GLOBAL TEMPERATURE TRENDS

by

STEPHEN MORRIS

(Under the Direction of Lynne Seymour)

ABSTRACT

Trends in global temperature anomalies from the past half century are investigated. To account for the seasonality of the data, a statistical model that includes monthly intercept and slope terms will be developed. Atmospheric CO_2 concentrations, sunspot numbers, and various global pressure oscillations will also be included in the model. Least squares estimates of all parameters in the model will be derived, as will standard errors that account for temporal correlations in the data. To this end, standardized residuals will be modeled using an autoregressive moving average (ARMA) model. Derived standard errors will be based on the fitted ARMA model.

INDEX WORDS: Positive serial correlation, seasonality, global warming, CO_2, ordinary least squares estimation, standard error, autoregressive moving average.

GLOBAL TEMPERATURE TRENDS

by

STEPHEN MORRIS

BS in Statistics, University of Georgia, 2010

BA in Sociology, University of Georgia, 2010

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2014

GLOBAL TEMPERATURE TRENDS

by

Stephen Morris

Major Professor: Lynne Seymour
Committee: Jaxk Reeves
Kim Love-Myers

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

ACKNOWLEDGEMENTS

A special thanks to those who, years ago, encouraged me to go to grad school (Dr. Coverdill, Dr. Sriram, and Dr. Mandal) and to those who, more recently, guided me towards graduation (Dr. Seymour, Dr. Reeves, and Dr. Love-Myers). The friends I made along the way were also crucial to my success, as the journey seems less daunting when surrounded by familiar faces and kindred spirits. Our troubles were all the same.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In popular literature, the global temperature trend over the past millennium is often compared to the shape of a hockey stick – relatively flat for a long period of time, followed by an abrupt increase. This abrupt increase, depicted by the blade of the stick, represents the trend in global temperatures over the past century. To this end, the comparison is spot on. Global temperatures have clearly been trending upwards during the past century. Over the past fifteen years, however, the rate of increase appears to have slowed down. Some point to this as evidence that warming has stopped altogether, while others contend that this is not the case and warming will continue into the foreseeable future.

The objective of this research is to investigate global temperature trends since the mid-1900s. In doing so, we will weigh in on the point of contention mentioned in the previous paragraph. The data set we will use is NASA's Global Land-Ocean Temperature index, which reports average monthly global temperature anomalies (the difference between the average global temperature for a given month and the average global temperature over a 30 year base period). To account for the seasonality of the data, a model that includes monthly intercepts and monthly trend parameters will be considered. Some variables that have long been associated with global temperatures will also be considered: atmospheric CO_2 concentrations, sunspot numbers, and a few global pressure oscillations.

It is not uncommon for the error terms of models based on climate data to be correlated. However, this correlation violates one of the basic assumptions of the general linear model.

When the residuals show positive serial correlation, standard errors of the parameter estimates will be underestimated, as will significance levels of the parameter effects (Lund, Seymour, and Kafadar, 2001). With artificially small significance levels, statistical significance might be inferred incorrectly. After arriving at a final model, standard errors that take residual correlation into account will need to be estimated. The new standard errors will then be used to calculate accurate test statistics and significance levels. If necessary, these new significance levels will be used to adjust the model and the process of calculating accurate standard errors will be repeated.

CHAPTER 2

DATA

2.1 Temperature Anomalies

A temperature anomaly measures the difference between the mean global temperature for a given month (in 0.01 degrees Celsius) and the mean global temperature for that month from 1951 to 1980. NASA reports that mean global temperature anomalies are computed from station temperature anomalies and the station temperature anomalies come from a regularly spaced array of virtual stations covering both land and ocean. At each station, an average daily temperature is calculated by averaging the daily high and the daily low (Dunbar, 2005). These average daily temperatures can be averaged over the course of a month to yield the average monthly temperature. With the average monthly temperature, a temperature anomaly can be calculated. A positive temperature anomaly indicates that the average global temperature for the month is higher than the average global temperature for that month during the 30 year base period; a negative temperature anomaly indicates the reverse. Figure 1 depicts the anomalies for January 1880 through January 2014. From the late 1970s to the present, it is clear that the temperature anomalies have been increasing, but the rate of increase appears to have slowed in recent years.

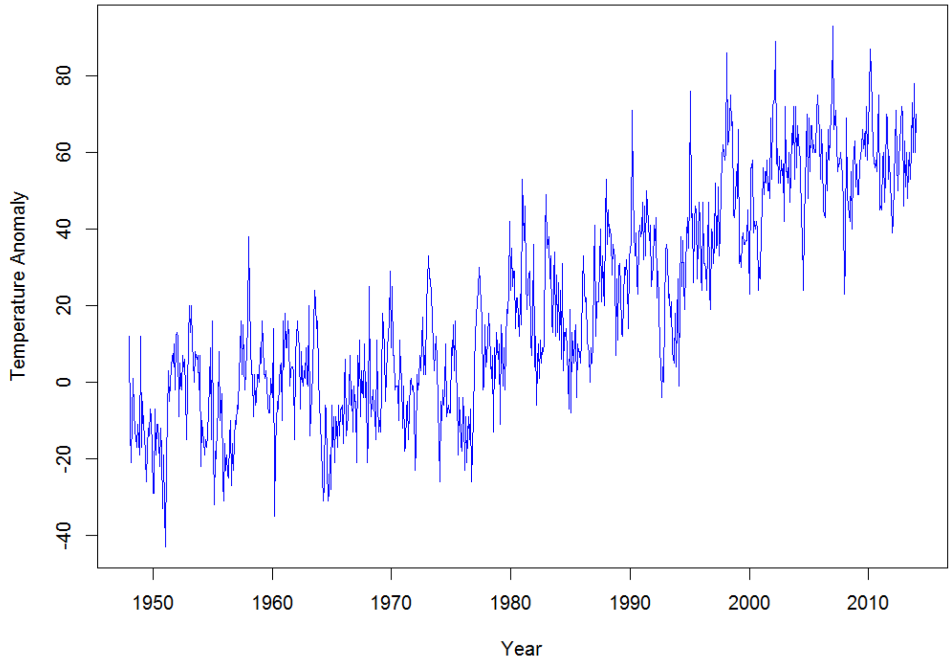

Figure 1: Average monthly temperature anomalies from January 1880 to January 2014

The temperature anomaly data set used in this paper can be found online at http://data.giss.nasa.gov. The data set reports monthly global temperature anomalies for every month from January, 1880, to the present. The data are adjusted for station moves, equipment updates, and any other non-climatic temperature jumps.

2.2 Atmospheric CO_2 Concentrations

Greenhouse gases are known to affect temperatures on the Earth. As CO_2 is one of the primary greenhouse gases, it is imperative to consider the role of atmospheric CO_2 concentrations when studying global temperatures.

Since the late 1950s, daily atmospheric CO_2 concentrations have been taken at the Mauna Loa Observatory in Hawaii. As monthly means, these data are available online at co2now.org. The website reports that, "Monthly mean CO_2 concentrations are determined from daily averages for the number of CO_2 molecules in every one million molecules of dried air (water vapor removed)." The data, recorded in parts per million (ppm), are graphed in Figure 2.

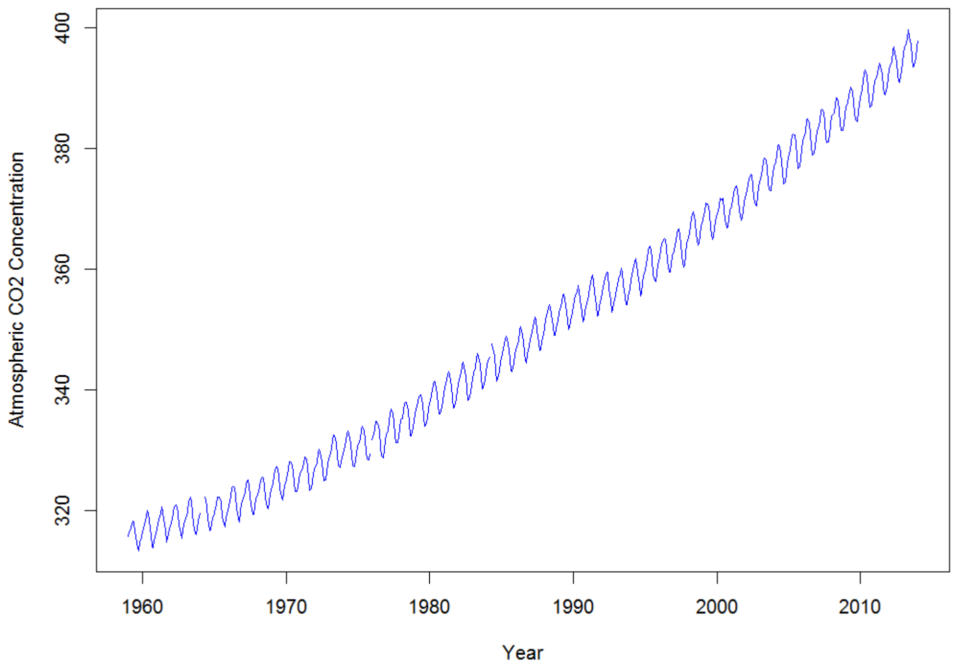

Figure 2: Average monthly atmospheric CO_2 concentration from January 1959 to January 2014

Clearly, atmospheric CO_2 concentrations have been increasing since the Mauna Loa record began in the late 1950s. It is also worth noting that CO_2 concentrations are known to be well mixed in the atmosphere, so using concentrations from just one site should not introduce any bias.

