STAT 479 Project : Time Series Analysis of Annual Temperature Anomalies from 1850 to 2009

Abhimanyu Nag (1673217) and Cody Soh (1709259)

Abhimanyu* **- Code, Document, Data Analysis,**
Cody* **- Code, Document, Data Analysis**

$*$**Equal Contribution**

Abstract

Climate change is one of the most pressing challenges of our time, with profound implications for ecosystems, societies, and economies worldwide. In this research project, we conduct a comprehensive time series analysis of annual temperature anomalies from the average over the period 1850-2009. Our study aims to elucidate long-term trends, detect patterns of variability, and provide insights into the different properties of the time series dataset. Using *R*, We will be conducting a thorough analysis of the *global.temp* dataset in the *'tswge'* library using techniques taught in STAT 479 class conducted in Winter 2024. The main aim of this paper is to demonstrate an understanding of the ideas and concepts taught in class as well as gain an insight into the climate change patterns through analysis of changing patterns in temperature.

Keywords : Climate Change · Time series · R · Data Analysis

1 Introduction

Climate change is one of the most pressing issues facing humanity in the 21st century, with far-reaching implications for ecosystems, economies, and societies worldwide [1]. Understanding the dynamics of global temperature anomalies is crucial for assessing the extent and impact of climate change over time. In this study, we focus on analyzing the *global.temp* dataset [2], which provides annual temperature anomalies from the average for the years 1850 to 2009. This dataset serves as a valuable resource for exploring historical climate trends and patterns, shedding light on the complex interactions driving long-term temperature variability.

(i) Statement of the Research Problem:

The selected dataset presents an opportunity to address several key research questions, including the identification of long-term temperature trends and the detection of periodic patterns using concepts taught in STAT 479 class in Winter 2024 [3, 4]. By analyzing this dataset, we aim to gain insights into the underlying mechanisms driving temperature changes over the past 160 years and their implications for future climate projections.

(ii) Literature Review:

Extensive research has been conducted on global temperature trends and climate variability, with tests utilizing concepts and methods learnt in STAT 479 class (see [3, 4]) . We cover ideas such as stationarity [4], trend analysis [4], Box Pierce Test [4], ACF-PACF plots [3] and the best fitted model [3]. We also looked into similar analyses (see [5, 6]) on analysis of time series for global climate change, which used methods such as satellite images [6] and monte carlo simulations [5].

(iii) Specific Objectives and Hypotheses:

The specific objectives of this study include:

1. *Testing for Stationarity :* We conduct an Augmented Dickey Fuller root test with $H_o =$ A unit root is present and $H_a =$ A unit root is not present
2. Checking for Trends : We plot the data and check for trends
3. *Autocorrelation :* We conduct the Box Pierce test to test for autocorrelation of the time series data with $H_o =$ There is no autocorrelation and $H_a =$ There is autocorrelation
4. *ACF and PACF Plots :* We plot the ACF and PACF plots of the first difference of the dataset and predict the behaviour of the time series data.
5. *Best Fitted Model :* We derive and fit the best fitted model from the conclusions of the above tests and test for invertibility.

(iv) Significance and Importance of the Study:

Understanding historical climate variability is crucial for informing climate science, policy-making, and adaptation strategies in the face of ongoing climate change. By analyzing the global.temp dataset, this study aims to contribute valuable insights into the drivers of temperature change over the past 160 years, aiding in the development of more accurate climate models and projections. Moreover, the findings of this study have implications for mitigating the impacts of climate change and promoting sustainable development practices globally.

2 Data Analysis

```
#Load the "tswge" package
library(tswge)
#Load the "global.temp" dataset
data(global.temp)
#Check if the process is stationary using a unit root test, i.e., the ADF Test
library(tseries)
adf.test(global.temp)
```

```
        Augmented Dickey-Fuller Test

data:  global.temp
Dickey-Fuller = -2.0365, Lag order = 5, p-value = 0.5611
alternative hypothesis: stationary
```

Figure 1: Output for ADF Test for global.temp datatset

P-Value for the ADF Test for the global.temp dataset = 0.5611 > 0.05. Therefore, we fail to reject the null hypothesis and conclude that the data is not stationary. As the global.temp dataset was not stationary, we can attempt differencing the global.temp dataset as differencing a time series may remove trends.

```
#Compute the first difference
diff1 <- diff(global.temp)
#Check if the first difference is stationary
adf.test(diff1)
```

```
Warning: p-value smaller than printed p-value
        Augmented Dickey-Fuller Test

data:  diff1
Dickey-Fuller = -7.2421, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

Figure 2: Output for ADF Test for First Difference

P-Value for the ADF Test of the first difference of the global.temp dataset < 0.01 < 0.05. Therefore, we reject the null hypothesis and conclude that the data is stationary.

