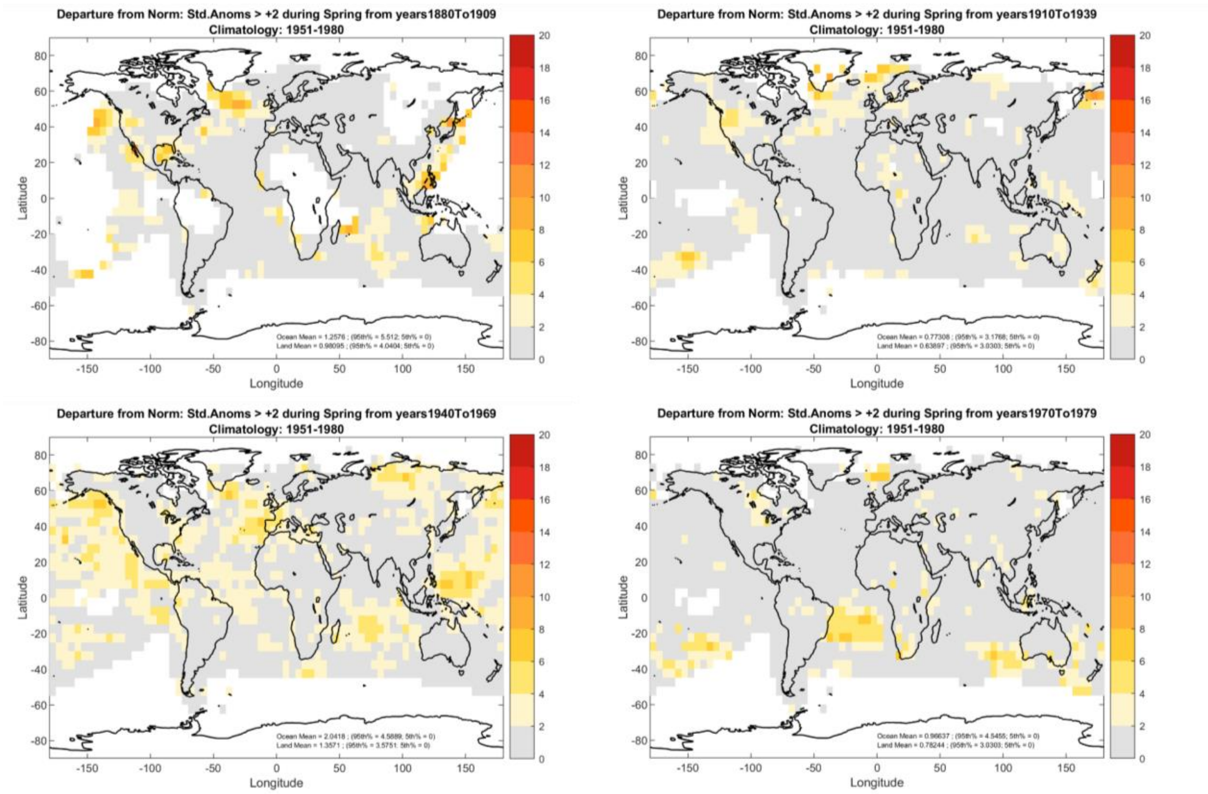

Figure 3.27. As per Figure 3.23, but for Spring (MAM) from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

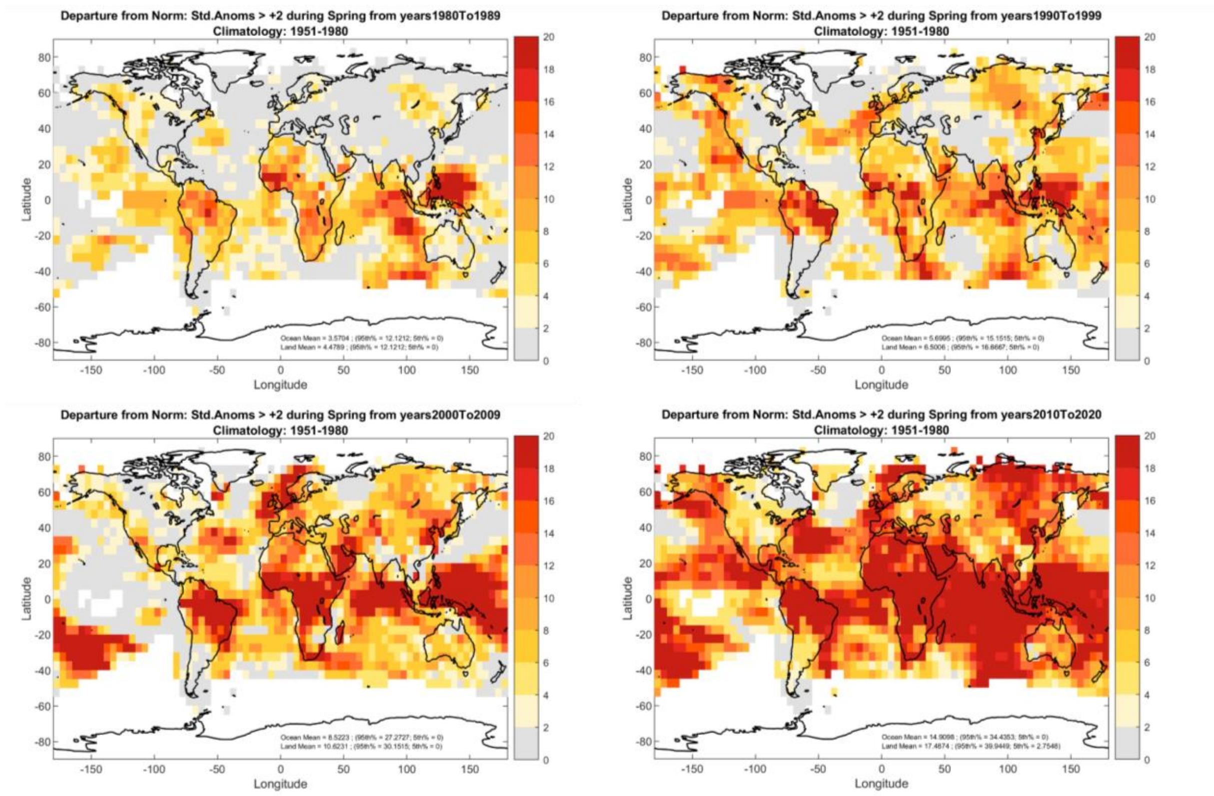

Figure 3.28. As per Figure 3.23, but for Spring (MAM) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.29 and 3.30 display Summer (JJA) frequency departure values of 2-sigma events from 1880 to 2020. 2-sigma land and ocean departure values average 1.12 from 1880 to 1979 and average 9.93 from 1980 to 2020. What this means is that from 1880 to 1979, 2-sigma events occurred about the same as statistical expectation, where 2-sigma events occurred around 9 times statistical expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.30) suggests 5% of land and ocean areas experienced 2-sigma events about 11-12 times more frequently than expected, but from 2010 to 2020, that number increased to 40-43 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 1.38 (Figure 3.30), implying that greater than 95% of land and ocean area experienced extreme (2-sigma) warmth beyond statistical expectation.

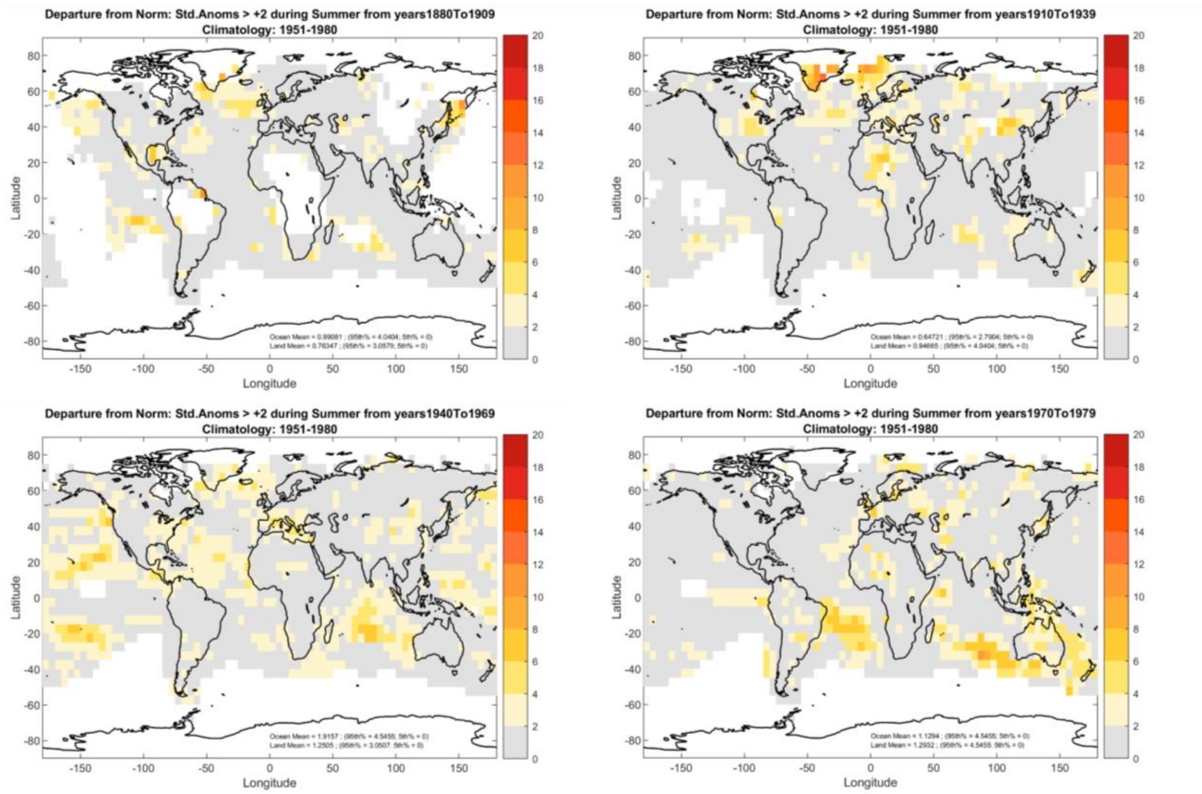

Figure 3.29. As per Figure 3.23, but for Summer (JJA) from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

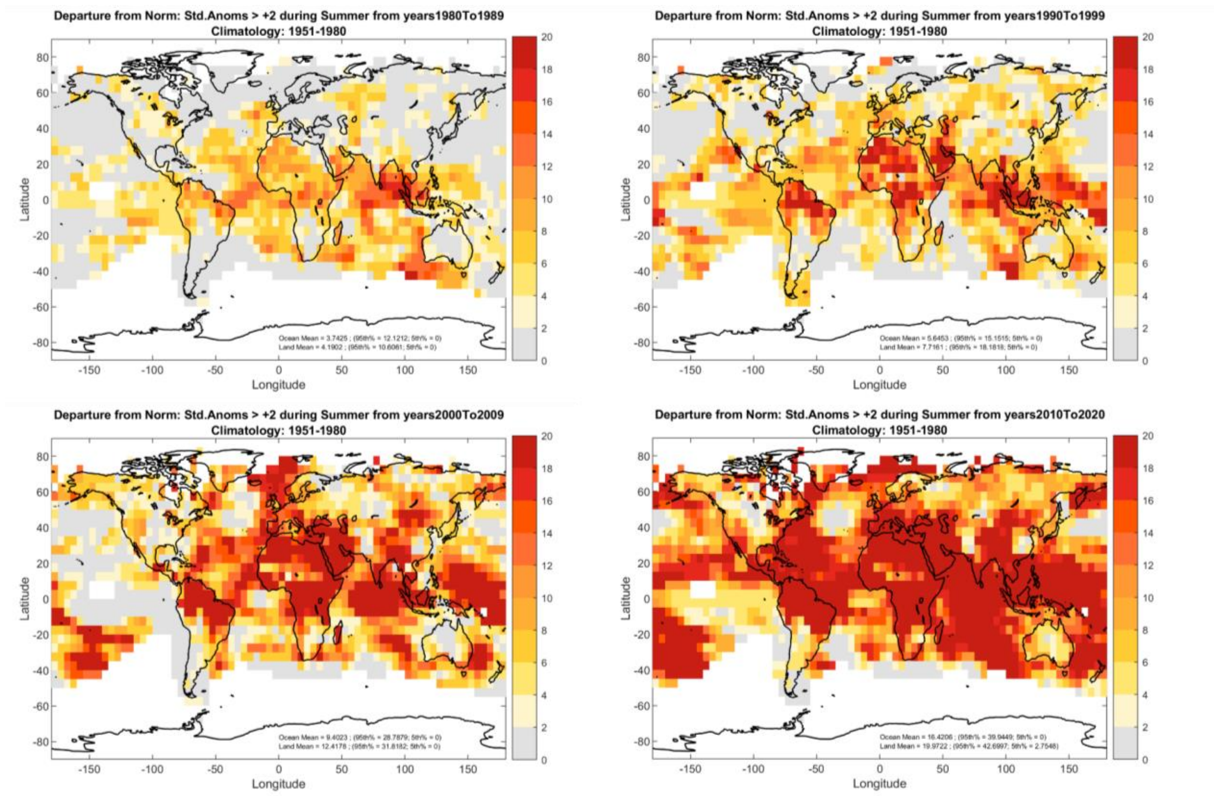

Figure 3.30. As per Figure 3.23, but for Summer (JJA) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figure 3.31 shows spatially-averaged departure frequencies relative to statistical expectation for 2-sigma events for ocean and land across all seasons and time frames. Based on this figure, 2-sigma events remain roughly equal in frequency until the 1980s, where the 2-sigma event frequency increases exponentially between 1980 and 2020. Land frequency departures exceed those of the ocean, perhaps related to the slower warming rate of oceans overall due to the lapse rate feedback (Colman, 2003; Ceppi et al., 2017). However, both ocean and land results highlight the overall increase in 2-sigma event frequency relative to the climatology.

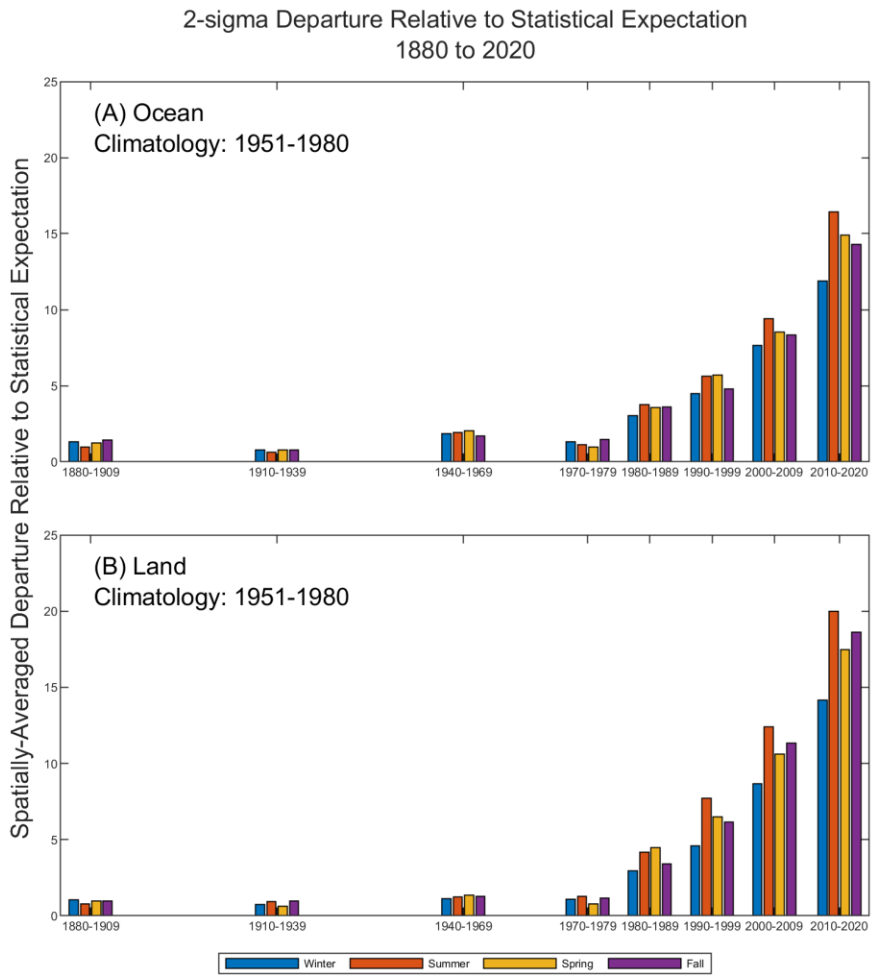

Figure 3.31. Average (a) Ocean and (b) Land spatially-averaged departure frequencies relative to statistical expectation for 2-sigma events applied to distributions of monthly standardized temperature anomalies from 1880 to 2020.

3.4: Frequency of 3-sigma Events

To assess how the location of the monthly standardized temperature distributions are changing over time, we calculated the frequency of 3-sigma events, or monthly standardized temperature anomalies above +3 standard deviations. This value was divided by the expected 0.1% and plotted on global maps, in order to show the spatial distribution of frequency departures, or frequencies divided by expected frequency of a Gaussian distribution centered on zero (i.e. a climatology period). Within a Gaussian distribution, 3-sigma events are rare, but what the results show is that rare events are becoming more common.

Figures 3.32 and 3.33 display Fall (SON) frequency departure values of 3-sigma events from 1880 to 2020. 3-sigma land and ocean departure values average 3.45 from 1880 to 1979 and average 65.51 from 1980 to 2020. What this means is that from 1880 to 1979, 3-sigma events occurred 3 times more frequently than statistical expectation, and occurred 66 times more frequently than expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.33) suggests 5% of land and ocean areas experienced 3-sigma events about 67 times more frequently than expected, but from 2010 to 2020, that number increased to 485-697 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 0 (Figure 3.33), implying that at least 5% of land and ocean area experienced no 3-sigma events or that less than 95% of land and ocean area experienced 3-sigma events beyond statistical expectation.

