Forecasting Global warming and its repercussions on our ecosystem.

Lormel Bationo

Abstract

Temperature anomalies can be thought of as the deviations from a long-term average temperature for a specific location and time period.This parameter is commonly used in environmental science research topics regarding global temperature. Temperature anomnalies play a vital role in helping scientists and policymakers, to understand, monitor, and respond to changes in temperature patterns over time. This reports aims to analyze and forecast the trend patterns of global temperature using temperature anomalies. We use data gathered by the National Center for Environmental information. Our two models are used to assess the trend and check for potential seasonality which can reveal factors that could be key to explaining the current trend or the different fluctuations of the global temperature. Our models will also be used for forecasting in order to predict the global temperature trends and asses the impact that the ongoing efforts to address and adapt to climate change have had on our ecosystem. The forecasting results should help us suggesting important policies regarding our findings. Later on, we will use the Root Mean Squares Estimates combined with the Mariano Diebold-test to determine the accuracy of our models predictions.

Background and significance

For years now, the rise of global temperature has grown to become a capital issue for the future of our planet. According to experts, the constant rise of the global temperature can lead us to catastrophic economic, social and ecological consequences that could result in the extinction of many species including plants, animals and even humans. Global warming has grown to become a worldwide matter. Regarding the constant rise of the global temperature, politics around the world all agree that actions need to be taken to drastically reduce the temperature of the globe. According to experts, if we pass the ceiling of 1.5 °C in temperature anomaly, the planet could have irreversible consequences. Considering all these previsions, it is logical to wonder what the global temperature will be in the upcoming year and how is it going to impact life on earth? Therefore, it becomes necessary to develop models that can asses this phenomenon and also make predictions about the changes that would occur in the future. Being able to understand and predict those aspects is determinant as it will help us to take the appropriate measures in an effort to protect and preserve our ecosystem equilibrium. While it may seem intuitive to use temperature averages to make predictions about temperature, we will be using another variable called temperature anomalies. In fact, temperature anomalies are a more reliable estimator than temperature averages. They are considered against baseline temperatures and are computed over an average of at least 30 years. Therefore, temperature anomalies can be a better choice over average temperatures when the goal is to highlight trends and changes over time. This approach will help emphasizing the changes in temperature patterns rather than focusing on specific temperature values. In this project, our goal is to analyze and understand the changes of global temperature over time and make predictions about changes in the future.To finish, we will try to assess their repercussions on our ecosystem.

Methods

In order to build an appropriate model and have accurate forecasting, I wanted to make sure that I would be using a dataset that is large enough, consistent over time and that would be coming from a reliable source. For the sake of accuracy and reliability, we use data compiled by the National Center for Environmental Information (NCEI) to reduce the chances of having measurements errors. The gathered data from the NCEI range from 1895 to 2023 . In order to make the study more relevant since our estimations were based on the past century, the dataset was reduced and only the data from the past century were considered.This data transformation made the dataset more manageable and more relevant allowing for better insight into the temperature anomaly for the past century and then increasing the chances of having better forecasting performances (considering the way that temperature anomalies are computed and also the different factors that impacted global warming over the past century). In our case, data transformation to achieve normality was not necessary because this approach was not relevant regarding the goal of this research. We kept the dataset as it was downloaded from the original source and started with the time series analysis. We then proceeded to decompose our data into trend, seasonal components and residuals. The trend was useful in identifying the long term patterns. The seasonal components were also useful in revealling recurring patterns over certain period. Then, we can proceeded to use our models to asses the residuals and check for autocorellation.

The auto-regressive Moving Average Model(ARMA)

The ARMA (autoregressive moving average) model is a valuable statistical method that combines Autoregressive (AR) and moving average (MA) components.The autoregressive component uses past values in order to predict future values while the moving average component uses past errors to predict future values.