```
#Compute the second difference
diff2 <- diff(diff1)
```

```
#Check if the second difference is stationary
adf.test(diff2)
```

```
Warning: p-value smaller than printed p-value
        Augmented Dickey-Fuller Test

data:  diff2
Dickey-Fuller = -10.054, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

Figure 3: Output for ADF Test for Second Difference

P-Value for the ADF Test of the second difference of the global.temp dataset $< 0.01 < 0.05$. Therefore, we reject the null hypothesis and conclude that the data is stationary.

```
#Plot the time series data to visualize trends
plot(global.temp, type = "l", main = "Plot of global.temp")
```

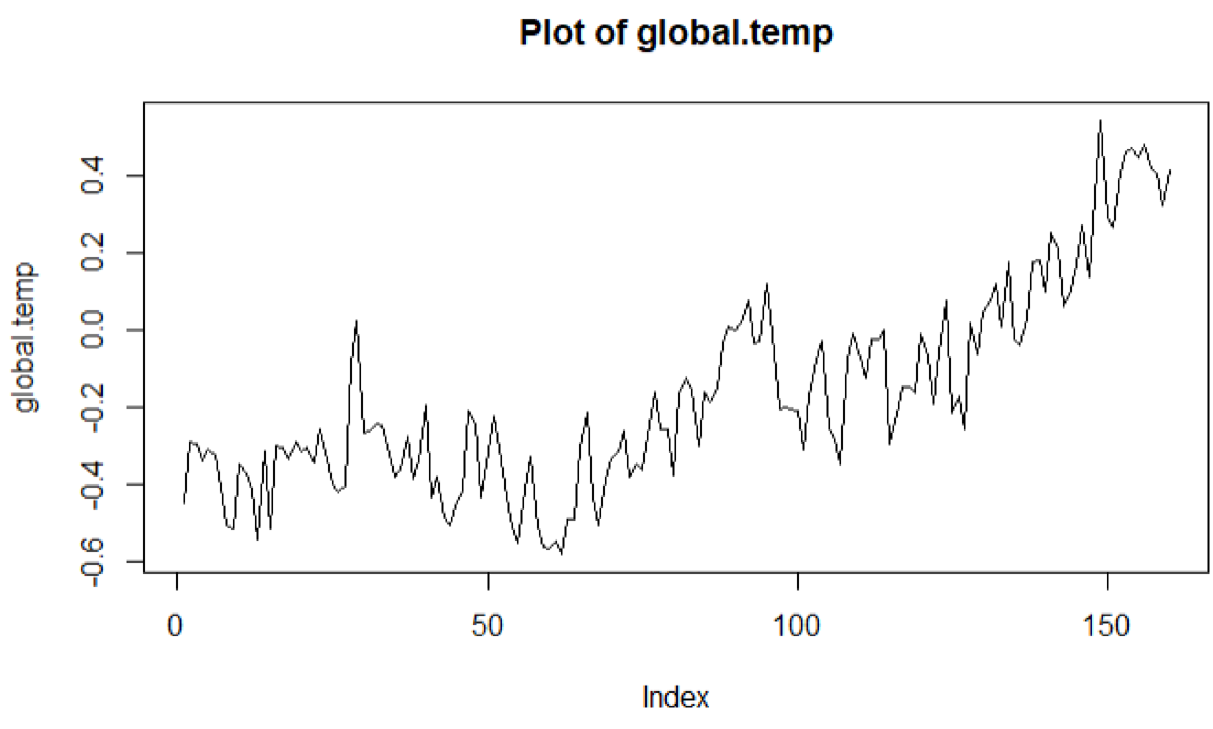

Figure 4: Plot of global.temp Dataset

Global Temperature seems to be approximately constant for the first 60 years of the dataset. After that, the Global Temperature seems to be increasing.