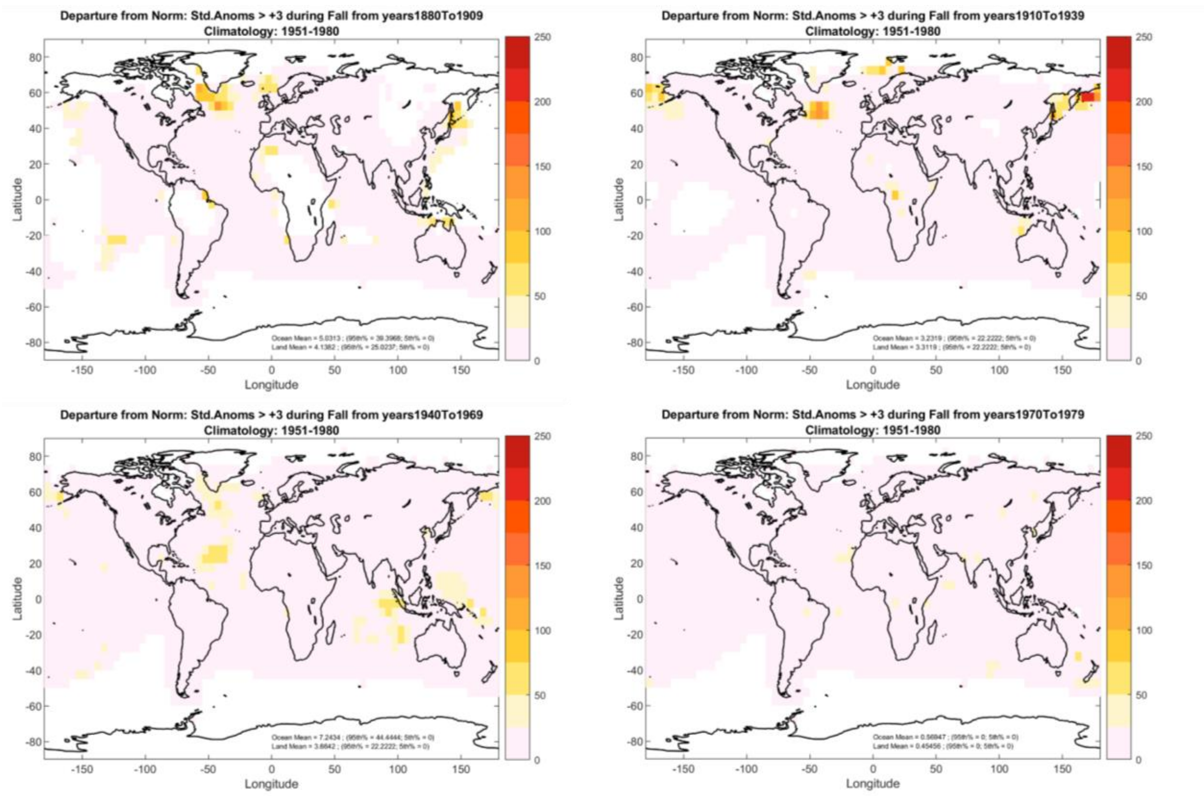

Figure 3.32. Maps depict the departure of the frequency of 3-sigma events in Fall (SON) from statistical expectation of 0.1% as the ratio (observed frequency divided by expectation frequency) for (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

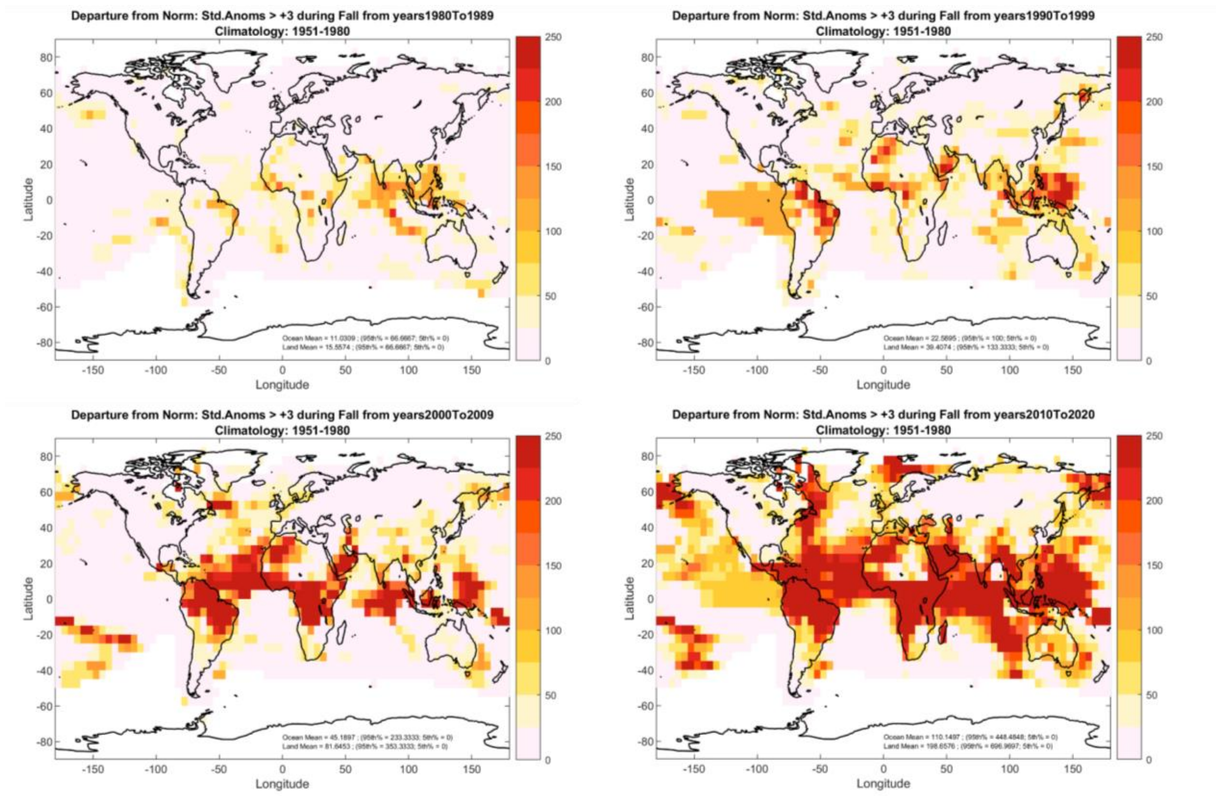

Figure 3.33. As per Figure 3.32, but for 3-sigma events in Fall (SON) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.34 and 3.35 display Winter (DJF) frequency departure values of 3-sigma events from 1880 to 2020. 3-sigma land and ocean departure values average 3.85 from 1880 to 1979 and average 46.30 from 1980 to 2020. What this means is that from 1880 to 1979, 3-sigma events occurred 4 times more frequently than statistical expectation, and occurred 46 times more frequently than expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.35) suggests 5% of land and ocean areas experienced 3-sigma events about 67 times more frequently than expected, but from 2010 to 2020, that number increased to 364-485 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 0 (Figure 3.35), implying that at least

5% of land and ocean area experienced no 3-sigma events or that less than 95% of land and ocean area experienced 3-sigma events beyond statistical expectation.

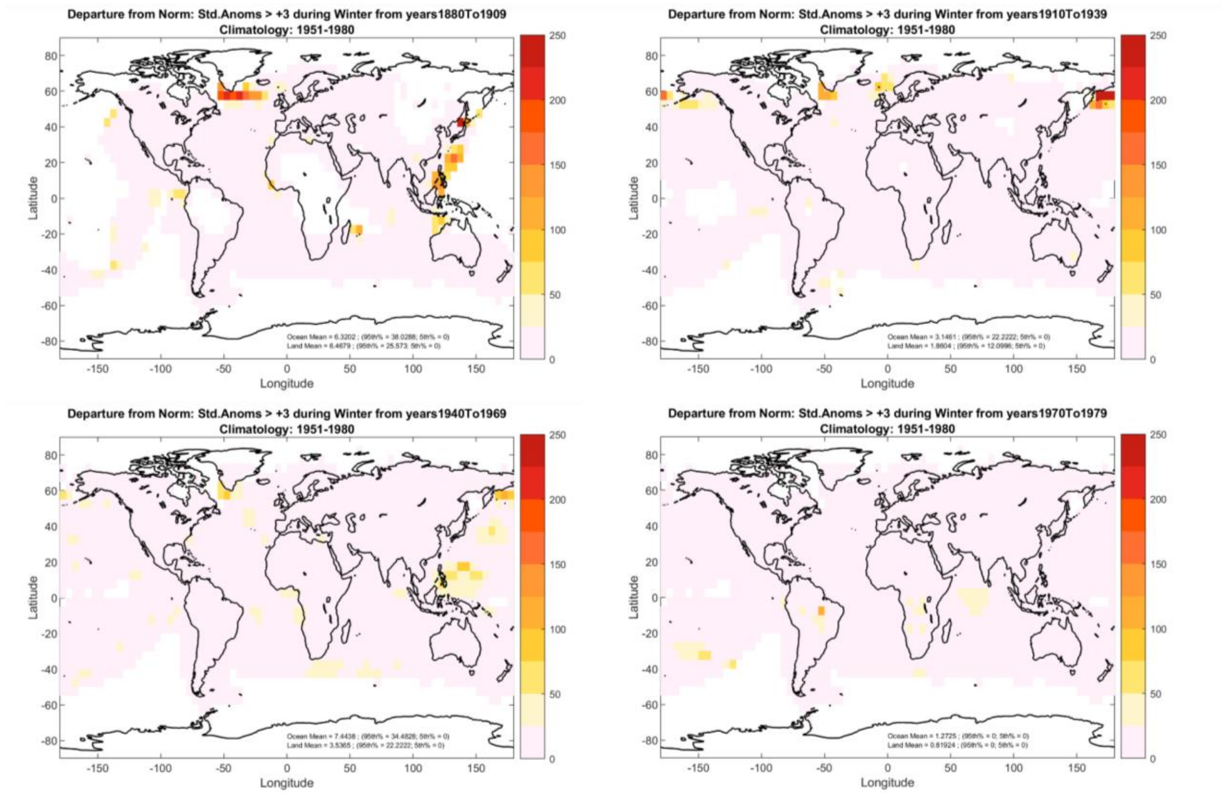

Figure 3.34.As per Figure 3.32, but for 3-sigma events in Winter (DJF) from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

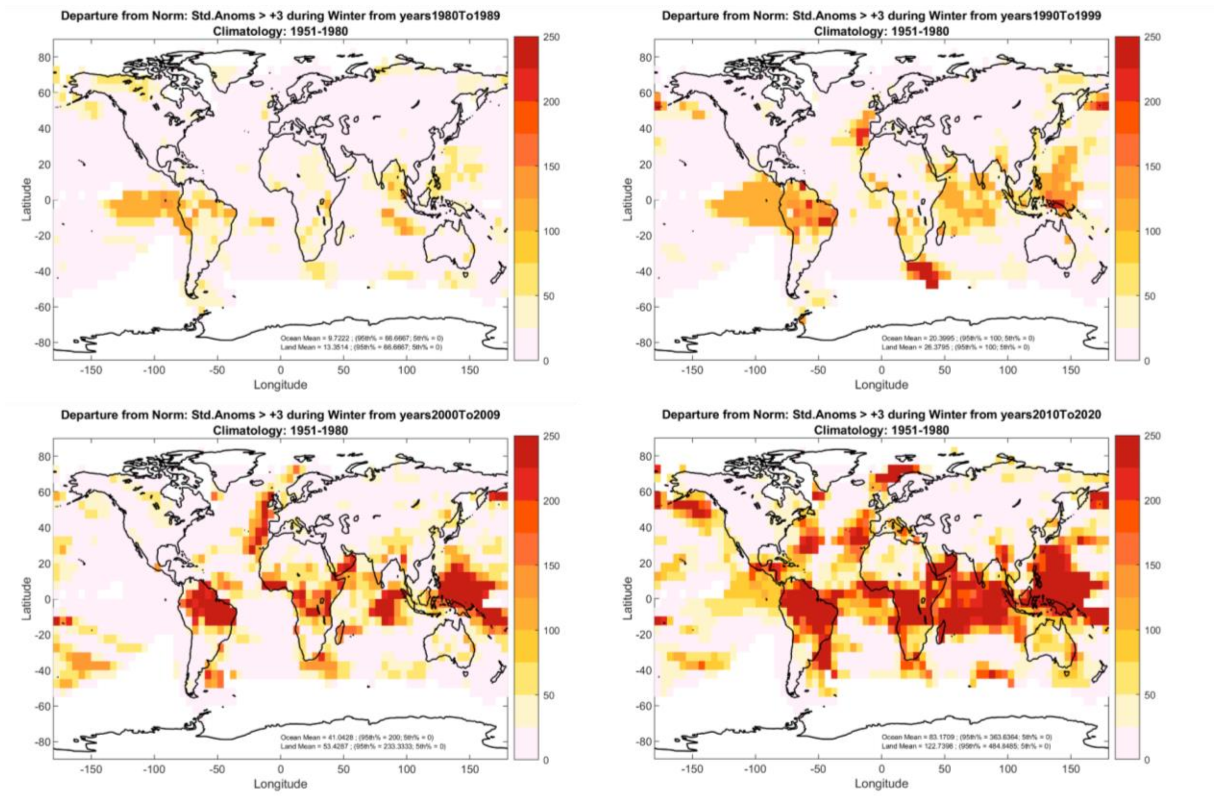

Figure 3.35. As per Figure 3.32, but for 3-sigma events in Winter (DJF) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.36 and 3.37 display Spring (MAM) frequency departure values of 3-sigma events from 1880 to 2020. 3-sigma land and ocean departure values average 3.31 from 1880 to 1979 and average 63.9 from 1980 to 2020. What this means is that from 1880 to 1979, 3-sigma events occurred 3 times more frequently than statistical expectation, and occurred around 64 times more frequently than expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.37) suggests 5% of land and ocean areas experienced 3-sigma events about 100 times more frequently than expected, but from 2010 to 2020, that number increased to 455-636 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 0 (Figure 3.37), implying that at least 5% of land and ocean area experienced no 3-sigma events or that less than 95% of land and ocean area experienced 3-sigma events beyond statistical expectation.

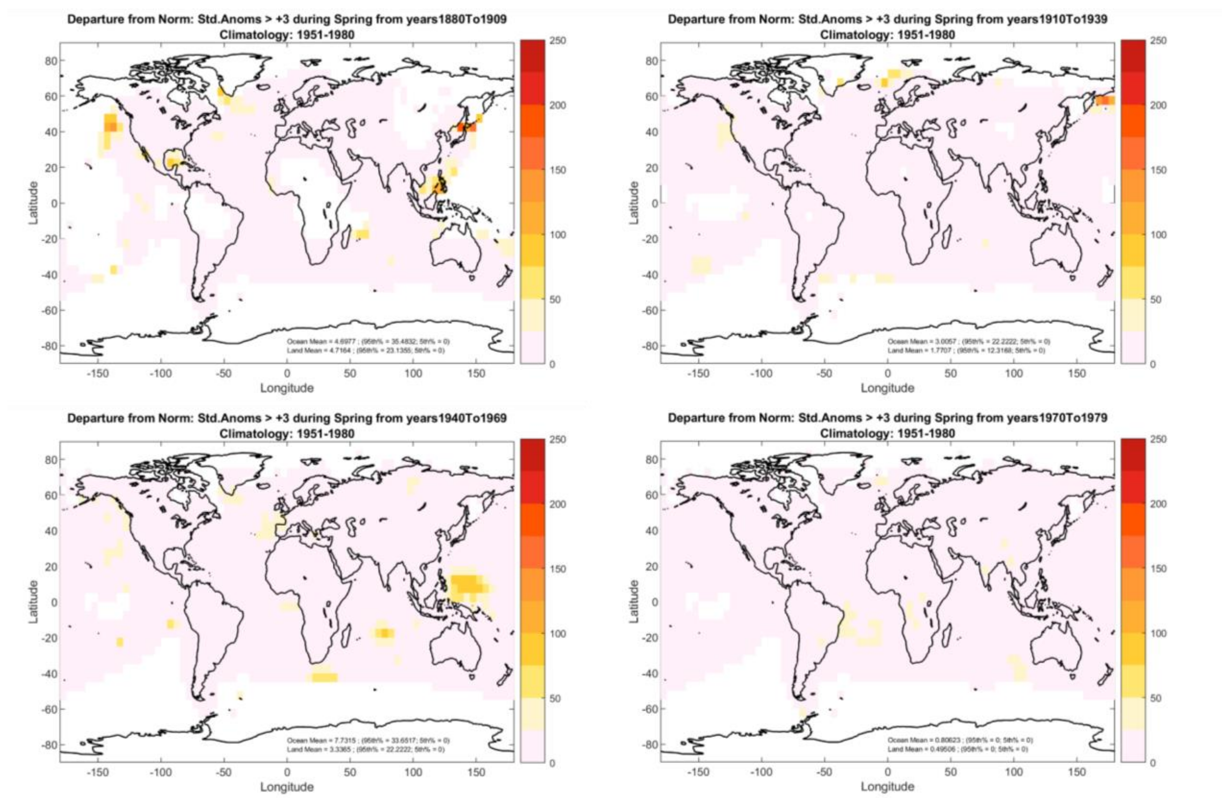

Figure 3.36. As per Figure 3.32, but for 3-sigma events in spring (MAM) from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

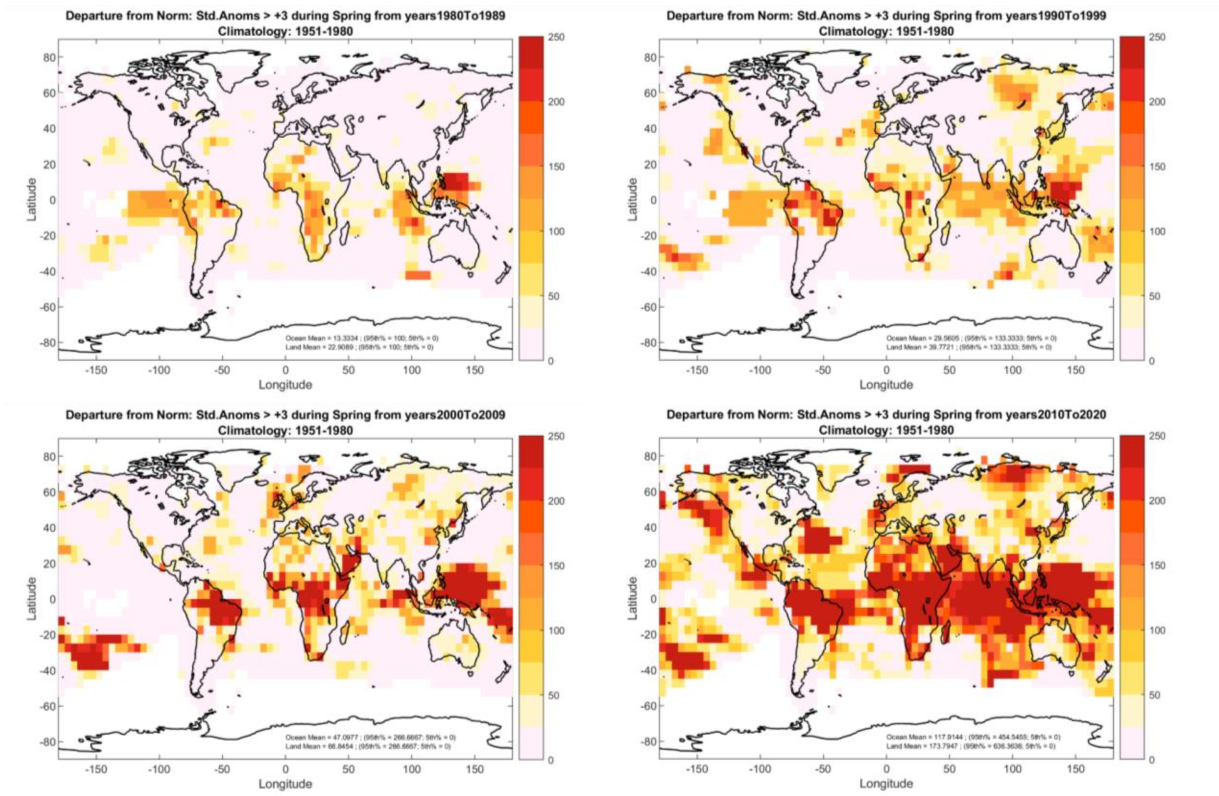

Figure 3.37.As per Figure 3.32, but for 3-sigma events in spring (SON) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figures 3.38 and 3.39 display Summer (JJA) frequency departure values of 3-sigma events from 1880 to 2020. 3-sigma land and ocean departure values average 3.0 from 1880 to 1979 and average 75.0 from 1980 to 2020. What this means is that from 1880 to 1979, 3-sigma events occurred 3 times more frequently than statistical expectation, and occurred 75 times more frequently than expectation from 1980 to 2020. The 95th percentile of frequency departures in the 1980 to 1989 time frame (Figure 3.39) suggests 5% of land and ocean areas experienced 3-sigma events about 100 times more frequently than expected, but from 2010 to 2020, that number increased to 545-697 times expectation. Conversely, from 2010 to 2020, the 5th percentile of frequency departures averaged over land and ocean was 0 (Figure 3.39), implying that at least 5% of land and ocean area experienced no 3-sigma events or that less than 95% of land and ocean area experienced 3-sigma events beyond statistical expectation.