2.3 Sunspot Numbers

A sunspot is a visibly dark spot that appears on the surface of the sun. These spots are dark because they are cooler than the sun's surrounding surface area – this is a product of magnetic activity. Whether or not sunspots affect Earth's climate is still up for debate. However, periods of low solar activity have occurred in conjunction with cooler global temperatures in the past. The Maunder Minimum, for example, was a period of decreased solar activity that coincided with part of the Little Ice Age (1645 – 1715). Thus, there has been speculation that sunspots have some effect on the global climate.

Sunspots can occur individually or in groups. A sunspot group will contain an average of ten spots. The sunspot number can be calculated by summing the number of individual spots and ten times the number of groups. NASA provides sunspot number data online at solarscience.msfc.nasa.gov. Monthly means are available from 1749 to the present. Means from January 1948 to January 2014 are plotted in Figure 3. The cycle seen in the data is roughly eleven years.

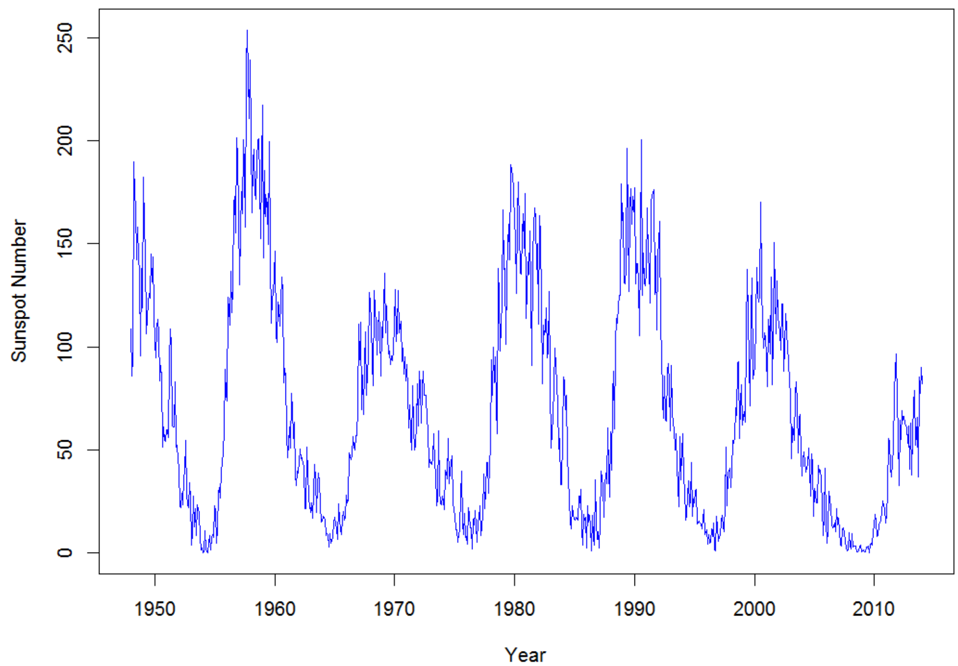

Figure 3: Average monthly sunspot number from January 1948 to January 2014

2.4 Global Pressure Oscillations

A global pressure oscillation is an oscillation in surface air pressure between two or more poles. These oscillations can have pronounced effects on the global climate, though these effects can be complex. One global pressure oscillation discussed herein is the El Niño Southern Oscillation (ENSO). According to the NOAA, "The ENSO cycle refers to the coherent and sometimes very strong year-to-year variations in sea surface temperatures, convective rainfall, surface air pressure, and atmospheric circulation that occur across the equatorial Pacific Ocean. El Niño and La Niña represent opposite extremes in the ENSO cycle."

The strength of a pressure oscillation is usually reported as a function of the difference in air pressure between the poles of the oscillation. Or, perhaps more precisely, the oscillation is a function of the heights at which a certain level of air pressure occurs. The ENSO, for example, is a function of the 500 millibar (mb) heights in Tahiti and Darwin, Australia. The 500mb height is the height at which 500mb of air pressure occurs. (As a reference, air pressure at sea level hovers around 1000mb and decreases as elevation increases.) In this paper, 500mb heights are used instead of surface pressure measurements because the poles may be located at different heights above sea level. Thus, the surface pressure at one of the poles might always be higher than the surface pressure at another pole. In a manner of speaking, comparing the 500mb heights instead of surface pressure serves to level the playing field. Figure 4 depicts monthly averages of the 500mb height in Darwin from January 1948 to January 2014.

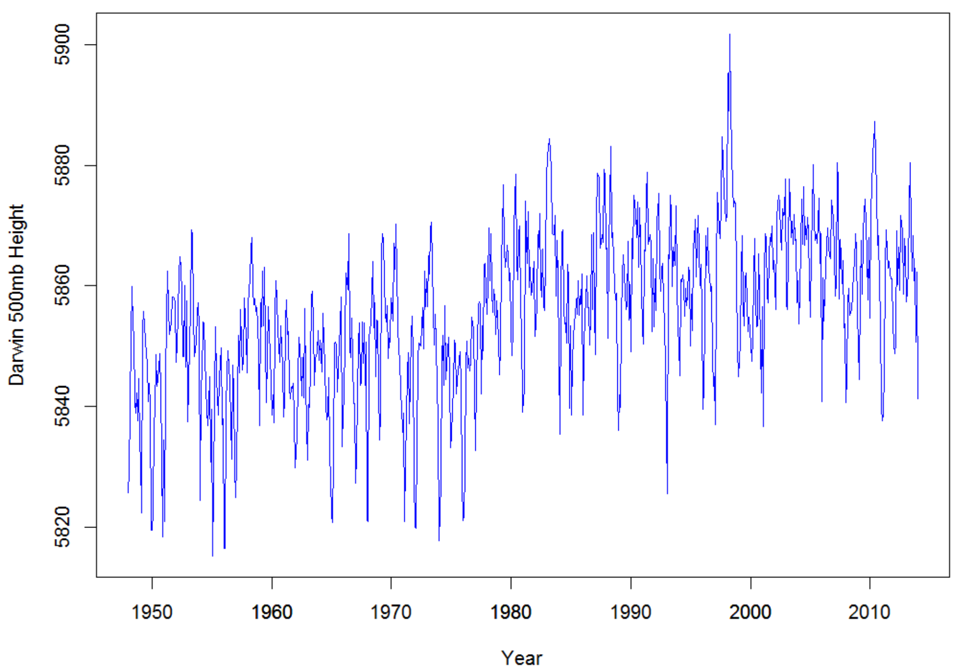

Figure 4: Average monthly 500mb height in Darwin from January 1948 to January 2014

Two other pressure oscillations will be considered in this paper: the North Atlantic Oscillation (NAO) and the Interdecadal Pacific Oscillation (IPO). The NAO has poles in Azores and Iceland. Concerning the positive and negative phases of the NAO, the NOAA reports that, "Both are associated with changes in the intensity and location of the North Atlantic jet stream and storm track, and in large-scale modulations of the normal patterns of zonal and meridional heat and moisture transport, which in turn results in changes in temperature and precipitation patterns often extending from eastern North America to western and central Europe." Figure 5 depicts monthly averages of the 500mb heights in Iceland from January 1948 to January 2014.