There are no covariates to consider in the global.temp dataset.

```r
#Test for Autocorrelation
box.pierce <- Box.test(global.temp, lag = 20, type = "Box-Pierce")
print(box.pierce)
```

```
        Box-Pierce test

data:  global.temp
X-squared = 1171.7, df = 20, p-value < 2.2e-16
```

Figure 5: Output for Box-Pierce Test

P-Value for the Box-Pierce Test < 2.2e-16 < 0.05. Therefore, we can reject the null hypothesis and conclude that autocorrelation exists within the global.temp dataset.

```r
#ACF and PACF Plots
acf(diff1)
pacf(diff1)
```

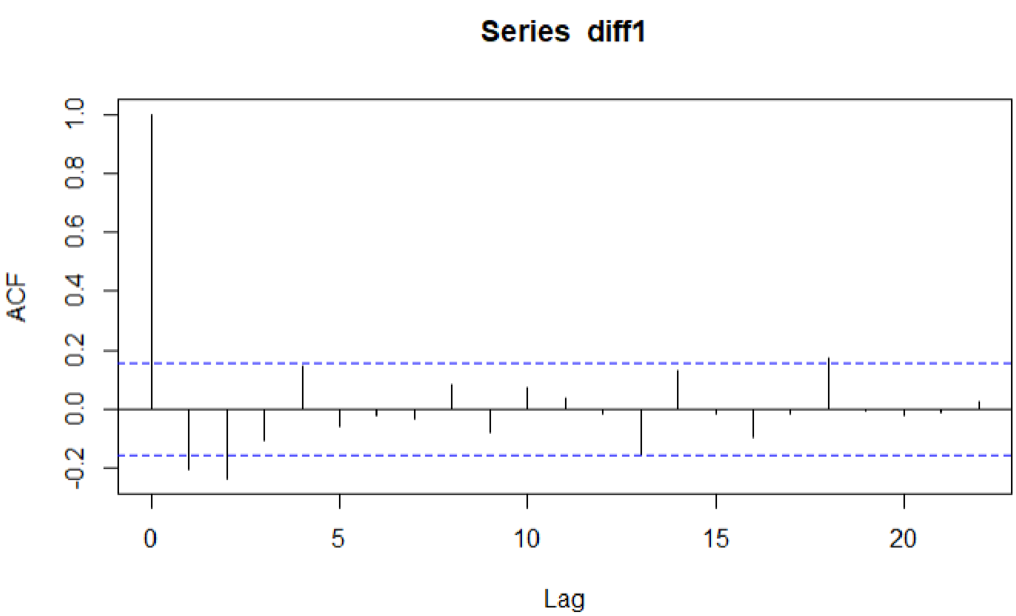

Figure 6: ACF Plot

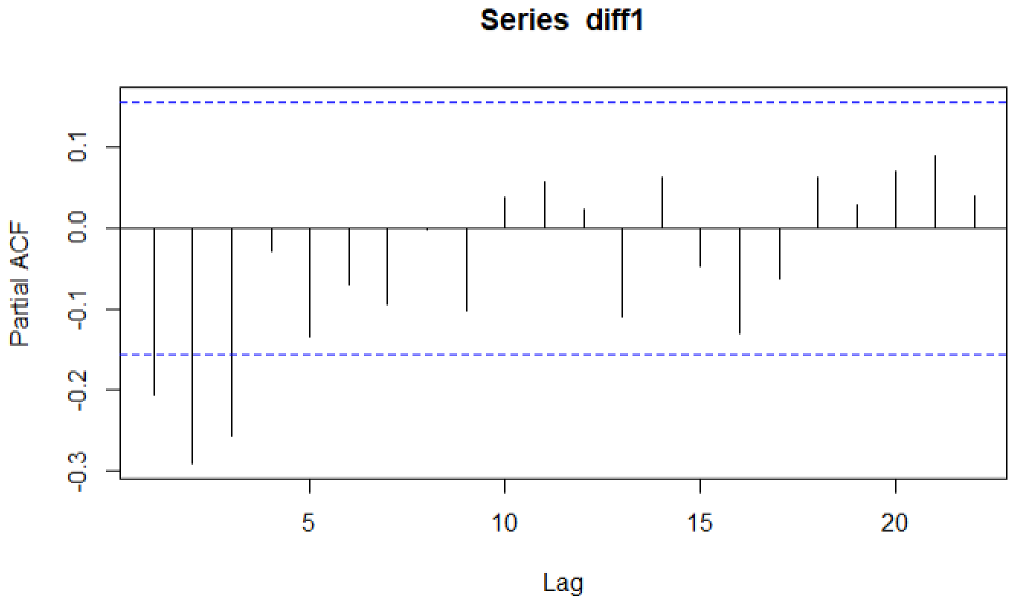

Figure 7: PACF Plot

We checked the ACF and PACF for the first difference of the global.temp dataset as it is stationary. Based on the ACF and PACF plots of the first difference of the global.temp datasest, the ACF = 0 for lags > 2 and the PACF seems to decrease geometrically from lag = 2. Therefore, it seems that the first difference of the global.temp series follows the behaviour of an MA(2) process.

```
library(forecast)
#Best Fitted Model
auto.arima(global.temp)
```

```
ARIMA(0,1,2) with drift

Coefficients:
          ma1      ma2   drift
      -0.3979  -0.2692  0.0051
s.e.   0.0723   0.0684  0.0028

sigma^2 = 0.0109:  log likelihood = 134.94
AIC=-261.89   AICc=-261.63   BIC=-249.61
```

Figure 8: Output for auto.arima()

Based on the auto.arima() function, it seems that the global.temp datatset is best fitted to an ARIMA (0, 1, 2) model, i.e., an MA(2) model with an order of differencing of 1.
The equation for the best fitted MA(2) model seems to be:

$$X_t = \mu + w_t - 0.3979w_{t-1} - 0.2692w_{t-2}$$

```
#Testing for Invertibility
#Checking the roots of the Characteristic Equation
roots <- polyroot(c(1, 0.3979, 0.2692))
abs_roots <- abs(roots)
abs_roots
```

```
[1] 1.927358 1.927358
```

Figure 9: Roots of the Characteristic Polynomial

The roots of the characteristic polynomial = $(1.927358, 1.927358)$, which are both greater than 1. Therefore, the MA(2) process is invertible. As the best fitted model is an MA model, we only check for invertibility.