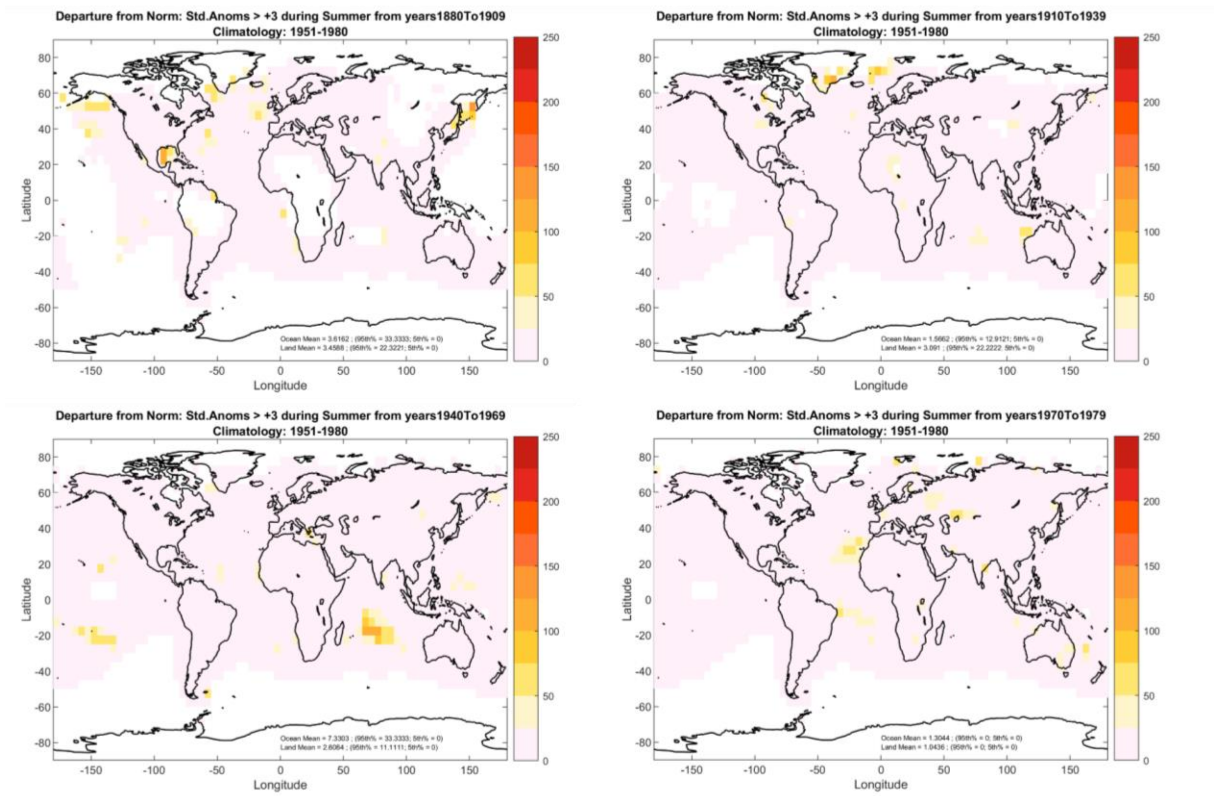

Figure 3.38. As per Figure 3.32, but for 3-sigma events in summer (JJA) from (a) 1880-1909, (b) 1910-1939, (c) 1940-1969 and (d) 1970-1979.

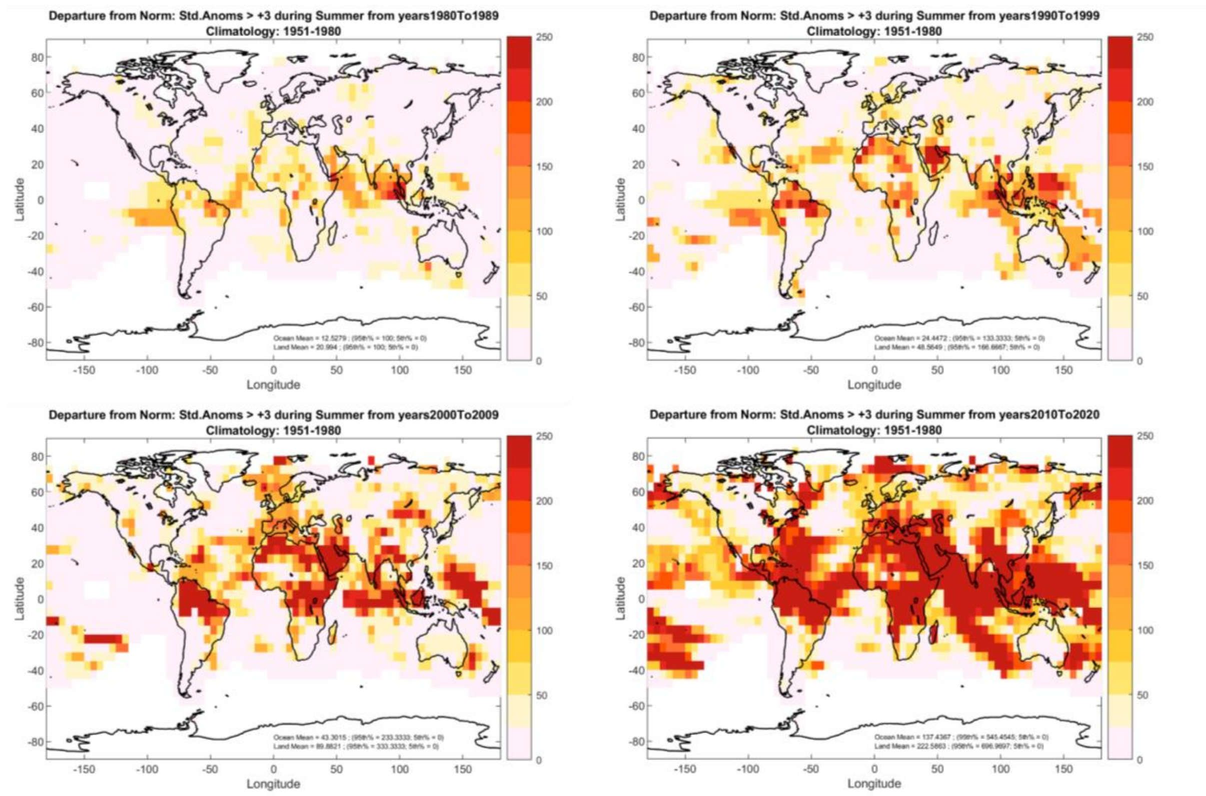

Figure 3.39. As per Figure 3.32, but for 3-sigma events in summer (JJA) from (a) 1980-1989, (b) 1990-1999, (c) 2000-2009 and (d) 2010-2020.

Figure 3.40 shows spatially averaged departure frequencies relative to statistical expectation for 3-sigma events for ocean and land across all seasons and time frames, showing that 3-sigma event frequency departures vary until 1980, but increase exponentially from 1980 to 2020. Like 2-sigma events, land frequency departures exceed those of the ocean, perhaps related to the slower warming rate of oceans overall due to the lapse rate feedback (Colman, 2003; Ceppi et al., 2017). However, both ocean and land results highlight the overall increase in 3-sigma events over time.

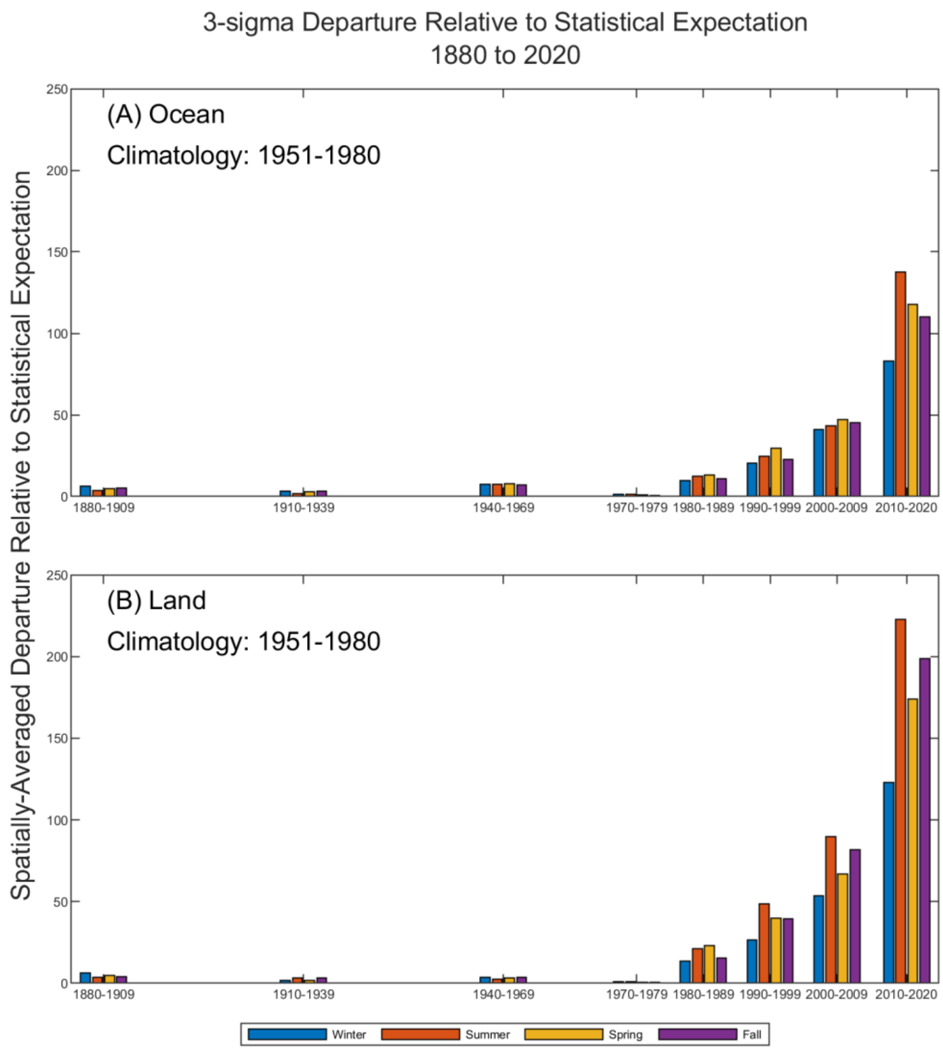

Figure 3.40. Average (a) Ocean and (b) Land spatially-averaged departure frequencies relative to statistical expectation for 3-sigma events applied to distributions of monthly standardized temperature anomalies from 1880 to 2020.

CHAPTER 4: SUMMARY

Within a Gaussian distribution, a 2-sigma event would be roughly considered as a 1 in a 44-year event and a 3-sigma event would be roughly considered as a 1 in a 740-year event. As temperature distributions are positively shifting, the return times of these events are becoming smaller. For example, Figure 2.5 from Chapter 2 shows an example of two distributions of standardized temperature anomalies from one grid box in the southeastern United States for monthly data from each of the Summer months. From 1951-1980 to 2010-2020, the return time for 2-sigma events (a month within the three months of the season, or Summer, in this case) decreased from 44 years to 2 years, and the return time for 3-sigma events decreased from 740 years to 5 years. This means that within that grid cell, 2-sigma events have become a 1 in every 2-year event and 3-sigma events have become a 1 in every 5-year event.

The results of my global analysis of NOAAv5 temperature data show a dramatic increase in the frequency of 2-sigma and 3-sigma temperature events between 1880 and 2020, as well as a corresponding decline of -2-sigma events over the same time period. This is in general agreement with Hansen et al. (2012), but my results extend the analysis an additional 10 years to nearly the present day and are mapped at the spatial resolution of the NOAAv5 data. I also showed that the climatologies of different time frames (10 year and 30 year) remained approximately Gaussian, suggesting that the increases in 2-sigma and 3-sigma frequencies were largely due to an increase in the mean temperature across time frames rather than a change in the skewness of the distributions.

Spatially, observed changes were most intense along the low/mid latitudes, equator and over land surfaces, however, the increase in 2-sigma and 3-sigma events was widespread across the entire globe. Across the equatorial and mid-latitude regions, the frequency of 2-sigma and 3-

sigma temperature events are largest during the last 40 years, with the largest increase in frequency occurring from 2000-2009 to 2010-2020, evident both spatially (Figures 3.24, 3.26, 3.28, 3.30, 3.33, 3.35, 3.37, 3.39) and temporally (Figures 3.31, 3.40).

Across the United States, high frequencies of 2-sigma and 3-sigma events are occurring most over the coastal regions of the country, specifically from 2010 to 2020 (Figures 3.30, 3.39). Over South America, 2-sigma events are widespread in the last 40 years, where 3-sigma events occur mostly in the northern portion of the country. Over Africa and the Middle Eastern region, 2-sigma and 3-sigma events are widespread across the entire continent, especially from 2010-2020. Within Europe, 2-sigma events occurred largely in the last 40 years, however, 3-sigma events were most exacerbated in western Europe from 2010 to 2020. There are larger pockets of 2-sigma and 3-sigma events occurring in Asia, specifically southeast Asia, with the largest increase occurring from 2000-2009 to 2010-2020. Additionally, there are increases in the frequency of 2-sigma and 3-sigma events across the Arctic, perhaps connected to observed Arctic amplification (Serreze et al., 2011). Russia, northern North America, southern South America and Australia seem to tell a different story in that 2-sigma and 3-sigma events occur much less than the rest of the globe.

Across the Pacific Ocean, there are oscillations in the frequency of monthly standardized temperature anomalies, where changes are not consistent spatially (Figures 3.34, 3.30, 3.33, 3.39). This potentially comes from ENSO events that generate colder or warmer sea surface temperatures in the tropical Pacific depending on a La Niña or El Niño year, which is something I could test for with an ENSO index (e.g., Webb and Magi, 2022) by removing years with significant/strong ENSO state from my analysis.

Generally, ocean temperatures are warming at a slower rate than land due to the lapse rate feedback (Colman, 2003; Ceppi et al., 2017). This is evident in the frequency of 2-sigma and 3-sigma events, where the overall land averages are larger than ocean averages across all seasons and time frames (Figures 3.31, 3.40). Because there is uncertainty here, this is something I could work on in the future in order to better understand the influence of the location of the mean temperature of the oceans, which always dominates relative to the spread of the distribution of ocean temperatures.

Regions in the Pacific, Atlantic, and Indian ocean, as well as across Africa, eastern/central Europe, northern/southern South America, and southern North America show low KS p-values, suggesting that those regions temperature distributions are not approximately Gaussian (Figures 3.3, 3.5, 3.7, 3.9). Results for 2-sigma and 3-sigma events are more difficult to interpret when the multiyear climatology is not Gaussian, so the KS test could be used to filter out results where the distributions are no longer sufficiently Gaussian, or weight the averages to favor the grid boxes with distributions that are closer to Gaussian.

My analysis presented the entire planet, or land and ocean surfaces over the entire planet, but future work might help sharpen my analysis by considering seasonal warming in the Northern and Southern Hemispheres separately. This would help isolate the different characteristics of the warm and cold seasons of each hemisphere.

Additionally, future work could analyze sub-regions across the globe. For example, in the last 40 years, Summer months show that there are large pockets of 2-sigma and 3-sigma events occurring across Venezuela and Columbia, as well as China, Vietnam, the Philippines, most countries in Africa and the Middle East (Figures 3.30, 3.39). Looking further into these smaller spatial scales would allow a more in-depth analysis of regional heat extreme patterns.

My results show an increase in the frequency of 2-sigma and 3-sigma events across all seasons, but Summer frequency departures ranged from 4 to 223 (Figures 3.30, 3.39), and Fall from 4 to 199 (Figures 3.24, 3.33), while Spring and Winter frequency departures ranged from 3 to 174 (Figures 3.26, 3.28, 3.35, 3.37). Although Hansen et al. (2012) focused on changes in the distributions of standardized temperature anomalies for Winter and Summer, there are similarities in results. Hansen et al. (2012) observed that there was a significant positive shift during winter and summer. In this study, we do see that shift for both seasons, with Summer being the most extreme. This could be due to Summer climatologies typically having a smaller standard deviation due to less variability in temperature. For -2 -sigma events, Summer experienced the largest decline in frequency (Figures 3.20-3.21). The Summer months are already the warmest months in the Northern Hemisphere, and my results show that the Summer months are now almost exclusively experiencing high sigma events compared to the 1951-1980 climatological base period.