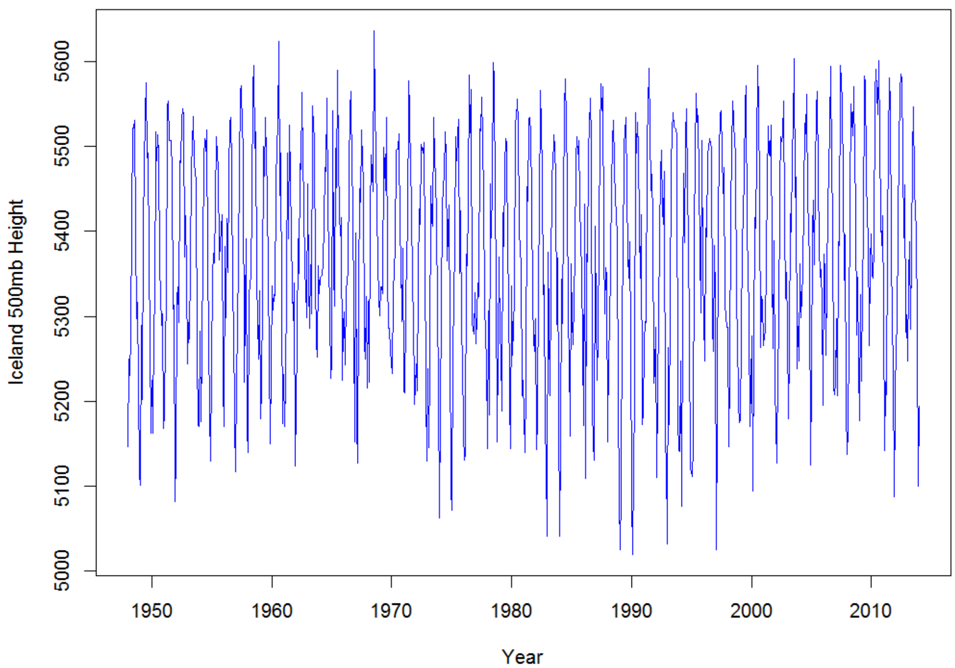

Figure 5: Average monthly 500mb height in Iceland from January 1948 to January 2014

The IPO has poles in the North Central Pacific Ocean (NCP) and the Gulf of Alaska (GOA). Like ENSO and NAO, phases of the IPO can affect the climate both globally and locally. Average monthly 500mb heights in NCP from January 1948 to January 2014 are shown in Figure 6.

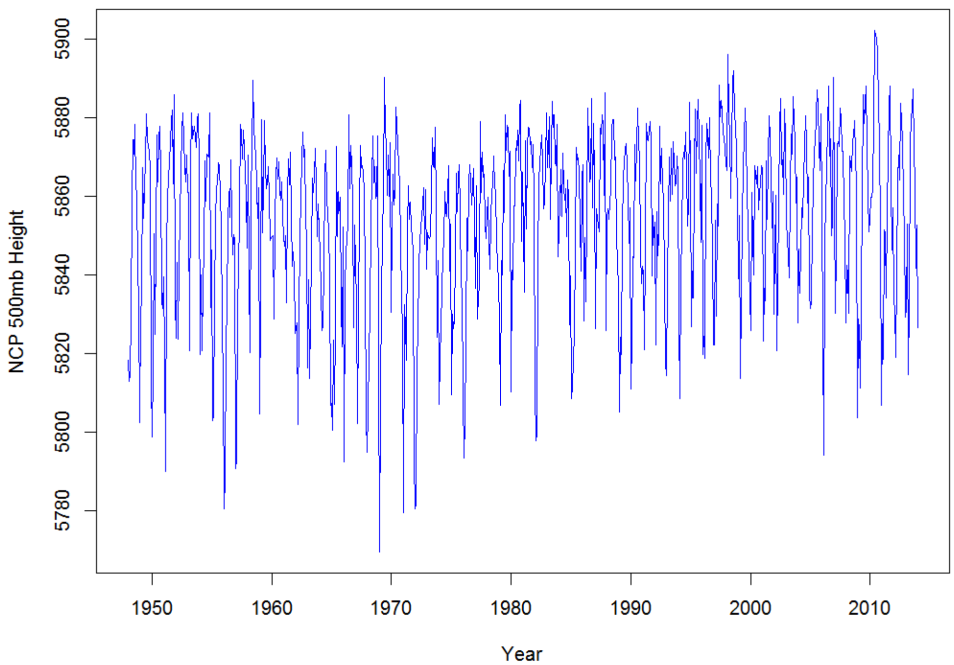

Figure 6: Average monthly 500mb height in NCP from January 1948 to January 2014

The average monthly 500mb heights of the poles of each of these oscillations are available from the NOAA website. GPS coordinates of each pole are shown in Table 1. The data run from January 1948 to January 2014.

Table 1: GPS Coordinates			
Pole	**Coordinates**	**Pole**	**Coordinates**
Darwin (ENSO)	12 S, 131 E	Tahiti (ENSO)	18 S, 150 W
Azores (NAO)	39 N, 24 W	Iceland (NAO)	64 N, 24 W
NCP (IPO)	20 N, 170 W	GOA (IPO)	55 N, 150 W

Summary statistics for all nine variables discussed in this chapter are shown in Table 2. It is worth noting that the CO_2 data do not extend as far back in time as the other data sets do. In the sections that follow, all models will use data from January 1959 to January 2014. For this reason, the summary statistics shown in Table 2 only reflect data from January 1959.

Table 2: Summary Statistics					
Variable	**n**	**Mean**	**SD**	**Minimum**	**Maximum**
TA	661	25.12	25.58	-35.00	93.00
CO_2	656	350.13	24.05	313.26	399.76
SSN	661	65.29	49.97	0.00	217.40
Darwin	661	5857.17	12.75	5817.74	5901.80
Tahiti	661	5855.47	13.39	5806.23	5903.43
Azores	661	5740.98	90.51	5478.21	5910.84
Iceland	661	5367.88	132.12	5019.14	5636.42
NCP	661	5854.56	21.63	5769.45	5902.20
GOA	661	5422.82	140.22	5132.61	5738.42

2.5 Multicollinearity

Since each of the eight independent variables mentioned in this section is climate-related, an investigation of their correlations is warranted. Correlations between the independent variables and the temperature anomalies are shown in Table 3. Note that the eight independent variables have not been seasonally adjusted, so each probably features a seasonal component. Because of this, these correlations are likely inflated. All correlations were calculated based on data from January 1959 to January 2014.

Table 3: Correlation Matrix									
	TA	**CO_2**	**SSN**	**Darwin**	**Tahiti**	**Azores**	**Iceland**	**NCP**	**GOA**
TA	1.00	0.86	-0.06	0.57	0.52	0.04	-0.01	0.27	-0.03
CO_2	0.86	1.00	-0.19	0.48	0.49	0.04	0.03	0.18	0.03
SSN	-0.06	-0.19	1.00	0.08	-0.03	0.03	-0.05	0.03	0.03
Darwin	0.57	0.48	0.08	1.00	0.67	0.21	0.29	0.58	0.18
Tahiti	0.52	0.49	-0.03	0.67	1.00	0.02	0.18	0.30	0.09
Azores	0.04	0.04	0.03	0.21	0.02	1.00	0.48	0.60	0.71
Iceland	-0.01	0.03	-0.05	0.29	0.18	0.48	1.00	0.51	0.72
NCP	0.27	0.18	0.03	0.58	0.30	0.60	0.51	1.00	0.43
GOA	-0.03	0.03	0.03	0.18	0.09	0.71	0.72	0.43	1.00

The variable showing the strongest correlation with the temperature anomalies is atmospheric CO_2 concentrations ($r = 0.86$). As expected, the poles of each oscillation show correlation with each other ($r = 0.67$ for ENSO, $r = 0.48$ for NAO, and $r = 0.43$ for IPO). We also see a good deal of correlation across oscillations – GOA and Iceland ($r = 0.72$), GOA and Azores ($r = 0.71$), and NCP and Azores ($r = 0.60$), to name a few. Again, recall that these correlations could be inflated. The independent variables could also show strong correlation with the monthly intercept and slope terms. All correlations will be taken into consideration when fitting models, as multicollinearity can inflate the variance (and therefore standard error) of the coefficient estimates. With inflated standard errors, picking which predictors to include can be difficult and the exact effect of each predictor might be unclear.

CHAPTER 3

STATISTICAL METHODS

3.1 Regression Model

The regression model we will consider is:

$$\begin{aligned} X_{nT+v} = \beta_0 + \beta_1 Darwin_{nT+v} + \beta_2 Tahiti_{nT+v} + \beta_3 Iceland_{nT+v} + \beta_4 Azores_{nT+v} \\ + \beta_5 NCP_{nT+v} + \beta_6 GOA_{nT+v} + \beta_7 CO2_{nT+v} + \beta_8 SSN_{nT+v} \\ + \mu_1 + \ldots + \mu_7 + \mu_9 + \ldots + \mu_{12} + \alpha(nT+v) + \alpha_1(nT+v)I_1 + \ldots \\ + \alpha_7(nT+v)I_7 + \alpha_9(nT+v)I_9 + \ldots + \alpha_{12}(nT+v)I_{12} + \varepsilon_{nT+v} \end{aligned} \tag{3.1}$$

where X_{nT+v} represents the average global temperature anomaly during the v^{th} month of year $n \geq 0$. Time will be scaled such that $n = 0$ represents January, 1959. Naturally, $v = 1$ corresponds to January, $v = 2$ corresponds to February, and so on. The period of the data, T, is 12. The expression $(nT + v)$ can be thought of as an index.