An ARMA of order (p,q) is given by $X_t = \phi_1 + ... + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}, t \in \mathbb{Z}$

The model is able to capture the relationship between current and past values in time series as well as their influence on past error terms making it a great choice for predicting patterns in time series. Selecting the appropriate order of an ARMA model is important especially for forecasting purposes. In fact when using ARMA, model selection is very important. In fact, certasin models might not capture all the patterns within the data or they might have too many parameters and then overfit the data. These are errors that could discredit our forecasting. In order to select an appropriate model, we used three common information criterion that are **"AIC"**, **"AICC"** and **"BIC"** which stand respectively **for Akaike Information Criterion**, **Corrected Akaike Information Criterion** and **Bayesian Information Criterion**. These information criterion may perform differently and are very useful in selecting an appropriate model. In our case, all these models performed almost similarly. Therefore, we choose the ARMA **"BIC"** as our model. In fact the 'BIC" model is the most simple compared to the two previous models and can helps us avoid overfitting.

Singular Spectrum Analysis(SSA)

The singular spectrum analysis is a technique used for the analysis and decomposition of time series data into fundamental components such as trend, seasonal components and noise. It is particularly useful for extracting underlying patterns from noisy and complex time series. It comports a decomposition stage that allows a clear understanding of the singular components. We obtain eigenvalues and pairs of eigenvectors that help to capture the different patterns represented in the time series. The eigenvectors are then grouped based on their corresponding values. To finish, we do a reconstruction stage that consists in grouping the different components to obtain an approximation of the original time series. SSA does not assume stationary and thus does not require data transformation (Hossein Hassani, A bried introduction to SSA). SSA can provide forecasts for individual components within the time series. This can help predict specific aspects of the time series such as trends and seasonal patterns making it a great choice for time series modelling and forecasting.

Results of the analysis

Our original time series plot displayed without surprise a growing temperature anomaly over time which confirms the idea that the global temperature has been increasing for the past century(figure 1). In fact, positive temperature anomalies suggest that the observed temperatures are warmer than the long term average. We used polynomial fitting to fit and remove the trend which did not help us achieving stationarity (figure 3). Stationarity is very important concept in time series analysis especially for implementing the models like ARMA. Therefore, we used differencing which is a very powerful method that can help removing the trend and seasonal component to make our series stationary. We applied a differencing opperator twice which successfully captured the trend and made our series stationary (figure 4). Since our ARMA models perfomed similarly, we selected the ARMA "BIC" to make sure that we have a simple model and avoid complications. The selected arma model("BIC") fitted the model correctly and allowed us to perform forecasting. The results of the forecasting using the ARMA model showed that temperatures anomalies that were in a long term upward trend started to slow down before decreasing linearly over time (figure 5). The singular spectrum analysis gave us a deeper insight of the data. The time series was decomposed into several components that displayed the presence of a linear upward trend. It revealed the presence of seasonal components that were not very apparent in the ARMA model(figure 7b) . The correlation matrix plot coupled with the pairs of eigenvectors show the presence of seasonal components which was very helpful for grouping(figure 8). The forecast results using the Singular Spectrum Analysis model also predicts a downward trend for the temperature anomalies for the next 12 years (figure 11). However, the slope is less steep compared to the ARMA model prediction. Both the ARMA and the SSA model predicted a decrease in temperature anomalies but with different variations. We then assessed the forecasting performances by calculating the Root Mean Squared Errors for the two models. The SSA model showed a lower RMSE. Since forecasting was a very esential part of this research we also performed the Diebold-mariano test which can be used to compare the forecasting performances of our two models. The obtained test statistics was very low (-4.3257) which is indicative of forecast difference between the two models. This is later confirmed by the very low p-value that makes us reject the null hypothesis that there is no difference in forecast accuracy between the two models. We then conclude that there is enough statistical evidence to support that the Singular Spectrum Analysis (SSA) model provides significantly better forecast accuracy than the AutoRegressive Moving Average (ARMA) model over the 12-time point horizon(figure12). Therefore even though both models predicted a future downward trend for temperature anomalies, the SSA prediction is more reliable.