In this project, I used the NOAAv5 Ocean and Land Temperature Anomaly dataset (Huang et al., 2020; Zhang et al., 2019), but my methodology could be applied to any gridded temperature dataset, like the Berkeley Earth global temperature data (Rohde and Hausfather, 2020) or even a temperature reanalysis dataset such as ERA5 (Hersbach et al., 2020). Berkeley Earth has a higher spatial resolution, uses different computational methods to fill in data gaps, and extends further back in time, thus, would be a good tool to use to further replicate the shift in temperature distributions, as well as could be more useful in interpreting the frequency increases of 2-sigma and 3-sigma events given a higher resolution that would refine any spatial patterns. ERA5 is based on observational data, but uses a model to derive sub-daily (reanalysis) temperatures that could reveal trends in daily or day-night temperature extremes, or be combined with analysis of the trends in relative humidity to understand changes in the heat index.

Global warming causes the mean temperature of a given climatology to increase, increasing the chance of extreme heat events relative to the climatology of the past. As shown in this project, the frequency of 2-sigma and 3-sigma events has increased exponentially since 1980. Extreme heat events were once a rare occurrence, but relative to the climatology of the mid-20th century, extreme heat is now a regular occurrence in every season and across most of the world.

REFERENCES

Bland, J.M., Altman, D.G. (1996). "Statistics notes: measurement error". BMJ. 312 (7047): 1654. doi:10.1136/bmj.312.7047.1654. PMC 2351401. PMID 8664723.

Ceppi, P., Brient, F., Zelinka, M.D. and Hartmann, D.L. (2017), Cloud feedback mechanisms and their representation in global climate models. WIREs Clim Change, 8: e465. https://doi.org/10.1002/wcc.465.

Cohen, J., Pfeiffer, K., & Francis, J. A. (2018). Warm Arctic episodes linked with increased frequency of extreme winter weather in the United States. Nature communications, 9(1), 1-12.

Colman, R. (2003). A comparison of climate feedbacks in general circulation models. Climate Dynamics 20, 865–873. https://doi.org/10.1007/s00382-003-0310-z.

Egan, P., Mullin, M. (2016). Recent improvement and projected worsening of weather in the United States. Nature 532, 357–360. https://doi.org/10.1038/nature17441.

Gärtner, L., Schoen, H. (2021). Experiencing climate change: revisiting the role of local weather in affecting climate change awareness and related policy preferences. Climatic Change 167, 31. https://doi.org/10.1007/s10584-021-03176-z.

Gulev, S.K., P.W. Thorne, J. Ahn, F.J. Dentener, C.M. Domingues, S. Gerland, D. Gong, D.S. Kaufman, H.C. Nnamchi, J. Quaas, J.A. Rivera, S. Sathyendranath, S.L. Smith, B. Trewin, K. von Schuckmann, and R.S. Vose. (2021). Changing State of the Climate System. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 287–422, doi: 10.1017/9781009157896.004.

Hansen J, et al. (1988). Global climate changes as forecast by Goddard Institute for Space Studies 3-dimensional model. J Geophys Res Atmos 93:9341–9364.

Hansen, J., Sato, M., & Ruedy, R. (2012). Perception of climate change. Proceedings of the National Academy of Sciences, 109(37), E2415-E2423.

Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. Q J R Meteorol Soc. 2020; 146: 1999– 2049. https://doi.org/10.1002/qj.3803.

Howe, Peter D., Matto Mildenberger, Jennifer R. Marlon, and Anthony Leiserowitz. (2015). “Geographic variation in opinions on climate change at state and local scales in the USA.” Nature Climate Change, doi:10.1038/nclimate2583.

Huang, B., M. J. Menne, T. Boyer, E. Freeman, B. E. Gleason, J. H. Lawrimore, C. Liu, J. J. Rennie, C. Schreck, F. Sun, R. Vose, C. N. Williams, X. Yin, H.-M. Zhang. (2020). Uncertainty estimates for sea surface temperature and land surface air temperature in NOAAGlobalTemp version 5., J. Climate, 33, 1351-1379, DOI: 10.1175/JCLI-D-19-0395.1.

Jay, A., D.R. Reidmiller, C.W. Avery, D. Barrie, B.J. DeAngelo, A. Dave, M. Dzaugis, M. Kolian, K.L.M. Lewis, K. Reeves, and D. Winner. (2018). Overview. In Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II. U.S. Global Change Research Program, Washington, DC, USA, pp. 33–71. doi: 10.7930/NCA4.2018.CH1.

Kunkel, K.E., D.R. Easterling, A. Ballinger, S. Bililign, S.M. Champion, D.R. Corbett, K.D. Dello, J. Dissen, G.M. Lackmann, R.A. Luettich, Jr., L.B. Perry, W.A. Robinson, L.E. Stevens, B.C. Stewart, and A.J. Terando. (2020). North Carolina Climate Science Report. North Carolina Institute for Climate Studies, 233 pp. https://ncics.org/nccsr.

McBean, G. (2004). Climate change and extreme weather: a basis for action. Natural Hazards, 31(1), 177-190.

Mumenthaler, C., Renaud, O., Gava, R., & Brosch, T. (2021). The impact of local temperature volatility on attention to climate change: Evidence from Spanish tweets. Global environmental change, 69, 102286.

NOAA NCEI, https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/, accessed March 2023.

NOAA National Centers for Environmental information. (2023). Climate at a Glance: Regional Time Series, retrieved on March 21, 2023 from https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/regional/time-series.

Painter, J., & Hassol, S. J. (2020). Reporting extreme weather events. In Research handbook on communicating climate change (pp. 183-195). Edward Elgar Publishing.

Perkins-Kirkpatrick, S.E., Lewis, S.C. (2020). Increasing trends in regional heatwaves. Nat Commun 11, 3357. https://doi.org/10.1038/s41467-020-16970-7.

Robinson, A., Lehmann, J., Barriopedro, D. et al. (2021). Increasing heat and rainfall extremes now far outside the historical climate. npj Clim Atmos Sci 4, 45. https://doi.org/10.1038/s41612-021-00202-w.

Rohde, R. A. and Hausfather, Z. (2020). The Berkeley Earth Land/Ocean Temperature Record, Earth Syst. Sci. Data, 12, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020.

Serreze, M. C., & Barry, R. G. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and planetary change*, *77*(1-2), 85-96.

Simolo, C., Corti, S. (2022). Quantifying the role of variability in future intensification of heat extremes. Nat Commun 13, 7930. https://doi.org/10.1038/s41467-022-35571-0.

Sisco, M. R., Bosetti, V., & Weber, E. U. (2017). When do extreme weather events generate attention to climate change?. Climatic change, 143(1), 227-241.

USGCRP. (2018). Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II [Reidmiller, D.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, K.L.M. Lewis, T.K. Maycock, and B.C. Stewart (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, 1515 pp. doi: 10.7930/NCA4.2018.

Webb, E.J. and B. I. Magi (2022). The Ensemble Oceanic Nino Index, International Journal of Climatology, https://doi.org/10.1002/joc.7535.

Wilks, D. S. (2011). Statistical methods in the atmospheric sciences (Vol. 100). Academic press.

Zhang, Huai-Min, Jay H. Lawrimore, Boyin Huang, Matthew J. Menne, Xungang Yin, Ahira Sȧnchez-Lugo, Byron E. Gleason, Russell Vose, Derek Arndt, J. Jared Rennie, and Claude N. Williams. (2019). "Updated Temperature Data Give a Sharper View of Climate Trends." Eos, 100, https://doi.org/10.1029/2019EO128229. Published on 19 July 2019.

Zhang, H.-M., B. Huang, J. Lawrimore, M. Menne, Thomas M. Smith, NOAA Global Surface Temperature Dataset (NOAAGlobalTemp), Version 5. NOAA National Centers for Environmental Information. doi:10.25921/9qth-2p70, accessed September 2021.

APPENDIX A: ANALYSIS CODE

Below is a matlab script used to retrieve and process NOAAv5 data. In this case, script A is used in conjunction with Script B and Script C, which will follow. Script B is used for any calculations, while Script C creates global maps/figures.

Script A:

```
function [output] = A_GetAndProcessNOAAV5Data();
%
% [NOAAV5Data] = A_GetAndProcessNOAAV5Data;
% save('NOAAV5Data.mat','NOAAV5Data','-mat');
% load('NOAAV5Data.mat'); % loads structure NOAAV5Data into workspace
%
%Declaring Variables from NOAA dataset
% % % fileName =
'W:\data\climate\NOAAGlobalTemp\NOAAGlobalTemp_v5.0.0_gridded_s188001_e202
209_c20221008T133300.nc';
% % % time = ncread(fileName, 'time');
% % % lat = ncread(fileName, 'lat');
% % % lon = ncread(fileName, 'lon');
% % % global_temp_anom = ncread(fileName, 'anom');
time = ncread('NOAAGlobalTemp.nc', 'time');
lat = ncread('NOAAGlobalTemp.nc', 'lat');
lon = ncread('NOAAGlobalTemp.nc', 'lon');
global_temp_anom = ncread('NOAAGlobalTemp.nc', 'anom');

%Permute global anomalies to move z to the 4th dimension
tempAnom = permute(global_temp_anom, [4,2,1,3]);

% Attempt to transfer data from gregorian time
decimalMonth = double(time);
yearRange = [1880:2021]';
% % % yearRange = [1880:2022]';
[timeMatrix] = BuildTimeMatrix(yearRange,'monthly');
yyyymm = [timeMatrix(1:length(decimalMonth),:)];
decimalYears = yyyymm(:,1)+(yyyymm(:,2)-1)/12;

%Flip Latitude for the data and the coordinate
tempAnom = flipdim(tempAnom,2);
lat = flipud(lat);

%Reconfigure Longitude for data and the coordinate
halfway = size(tempAnom,3)/2;
wholeway = size(tempAnom,3);
east_Hem = tempAnom(:,:,1:halfway);
```

```
west_Hem = tempAnom(:,:,halfway+1:wholeway);
tempAnom = cat(3,west_Hem,east_Hem);
temperature.gridded = tempAnom;
lonEH = lon(1:halfway);
lonWH = lon(halfway+1:wholeway)-360;

%creating final lon vector that spans -180 to 180
lon = cat(1,lonWH,lonEH);

load('latlongrids.mat');  % load('C:/Users/Bill.VanOrmer/Dropbox/Graduate
Research/latlongrids.mat');
gridValues = latlongrids.r36x72;
landarea = latlongrids.r36x72.larea; % km2
gridarea = latlongrids.r36x72.garea; % km2

climYears = [1951,1980]; %%%
startClim = find(yyyymm(:,1)==min(climYears)&yyyymm(:,2)==1); %%%
stopClim = find(yyyymm(:,1)==max(climYears)&yyyymm(:,2)==12); %%%
subsetClim = temperature.gridded(startClim:stopClim,:,:);

% cycle over months and extra the month's climatology
for ii=1:12;
   subsetMonth = subsetClim(ii:12:end,:,:);
   temperature.climatology.mean(ii,:,:) = mean(subsetMonth,1); %%%
   temperature.climatology.std(ii,:,:) = std(subsetMonth,1);
end;  % for ii loop
temperature.climatology.climYears = climYears; %%%
temperature.climatology.startClim = startClim; %%%
temperature.climatology.stopClim = stopClim; %%%
temperature.climatology.subsetClim = subsetClim; %%%
temperature.climatology.originalData = temperature.gridded; %%%

anomReconfig = NaN.*tempAnom;
stdanomReconfig = NaN.*tempAnom;
%Loop to standardize anoms
for kk = 1:12;
   oneMonthTSOriginal = temperature.climatology.originalData(kk:12:end,:,:);
   % re-configure and re-zero to a new reference period per Boyin
   % Huang's emails (subtract off the monthly climatology)
   oneMonthTSAnomReconfig = oneMonthTSOriginal-
temperature.climatology.mean(kk,:,:);
   oneMonthTSStdAnomReconfig =
oneMonthTSAnomReconfig./temperature.climatology.std(kk,:,:);
   anomReconfig(kk:12:end,:,:) = oneMonthTSAnomReconfig;
   stdanomReconfig(kk:12:end,:,:) = oneMonthTSStdAnomReconfig;
end;
```

```
% % % std_anoms = NaN*tempAnom;
% % % %Loop to standardize anoms
% % % for kk = 1:12;
% % %     month_one_ = tempAnom(kk:12:end,:,:);
% % %     % re-configure and re-zero to a new reference period per Boyin
% % %     % Huang's emails (subtract off the monthly climatology)
% % %     month_one_ = month_one_ - temperature.climatology.mean(kk,:,:); %%%
% % %     tempAnom(kk:12:end,:,:) = month_one_; %%%
% % %     month_one_ = month_one_./temperature.climatology.std(kk,:,:);
% % %     std_anoms(kk:12:end,:,:) = month_one_ ;
% % % end;
% % % std_anoms = NaN*tempAnom;
% % % %Loop to standardize anoms
% % % for ii = 1:length(lat); %cycle over latitudes
% % %     for jj = 1:length(lon); %cycle over longitudes
% % %         for kk = 1:12;
% % %             month_one_ = tempAnom(kk:12:end,ii,jj);
% % %             % re-configure per Boyin Huang's emails
% % %             month_one_ = month_one_ - temperature.climatology.mean(kk,ii,jj);
% % %             month_one_ = month_one_/temperature.climatology.std(kk,ii,jj);
% % %             std_anoms(kk:12:end,ii,jj) = month_one_ ;
% % %         end;
% % %     end;
% % % end;
output.time = time;
output.lat = lat;
output.lon = lon;
output.tempAnom = anomReconfig; % tempAnom; %%%
output.std_anoms = stdanomReconfig; % std_anoms;  %%%
output.yyyymm = yyyymm;
output.landarea = landarea;
output.gridarea = gridarea;
output.climatology = temperature.climatology; %%%
end % main

function [output] = BuildTimeMatrix(yearValues,timeResolution);
%
%  Build a matrix of month, day, daynumber, hour of day.  The dimensions
%  are rows = day number with all relevant time information and for hourly
%  then the rows are day number with decimal hour
%
leapYears = 1860:4:2040;
% cycle over the years
for ii=1:length(yearValues);
   clear yyyymm yyyymmdd yyyymmddhh
```

```
    % check to see if we're in a leap year
    if isempty(find(leapYears==yearValues(ii)));
        % not a leap year
        daysinmonth = [31,28,31,30,31,30,31,31,30,31,30,31];
    else
        % leap year
        daysinmonth = [31,29,31,30,31,30,31,31,30,31,30,31];
    end;  % if isempty loop
    daytotal = [1,cumsum(daysinmonth)];
    daycount = 1;
    dayhourcount = 1;
    % cycle over months
    for jj=1:length(daysinmonth);
        yyyymm(jj,:) = [yearValues(ii),jj];
        % cycle over days in that month
        for kk=1:daysinmonth(jj);
            % yyyymmdd(daycount,:) = [daycount,yearValues(ii),jj,kk,0,0,0];
            yyyymmdd(daycount,:) = [daycount,yearValues(ii),jj,kk];
            % cycle over hours in day
            for mm=0:23;
                decimalday = daycount+mm/24;
                yyyymmddhh(dayhourcount,:) = [decimalday,yearValues(ii),jj,kk,mm];
                % yyyymmddhh(dayhourcount,:) = [decimalday,yearValues(ii),jj,kk,mm,0,0];
                dayhourcount = dayhourcount+1;
            end;  % for mm loop
            daycount = daycount+1;
        end;  % for kk loop
    end;  % for jj loop
    % % %     % cycle over seasons
    % % %     seasons = {(12,1,2),(3:5),(6:8),(9:11)};
    yyyy(ii,1) = yearValues(ii);
    if ii==1;
        timeMatrices.yyyymm = yyyymm;
        timeMatrices.yyyymmdd = yyyymmdd;
        timeMatrices.yyyymmddhh = yyyymmddhh;
    else
        timeMatrices.yyyymm = cat(1,timeMatrices.yyyymm,yyyymm);
        timeMatrices.yyyymmdd = cat(1,timeMatrices.yyyymmdd,yyyymmdd);
        timeMatrices.yyyymmddhh = cat(1,timeMatrices.yyyymmddhh,yyyymmddhh);
    end;  % if ii loop
end;
switch timeResolution
    case 'hourly'
        output = timeMatrices.yyyymmddhh;
    case 'daily'
        output = timeMatrices.yyyymmdd;
```

```
    case 'monthly'
        output = timeMatrices.yyyymm;
    case 'seasonal'
        disp('   Seasonal time matrix is not built, but seasonal averages can be calculated
within scripts');
        output = [];
    case 'annual'
        output = yyyy;
    otherwise
        disp('   Data is sub-hourly. Time matrix will not be built at finer scales than hourly');
        output = [];
end;  % switch loop
end % function
```