$Darwin_{nT+v}$ represents the average 500mb height in Darwin for the v^{th} month of year n, and β_1 represents the expected change in the mean global temperature anomaly for a one meter increase in the average 500mb height in Darwin while all other variables are held constant. Similar definitions can be provided for the other explanatory variables and their coefficients.

The monthly intercept terms $(\mu_1, \ldots, \mu_7, \mu_9, \ldots, \mu_{12})$ represent deviations from the average August temperature anomaly during month v without trend. The monthly slope terms $(\alpha_1, \ldots, \alpha_7, \alpha_9, \ldots, \alpha_{12})$ represent deviations from the average August temperature anomaly change rate during month v. Thus, α_1 represents the difference between the year to year temperature anomaly change rates (in 0.01 degrees Celsius) during the months of January and

August. This rate of change might vary from month to month, so terms are included for each month. The I_v variables are indicator variables that will equal 1 during month v and 0 otherwise. By itself, α can be thought of as a baseline change rate.

The error terms, ε_{nT+v}, are assumed to be normally distributed with a mean of zero and a constant variance. Further, these terms are often assumed to be independently distributed. As noted in the introduction, it is this second assumption that we anticipate will be violated. If this proves to be the case, the standard errors of the model parameters will be underestimated, test statistics will be inflated, and significance levels will be artificially small. Accurate standard errors that account for residual correlation will need to be derived. This process will be described in a later section.

3.2 Parameter Estimation

Parameter estimates will be calculated according to the method of ordinary least squares. Using matrix notation, (3.1) can be represented as:

$$\mathbf{X} = \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{X} = (X_1, \ldots, X_m)'$ is a vector containing the average monthly global temperature anomalies from January 1959 to January 2014, $\mathbf{D}$ is the $m \times 32$ design matrix (described below), $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_8, \mu_1, \ldots, \mu_7, \mu_9, \ldots, \mu_{12}, \alpha, \alpha_1, \ldots, \alpha_7, \alpha_9, \ldots, \alpha_{12})'$ is the parameter vector, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_m)'$ is a vector of regression errors.

The first column of **D** will correspond to the intercept term, so it will be a vector of 1s. The next six columns of the design matrix **D** will correspond to the average 500mb heights at Darwin, Tahiti, Iceland, Azores, NCP (North-Central Pacific), and GOA (Gulf of Alaska). Thus, the entry in the first row of the second column of **D** is the average 500mb height in Darwin for January 1959. The entry in the last row of the second column is the average 500mb height in

Darwin for January 2014. The eighth and ninth columns will correspond to the average atmospheric CO_2 levels and the average sunspot number, respectively. Columns ten through twenty correspond to the eleven monthly intercept terms. Again, note that these terms represent deviations from the average August temperature anomaly during month *v* without trend. An entry in column ten will be 1 if the data in that row are from January, 0 otherwise. Using the remaining ten months, similar statements can be made about entries in columns eleven through twenty. Column 21 will be equal to the index, $nT + v$. Columns 22 through 32 will correspond to the eleven monthly slope deviation terms. An entry in column 22 will be $nT + v$ if the data in that row are from January, 0 otherwise. The first nine rows of columns ten through 32 will look like:

$$\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0
\end{matrix}$$

As the model is adjusted, of course, **D** will need to be adjusted as well.

An estimator of $\boldsymbol{\beta}$ can be computed using the ordinary least squares formula:

$$\hat{\boldsymbol{\beta}} = (\mathbf{D'D})^{-1}\mathbf{D'X}.$$

The standard error of the i^{th} parameter estimate in $\hat{\boldsymbol{\beta}}$ will be the square root of the $(i, i)^{th}$ entry in the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$:

$$\text{var}\left(\hat{\boldsymbol{\beta}}\right) = \sigma^2(\mathbf{D'D})^{-1} \tag{3.2}$$

where σ^2 can be estimated by the mean squared error:

$$\hat{\sigma}^2 = \text{MSE} = \frac{\sum(\varepsilon_{nT+v})^2}{m - p}$$

where p is the number of parameters in the model. Again, note that the standard errors described above will be underestimated if the model error terms are found to be positively correlated. As a result, parameter significance levels will also be underestimated. The process of calculating accurate standard errors and significance levels is described in the next section. For ease, the underestimated significance levels will be used until we arrive at a final model. Note that if a parameter is found to be statistically insignificant with an underestimated significance level, it will also be statistically insignificant when a more accurate significance level is calculated. Thus, using the underestimated significance levels will not increase the chances of erroneously tossing a statistically significant variable.

In progressing towards a final model, we will first remove insignificant variables by backwards elimination ($\alpha = 0.05$). After all remaining variables are significant, variance inflation factors (VIF) will be used to address any issues with multicollinearity. The VIF measures how much the variances of estimated regression coefficients are inflated when compared to having uncorrelated predictors, thus quantifying the severity of multicollinearity in a model. A VIF above 10 indicates that multicollinearity may have pronounced effects on the least squares estimates of the regression coefficients (Kutner, Nachtsheim, Neter, and Li, 2005). The variance inflation factor for the j^{th} variable can be calculated:

$$\mathrm{VIF}_j = \frac{1}{1 - R_j^2}$$

where R_j^2 is the R^2 value of a model with response variable j and with explanatory variables which are the remaining explanatory variables from the original model. A backwards elimination approach with respect to VIFs will be used with a cutoff of 10. After examining variance inflation factors, residual plots will be used to judge the validity of the model. If the residual plot shows that the fit is inadequate, the model will need to be adjusted.

3.3 Correlated Residuals

Upon arriving at a final model, we will investigate the residuals further for correlation. To this end, we will look at an ACF plot of the residuals and also run the Durbin-Watson test. The null hypothesis for this test is that autocorrelation is not present. Thus, if the Durbin-Watson test returns a p-value less than 0.05, we can say that the residuals are correlated. The consequences of correlation in the residuals have already been noted. If the assumption that residuals are independently distributed is found to be violated, the variance-covariance matrix in (3.2) will need to be recalculated. The adjusted variance-covariance matrix will be calculated as:

$$\mathrm{var}\left(\hat{\boldsymbol{\beta}}\right) = (\mathbf{D'D})^{-1}(\mathbf{D'\Gamma D})(\mathbf{D'D})^{-1} \tag{3.3}$$

where $\mathbf{\Gamma}$ is the variance-covariance matrix of the error terms. Because it is common for climate data (temperature data especially) to show seasonal cycles in variability, we may not want to estimate $\mathbf{\Gamma}$ directly from the sample. Lund *et al.* (2001) suggest that the error terms may be modeled as the product of a periodic function and a stationary series:

$$\varepsilon_m = \sigma(m)A_m$$

where σ is a non-negative periodic function with period T (in our case, $T = 12$), and $\{A_m\}$ is a mean zero stationary series with autocovariance $\gamma_A(h) = \mathrm{Cov}(A_{t+h}, A_t)$ at lag h. The parameters $\sigma(v)$ can be estimated as monthly standard deviations, and $\{A_m\} = \{\sigma(m)^{-1}\varepsilon_m\}$ can be modeled with the autoregressive moving average (ARMA) family of models. With small samples, the AICc statistic will select a more parsimonious ARMA model than will the AIC statistic (Hurvich and Tsai, 1989). Though this distinction may be lost in larger samples, we will use the AICc statistic in selecting a best model. Because it is not computationally difficult, the AIC statistic and the BIC statistic will also be considered. After fitting an ARMA model to $\{A_m\}$, an estimate for $\gamma_A(h)$ can be calculated for $h = (0, 1, \ldots, m-1)$. Using these estimates, we can estimate $\mathbf{\Gamma}$:

$$\widehat{\mathbf{\Gamma}}_{i,j} = \hat{\sigma}\left(i - T\left\lfloor\frac{i}{T}\right\rfloor\right)\hat{\sigma}\left(j - T\left\lfloor\frac{j}{T}\right\rfloor\right)\hat{\gamma}_A(|i-j|) \tag{3.4}$$

where i and j each range from 1 to m, and where $\lfloor\cdot\rfloor$ represents the greatest integer function. (Thus, each index of $\hat{\sigma}$ in (3.4) will range from 1 to 12.)