Discussion of the research

In examining temperature anomalies over the past century using ARMA and SSA, our models, despite limited insight into the causes of global warming, have enabled us to make substantial discoveries about the Earth's temperature patterns. Since the beginning of the century, temperature anomalies have been on the rise. Both our models predicted an upcoming decrease of temperature anomaly for the future. However, it is important to note that the perfomance evaluation attributes better forecasting performances to our SSA model. When we look at the decreasing slope of temperature anomalies, the slope or the ARMA model is very steep compared to the SSA model. This suggest that the ARMA model predicts a very rapid decrease of temperature anomalies for the next 12 years. However the SSA model is also predicting a decrease but slower and with different fluctuations which confirms that the SSA models captures the seasonality present in our time series. This also confirms the results of our forecasting performance evaluation. In fact, climate change can take a lot of time to change patterns and the SSA forecast seems to have a more realistic prediction compared to the ARMA model. This unexpected downward trend could have significant optimistic implications for the future of the planet. We could attribute this to strong climate actions taken to reduce the impacts of global warming. According to IPCC (intergovernmental panel on climate change) chair Hoesung Lee, **"We are at a crossroads. The decisions we make now can secure a livable future. I am encouraged by climate action being taken in many countries"**. We know that global warming has been the topic of multiple climate forums because, the rise of global temperature, if constant, can represents a real threat to life on earth.The different actions started to limit the rising of Global temperature are starting to pay off but are not yet sufficient. One example of that is the **Paris Climate Agreement which aims to limit temperature anomalies to 1.5°C**. In fact, there are more countries around the world that are amending laws to encourage the switch to renewable energy thus, reducing greenhouse gas emissions. For example the United States department of energy plans to finance the auto transition to electric vehicles with over $12 billions in loans and grants. Therefore, we can confidently say that these results are the consequences on years of joints efforts of countries around the world to limit the impacts of global warming.

While the observed decrease in temperature anomalies appears promising for the future, it is essential to acknowledge the limitations of our research. Even though our model accuracy is not questioned in this context, a time series analysis alone may not provide a comprehensive understanding of the different aspects that are influencing global temperatures.There are multiple factors that directly or indirectly contributing to global warming. Therefore, a better approach to this would be for example coupling this time series analysis with a multiple linear regressions model with different factors such as deforestation, industrial activity, land use or the use of fossil fuels. We can even include the different actions taken to reduce the global temperatures in our model. This would allow us to work directly on understanding the causes of global warming and the areas of interest that need improvements. We will then be able to make a more comprehensive model for forecasting by combining the time series analysis and the effect of the different factors in our regression model. This will allow us to build more reliable models that weight the impact of the different factors that influence global temperature and provide more accurate forecasting for the future.It will also allow for better policy to directly have an impact on the significant factors.

Another aspect to consider is that the decreasing trend might not be continuous because we notice the presence of seasonal components in our time series. these observed oscillating fluctuations could be explained by el nino and la nina phenomenon. In fact el nino and la nina are two opposing climate patterns that break from normal conditions. They are part of a larger phenomenon called the el nino-Southern Oscillation(ENSO) cyle. el nino is the warm phase while la nina is the cool phase. For example during an el nino phase, winters are warmer than usual while summers are hotter and during la nina phases, summers are less hot while winters are less cold. These phenomenon that are occurring in the pacific ocean alternate and have a significant impact on global weather. Therefore, even though our forecasting results suggest that the temperature anomalies could decrease and be maintained to 1.5°C , they are still high since temperature anomalies that are positive suggest that the current year is still warmer than the previous one.Therefore the different actions taken over the past 5 decades should continue and even increase to make the transition to renewable energy more efficient and then protect our planet from terrible consequences.

References

1. Rothschild, E. (2023). Wartime in the history of economic thought: episodes in European history. *European Journal of the History of Economic Thought*, 1–13. https://doi.org/10.1080/09672567.2023.2265518

2. NCEI.Monitoring.Info@noaa.gov. (n.d.). National Time Series — Climate at a Glance — National Centers for Environmental Information (NCEI). https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/national/time-series

3. Howe, P. D. Leiserowitz, A. Who remembers a hot summer or a cold winter? The asymmetric effect of beliefs about global warming on perceptions of local climate conditions in the U. S. Glob. Environ. Change 23, 1488–1500 (2013).