Script B:

```
function [output] = B_EvaluateStandardizedAnomalies(input);
%
% load('NOAAV5Data.mat'); % loads structure NOAAV5Data into workspace
% [NOAAV5Data] = B_EvaluateStandardizedAnomalies(NOAAV5Data);
%
output = input;
if nargin==0;
    error('More efficient to run A_GetAndProcessNOAAV5Data.m separately, save the
output, and compile this script with pre-processed data')
    [NOAAV5Data] = A_GetAndProcessNOAAV5Data;
else
    NOAAV5Data = output;
    time = NOAAV5Data.time;
    lat = NOAAV5Data.lat;
    lon = NOAAV5Data.lon;
    tempAnom = NOAAV5Data.tempAnom;
    std_anoms = NOAAV5Data.std_anoms;
    yyyymm = NOAAV5Data.yyyymm;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
    % Brian notes: the ocean and land mask may be useful down the road when
    % we are looking at difference in seasonal changes between major parts
    % of the world (ala Hansen et al 2012 approach)
    landarea = NOAAV5Data.landarea;
    gridarea = NOAAV5Data.gridarea;
    oceanmask2D = ones(length(lat),length(lon));
    oceanmask2D(find(landarea>0)) = NaN;
    landmask2D = ones(length(lat),length(lon));
    landmask2D(find(landarea==0)) = NaN;
```

```
%     std_anoms_land =
std_anoms.*repmat(permute(oceanmask2D,[3,1,2]),[length(yyyymm),1,1]);
%     std_anoms_ocean =
std_anoms.*repmat(permute(landmask2D,[3,1,2]),[length(yyyymm),1,1]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
end; % if nargin

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
% Brian added this code
% this if statement tests for the existence of a field in the input
% structure that is called 'analysis' and then removes it so that each time
% you run this script, it's a fresh slate.
if isfield(NOAAV5Data,'analysis');
   disp('removing the analysis fields in the input structure so it can be re-written');
   NOAAV5Data = rmfield(NOAAV5Data,'analysis');
end;  % if isfield
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

%%%%%%% Beginning of loop/configuration for std anom, region, year, season

% Brian notes: region names may be unneeded going forward since we are
% studying global maps
% % % regionNames = {'US','SEUS','NEUS','SWUS','NWUS','NSA','SSA'};

%Season Declaration
seasonNames = {'Winter','Summer','Spring','Fall'};
seasonbox.Winter = [12,1,2];
seasonbox.Summer = [6,7,8];
seasonbox.Spring = [3,4,5];
seasonbox.Fall = [9,10,11];

%%% Distribution visual loop ------
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
randlat = randperm(length(lat));
randlon = randperm(length(lon));
randlat = [12];
randlon = [19];
% % % disp(lat(randlat));
% % % disp(lon(randlon));
% % % pause;
```

```
%Start of for loop
for ii=1:length(randlat);
   for mm=1:3; % length(randlon); % run for 1st 3 random longitudes (or all longitude if
you use length(randlon))
% % % for ii=1:length(lat);
% % %     for mm=1:length(lon);
        %Cycle over seasons
        for jj=2; %1:length(seasonNames);
            seasonName = seasonNames{jj};
            season1 = seasonbox.(seasonName)(1);
            season2 = seasonbox.(seasonName)(2);
            season3 = seasonbox.(seasonName)(3);
            yearChunks =
[1880,1909;1910,1939;1940,1969;1970,1979;1980,1989;1990,1999;2000,2009;2010,202
0];
            yearRanges =
{'years1880To1909','years1910To1939','years1940To1969','years1970To1979','years198
0To1989','years1990To1999','years2000To2009','years2010To2020'};
% % %             % confirms that the distribution for the chosen base period is
% % %             % indeed mean of 0 and std of 1 (per script A_)
% % %             yearChunks = [1951,1980];
% % %             yearRanges = {'years1951To1980'};
            %Cycle over years
            for kk=1:length(yearChunks(:,1));
                yearChunk = yearChunks(kk,:);
                numYearsInYearChunk = max(yearChunk)-min(yearChunk)+1; %%% number
of years in year chunk
                numMonthsInYearChunk = numYearsInYearChunk*3; %%% always assuming
3 month season!
                yearRange = yearRanges{kk};
                dateMatches = find((yyyymm(:,2)==season1 | yyyymm(:,2)==season2
|yyyymm(:,2)==season3) &
yyyymm(:,1)>=yearChunk(1)&yyyymm(:,1)<=yearChunk(2));
std_anomsMatches = std_anoms(dateMatches,randlat(ii),randlon(mm));
                analysis.(seasonName).(yearRange).std_anoms = std_anomsMatches(:);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
                % Brian added this code
                % need to keep track of the number of months that have data
                % within the season subset of the time chunk as the
                % normalization constant per lat-lon gridbox so that places
                % in the world that have no data or little data can be
                % excluded from the analysis. The functions in the line are
                % using isnan(), find(), and length() in combination to
                % count the number of months that are *not* NaN (ie. months
```

```
		% with data)
		numMonthsWithDataInTimeChunk =
length(find(isnan(analysis.(seasonName).(yearRange).std_anoms)==0));
		fracMonthsWithDataInTimeChunk =
numMonthsWithDataInTimeChunk/numMonthsInYearChunk;  %%% fraction of months
with data (in time chunk)
% % %
analysis.(seasonName).(yearRange).numMonthsWithDataInTimeChunk(ii,mm) =
numMonthsWithDataInTimeChunk;
% % %
analysis.(seasonName).(yearRange).fracMonthsWithDataInTimeChunk(ii,mm) =
fracMonthsWithDataInTimeChunk;  %%% preserve the output

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

		%Anomalies greather than 2
		anoms_greater_two_field =
find(analysis.(seasonName).(yearRange).std_anoms>2);
		anoms_greater_two_values =
analysis.(seasonName).(yearRange).std_anoms(anoms_greater_two_field);
		anoms_greater_two_count = length(anoms_greater_two_values);
% % %		analysis.(seasonName).(yearRange).anoms_greater_two_count(ii,mm)
= anoms_greater_two_count;
		%Anomalies less than -2
		anoms_less_two_field = find(analysis.(seasonName).(yearRange).std_anoms<-
2);
		anoms_less_two_values =
analysis.(seasonName).(yearRange).std_anoms(anoms_less_two_field);
		anoms_less_two_count = length(anoms_less_two_values);
% % %		analysis.(seasonName).(yearRange).anoms_less_two_count(ii,mm) =
anoms_less_two_count;
		%Anomalies greater than + 3
		anoms_greater_three_field =
find(analysis.(seasonName).(yearRange).std_anoms>3);
		anoms_greater_three_values =
analysis.(seasonName).(yearRange).std_anoms(anoms_greater_three_field);
		anoms_greater_three_count = length(anoms_greater_three_values);
% % %
analysis.(seasonName).(yearRange).anoms_greater_three_count(ii,mm) =
anoms_greater_three_count;
		%Calculating the IQR for the dataset
		iqr_anom = iqr(analysis.(seasonName).(yearRange).std_anoms);
% % %		analysis.(seasonName).(yearRange).iqr_anom(ii,mm) = iqr_anom;
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
        % Brian added this code
        % Normalization of all variables to a frequency in fraction
        % of a season with condition X within the time chunk; The
        % first normalization accounts for the variable number of
        % months at a particular lat/lon with data, and the 2nd
        % normalization accounts for the subset of a year that we
        % are studying. frequency of X = X/N * 3 months/season * 1
        % season/year;
        %
        % a 2nd layer is deciding how many months within a time
        % chunk need data for a "valid" assessment. For example, 20
        % years with 3 month seasons means 60 possible months with
        % an analysis condition X. What if only 5 months have data?
        % Is that a fair comparison with a time chunk that has 60
        % months of data? We might have to see about this, but 15
        % is a round number starting point, such that any time
        % chunk with fewer than that many months of data is
        % excluded from the analysis (turned into a NaN in the
        % frequence calculation). 15 is also 5 years of 3-month
        % seasons, hence "round number" but it might be too small a
        % minimum, i'm not sure.

        minNumMonthsWithDataInTimeChunk = 50;
        minPercentMonthsWithDataInTimeChunk = 0.80;  %%% new percent based
threshold

        % CHANGE THIS TO BE BASED ON fracMonthsWithDataInTimeChunk
        % vs minPercentMonthsWithDataInTimeChunk

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        if
fracMonthsWithDataInTimeChunk>=minPercentMonthsWithDataInTimeChunk;


% % %
analysis.(seasonName).(yearRange).frequency_anoms_greater_two_count(ii,mm) = ...
% % %              anoms_greater_two_count/numMonthsWithDataInTimeChunk;
% % %
analysis.(seasonName).(yearRange).frequency_anoms_less_two_count(ii,mm) = ...
% % %              anoms_less_two_count/numMonthsWithDataInTimeChunk;
% % %
analysis.(seasonName).(yearRange).frequency_anoms_greater_three_count(ii,mm) = ...
% % %              anoms_greater_three_count/numMonthsWithDataInTimeChunk;
```

```
%Metric calculation
kstestInput = analysis.(seasonName).(yearRange).std_anoms-
nanmean(analysis.(seasonName).(yearRange).std_anoms);
[h,p] = kstest(kstestInput);
analysis.(seasonName).(yearRange).p(ii,mm) = p;
            %Calculating mean/median ratio
            mean_data = nanmean(analysis.(seasonName).(yearRange).std_anoms);
            median_data = nanmedian(analysis.(seasonName).(yearRange).std_anoms);
            gaussian_check = abs(mean_data / median_data);
% % %               analysis.(seasonName).(yearRange).gaussian_check(ii,mm) =
gaussian_check;
disp(iqr_anom)
disp(p)
disp((anoms_greater_three_count/numMonthsWithDataInTimeChunk)/0.0013)
disp((anoms_greater_two_count/numMonthsWithDataInTimeChunk)/0.022)

[hNonZeroedMean,pNonZeroedMean] =
kstest(analysis.(seasonName).(yearRange).std_anoms);
%
% this commandline execution shows how the distributions are not
% always zero mean with 1 std before the data is re-centered in script A_
% according to email from Boyin Huang
% % % monthNum = 3; figure(2); clf reset; subplot(1,2,1);
pcolor(NOAAV5Data.lon,NOAAV5Data.lat,permute(NOAAV5Data.climatology.mean(
monthNum,:,:),[2,3,1])); shading flat; colorbar; subplot(1,2,2);
pcolor(NOAAV5Data.lon,NOAAV5Data.lat,permute(NOAAV5Data.climatology.std(mo
nthNum,:,:),[2,3,1])); shading flat; colorbar;
%
% this commandline execution shows the return time of different sigma
% events for a normal distribution. 3sigma ~ 1000 year event, for example
% % % XVals = [-6:0.1:0]; blah = normcdf(XVals); figure(1); clf reset;
semilogy(XVals,1./blah);
%
% this command shows how likely it would be to have an event occur at
% 2 (2sigma in the case of stdanom distribution around a common base
% period) or greater for a dist with mean zero and std 1
% % % blah = normcdf(2,0,1,'upper');
%
% I wonder if it'd be *better* to discuss the frequency changes based on
% the distribution *indicated* by the obs rather than the obs themselves.
% Wish I had a statistician I could rely on. Specifically, rather than
% relying on the relatively limited seasonal obs over the time chunk, would
% it be better to use the distribution itself as given by the mean, std,
% and an assessment of how "gaussian" the sample population actually is?
% And then assess how population distribution indicates how much things
```

```
% have shifted into higher sigma anomalies?
%
std_data = nanstd(analysis.(seasonName).(yearRange).std_anoms);
min_data = min(analysis.(seasonName).(yearRange).std_anoms);
max_data = max(analysis.(seasonName).(yearRange).std_anoms);
figure(50); clf reset;
Hh =
histogram(analysis.(seasonName).(yearRange).std_anoms,'normalization','probability','bi
nwidth',0.5,'binlimits',[-8,8]);
HhN = histogram('BinEdges',Hh.BinEdges, 'BinCounts',Hh.Values./max(Hh.Values));
set(HhN,'facecolor',[1,0.8,0.8],'edgecolor',[1,0.7,0.7]);
hold on;
plotXRange = [-6:0.1:6];
plotYValues = normpdf(plotXRange,mean_data,std_data);
plot(plotXRange,plotYValues./max(plotYValues),'linewidth',2,'color','r');
plotBasePeriodValues = normpdf(plotXRange,0,1);
plot(plotXRange,plotBasePeriodValues./max(plotBasePeriodValues),'linewidth',2,'color','
k');
Ht = title(cat(2,seasonName,' from ',yearRange,' for (lat,lon) =
(',num2str(lat(randlat(ii))),', ',num2str(lon(randlon(mm))),')'));
Hx = xlabel('Standardized Temperature Anomaly');
Hy = ylabel('Normalized Frequency');
% textLine2 = (cat(2,'3-Sigma Events =
',num2str((anoms_greater_three_count/numMonthsWithDataInTimeChunk)/0.0013)));
% textLine1 = (cat(2,'2-Sigma Events =
',num2str((anoms_greater_two_count/numMonthsWithDataInTimeChunk)/0.022)));
%
% textLine3 = (cat(2,'KS Test P-Val = ',num2str(p)));
% textLine4 = (cat(2,'IQR = ',num2str(iqr_anom)));
% Htext = text(-5.6,0.9,[{textLine1};{textLine2};{textLine3};{textLine4}]);
set([Ht,Hx,Hy],'fontsize',14);
GT2Frequency = normcdf(2,mean_data,std_data,'upper');
GT2FrequencyBase = normcdf(2,0,1,'upper');
GT3Frequency = normcdf(3,mean_data,std_data,'upper');
GT3FrequencyBase = normcdf(3,0,1,'upper');
blah = 1/normcdf(1.28155,0,1,'upper') %is a 1 in a 10 year frequency (1.28155 sigma)
blah = 1/normcdf(2.32635,0,1,'upper') %is a 1 in a 100 year frequency (2.32635 sigma)
blah = 1/normcdf(3.09024,0,1,'upper') %is a 1 in a 1000 year frequency (3.09024 sigma)
RT10Frequency = normcdf(1.28155,mean_data,std_data,'upper');
RT10FrequencyBase = normcdf(1.28155,0,1,'upper');
RT100Frequency = normcdf(2.32635,mean_data,std_data,'upper');
RT100FrequencyBase = normcdf(2.32635,0,1,'upper');
RT1000Frequency = normcdf(3.09024,mean_data,std_data,'upper');
RT1000FrequencyBase = normcdf(3.09024,0,1,'upper');
textString = cat(1,{cat(2,'KS p = ',num2str(p))}, ...
    {' '}, ...
```

```
    {cat(2,'>2Sigma Events = ',num2str(GT2Frequency/GT2FrequencyBase))}, ...
%     {cat(2,'>2Sigma return time in base period = ',num2str(1/GT2FrequencyBase))}, ...
%     {cat(2,'>2Sigma return time = ',num2str(1/GT2Frequency))}, ...
    {' '}, ...
    {cat(2,'>3Sigma Events = ',num2str(GT3Frequency/GT3FrequencyBase))}, ...
%     {cat(2,'>3Sigma return time in base period = ',num2str(1/GT3FrequencyBase))}, ...
%     {cat(2,'>3Sigma return time = ',num2str(1/GT3Frequency))}, ...
%     {' '}, ...
%     {cat(2,'10yr return time = ',num2str(1/RT10Frequency))}, ...
%     {cat(2,'100yr return time = ',num2str(1/RT100Frequency))}, ...
%     {cat(2,'1000yr return time = ',num2str(1/RT1000Frequency))}, ...
%     {' '}, ...
%     {cat(2,'mean and std = ',num2str(mean_data),' \pm ',num2str(std_data))});
Htxt = text(-5.7,0.8,textString);
set(gca,'xlim',[min(plotXRange),max(plotXRange)]); % ,'xtick',[1860:20:2020]); %
,'xdir','normal'
set(gca,'ylim',[0,1.1],'ytick',[0:0.2:1]);
disp(cat(2,' p = ',num2str(p),' (pNonZeroedMean = ',num2str(pNonZeroedMean),') '));
disp(cat(2,' mean = ',num2str(mean_data),', std = ',num2str(std_data),', min =
',num2str(min_data),', max = ',num2str(max_data),' (pausing)'));
pause;

                else
% % %
analysis.(seasonName).(yearRange).frequency_anoms_greater_two_count(ii,mm) = NaN;
% % %
analysis.(seasonName).(yearRange).frequency_anoms_less_two_count(ii,mm) = NaN;
% % %
analysis.(seasonName).(yearRange).frequency_anoms_greater_three_count(ii,mm) =
NaN;
% % %
% % %                     analysis.(seasonName).(yearRange).iqr_anom(ii,mm)= NaN;
% % %                     analysis.(seasonName).(yearRange).gaussian_check(ii,mm)= NaN;
% % %                     analysis.(seasonName).(yearRange).p(ii,mm) = NaN;


                end; % if numMonthsWithDataInTimeChunk

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

            end; % year chunks
        end; % season
    end; % lon
end; % lat
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%Start of for loop for normal analysis
for ii=1:length(lat);
    for mm=1:length(lon);