Before inserting (3.4) into (3.3), the residuals of the ARMA model will need to be examined – both a residual plot and an ACF plot of the residuals will be considered, as will a test for serial correlation. The Durbin-Watson test is known to be biased for ARMA models, so the Ljung-Box test will be used in its place. Like the Durbin-Watson test, the null hypothesis for the Ljung-Box test is that autocorrelation is not present. The Ljung-Box test, however, tests for overall randomness based on a certain number of lags instead of testing for autocorrelation at each distinct lag. There does not seem to be a consensus on the number of lags to use when running this test. Tsay (2005) suggests using ln(N) lags (where N is the length of the series) to provide the best power performance. Alternatively, if the data are seasonal, Tsay (2005) suggests using some multiple of the period. To this end, we will run the test using 12 lags, 24 lags, 36 lags, and also 48 lags. Ideally, we will fail to reject the null hypothesis in each case and conclude that the residuals of the ARMA model are essentially white noise. If this is the case, then the autocovariance structure of the model can be used to estimate $\mathbf{\Gamma}$. Using this estimate, accurate standard errors will be calculated with (3.3). With accurate standard errors, accurate test statistics and significance levels can be calculated. If some of the new significance levels are found to be greater than 0.05, this entire process will need to be repeated with an updated model.

CHAPTER 4

RESULTS

4.1 Fitting a Model

After using backwards selection and removing significant multicollinearity, (3.1) reduced to:

$$
\begin{aligned}
X_{nT+v} = {} & \beta_0 + \beta_1 Darwin_{nT+v} + \beta_3 Iceland_{nT+v} + \beta_4 Azores_{nT+v} + \beta_5 NCP_{nT+v} \\
& + \beta_7 CO2_{nT+v} + \beta_8 SSN_{nT+v} + \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_9 + \mu_{10} \\
& + \mu_{11} + \mu_{12} + \varepsilon_{nT+v}.
\end{aligned} \tag{4.1}
$$

Model (4.1) produces an RMSE of 11.39 and an R^2 value of 0.81, suggesting this model can account for 81% of the variation in the temperature anomalies. Parameter estimates, standard errors, test statistics, and significance levels are provided in Table 4.

Table 4: Model Results for (4.1)				
Parameter	**Estimate**	**Standard error**	**Test statistic**	**P-value**
β_0	-4338.089	294.569	-14.727	< 0.001
β_1 (Darwin)	0.498	0.057	8.663	< 0.001
β_3 (Iceland)	0.014	0.007	2.196	0.028
β_4 (Azores)	0.025	0.008	3.013	0.003
β_5 (NCP)	0.160	0.035	4.585	< 0.001
β_7 (CO_2)	0.795	0.024	33.528	< 0.001
β_8 (SSN)	0.025	0.010	2.610	0.010
μ_1	24.731	3.105	7.964	< 0.001
μ_2	25.305	3.123	8.104	< 0.001
μ_3	21.360	2.964	7.207	< 0.001
μ_4	9.630	2.459	3.917	< 0.001
μ_9	4.197	1.787	2.349	0.019
μ_{10}	7.591	2.055	3.694	< 0.001
μ_{11}	12.783	2.443	5.233	< 0.001
μ_{12}	16.193	2.867	5.648	< 0.001

Observed and fitted values for (4.1) are shown in Figure 7.

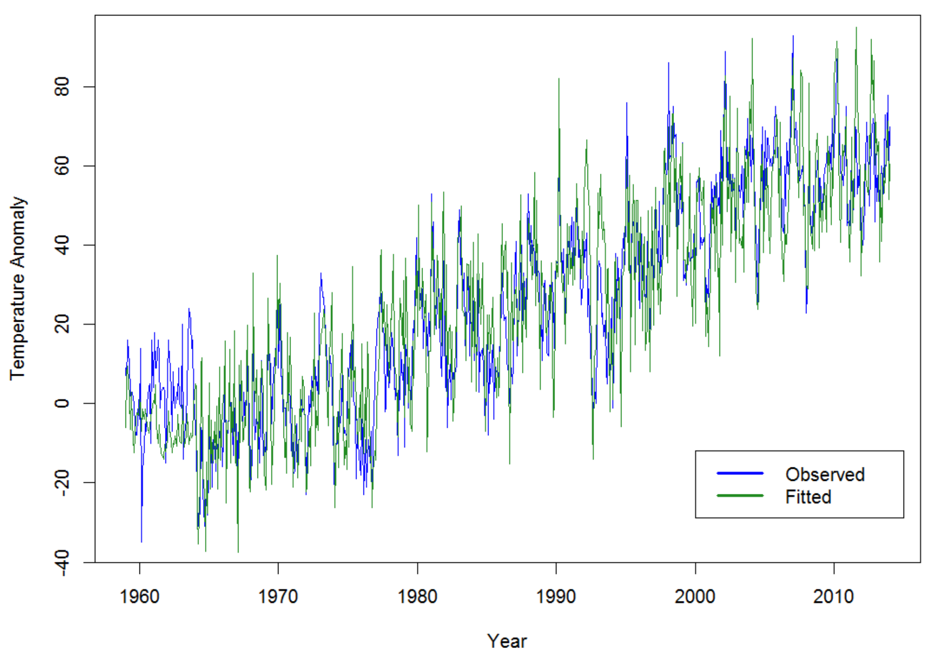

Figure 7: Observed and fitted values for (4.1)

The model does seem to capture the overall trend of the temperature anomalies, though it looks like the model underestimates observed values for the first few years of the record. A residual plot for (4.1) is shown in Figure 7. A reference line at 0 is included.

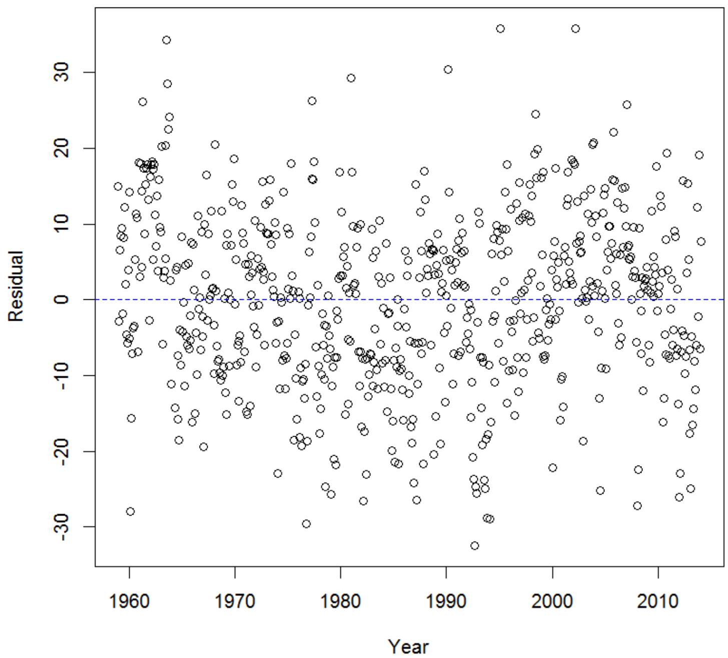

Figure 8: Residuals of (4.1)

Ideally, residuals will be evenly scattered around 0. As seen in Figure 7, the model tends to underestimate in the first few years of the record. This pattern is also observable in Figure 8. Further, it appears the residuals follow a slight U-shape. The U-shape suggests that the linear fit may not be adequate. To address this issue, a quadratic term was added to the model:

$$X_{nT+v} = \beta_0 + \beta_1 Darwin_{nT+v} + \beta_3 Iceland_{nT+v} + \beta_4 Azores_{nT+v} + \beta_5 NCP_{nT+v}$$
$$+ \beta_7 CO2_{nT+v} + \beta_8 SSN_{nT+v} + \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_9 + \mu_{10} \qquad (4.2)$$
$$+ \mu_{11} + \mu_{12} + \lambda(nT+v)^2 + \varepsilon_{nT+v}.$$

Here, the interpretation for λ is not as simple as the interpretations for the other parameters. However, note that the index, $(nT+v)$, is essentially a function of time. Thus, $(nT+v)^2$ could be considered an increasing function of time. To a certain extent, we can say λ is capturing the effect of time. Parameter estimates, standard error estimates, test statistics, and significance levels for (4.2) are shown in Table 5.