3 Conclusion

Using the ADF test, we concluded that the global.temp dataset is not stationary. However, the first and second differences are stationary. Based on the plot of global.temp dataset, there seems to be an increasing trend. There are no covariates to consider in this datatset. Using the Box-Pierce Test, we concluded that autocorrelation exists within the dataset. Based on the ACF and PACF plots for the first difference of the dataset, we concluded that it follows the behaviour of an MA(2) process. Using the auto.arima() function, we concluded that the best fitted model for the dataset is an MA(2) model with an order of differencing of 1. The equation for the best fitted model seems to be: $X_t = \mu + w_t - 0.3979 w_{t-1} - 0.2692 w_{t-2}$. Lastly, by checking the roots of the characteristic polynomial, we concluded that the best fitted model is invertible.

References

[1] NOAA National Centers for Environmental Information, Monthly Global Climate Report for Annual 2023, published online January 2024, retrieved on March 6, 2024 from https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202313.

[2] "Applied Time Series Analysis with R, 2nd edition" by Woodward, Gray, and Elliott

[3] Pandher, Sharandeep. (2024, February). Statistical Models for Time Series [Lecture notes]. University of Alberta. https://eclass.srv.ualberta.ca/pluginfile.php/10793731/mod_resource/content/1/Chapter-2%28Statistical%20Models%20for%20Time%29-Final-2.pdf.

[4] Pandher, Sharandeep. (2024, January). Characteristics of Time Series [Lecture notes]. University of Alberta. https://eclass.srv.ualberta.ca/pluginfile.php/10754008/mod_resource/content/8/Chapter-1%28479%29.pdf

[5] Rahmstorf, S., Foster, G., & Cahill, N. (2017). Global temperature evolution: recent trends and some pitfalls. Environmental Research Letters, 12(5), 054001.

[6] Spencer, R. W., & Christy, J. R. (1990). Precise monitoring of global temperature trends from satellites. Science, 247(4950), 1558-1562.