        %Cycle over seasons
        for jj=1:length(seasonNames);
            seasonName = seasonNames{jj};


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
            % Brian added this code
            % assumes that there are always 3 months in a season; would
            % need to change the code if that changes, but add this element
            % to double-check and remind of this "assumption"
            if length(seasonbox.(seasonName))~=3;
                error('code requires 3 months in a season');
            else
                % this may be useful at some point in the code but maybe
                % not
                numMonthsInSeason = length(seasonbox.(seasonName));
            end; % if length

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

            season1 = seasonbox.(seasonName)(1);
            season2 = seasonbox.(seasonName)(2);
            season3 = seasonbox.(seasonName)(3);
            yearChunks =
[1880,1909;1910,1939;1940,1969;1970,1979;1980,1989;1990,1999;2000,2009;2010,202
0];
            yearRanges =
{'years1880To1909','years1910To1939','years1940To1969','years1970To1979','years198
0To1989','years1990To1999','years2000To2009','years2010To2020'};
            %Cycle over years
            for kk=1:length(yearChunks);
                yearChunk = yearChunks(kk,:);
                numYearsInYearChunk = max(yearChunk)-min(yearChunk)+1; %%% number
of years in year chunk
                numMonthsInYearChunk = numYearsInYearChunk*3; %%% always assuming
3 month season!
```

```
            yearRange = yearRanges{kk};
            dateMatches = find((yyyymm(:,2)==season1 | yyyymm(:,2)==season2
|yyyymm(:,2)==season3) &
yyyymm(:,1)>=yearChunk(1)&yyyymm(:,1)<=yearChunk(2));
            std_anomsMatches = std_anoms(dateMatches,ii,mm); % std_anoms_land for
just land
%               std_anomsMatchesLand = std_anoms_land(dateMatches,ii,mm);
%               std_anomsMatchesOcean = std_anoms_ocean(dateMatches,ii,mm);


            analysis.(seasonName).(yearRange).std_anoms = std_anomsMatches(:);
%               analysis.(seasonName).(yearRange).std_anoms_land =
std_anomsMatchesLand(:);
%               analysis.(seasonName).(yearRange).std_anoms_ocean =
std_anomsMatchesOcean(:);
%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
            % Brian added this code
            % need to keep track of the number of months that have data
            % within the season subset of the time chunk as the
            % normalization constant per lat-lon gridbox so that places
            % in the world that have no data or little data can be
            % excluded from the analysis. The functions in the line are
            % using isnan(), find(), and length() in combination to
            % count the number of months that are *not* NaN (ie. months
            % with data)
            numMonthsWithDataInTimeChunk =
length(find(isnan(analysis.(seasonName).(yearRange).std_anoms)==0));
            fracMonthsWithDataInTimeChunk =
numMonthsWithDataInTimeChunk/numMonthsInYearChunk;  %%% fraction of months
with data (in time chunk)
            % snippet below (% % %) can be deleted - it was mostly a
            % numerical check to make sure the script was counting what
            % i thought it was counting
% % %             if numMonthsWithDataInTimeChunk>100;
% % %                 disp(cat(2,'num good months =
',num2str(numMonthsWithDataInTimeChunk), ...
% % %                     ' and num months = ',num2str(length(dateMatches)), ...
% % %                     ' for ',seasonName,' in ',yearRange));
% % %                 pause;
% % %             end; % if
            analysis.(seasonName).(yearRange).numMonthsWithDataInTimeChunk(ii,mm)
= numMonthsWithDataInTimeChunk;
```

```
            analysis.(seasonName).(yearRange).fracMonthsWithDataInTimeChunk(ii,mm)
= fracMonthsWithDataInTimeChunk;  %%% preserve the output

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

            %Anomalies greather than 2
            anoms_greater_two_field =
find(analysis.(seasonName).(yearRange).std_anoms>2);
            anoms_greater_two_values =
analysis.(seasonName).(yearRange).std_anoms(anoms_greater_two_field);
            anoms_greater_two_count = length(anoms_greater_two_values);
            analysis.(seasonName).(yearRange).anoms_greater_two_count(ii,mm) =
anoms_greater_two_count;
            %Anomalies less than -2
            anoms_less_two_field = find(analysis.(seasonName).(yearRange).std_anoms<-
2);
            anoms_less_two_values =
analysis.(seasonName).(yearRange).std_anoms(anoms_less_two_field);
            anoms_less_two_count = length(anoms_less_two_values);
            analysis.(seasonName).(yearRange).anoms_less_two_count(ii,mm) =
anoms_less_two_count;
            %Anomalies greater than + 3
            anoms_greater_three_field =
find(analysis.(seasonName).(yearRange).std_anoms>3);
            anoms_greater_three_values =
analysis.(seasonName).(yearRange).std_anoms(anoms_greater_three_field);
            anoms_greater_three_count = length(anoms_greater_three_values);
            analysis.(seasonName).(yearRange).anoms_greater_three_count(ii,mm) =
anoms_greater_three_count;
            %Calculating the IQR for the dataset
            iqr_anom = iqr(analysis.(seasonName).(yearRange).std_anoms);
            analysis.(seasonName).(yearRange).iqr_anom(ii,mm) = iqr_anom;



%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
            % Brian added this code
            % Normalization of all variables to a frequency in fraction
            % of a season with condition X within the time chunk; The
            % first normalization accounts for the variable number of
            % months at a particular lat/lon with data, and the 2nd
            % normalization accounts for the subset of a year that we
            % are studying. frequency of X = X/N * 3 months/season * 1
            % season/year;
```

```
        %
        % a 2nd layer is deciding how many months within a time
        % chunk need data for a "valid" assessment. For example, 20
        % years with 3 month seasons means 60 possible months with
        % an analysis condition X. What if only 5 months have data?
        % Is that a fair comparison with a time chunk that has 60
        % months of data? We might have to see about this, but 15
        % is a round number starting point, such that any time
        % chunk with fewer than that many months of data is
        % excluded from the analysis (turned into a NaN in the
        % frequence calculation). 15 is also 5 years of 3-month
        % seasons, hence "round number" but it might be too small a
        % minimum, i'm not sure.

        minNumMonthsWithDataInTimeChunk = 50;
        minPercentMonthsWithDataInTimeChunk = 0.80;  %%% new percent based
threshold

        % CHANGE THIS TO BE BASED ON fracMonthsWithDataInTimeChunk
        % vs minPercentMonthsWithDataInTimeChunk

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        if
fracMonthsWithDataInTimeChunk>=minPercentMonthsWithDataInTimeChunk;

analysis.(seasonName).(yearRange).frequency_anoms_greater_two_count(ii,mm) = ...
            anoms_greater_two_count/numMonthsWithDataInTimeChunk;

analysis.(seasonName).(yearRange).frequency_anoms_less_two_count(ii,mm) = ...
            anoms_less_two_count/numMonthsWithDataInTimeChunk;

analysis.(seasonName).(yearRange).frequency_anoms_greater_three_count(ii,mm) = ...
            anoms_greater_three_count/numMonthsWithDataInTimeChunk;

          %Metric calculation
          kstestInput = (analysis.(seasonName).(yearRange).std_anoms) -
(nanmean(analysis.(seasonName).(yearRange).std_anoms));
          [h,p] = kstest(kstestInput);
%             [hNonCentered,pNonCentered] =
kstest(analysis.(seasonName).(yearRange).std_anoms);
%             figure(50); clf reset; histogram(kstestInput);
%             disp(cat(2,' p = ',num2str(p),' and pNonCentered =
',num2str(pNonCentered),' (pausing)')); pause;
          analysis.(seasonName).(yearRange).p(ii,mm) = p;
          %Calculating mean/median ratio
          mean_data = nanmean(analysis.(seasonName).(yearRange).std_anoms);
```

```
                    median_data = nanmedian(analysis.(seasonName).(yearRange).std_anoms);
                    gaussian_check = abs(mean_data / median_data);
                    analysis.(seasonName).(yearRange).gaussian_check(ii,mm) =
gaussian_check;

                else

analysis.(seasonName).(yearRange).frequency_anoms_greater_two_count(ii,mm) = NaN;

analysis.(seasonName).(yearRange).frequency_anoms_less_two_count(ii,mm) = NaN;

analysis.(seasonName).(yearRange).frequency_anoms_greater_three_count(ii,mm) =
NaN;

                    analysis.(seasonName).(yearRange).iqr_anom(ii,mm)= NaN;
                    analysis.(seasonName).(yearRange).gaussian_check(ii,mm)= NaN;
                    analysis.(seasonName).(yearRange).p(ii,mm) = NaN;


                end; % if numMonthsWithDataInTimeChunk

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

            end; % year chunks
        end; % season
    end; % lon
end; % lat

analysisFields =
{'frequency_anoms_greater_two_count','frequency_anoms_less_two_count','frequency_a
noms_greater_three_count','iqr_anom','gaussian_check','p'};
% Cycle over seasons
for jj=1:length(seasonNames);
    seasonName = seasonNames{jj};
    % Cycle over years
    for kk=1:length(yearChunks);
        yearRange = yearRanges{kk};
        % cycle over analysis fields
        for mm=1:length(analysisFields);
            analysisField = analysisFields{mm};
            % define masked versions of the maps
            analysisValuesAll = analysis.(seasonName).(yearRange).(analysisField); % the
map of values
            analysisValuesOcean = analysisValuesAll.*oceanmask2D; % turns land into NaN
            analysisValuesLand = analysisValuesAll.*landmask2D; % turns ocean into NaN
```

```
% account for where there is no data by applying a mask to
% the gridarea that turns all grid boxes with no data (at that
% particular year range and for that particular season) into a
% NaN and then mask these values from the calculation of the
% weighted mean.
noDataAll = find(isnan(analysisValuesAll)==1);
noDataOcean = find(isnan(analysisValuesOcean)==1); % for excluding missing
data and land data
noDataLand = find(isnan(analysisValuesLand)==1); % for excluding missing data
and ocean data
gridareaAll = gridarea; % initialize to the gridarea before masking
gridareaOcean = gridarea; % initialize to the gridarea before masking
gridareaLand = gridarea; % initialize to the gridarea before masking
gridareaAll(noDataAll) = NaN;
gridareaOcean(noDataOcean) = NaN;
gridareaLand(noDataLand) = NaN;

% diagnostic figures that you can uncomment and see what these
% masks look like (i find visuals a useful confirmation that
% what i do in code is working like i expect)
%       figure(11); clf reset; pcolor(lon,lat,gridareaAll); shading flat;
%       pause
%       figure(11); clf reset; pcolor(lon,lat,gridareaOcean); shading flat;
%       pause
%       figure(11); clf reset; pcolor(lon,lat,gridareaLand); shading flat;
%       pause

% next command makes a 3x1 vector of gridarea-weighted mean values from
% all data, ocean data, and land data; the values are stored in a new field linked
% to the name of the analysisField+_means

analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means')) = ...
   [nansum(analysisValuesAll.*gridareaAll)/nansum(gridareaAll); ...
   nansum(analysisValuesOcean.*gridareaOcean)/nansum(gridareaOcean); ...
   nansum(analysisValuesLand.*gridareaLand)/nansum(gridareaLand)];
analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95')) = ...
   [prctile((analysisValuesAll(:)),95); ...
   prctile((analysisValuesOcean(:)),95); ...
   prctile((analysisValuesLand(:)),95)];
analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5')) = ...
   [prctile((analysisValuesAll(:)),5); ...
   prctile((analysisValuesOcean(:)),5); ...
   prctile((analysisValuesLand(:)),5)];

% this puts the calculations you make above into a tabular
% format for simpler translation to a scatterplot figure; the
```

```
        % rows are yearChunks and the columns are the 4 seasons
        % (winter, spring, summer, fall) and there is one table for
        % each statistical measure (mean, 95th, 5th, so far, but can be
        % easily modified to include tables of other statistics like
        % median, 25th, 75th, or whatever) for each spatial domain
        % (all, ocean, land; again can be easily modified to be a table
        % for additional spatial domains such as USA or Australia or
        % something). Access the analysis.tables data by selecting the
        % analysis.tables.data.(ANALYSIS FIELD NAME).(DOMAIN) for a
        % table of values with rows that are yearChunks (8 currently)
        % and columns that are seasons (4 currently).
        analysis.tables.data.(cat(2,analysisField,'_means')).all(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(1);
        analysis.tables.data.(cat(2,analysisField,'_means')).ocean(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(2);
        analysis.tables.data.(cat(2,analysisField,'_means')).land(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(3);
        % gather the tabular version of the 95th percentile of the
        % different analysis fields and spatial domains
        analysis.tables.data.(cat(2,analysisField,'_95')).all(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(1);
        analysis.tables.data.(cat(2,analysisField,'_95')).ocean(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(2);
        analysis.tables.data.(cat(2,analysisField,'_95')).land(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(3);
        % gather the tabular version of the 5th percentile of the
        % different analysis fields and spatial domains
        analysis.tables.data.(cat(2,analysisField,'_5')).all(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(1);
        analysis.tables.data.(cat(2,analysisField,'_5')).ocean(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(2);
        analysis.tables.data.(cat(2,analysisField,'_5')).land(kk,jj) = ...
          analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(3);
        % include metadata explaining (to yourself) what the rows and
        % columns of the tabular version of the output correspond to
        % physically, connecting it to what is in the code above
        analysis.tables.seasonNames = seasonNames;
        analysis.tables.yearRanges = yearRanges;
        analysis.tables.yearChunks = yearChunks;
        analysis.tables.rows = 'seasonNames';
        analysis.tables.cols = 'yearChunks';



    end; % mm
end; % kk
```

```
end; % jj


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
% Brian added this code
% store analysis-specific constants so that other scripts can use what you
% did to make the analysis. In other words, this "saves" your work as a new
% field in the structure created in script A_*.m
analysis.numMonthsInSeason = numMonthsInSeason;
analysis.minNumMonthsWithDataInTimeChunk =
minNumMonthsWithDataInTimeChunk;
analysis.seasonNames = seasonNames;
analysis.seasonbox = seasonbox;
analysis.yearChunks = yearChunks;
analysis.yearRanges = yearRanges;
% store the results of this script as a new field in the input structure
output.analysis = analysis;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

% Brian notes: these can be deleted
% % % %Important outputs for main function
% % % output.NOAAV5Data = NOAAV5Data;
% % % output.anoms = anoms;

end % (main) function
```