Table 5: Model Results for (4.2)

Parameter	Estimate	Standard error	Test statistic	P-value
β_0	-4476.958	291.578	-15.354	< 0.001
β_1 (Darwin)	0.579	0.059	9.774	< 0.001
β_3 (Iceland)	0.009	0.007	1.377	0.169
β_4 (Azores)	0.022	0.008	2.686	0.007
β_5 (NCP)	0.146	0.034	4.224	< 0.001
β_7 (CO_2)	0.153	0.141	1.087	0.278
β_8 (SSN)	0.030	0.009	3.161	0.002
μ_1	22.830	3.085	7.400	< 0.001
μ_2	23.831	3.091	7.710	< 0.001
μ_3	19.760	2.938	6.725	< 0.001
μ_4	8.723	2.429	3.591	< 0.001
μ_9	1.333	1.865	0.714	0.475
μ_{10}	3.882	2.177	1.783	0.075
μ_{11}	9.544	2.506	3.809	< 0.001
μ_{12}	13.505	2.883	4.685	< 0.001
λ	1.164e-04	2.524e-05	4.612	< 0.001

Note that after introducing the quadratic piece to the model, several variables are no longer significant. Further, the test statistic for CO_2 decreased from 33.528 to 1.087 (while its standard error increased from 0.024 to 0.141). This is likely due to multicollinearity, as CO_2 and λ had VIFs of 58.59 and 55.32 respectively. These two variables may essentially be measuring the same thing. After removing the insignificant parameters and multicollinearity from (4.2), the model reduced to:

$$X_{nT+v} = \beta_0 + \beta_1 Darwin_{nT+v} + \beta_4 Azores_{nT+v} + \beta_5 NCP_{nT+v} + \beta_8 SSN_{nT+v}$$
$$+ \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_{11} + \mu_{12} + \lambda(nT+v)^2 + \varepsilon_{nT+v}. \quad (4.3)$$

This model produced an RMSE of 11.33 and an R^2 value of 0.82, suggesting that it can account for about 82% of the variation in the temperature anomalies. Parameter estimates, standard error estimates, test statistics, and significance levels are shown in Table 6. Note that for (4.3), the matrix **D′D** (where **D** is the design matrix) was found to be singular.

Table 6: Model Results for (4.3)

Parameter	Estimate	Standard error	Test statistic	P-value
β_0	-4431.779	279.194	-15.873	< 0.001
β_1 (Darwin)	0.595	0.055	10.884	< 0.001
β_4 (Azores)	0.018	0.008	2.354	0.019
β_5 (NCP)	0.143	0.034	4.188	< 0.001
β_8 (SSN)	0.033	0.009	3.510	< 0.001
μ_1	19.412	2.144	9.055	< 0.001
μ_2	20.220	2.188	9.242	< 0.001
μ_3	16.202	2.113	7.668	< 0.001
μ_4	5.818	2.021	2.879	0.004
μ_{11}	6.542	1.788	3.659	< 0.001
μ_{12}	9.924	1.942	5.109	< 0.001
λ	1.455e-04	4.160e-06	34.973	< 0.001

Observed and fitted values for this model are shown in Figure 9.

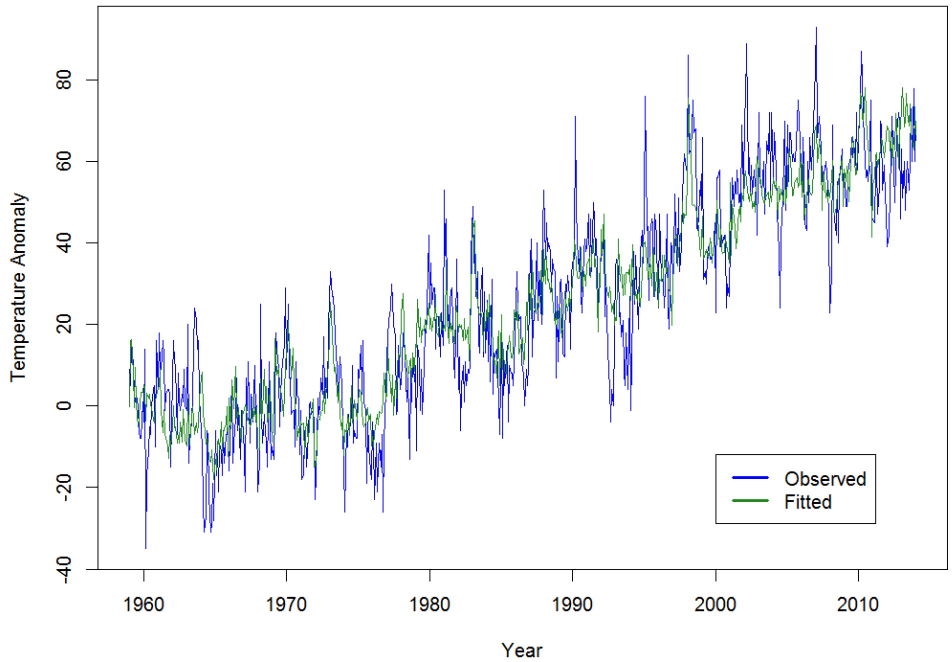

Figure 9: Observed and fitted values for (4.3)

Like in (4.1), this model seems to capture the overall trend of the temperature anomalies well, and we see less variation in the fitted values. Residuals for (4.3) are shown in Figure 10.

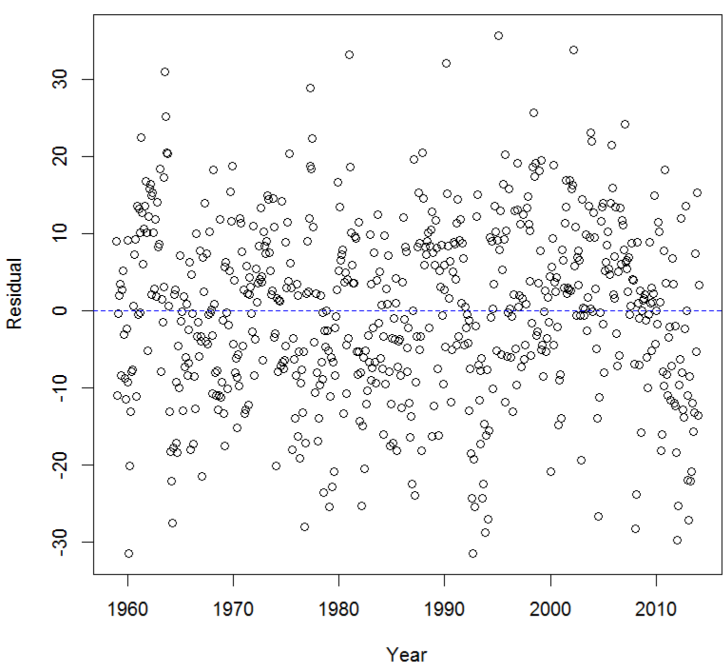

Figure 10: Residuals of (4.3)

Including the quadratic term did dampen the U-shape, though it doesn't look like the trend has entirely been removed. This leaves us at a bit of a crossroads – continue with (4.1) or continue with (4.3)? By RMSE (11.39, 11.33) and R^2 (0.81, 0.82), the models are very similar. The main difference between the two is that (4.1) includes CO_2 and (4.3) includes the quadratic parameter, λ, which is not easy to interpret and which might be masking the effect of CO_2. As far as climate goes, (4.1) certainly makes more sense than (4.3), as it is not safe to assume the quadratic behavior will extend beyond the range of the data. Both models have some deficiencies, but we will continue with (4.1). Turning our heads to positive serial correlation, an ACF plot of the residuals of (4.1) is shown in Figure 11.

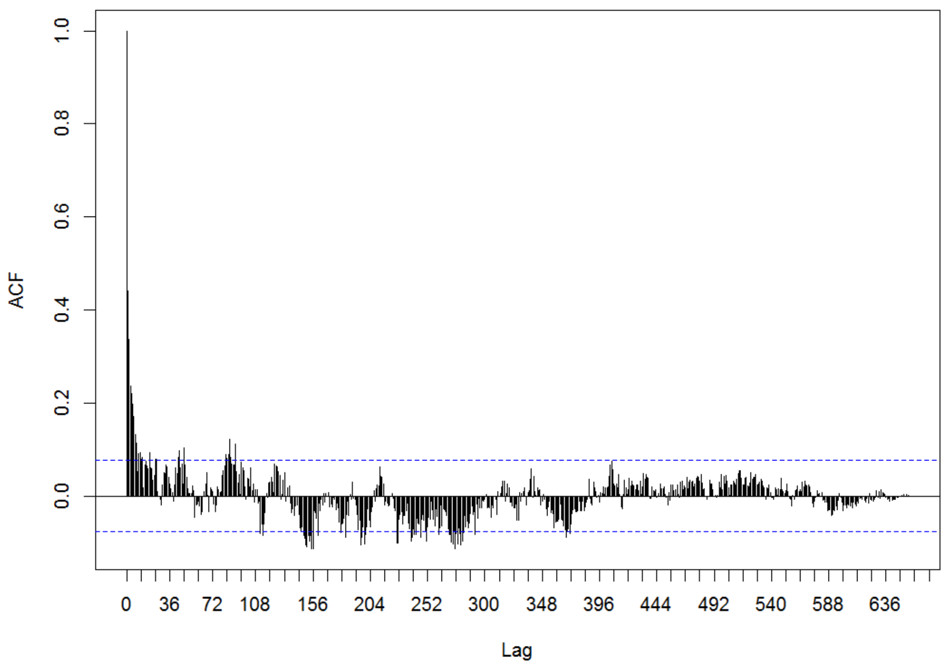

Figure 11: ACF plot of the residuals for (4.1)

By chance, a small amount of correlation in the residuals should be expected. However, the degree of correlation does not appear insignificant. Using the residuals from (4.1), the Durbin-Watson test returns a p-value of essentially 0, confirming that autocorrelation is present.