4. Svs. (2023, September 14). Summer 2023 Record high global temperatures. SVS. https://svs.gsfc.nasa.gov/5161/ —item Libretexts. (2022, September 7). 3.1: Introduction to autoregressive Moving Average (ARMA) processes. Statistics LibreTexts. https://stats.libretexts.org/Bookshelves/Advanced_Statistics/Time_Series_Analysis_(Aue)/33A_ARMA_Processes/3.13A_Introduction_to_Autoregressive_Moving_Average_(ARMA)_Processes

5. Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of Time Series Structure: SSA and related techniques, Chapman Hall/CRC, New York -

Appendix

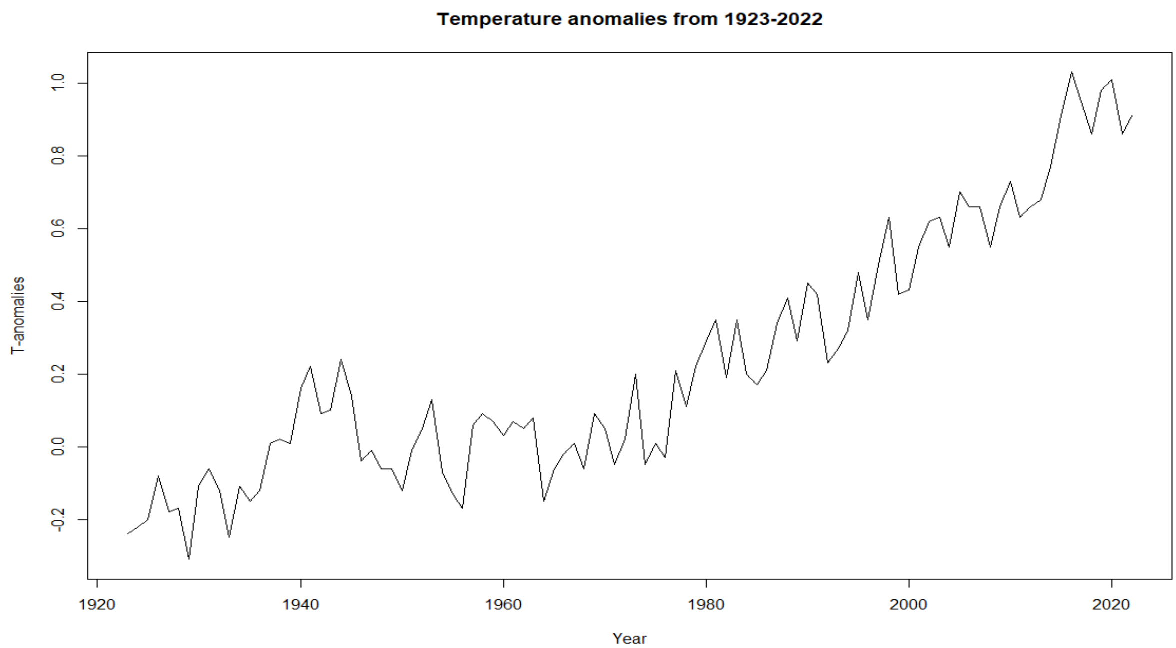


Figure 1: Original time series plot

```
Call:
lm(formula = Anomaly ~ pdata$Year + ys, data = pdata)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.219041 -0.082482 -0.002362  0.073538  0.314220 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  4.904e+02  5.752e+01   8.525 2.03e-13 ***
pdata$Year  -5.077e-01  5.833e-02  -8.703 8.41e-14 ***
ys           1.314e-04  1.479e-05   8.885 3.42e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1102 on 97 degrees of freedom
Multiple R-squared:  0.8972,	Adjusted R-squared:  0.8951 
F-statistic: 423.2 on 2 and 97 DF,  p-value: < 2.2e-16

> 
```