Script C:

```
function [] = C_CreatePlotsFromStandardizedAnomalies(input);
%
% C_CreatePlotsFromStandardizedAnomalies(NOAAV5Data);
%
if nargin==0;
   error('Run A_GetAndProcessNOAAV5Data.m and
B_EvaluateStandardizedAnomalies.m first')
else
   NOAAV5Data = input;
   time = NOAAV5Data.time;
   lat = NOAAV5Data.lat;
   lon = NOAAV5Data.lon;
   yyyymm = NOAAV5Data.yyyymm;
    tempAnom = NOAAV5Data.tempAnom;
   std_anoms = NOAAV5Data.std_anoms;
   analysis = NOAAV5Data.analysis;
```

```
   numMonthsInSeason = NOAAV5Data.analysis.numMonthsInSeason;
   minNumMonthsWithDataInTimeChunk =
NOAAV5Data.analysis.minNumMonthsWithDataInTimeChunk;
   seasonNames = NOAAV5Data.analysis.seasonNames;
   seasonbox = NOAAV5Data.analysis.seasonbox;
   yearChunks = NOAAV5Data.analysis.yearChunks;
   yearRanges = NOAAV5Data.analysis.yearRanges;
   gridarea = NOAAV5Data.gridarea;
end; % if nargin

coastlines = load('coastlines_180_180.dat');

% Brian notes: set plotGlobalMapsOfStdAnom to 1 to run this if-statement
% and produce maps of the standardized anomalies
plotGlobalMapsOfStdAnom = 0;
savefig = 0;
printQuality = '-r300';
if plotGlobalMapsOfStdAnom;
   %For Loop to Plot monthly Standardized anomalies --- with colored MAP
   for pp= 1:100:1000;
      figure(1); clf reset;
      std_anom = permute(std_anoms(pp,:,:),[2,3,1]);
      % % %        pcolor(lon,lat,std_anom);
      % the lon and lat inputs to sanePColor() must be row vectors and
      % not column vectors
      sanePColor(lon',lat',std_anom);
      shading flat;
      cptcmap('temperature');
      colorbar;
      caxis([-3,3]);

      hold on;
      plot(coastlines(:,1),coastlines(:,2),'k-');
      Ht = title(cat(2,'Global Map of Standardized Anomlies'));
      Hx = xlabel('Longitude');
      Hy = ylabel('Latitude');
      set([Hx,Hy,Ht,gca],'fontsize',8);
      pause;

      if savefig; %if statement to create plots
         ChangePlotDimensions(6.5,4);  % width, height
         figname = cat(2,'GlobalAnomMap.png');
         disp(cat(2,'** Overwriting fig to ',figname));
         print('-dpng',printQuality,figname);
      end;
   end;
```

```
end; % if plotGlobalMapsOfStdAnom

%Plot of Nondepartures anaylsis fields
plotGlobalMapsOfNonDepartures = 0;
if plotGlobalMapsOfNonDepartures;
    %Loop that will create global maps of std. anomlies less than -2 deviations
    savefig = 1;
    printQuality = '-r300';
    analysisFields =
{'frequency_anoms_greater_two_count','frequency_anoms_less_two_count','frequency_a
noms_greater_three_count','iqr_anom','gaussian_check','p'};
    %Cycle over seasons
    for ii=1:length(seasonNames);
        seasonName = seasonNames{ii};
        %Cycle over years
        for jj=1:length(yearRanges);
            yearRange = yearRanges{jj};
            for mm=4:6; %1:length(analysisFields);
                analysisField = analysisFields{mm};
                figure(mm); clf reset;

                %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                % Brian added this part
                % the lon and lat inputs to sanePColor() must be row vectors and
                % not column vectors
                plotValues = analysis.(seasonName).(yearRange).(analysisField); %%%
                plotValues = plotValues;
                sanePColor(lon',lat',plotValues); %%%
                shading flat;
                colorbar;

                %Set Colormaps for each metric
                if strcmp(analysisField,'frequency_anoms_greater_two_count');
                    cptcmap('sunshine_9lev_custom.cpt');
                elseif strcmp(analysisField,'frequency_anoms_less_two_count');
                    cptcmap('precip_11lev_custom.cpt');
                elseif strcmp(analysisField,'frequency_anoms_greater_three_count');
                    cptcmap('sunshine_9lev_custom.cpt');
                elseif strcmp(analysisField,'iqr_anom');
                    cptcmap('7lev_green_cont.cpt');
                elseif strcmp(analysisField,'gaussian_check');
                    cptcmap('3lev_symm.cpt');
                elseif strcmp(analysisField,'p');
                    cptcmap('test2.cpt');
                end; % strcmp
```

```
		textLine1 = (cat(2,'Ocean Mean = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(2),3),' ; 
(95th% = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(2),3),'; 5th% = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(2),3),')'));
		textLine2 = (cat(2,'Land Mean = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(3),3),' ; 
(95th% = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(3),3),'; 5th% = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(3),3),')'));
		%		textLine3 = (cat(2,'95th Percentile = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(1),3)));
		%		textLine4 = (cat(2,'5th Percentile = 
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(1),3)));
		Htext = text(0,-80,[{textLine1};{textLine2}]);

		%If statement to control title for plots
		if (mm == 1 );

		Ht = title({(cat(2,'Std.Anoms > +2 during ',seasonName,' from 
',yearRange)),'Climatology: 1951-1980'});
		   Hx = xlabel('Longitude');
		   Hy = ylabel('Latitude');
		   set([Hx,Hy,Ht,gca],'fontsize',8);
		   set([Htext],'fontsize',5);

		elseif (mm == 2);

		   Ht = title({(cat(2,'Std.Anoms < -2 during ',seasonName,' from 
',yearRange)),'Climatology: 1951-1980'});
		   Hx = xlabel('Longitude');
		   Hy = ylabel('Latitude');
		   set([Hx,Hy,Ht,gca],'fontsize',8);
		   set([Htext],'fontsize',5);

		elseif (mm == 3);

		   Ht = title({(cat(2,'Std.Anoms > +3 during ',seasonName,' from 
',yearRange)),'Climatology: 1951-1980'});
		   Hx = xlabel('Longitude');
		   Hy = ylabel('Latitude');
		   set([Hx,Hy,Ht,gca],'fontsize',8);
		   set([Htext],'fontsize',5);

		elseif (mm == 4);
```

```
            Ht= title({(cat(2,'IQR during ',seasonName,' from
',yearRange)),'Climatology: 1951-1980'});
            Hx = xlabel('Longitude');
            Hy = ylabel('Latitude');
            caxis([1,3]);
            set([Hx,Hy,Ht,gca],'fontsize',8);
            set([Htext],'fontsize',5);

         elseif (mm == 5);
            Ht= title({(cat(2,'Mean/Median during ',seasonName,' from
',yearRange)),'Climatology: 1951-1980'});
            Hx = xlabel('Longitude');
            Hy = ylabel('Latitude');
            caxis([0,2]);
            set([Hx,Hy,Ht,gca],'fontsize',8);
            set([Htext],'fontsize',5);

         elseif (mm == 6);
            Ht= title({(cat(2,'P-Values during ',seasonName,' from
',yearRange)),'Climatology: 1951-1980'});
            Hx = xlabel('Longitude');
            Hy = ylabel('Latitude');
            caxis([0,0.5]);
            set([Hx,Hy,Ht,gca],'fontsize',8);
            set([Htext],'fontsize',5);

         end;

         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
         % Brian added this part
         %caxis([-1,1]);
         hold on;
         plot(coastlines(:,1),coastlines(:,2),'k-','linewidth',1);
         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
         %pause;
         if savefig %if statement to create plots
            ChangePlotDimensions(6.5,4);  % width, height
            figname = cat(2,analysisField,'_',seasonName,'_',yearRange,'_80.png');
            disp(cat(2,'** Overwriting fig to ',figname));
            print('-dpng',printQuality,figname);
         end;
      end;
   end;
 end;
end; % plotGlobalMapsOfNonDepartures
```

```
%Loop to create map of Departures
plotGlobalMapsOfDepartures = 0;
if plotGlobalMapsOfDepartures;
    savefig = 1;
    printQuality = '-r300';
    analysisFields =
{'frequency_anoms_greater_two_count','frequency_anoms_less_two_count','frequency_a
noms_greater_three_count'};
    %Cycle over seasons
    for ii=1:length(seasonNames);
        seasonName = seasonNames{ii};
        %Cycle over years
        for jj=1:length(yearRanges);
            yearRange = yearRanges{jj};
            for mm=1:length(analysisFields);
                analysisField = analysisFields{mm};
                figure(mm); clf reset;

                if (mm == 1);
                    plotValues = analysis.(seasonName).(yearRange).(analysisField); %%
                    plotValues = plotValues/0.022; %%% anomaly relative to expectation of
2.2%
                    sanePColor(lon',lat',plotValues);
                    shading flat;
                    caxis([0,20]);%%%

                elseif (mm == 2);
                    plotValues = analysis.(seasonName).(yearRange).(analysisField); %
                    plotValues = plotValues/0.022; %%% anomaly relative to expectation of
2.2%
                    sanePColor(lon',lat',plotValues);
                    shading flat;
                    caxis([0,20]);%%%

                elseif (mm == 3);
                    plotValues = analysis.(seasonName).(yearRange).(analysisField); %%
                    plotValues = plotValues/0.001; %%% anomaly relative to expectation of
0.1%
                    sanePColor(lon',lat',plotValues);
                    shading flat;
                    caxis([0,250]);%%%
                end; %end if

                %Set Colormaps for each metric
                if strcmp(analysisField,'frequency_anoms_greater_two_count');
                    cptcmap('sunshine_9lev_custom.cpt');
```

```
            elseif strcmp(analysisField,'frequency_anoms_less_two_count');
                cptcmap('precip_11lev_custom.cpt');
            elseif strcmp(analysisField,'frequency_anoms_greater_three_count');
                cptcmap('sunshine_9lev.cpt');
            end; % strcmp

            colorbar;


            %If statement to control title for plots
            if (mm == 1 );
                textLine1 = (cat(2,'Ocean Mean =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(2)/0.022)),'
; (95th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(2)/0.022),';
5th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(2)/0.022),')',3));
                textLine2 = (cat(2,'Land Mean =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(3)/0.022)),'
; (95th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(3)/0.022),';
5th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(3)/0.022),')',3));
                %                textLine3 = (cat(2,'95th Percentile =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(1)/0.022)),3));
                %                textLine4 = (cat(2,'5th Percentile =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(1)/0.022)),3));

                Htext = text(0,-80,[{textLine1};{textLine2}]);
                Ht = title({(cat(2,'Departure from Norm: Std.Anoms > +2 during
',seasonName,' from ',yearRange)),'Climatology: 1951-1980'});
                Hx = xlabel('Longitude');
                Hy = ylabel('Latitude');
                set([Hx,Hy,Ht,gca],'fontsize',8);
                set([Htext],'fontsize',5);

            elseif (mm == 2)
                textLine1 = (cat(2,'Ocean Mean =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(2)/0.022)),'
; (95th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(2)/0.022),';
5th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(2)/0.022),')',3));
                textLine2 = (cat(2,'Land Mean =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(3)/0.022)),'
; (95th% =
```

```
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(3)/0.022),';
5th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(3)/0.022),')',3));
        %           textLine3 = (cat(2,'95th Percentile =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(1)/0.022)),3));
        %           textLine4 = (cat(2,'5th Percentile =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(1)/0.022)),3));
        %
        Htext = text(0,-80,[{textLine1};{textLine2}]);
        Ht = title({(cat(2,'Departure from Norm: Std.Anoms < -2 during
',seasonName,' from ',yearRange)),'Climatology: 1951-1980'});
        Hx = xlabel('Longitude');
        Hy = ylabel('Latitude');
        set([Hx,Hy,Ht,gca],'fontsize',8);
        set([Htext],'fontsize',5);

      elseif (mm == 3)
        textLine1 = (cat(2,'Ocean Mean =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(2)/0.001)),'
; (95th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(2)/0.001),';
5th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(2)/0.001),')',3));
        textLine2 = (cat(2,'Land Mean =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_means'))(3)/0.001)),'
; (95th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(3)/0.001),';
5th% =
',num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(3)/0.001),')',3));
        %           textLine3 = (cat(2,'95th Percentile =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_95'))(1)/0.001)),3));
        %           textLine4 = (cat(2,'5th Percentile =
',(num2str(analysis.(seasonName).(yearRange).(cat(2,analysisField,'_5'))(1)/0.001)),3));
        %
        Htext = text(0,-80,[{textLine1};{textLine2}]);
        Ht = title({(cat(2,'Departure from Norm: Std.Anoms > +3 during
',seasonName,' from ',yearRange)),'Climatology: 1951-1980'});
        Hx = xlabel('Longitude');
        Hy = ylabel('Latitude');
        set([Hx,Hy,Ht,gca],'fontsize',8);
        set([Htext],'fontsize',5);
      end;

      hold on;
      plot(coastlines(:,1),coastlines(:,2),'k-','linewidth',1);
      %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
            if savefig %if statement to create plots
                ChangePlotDimensions(6.5,4);  % width, height
                figname =
cat(2,'DepfromNorm_',analysisField,'_',seasonName,'_',yearRange,'_80.png');
                disp(cat(2,'** Overwriting fig to ',figname));
                print('-dpng',printQuality,figname);
            end;
        end;
    end;
  end;
end; % plotGlobalMapsOfDepartures

%Create Bar graphs to show trend of each metric
plotBar = 0;
savefig = 1;
analysisFields =
{'frequency_anoms_greater_two_count','frequency_anoms_less_two_count','frequency_a
noms_greater_three_count','iqr_anom','p'};
if plotBar;
for jj=1:length(yearRanges);
  yearRange = yearRanges{jj};
  xlabelString{jj} = cat(2,num2str(analysis.tables.yearChunks(jj,1)),'-
',num2str(analysis.tables.yearChunks(jj,2)));
  xtickVals(jj) = sum(analysis.tables.yearChunks(jj,1:2))/2;
 % switch
end;  % for jj loop

for mm=1:length(analysisFields);
  analysisField = analysisFields{mm};
  figure(1);
  clf reset;

  if (mm == 1 );
    VaL1 = analysis.tables.data.(cat(2,analysisField,'_means')).ocean/0.022;
    VaL2 = analysis.tables.data.(cat(2,analysisField,'_means')).land/0.022;
    t = tiledlayout('flow');
    t.TileSpacing = 'compact';
    nexttile
    Hb = bar(xtickVals,VaL1);
    ylim([0,25]);
    Htxt = text(1895,23,'(A) Ocean');
    Hlol = text(1895,21,'Climatology: 1951-1980')
    set([Htxt,Hlol],'fontsize',10);
    set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
    nexttile
    Hb = bar(xtickVals,VaL2);
```

```
        ylim([0,25]);
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        Htxt = text(1895,23,'(B) Land');
        Hlol = text(1895,21,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);

        Ht = title(t,{(cat(2,'2-sigma Departure Relative to Statistical Expectation')),'1880 to
2020'});
        Hy = ylabel(t,'Spatially-Averaged Departure Relative to Statistical Expectation');
        set([Hy,Ht,Hy],'fontsize',10);

legend('Winter','Summer','Spring','Fall','Location','southoutside','Orientation','horizontal');


    elseif (mm == 2)
        VaL1 = analysis.tables.data.(cat(2,analysisField,'_means')).ocean/0.022;
        VaL2 = analysis.tables.data.(cat(2,analysisField,'_means')).land/0.022;
        t = tiledlayout('flow');
        t.TileSpacing = 'compact';
        nexttile
        Hb = bar(xtickVals,VaL1);
        ylim([0,10]);
        Htxt = text(1895,9,'(A) Ocean');
        Hlol = text(1895,8,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        nexttile
        Hb = bar(xtickVals,VaL2);
        ylim([0,10]);
        Htxt = text(1895,9,'(B) Land');
        Hlol = text(1895,8,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);

        Ht = title(t,{(cat(2,'-2-sigma Departure Relative to Statistical Expectation')),'1880 to
2020'});
        Hy = ylabel(t,'Spatially-Averaged Departure Relative to Statistical Expectation');
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        set([Hy,Ht],'fontsize',10);

legend('Winter','Summer','Spring','Fall','Location','southoutside','Orientation','horizontal');


    elseif (mm == 3)
        VaL1 = analysis.tables.data.(cat(2,analysisField,'_means')).ocean/0.001;
        VaL2 = analysis.tables.data.(cat(2,analysisField,'_means')).land/0.001;
        t = tiledlayout('flow');
```

```
        t.TileSpacing = 'compact';
        nexttile
        Hb = bar(xtickVals,VaL1);
        ylim([0,250]);
        Htxt = text(1895,235,'(A) Ocean');
        Hlol = text(1895,210,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        nexttile
        Hb = bar(xtickVals,VaL2);
        ylim([0,250]);
        Htxt = text(1895,235,'(B) Land');
        Hlol = text(1895,210,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);

        Ht = title(t,{(cat(2,'3-sigma Departure Relative to Statistical Expectation')),'1880 to
2020'});
        Hy = ylabel(t,'Spatially-Averaged Departure Relative to Statistical Expectation');
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        set([Hy,Ht],'fontsize',10);

legend('Winter','Summer','Spring','Fall','Location','southoutside','Orientation','horizontal');


    elseif (mm == 4)
        t = tiledlayout('flow');
        t.TileSpacing = 'compact';
        nexttile
        Hb = bar(xtickVals,analysis.tables.data.(cat(2,analysisField,'_means')).ocean);
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        ylim([0,2.5]);
        Htxt = text(1895,2.3,'(A) Ocean');
        Hlol = text(1895,2.1,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);
        nexttile
        Hb = bar(xtickVals,analysis.tables.data.(cat(2,analysisField,'_means')).land);

        Ht = title(t,(cat(2,'IQR from 1880 to 2020')));
        Hy = ylabel(t,'Average IQR');
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        set([Hy,Ht],'fontsize',10);

legend('Winter','Summer','Spring','Fall','Location','southoutside','Orientation','horizontal');
        ylim([0,2.5]);
        Htxt = text(1895,2.3,'(B) Land');
        Hlol = text(1895,2.1,'Climatology: 1951-1980')
```

```
        set([Htxt,Hlol],'fontsize',10);

    elseif (mm == 5)
        t = tiledlayout('flow');
        t.TileSpacing = 'compact';
        nexttile
        Hb = bar(xtickVals,analysis.tables.data.(cat(2,analysisField,'_means')).ocean);
        ylim([0,1]);
        Htxt = text(1895,0.9,'(A) Ocean');
        Hlol = text(1895,0.8,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);
        nexttile
        Hb = bar(xtickVals,analysis.tables.data.(cat(2,analysisField,'_means')).land);
        ylim([0,1]);
        Htxt = text(1895,0.9,'(B) Land');
        Hlol = text(1895,0.8,'Climatology: 1951-1980')
        set([Htxt,Hlol],'fontsize',10);
        set(gca,'xtick',xtickVals,'xticklabel',xlabelString,'fontsize',5);

        Ht = title(t,{(cat(2,'P-Values from KS Test from 1880 to 2020')),'Climatology: 1951-
1980'});
        Hy = ylabel(t,'Average P-Val');
        set([Hy,Ht],'fontsize',10);

legend('Winter','Summer','Spring','Fall','Location','southoutside','Orientation','horizontal');


    end;

    if savefig %if statement to create plots
        ChangePlotDimensions(6,6);  % width, height
        % CHANGE THE NEXT LINETO A FIGURE NAME YOU WANT; I PUT BAR_
AS THE PREFIX JUST AS A PLACEHOLDER
        figname = cat(2,'BAR_',analysisField,'_.png');
        disp(cat(2,'** Overwriting fig to ',figname));
        print('-dpng',printQuality,figname); % set printQuality to -r600 for high resolution
maps and graphs, but -r300 should work well enough too
    end;

end;
end;
%End Bar graphs

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% start new code snippet for showing global warming from NOAAv5
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
plotGlobalWarming = 0;
savefig = 1;
printQuality = '-r900';
if plotGlobalWarming;
    % use temperature anomalies for this since we're trying to show global
    % warming, and since this is the basis for the standardized temperature
    % anomaly calculation
    analysisValuesAll = tempAnom;
    % Cycle over years
    for kk=1:length(yyyymm(:,1));
        analysisValuesAll_OneMonth = permute(analysisValuesAll(kk,:,:),[2,3,1]);
        % account for where there is no data by applying a mask to
        % the gridarea that turns all grid boxes with no data (at that
        % particular year range and for that particular season) into a
        % NaN and then mask these values from the calculation of the
        % weighted mean.
        noDataAll = find(isnan(analysisValuesAll_OneMonth)==1);
        gridareaAll = gridarea; % initialize to the gridarea before masking
        gridareaAll(noDataAll) = NaN;
        globalMeanMonthlyTemperature(kk,1) =
nansum(analysisValuesAll_OneMonth.*gridareaAll)/nansum(gridareaAll);
    end;  % for kk
    minYear = min(yyyymm(:,1));
    maxYear = max(yyyymm(:,1));
    years = minYear:1:maxYear;
    % calculate the annual average
    for kk=1:length(years);
        yearMatch = find(yyyymm(:,1)==years(kk));
        globalMeanAnnualTemperature(kk,1) =
mean(globalMeanMonthlyTemperature(yearMatch));
    end;  % for kk
    % build decades knowing that the last couple of years are not a decade
    % yet so stop in 2010 as opposed to 2020
    decades = 1880:10:2010;
    % calculate decadal average
    for kk=1:length(decades);
        startIndex = 1+10*(kk-1); % a simple way to increment by 10 based on kk
        globalMeanDecadalTemperature(kk,1) =
mean(globalMeanAnnualTemperature(startIndex:startIndex+9,1));
    end;  % for kk
    disp(table(decades',globalMeanDecadalTemperature));
    % make a figure
    figure(49);
    clf reset;
    bar(decades,globalMeanDecadalTemperature);
```

```
   hold on;
   % overlay the annual anomaly as a line plot, but because bar() plots at
   % the center point, subtract 4 from the years to "align" things
   % visually
   plot(years-4,globalMeanAnnualTemperature);
   Ht=title('NOAAv5 Average Global Temperature, 1880 to 2020');
   Hy= ylabel('Temperature Anomaly (deg C)');

   set([Ht,Hy],'fontsize',12);
   % the following set() command makes the figure box turn off and ticks
   % progress outward. Looks nicer IMO but up to you on the aesthetic
   set(gca,'Fontsize',8,'box','off','TickDir','out','TickLength',[0.006,0.006]);
   set([Ht,Hy],'fontsize',12);
   if savefig; %if statement to create plots
      ChangePlotDimensions(6,3);  % width, height
      figname = cat(2,'GlobalWarmingNOAAv5','.png');
      disp(cat(2,'** Overwriting fig to ',figname));
      print('-dpng',printQuality,figname);