Thus, we will need to derive accurate standard errors using (3.3). The process of deriving accurate standard errors warrants a section of its own.

4.2 Deriving Accurate Standard Errors

In this section, we will consider seasonally standardized residuals of:

$$X_{nT+v} = \beta_0 + \beta_1 Darwin_{nT+v} + \beta_3 Iceland_{nT+v} + \beta_4 Azores_{nT+v} + \beta_5 NCP_{nT+v} + \beta_7 CO2_{nT+v} + \beta_8 SSN_{nT+v} + \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_9 + \mu_{10} + \mu_{11} + \mu_{12} + \varepsilon_{nT+v}. \quad (4.1)$$

For simplicity, let $\{S_t\}$ represent seasonally standardized residuals of (4.1) with $t = (1, 2, \ldots, 661)$. As noted earlier, we will fit an ARMA model to $\{S_t\}$ in order to derive accurate standard errors. Using AICc as the selection criterion, a zero mean ARMA(1,1) model was deemed the best fit for $\{S_t\}$. Indeed, AIC and BIC also selected an ARMA(1,1) model. A zero mean ARMA(1,1) model takes the general form:

$$S_t = \phi_1(S_{t-1}) + \varepsilon_t + \theta_1\varepsilon_{t-1}.$$

The estimates for ϕ_1 and θ_1 are 0.78 and -0.45, respectively. The error terms, ε_t, are assumed to be normally distributed with a mean of zero and a constant variance, σ^2. An estimate for this variance is 0.14. To calculate an estimate for $\mathbf{\Gamma}$, we need to examine the autocovariance structure of this ARMA(1,1) model. Before doing so, we must verify that the residuals of this model are without trend. A residual plot is shown in Figure 12.

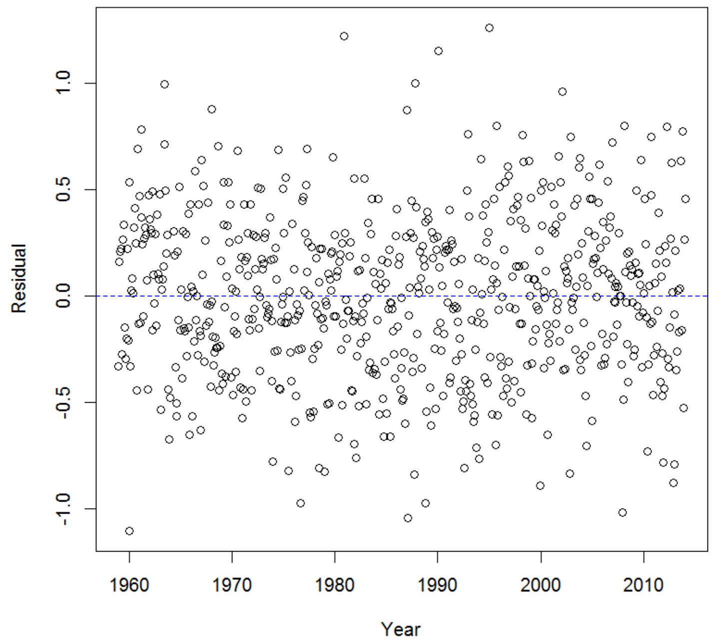

Figure 12: Residuals of ARMA(1,1) fit to {S_t}

For the most part, the residuals look evenly scattered around 0. There does seem to be a pinch in the plot around 1992. The ACF plot shows a few significant spikes but with over 600 lags, we should expect to see a small number of spikes (Figure 13).

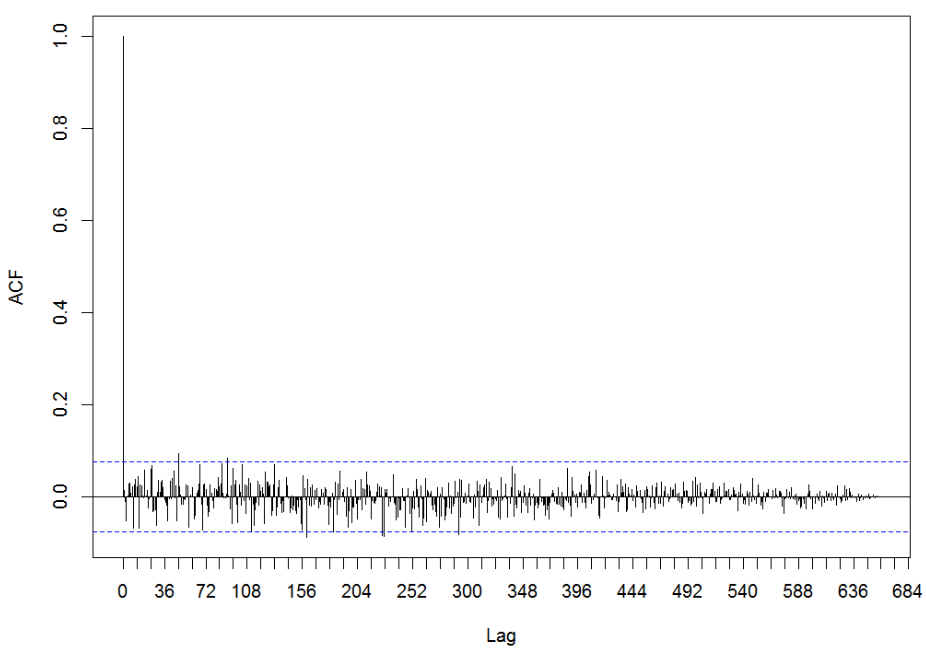

Figure 13: ACF plot of the residuals for the ARMA(1,1) fit to {S_t}

Using 12, 24, 36, and 48 lags, the Ljung-Box test returns p-values of 0.60, 0.64, 0.68, and 0.53, respectively. Lags of larger multiples of twelve were also looked at and all p-values were greater than 0.70. Thus, we have no reason to believe the residuals for this model are not zero mean white noise and the autocovariance structure of the model can be used to provide an estimate for $\mathbf{\Gamma}$. The autocovariance at lag i, $\gamma(i)$, can be computed via the tacvfARMA function in R. Then, using (3.4) and (3.3), accurate standard errors, test statistics, and significance levels can be calculated for (4.1). Results are shown in Table 7.

Table 7: Updated Model Results for (4.1)

Parameter	**Estimate**	**Standard error**	**Test statistic**	**P-value**
β_0	-4338.089	468.069	-9.268	< 0.001
β_1 (Darwin)	0.498	0.082	6.069	< 0.001
β_3 (Iceland)	0.014	0.007	2.149	0.032
β_4 (Azores)	0.025	0.008	3.084	0.002
β_5 (NCP)	0.160	0.038	4.208	< 0.001
β_7 (CO_2)	0.795	0.046	17.220	< 0.001
β_8 (SSN)	0.025	0.019	1.316	0.188
μ_1	24.731	3.234	7.646	< 0.001
μ_2	25.305	3.201	7.904	< 0.001
μ_3	21.360	2.810	7.601	< 0.001
μ_4	9.630	2.110	4.565	< 0.001
μ_9	4.197	1.505	2.789	0.005
μ_{10}	7.591	1.866	4.067	< 0.001
μ_{11}	12.783	2.334	5.474	< 0.001
μ_{12}	16.193	2.847	5.688	< 0.001

Note that the standard errors for the monthly intercepts actually decreased. This likely means that multicollinearity was inflating the standard errors shown in Table 4. Though (4.1) is without severe multicollinearity, VIFs for four of the monthly intercepts were greater than three, indicating some multicollinearity was present. All variables except for SSN remained statistically significant. The model was refit without SSN and the process of calculating accurate standard errors was repeated until all remaining variables were significant. At each stage, an ARMA(1,1) was deemed the best fit for the seasonally standardized residuals of the model and the residuals of the ARMA(1,1) model passed as white noise. The final model is:

$$X_{nT+v} = \beta_0 + \beta_1 Darwin_{nT+v} + \beta_4 Azores_{nT+v} + \beta_5 NCP_{nT+v} + \beta_7 CO2_{nT+v} + \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_{10} + \mu_{11} + \mu_{12} + \varepsilon_{nT+v}. \tag{4.1.2}$$

This model produced an RMSE of 11.51 and an R^2 of 0.81. These values are very close to those produced by both (4.1) and (4.3). Model results are shown in Table 8.