Figure 2: Regression output for least squares method

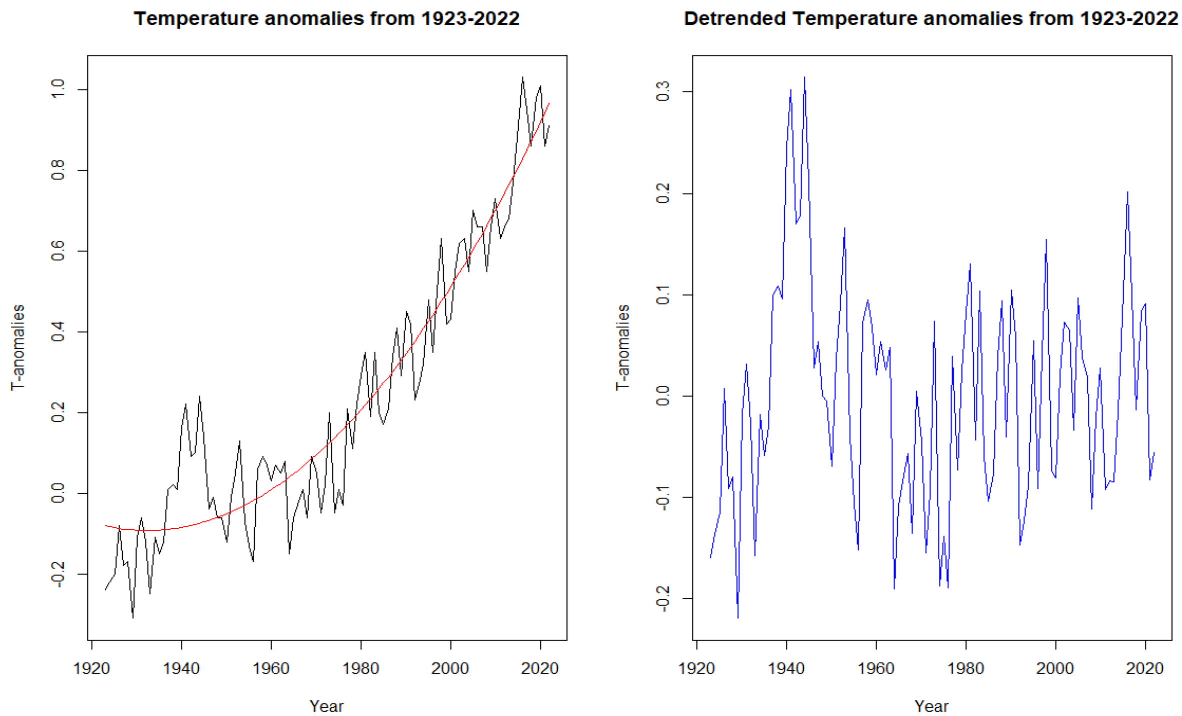


Figure 3: Original and detrended time series with least squares method

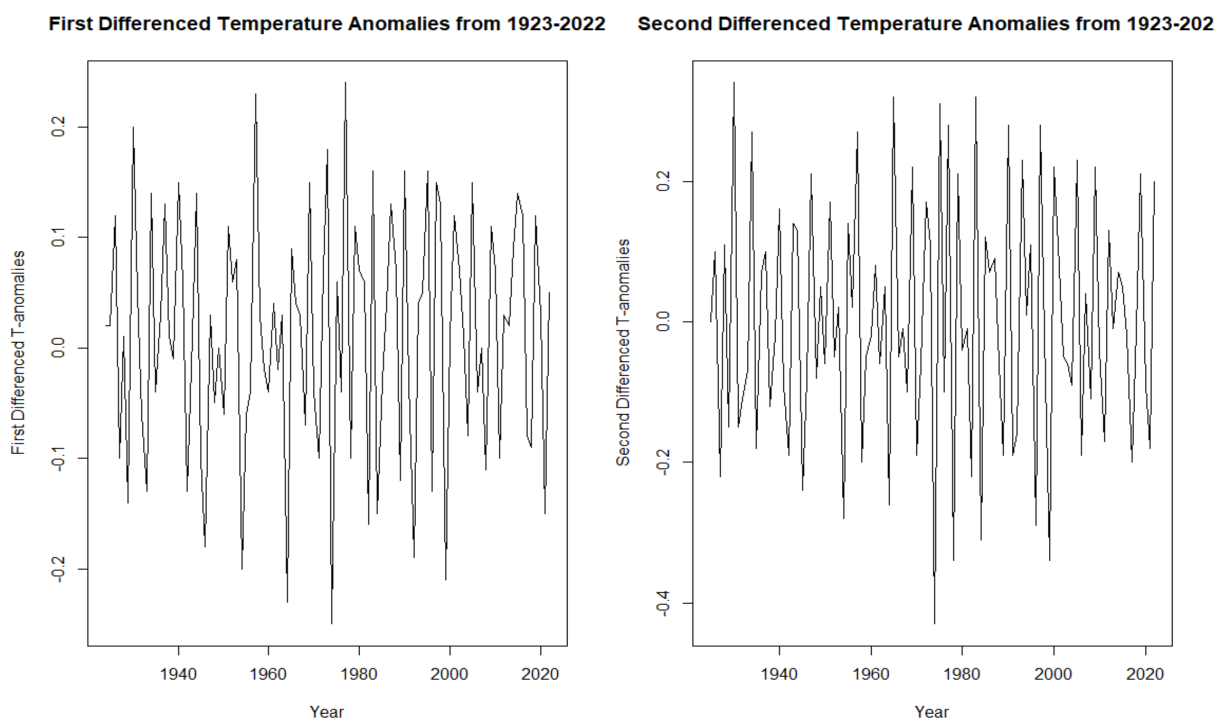