   end;
   pause
end; % if plotGlobalWarming
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% end new code snippet for showing global warming from NOAAv5
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% start new code snippet for showing maps of data coverage in NOAAv5
%%%%% through time "snapshots"
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Brian notes: set plotGlobalMapsOfStdAnom to 1 to run this if-statement
% and produce maps of the standardized anomalies
plotGlobalMapsOfStdAnom = 1;
savefig = 1;
printQuality = '-r900';
plotYearsMonths = [1880,1;1900,1;1930,1;1950,1;1980,1;2020,1];
if plotGlobalMapsOfStdAnom;
   %For Loop to Plot monthly Standardized anomalies --- with colored MAP
   % % %     for pp= 1:100:1000;
   for pp=1:length(plotYearsMonths);
      plotYear = plotYearsMonths(pp,1);
      plotMonth = plotYearsMonths(pp,2);
      dateMatch = find(yyyymm(:,1)==plotYear&yyyymm(:,2)==plotMonth);
      figure(70); clf reset;
      std_anom = permute(std_anoms(dateMatch,:,:),[2,3,1]);
      % % %        pcolor(lon,lat,std_anom);
      numNaN = length(find(isnan(std_anom(:))));
```

```
        percentMissing = numNaN/(length(lat)*length(lon))*100;
        % the lon and lat inputs to sanePColor() must be row vectors and
        % not column vectors
        sanePColor(lon',lat',std_anom);
        shading flat;
        cptcmap('temperature');
        colorbar;
        caxis([-3,3]);

        hold on;
        plot(coastlines(:,1),coastlines(:,2),'k-');
        % % %         Ht = title(cat(2,'Global Map of Standardized Anomlies'));
        Ht = title(cat(2,num2str(yyyymm(dateMatch,1)),' in January', ...
            ' (Fraction of Data Missing = ',num2str(percentMissing,3),'%)'));
        Hx = xlabel('Longitude');
        Hy = ylabel('Latitude');
        set([Hx,Hy,Ht,gca],'fontsize',14);
        if savefig; %if statement to create plots
            ChangePlotDimensions(6.5,4);  % width, height
            % % %             figname = cat(2,'GlobalAnomMap.png');
            figname = cat(2,'GlobalAnomMap','-',num2str(yyyymm(dateMatch,1)),'-
',num2str(yyyymm(dateMatch,2)),'.png');
            disp(cat(2,'** Overwriting fig to ',figname));
            print('-dpng',printQuality,figname);
        end;

    end;
end; % if plotGlobalMapsOfStdAnom
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% start new code snippet for showing maps of data coverage in NOAAv5
%%%%% through time "snapshots"
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
end % main function

function [] = MakeOlderSimpleMapWithRectangles(input);
% savefig = 0;
% printQuality = '-r600';
% for ii=1:length(regionNames);
%     regionName = regionNames{ii};
%     rectStartLat = latlonbox.(regionName)(1);
%     rectStartLon = latlonbox.(regionName)(3);
%     rectLonWidth = latlonbox.(regionName)(4)-latlonbox.(regionName)(3);
%     rectLatWidth = latlonbox.(regionName)(2)-latlonbox.(regionName)(1);
%
MakeSimpleMapWithRectangle(rectStartLon,rectStartLat,rectLonWidth,rectLatWidth);
%     % % %     set(gca,'ylim',[-60,55]);
```

```
%     if savefig; %if statement to create plots
%         ChangePlotDimensions(6.5,4);  % width, height
%         figname = cat(2,'map_',regionName,'.png');
%         disp(cat(2,'** Overwriting fig to ',figname));
%         print('-dpng',printQuality,figname);
%     end;  % if savefig
%     pause
% end; % for ii loop
end % function

%Function that creates a map with rectangles to display specific regions
   function [] =
MakeSimpleMapWithRectangle(rectStartLon,rectStartLat,rectLonWidth,rectLatWidth);
      %
      % Thrown together by B. Magi.
      %
      % Note that this has not been fully proofed so you
      % may need to adapt parts for your own application.
      %
      lat = 89.5:-1:-89.5;
      lon = -179.5:1:179.5;
      % create a dummy set of data just to plot something
      dummydata = NaN.*rand(180,360);
      figure(15);
      clf reset;
      pcolor(lon,lat,dummydata);
      shading flat;
      % only need the colorbar() and caxis() functions if you have actual data
      % % % colorbar;
      % % % caxis([-3,3]);
      % load an input file with coastline data for a rudimentary map (not
      % rasterized so that it can be re-projected)
      coastlines = load('coastlines_180_180.dat');
      hold on;
      plot(coastlines(:,1),coastlines(:,2),'k-');
      set(gca,'xlim',[-140,-60],'ylim',[-60,60]);
      % manually set rectangle dimensions
      % % % 28,48,-130,-76
      % % % rectStartLon = -130; % lon in degrees E (W is negative)
      % % % rectStartLat = 28; % lat in degrees N (S is negative)
      % % % rectLonWidth = 54; % 95 degrees east of the start lon
      % % % rectLatWidth = 20; % 20 degrees north of the start lat
      % load an input file which loads the structure ‘projectdata’
      % % % load('W:\UNCC\teaching\2021\ESCI4122-
ESCI5122\projects\data\monthly\Carolinas_monthly_stdanom.mat'); %
load('Carolinas_monthly_stdanom.mat');
```

```
        % % % % connect rectangle dimensions to projectdata
        % % % rectStartLon = projectdata.latlon(3);
        % % % rectStartLat = projectdata.latlon(1);
        % % % rectLonWidth = projectdata.latlon(4)-projectdata.latlon(3);
        % % % rectLatWidth = projectdata.latlon(2)-projectdata.latlon(1);

rectangle('Position',[rectStartLon,rectStartLat,rectLonWidth,rectLatWidth],'curvature',0.2
,'EdgeColor',[1,0,0],'LineWidth',3);
    end % function


    function [] = MakeOlderLinePlotsAndHistograms(input);
        %
        %
        %
        output = input;
        %Can choose to utilize this bit of code if I want to
        % % % disp('nothing works past this YET')
        doThis = 0;
        if doThis;


            % % %
            % % %         anoms.(regionName).(seasonName).(yearRange).std_anoms =
std_anomsMatches(:);
            % % %         anoms.(regionName).(seasonName).(yearRange).yyyymm =
yyyymm(dateMatches,:);
            % % %         anoms.(regionName).(seasonName).(yearRange).regionLat =
regionLat;
            % % %         anoms.(regionName).(seasonName).(yearRange).regionLon =
regionLon;
            % % %
            % % %         anoms_greater_two_field =
find(anoms.(regionName).(seasonName).(yearRange).std_anoms>2);
            % % %         anoms_greater_two_values =
anoms.(regionName).(seasonName).(yearRange).std_anoms(anoms_greater_two_field);
            % % %         anoms_greater_two_count = length(anoms_greater_two_values);
            % % %
            % % %         anoms_less_two_field =
find(anoms.(regionName).(seasonName).(yearRange).std_anoms<-2);
            % % %         anoms_less_two_values =
anoms.(regionName).(seasonName).(yearRange).std_anoms(anoms_less_two_field);
            % % %         anoms_less_two_count = length(anoms_less_two_values);
            % % %
            % % %         %Simple Calculations for Histogram
```

```
    % % %       anomMean =
nanmean((anoms.(regionName).(seasonName).(yearRange).std_anoms));
    % % %       anomMedian =
nanmedian((anoms.(regionName).(seasonName).(yearRange).std_anoms));
    % % %       anomSkewness =
skewness((anoms.(regionName).(seasonName).(yearRange).std_anoms));
    % % %       anomIQR =
iqr((anoms.(regionName).(seasonName).(yearRange).std_anoms));
    % % %       anomStd =
nanstd((anoms.(regionName).(seasonName).(yearRange).std_anoms));
    % % %       anoms.(regionName).(seasonName).(yearRange).anomMean =
anomMean;
    % % %       anoms.(regionName).(seasonName).(yearRange).anomMedian =
anomMedian;
    % % %       anoms.(regionName).(seasonName).(yearRange).anomSkewness
= anomSkewness;
    % % %       anoms.(regionName).(seasonName).(yearRange).anomIQR =
anomIQR;
    % % %       anoms.(regionName).(seasonName).(yearRange).anomStd =
anomStd;
    % % %
    % % %
    % % %
    % % %
    % % %       if ~isempty(anoms_greater_two_field);
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_greater_two_field =
anoms_greater_two_field;
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_greater_two_values =
anoms_greater_two_values;
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_greater_two_count =
anoms_greater_two_count;
    % % %           timeseries_greater_two_count(kk,1) =
length(anoms_greater_two_values);
    % % %       else
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_greater_two_field = 0;
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_greater_two_values = [];
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_greater_two_count = [];
    % % %           timeseries_greater_two_count(kk,1) = 0;
    % % %       end; % if ~isempty
    % % %
```

```
    % % %         if ~isempty(anoms_less_two_field);
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_less_two_field =
anoms_less_two_field;
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_less_two_values =
anoms_less_two_values;
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_less_two_count =
anoms_less_two_count;
    % % %           timeseries_less_two_count(kk,1) =
length(anoms_less_two_values);
    % % %         else
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_less_two_field = 0;
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_less_two_values = [];
    % % %
anoms.(regionName).(seasonName).(yearRange).anoms_less_two_count = [];
    % % %           timeseries_less_two_count(kk,1) = 0;
    % % %         end; % if ~isempty


    %Loop to create timeseries/lineplot of greater than 2 anom
    savefig = 0;
    printQuality = '-r300';
    for ii=1:length(regionNames);
      regionName = regionNames{ii};
      figure(ii); clf reset;

      for jj=1:length(seasonNames);
        seasonName = seasonNames{jj};

        switch seasonName
          case {'Winter','DJF'}
            colorPlot = [0.3010 0.7450 0.9330];
          case {'Spring','MAM'}
            colorPlot = [0.4660 0.6740 0.1880];
          case {'Summer','JJA'}
            colorPlot = [0.9290 0.6940 0.1250];
          case {'Fall','SON'}
            colorPlot = [0.8500 0.3250 0.0980];
        end; % switch

        subplot(2,1,1);
        % plotValues is in units of number total (regardless of years)
```

```
                plotValuesGT2 =
anoms.(regionName).(seasonName).timeseries_greater_two_count(2:end);
                plotValuesLT2 =
anoms.(regionName).(seasonName).timeseries_less_two_count(2:end);
                % plotValues is in units of number per year
                plotValuesGT2 =
anoms.(regionName).(seasonName).timeseries_greater_two_count(2:end)./anoms.(region
Name).(seasonName).numYearsInChunk(2:end);
                plotValuesLT2 =
anoms.(regionName).(seasonName).timeseries_less_two_count(2:end)./anoms.(regionNa
me).(seasonName).numYearsInChunk(2:end);
                %

plot(plotValuesGT2,'Color',colorPlot','Marker','o','MarkerFaceColor',colorPlot,'markersiz
e',7,'LineWidth',2);
                Ht = title(cat(2,'Std. Anoms > +2 in the ',regionName,' in ',seasonName,' '));
                ylimValues = get(gca,'ylim');
                ylimValues = [min(ylimValues),max(ylimValues)*1.15];
                set(gca,'ylim',ylimValues);
                set(gca,'xlim',[0.9,5.1],'xtick',[1:1:5],'XTickLabel',{'1880-1909','1910-
1939','1940-1969','1970-1999','2000-2020'});
                set([Ht,gca],'fontsize',12);
                set(gca,'box','off','TickDir','out','TickLength',[0.006,0.006]);
                %
                subplot(2,1,2);
                plot(plotValuesLT2,'Color',colorPlot,'Marker','o','MarkerFaceColor',
colorPlot,'markersize',7,'LineWidth',2);
                Ht = title(cat(2,'Std. Anoms < -2 in the ',regionName,' in ',seasonName,' '));
                % % %          Hx = xlabel('Time Range: 1880-2020');
                ylimValues = get(gca,'ylim');
                ylimValues = [min(ylimValues),max(ylimValues)*1.15];
                set(gca,'ylim',ylimValues);
                set(gca,'xlim',[0.9,5.1],'xtick',[1:1:5],'XTickLabel',{'1880-1909','1910-
1939','1940-1969','1970-1999','2000-2020'});
                set([Ht,gca],'fontsize',12);
                set(gca,'box','off','TickDir','out','TickLength',[0.006,0.006]);

                if savefig; %if statement to create plots
                   ChangePlotDimensions(6.5,4);  % width, height
                   figname = cat(2,'timeseries_',regionName,'_',seasonName,'.png');
                   disp(cat(2,'** Overwriting fig to ',figname));
                   print('-dpng',printQuality,figname);
                end;  % if savefig
            end;
        end;
```

```
%Start of histogram loop for start to end time period
savefig = 0;
printQuality = '-r600'; % set to –r600 for general res; use -r1200 for submission
for ii=1:length(regionNames);
    regionName = regionNames{ii};
    for jj=1:length(seasonNames);
        figure(jj+20); clf reset;
        seasonName = seasonNames{jj};
        plotYearRanges = [2,6];
        for kk=1:length(plotYearRanges);
            if kk==1;
                switch seasonName
                    case {'Winter','DJF'}
                        colorPlot = [0.3010 0.7450 0.9330];
                    case {'Spring','MAM'}
                        colorPlot = [0.4660 0.6740 0.1880];
                    case {'Summer','JJA'}
                        colorPlot = [0.9290 0.6940 0.1250];
                    case {'Fall','SON'}
                        colorPlot = [0.8500 0.3250 0.0980];
                end; % switch

            elseif kk==2;
                switch seasonName
                    case {'Winter','DJF'}
                        colorPlot = [0 0.4470 0.7410];
                    case {'Spring','MAM'}
                        colorPlot = [0 1 0];
                    case {'Summer','JJA'}
                        colorPlot = [1 1 0];
                    case {'Fall','SON'}
                        colorPlot = [0.6350 0.0780 0.1840];
                end; % switch

            end; % if kk
            yearRange = yearRanges{plotYearRanges(kk)};
            yearChunk = yearChunks(plotYearRanges(kk),:);
            numYearsInChunk = yearChunk(2)-yearChunk(1)+1;
            % % %         Hh =
histogram(anoms.(regionName).(seasonName).(yearRange).std_anoms,'FaceColor',color
Plot);
            % % %         Hh.Values = Hh.Values./numYearsInChunk;
            % % %         disp(Hh)
            % % %         pause
```

```
histogram((anoms.(regionName).(seasonName).(yearRange).std_anoms),'FaceColor',colo
rPlot,'Normalization','probability', ...
                         'BinWidth',0.25);
                    anomMean =
anoms.(regionName).(seasonName).(yearRange).anomMean;
                    anomMedian =
anoms.(regionName).(seasonName).(yearRange).anomMedian;
                    anomSkewness =
anoms.(regionName).(seasonName).(yearRange).anomSkewness;
                    anomIQR = anoms.(regionName).(seasonName).(yearRange).anomIQR;
                    anomStd = anoms.(regionName).(seasonName).(yearRange).anomStd;

                    hold on

                    % you might need this "title" for the text annotations. what i
                    % do here is convert the yearChunk numbers to a string and line
                    % them up so they say 1880-1909, etc. in the figure
                    textLine0 = cat(2,num2str(yearChunk(1)),'-',num2str(yearChunk(2)));
                    textLine1 = (cat(2,'Mean = ',num2str(anomMean,3)));
                    textLine2 = (cat(2,'Median = ',num2str(anomMedian,3)));
                    textLine3 = (cat(2,'Skew = ',num2str(anomSkewness,3)));
                    textLine4 = (cat(2,'Iqr = ',num2str(anomIQR,3)));
                    textLine5 = (cat(2,'Std = ',num2str(anomStd,3)));

                    xline(0,'r-');yline(0);
                    % set the y limit to something that universally captures
                    % the normalized range on the y axis. might need to
                    % tinker with it
                    set(gca,'ylim',[0,0.17]);
                    ylimValues = get(gca,'ylim'); % get the y limits of the current version of
the figure
                    set(gca,'xlim',[-8,8]); % freeze the xlimits to -8 to +8

                    if kk==1;
                         Htext = text(-
6.5,0.8*max(ylimValues),[{textLine0};{textLine1};{textLine2};{textLine3};{textLine4}
;{textLine5}]);
                    elseif kk==2;
                         Htext =
text(3.5,0.8*max(ylimValues),[{textLine0};{textLine1};{textLine2};{textLine3};{textLi
ne4};{textLine5}]);
                    end; % if kk

                    Hx = xlabel('Standardized Anomaly');
```

```
            % % %           Ht = title(cat(2,'Shift in ',regionName,' Temperatures in
',seasonName,' from 1880-1909 to 2000-2020'));
            Ht = [];
            set([Hx,Ht,gca,Htext],'fontsize',12);
            set(gca,'box','off','TickDir','out','TickLength',[0.006,0.006]);
            if savefig; %if statement to create plots
                ChangePlotDimensions(6.5,4);  % width, height
                figname =
cat(2,'histogram_',regionName,'_',seasonName,'_','1880to2020_','.png');
                disp(cat(2,'** Overwriting fig to ',figname));
                print('-dpng',printQuality,figname);
            end;
            % % %           pause(2);
        end;
        pause();
    end;
end;


%Loop for Histogram of every time period; Creates several different
%plots
% highlight code and press control+i will auto-justify/tab-indent
savefig = 0; % set to 1 to save figures (slooower), set to 0 to not save figure to a
file (faster!)
printQuality = '-r300'; % set to –r600 for general res; use -r1200 for submission
for ii=1:length(regionNames);
    regionName = regionNames{ii};
    figure(ii); clf reset;

    for jj=1:length(seasonNames);
        seasonName = seasonNames{jj};

        switch seasonName
            case {'Winter','DJF'}
                colorPlot = [0.3010 0.7450 0.9330];
            case {'Spring','MAM'}
                colorPlot = [0.4660 0.6740 0.1880];
            case {'Summer','JJA'}
                colorPlot = [0.9290 0.6940 0.1250];
            case {'Fall','SON'}
                colorPlot = [0.8500 0.3250 0.0980];
        end; % switch


        for kk=1:length(yearRanges);
            yearRange = yearRanges{kk};
```

```
histogram((anoms.(regionName).(seasonName).(yearRange).std_anoms),'FaceColor',colo
rPlot);
            anomMean =
anoms.(regionName).(seasonName).(yearRange).anomMean;
            anomMedian =
anoms.(regionName).(seasonName).(yearRange).anomMedian;
            anomSkewness =
anoms.(regionName).(seasonName).(yearRange).anomSkewness;
            anomIQR = anoms.(regionName).(seasonName).(yearRange).anomIQR;
            anomStd = anoms.(regionName).(seasonName).(yearRange).anomStd;

            textLine1 = (cat(2,'mean=',num2str(anomMean,3)));
            textLine2 = (cat(2,'median=',num2str(anomMedian,3)));
            textLine3 = (cat(2,'skew=',num2str(anomSkewness,3)));
            textLine4 = (cat(2,'iqr=',num2str(anomIQR,3)));
            textLine5 = (cat(2,'std=',num2str(anomStd,3)));

            xline(0,'r-');yline(0);
            ylimValues = get(gca,'ylim'); % get the y limits of the current version of
the figure
            set(gca,'xlim',[-8,8]); % freeze the xlimits to -8 to +8
            Htext =
text(3.5,0.8*max(ylimValues),[{textLine1};{textLine2};{textLine3};{textLine4};{textLi
ne5}]);
            Ht = title(cat(2,regionName,' ',seasonName,' ',yearRange));
            Hx = xlabel('Standardized Anomaly');
            set([Hx,Ht,gca,Htext],'fontsize',12);
            set(gca,'box','off','TickDir','out','TickLength',[0.006,0.006]);

            if savefig; %if statement to create plots
              ChangePlotDimensions(6.5,4);  % width, height
              figname = cat(2,regionName,'_',seasonName,'_',yearRange,'.png');
              disp(cat(2,'** Overwriting fig to ',figname));
              print('-dpng',printQuality,figname);
            end;
          end;
        end;
      end;


      % % % for ii=1:length(regionNames);
      % % %     for jj=1:length(seasonNames);
      % % %         figure(ii); clf reset;
```

```
        % % %
pcolor((regionLon),(regionLat),(anoms.(regionName).(seasonName).(yearRange)));
        % % %       shading flat;
        % % %       colorbar;
        % % %       caxis([-3,3]);
        % % %       coastlines = load('coastlines_180_180.dat');
        % % %       hold on;
        % % %       plot(coastlines(:,1),coastlines(:,2),'k-');
        % % %    end;
        % % % end;


        %WORK THAT I CAN RESUME IF I WANT
        doOldWork = 0;
        if doOldWork;
           %Plot histogram with SE US seasonal z-score
           figure(4); clf reset
           histogram(winter_anoms_se_us_1881);
           hold on
           histogram(winter_anoms_se_us_1991);
           title('Shift in Winter Temperature in the SE United States');
           legend('SE US Winter:1881-1910','SE US Winter:1991-2020');
           xlabel('Z-Score');
           ylabel('Frequency');
           figure(5);
           histogram(summer_anoms_se_us_1881);
           hold on
           histogram(summer_anoms_se_us_1991);
           disp(nanmean(summer_anoms_se_us_1881(:)));
           disp(nanmean(summer_anoms_se_us_1991(:)));
           title('Shift in Summer Temperature in the SE United States');
           legend('SE US Summer:1881-1910','SE US Summer:1991-2020');
           xlabel('Z-Score');
           ylabel('Frequency');
        end;
        % % % toc
     end; % if doThis
  end % function

%Function that corrects plot dimensions
  function [] = ChangePlotDimensions(width,height,orientation);
     %
     %  Change dimensions of a plot and (optionally) the paper/layout
     %  orientation.  Run as:
     %
     %  ChangePlotDimensions(width,height,orientation);
```

```
    %  ChangePlotDimensions(6,9);
    %  ChangePlotDimensions(6,9,'portrait');
    %  ChangePlotDimensions(9,6,'landscape');
    %
    %  width and height are in inches, so keep that in mind when creating a
    %  figure size.  width and height are wrt landscape of portrait
    %  orientation, where portrait is the default.  a good wxh for portrait is
    %  6x9, while for landscape is 9x6.
    %
    if nargin==2;
        orientation = 'portrait';
    end;  % if nargin loop
    set(gcf,'PaperPositionMode','manual');
    set(gcf, 'PaperUnits', 'inches');
    switch orientation
        case 'portrait'
            figPos = [0.5,0.5];
        case 'landscape'
            figPos = [0.5,0.5];
        otherwise
            error('!! Unknown paper orientation, dude');
    end;  % switch loop
    % PaperPosition is left, bottom, width, height but i think this flips in
    % landscape mode
    set(gcf,'PaperPosition',[figPos(1),figPos(2),width,height]);
    set(gcf,'PaperOrientation',orientation);
end % function
```