Table 8: Model Results for (4.1.2)

Parameter	Estimate	Standard error	Test statistic	P-value
β_0	-4338.089	476.484	-9.151	< 0.001
β_1 (Darwin)	0.498	0.085	6.087	< 0.001
β_4 (Azores)	0.025	0.008	2.671	0.008
β_5 (NCP)	0.160	0.039	4.207	< 0.001
β_7 (CO_2)	0.795	0.047	16.709	< 0.001
μ_1	24.731	2.411	8.308	< 0.001
μ_2	25.305	2.434	8.492	< 0.001
μ_3	21.360	2.095	8.015	< 0.001
μ_4	9.630	1.837	3.507	< 0.001
μ_{10}	7.591	1.467	3.243	0.001
μ_{11}	12.783	1.670	5.292	< 0.001
μ_{12}	16.193	1.979	5.849	< 0.001

Again, note that the standard error estimates for the monthly intercepts decreased. This can likely be explained by the removal of minor multicollinearity. In (4.1), VIFs for most of the monthly intercepts were greater than two. In (4.1.2), VIFs for all monthly intercepts included were less than two. With a decrease in VIFs (and therefore multicollinearity), we would expect the standard errors to decrease. It is also worth noting that the monthly intercepts are not statistically significant when the model is void of any other independent variables. This may mean that these terms are handling some of the seasonality of the other variables considered. Recall that these other variables have not been seasonally adjusted.

Observed and fitted values for (4.1.2) are shown in Figure 14.

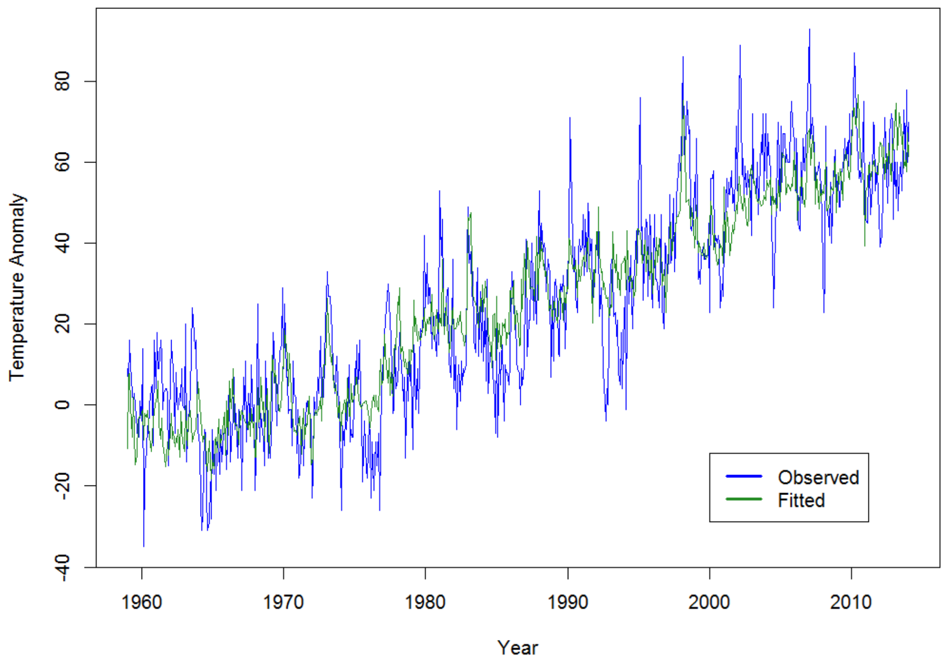

Figure 14: Observed and fitted values for (4.1.2)

As noted, this model does have its deficiencies (mainly, it looks like the model underestimates temperature anomalies early in the record and then again from 2002 – 2006). However, this model is preferred to (4.3) because it makes more sense from a climate perspective – it includes atmospheric CO_2 concentrations instead of a quadratic function of time with a murky interpretation. The quadratic behavior may not be expected to extend beyond the range of the data, but a continued relationship between atmospheric CO_2 concentrations and global temperature anomalies does not seem far-fetched.

CHAPTER 5

DISCUSSION

For reference, the final model is shown below:

$$X_{nT+v} = \beta_0 + \beta_1 Darwin_{nT+v} + \beta_4 Azores_{nT+v} + \beta_5 NCP_{nT+v} + \beta_7 CO2_{nT+v}$$
$$+ \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_{10} + \mu_{11} + \mu_{12} + \varepsilon_{nT+v}. \qquad (4.1.2)$$

This model returned an R^2 value of 0.81, suggesting that it can account for 81% of the variation in the global temperature anomaly data set. This model was selected over a model that included a quadratic function of time. It seems unlikely that time, by itself, would have such a statistically significant impact on temperature trends. If that were the case, then it stands to reason that our planet would be very, very hot by now (given its age). What's more likely is that time is capturing the effect of some other variable – maybe CO_2 concentrations, maybe atmospheric concentrations of another greenhouse gas (methane, nitrous oxide, water vapor, or ozone), or maybe something else not yet considered or understood.

The parameter estimates for Darwin, Azores, and NCP were all positive, suggesting that increases in the average 500mb heights at these locations are associated with increases in global temperatures. The average 500mb heights at each of these locations have been relatively stable since 1958, so it seems these variables can explain month to month or year to year variations in the temperature anomalies more than they can explain any long term trends.

The parameter estimate for CO_2 is relatively large (0.795), as is the test statistic (16.709). Unlike the other variables considered, atmospheric CO_2 concentrations have consistently and steadily increased since the late 1950s. Because atmospheric CO_2 concentrations have consistently increased, it seems likely that they can do more than just explain month to month or year to year variations in the global temperature anomalies. That is, the steady increase in atmospheric CO_2 concentrations could explain the long term trend observed in the temperature anomalies.

It is worth noting that a change point has been detected in the CO_2 series around 1991, though this change point was not taken into consideration in this research (Gallagher, Lund, and Robbins, 2014). The change point is due to the eruption of Mount Pinatubo (located in the Philippines) in June of 1991. This eruption was one of the largest eruptions of the 20th century.

As discussed in the introduction, the trend in global temperatures appears to have slowed over the past fifteen years. Some claim this is evidence that global warming has stopped entirely. The models considered in this paper do not support this claim. Over the past fifteen years, atmospheric CO_2 concentrations have increased by more than 25 parts per million. If increases in CO_2 concentrations are associated with increases in global temperatures (and we assert that they are), then it becomes difficult to argue that warming stopped more than a decade ago. That is, the evidence considered herein does not indicate that warming has stopped.

REFERENCES

"CO2 Now." *NOAA Mauna Loa CO2 Data*. Web. <http://co2now.org/Current-CO2/CO2-Now/noaa-mauna-loa-co2-data.html>.

Dunbar, Brian. "Earth Gets a Warm Feeling All Over." *NASA*. NASA, 09 Feb. 2005. Web. <http://www.nasa.gov/vision/earth/lookingatearth/earth_warm.html>.

"ENSO FAQ." *Climate Prediction Center*. Web. <http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensofaq.shtml#ENSO>.

Gallagher, Colin, Robert Lund, and Michael Robbins. "A General Regression Changepoint Test for Time Series Data." Web. <http://video.ucar.edu/mms/image/samsi2014_colon_gallagher.mp4>.

Hurvich, Clifford M., and Chih-Ling Tsai. "Regression and Time Series Model Selection in Small Samples." *Biometrika* 76.2 (1989): 297-307. Web.

Kutner, Michael H., Chris Nachtsheim, John Neter, and William Li. *Applied Linear Statistical Models*. 5th ed. Boston, Mass.: McGraw-Hill, 2005. Print.

Lund, Robert, Lynne Seymour, and Karen Kafadar. "Temperature Trends in the United States." *Environmetrics* 12.7 (2001): 673-90. Web.

"North Atlantic Oscillation (NAO)." *Climate Prediction Center*. Web. <http://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml>.

Tsay, Ruey S. *Analysis of Financial Time Series*. 2nd ed. New York: Wiley, 2005. Print.