Figure 4: Stationary time series after first and second differencing

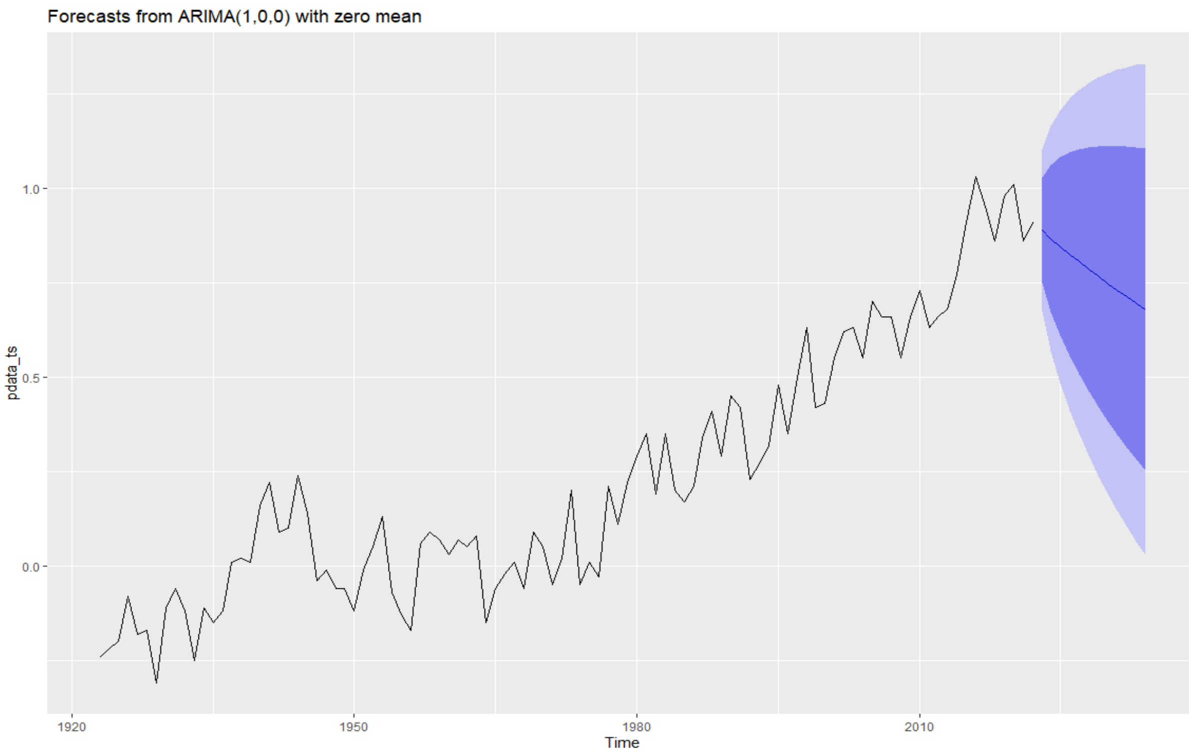


Figure 5: Forecasted temperature anomalies using ARMA with n.ahead =12

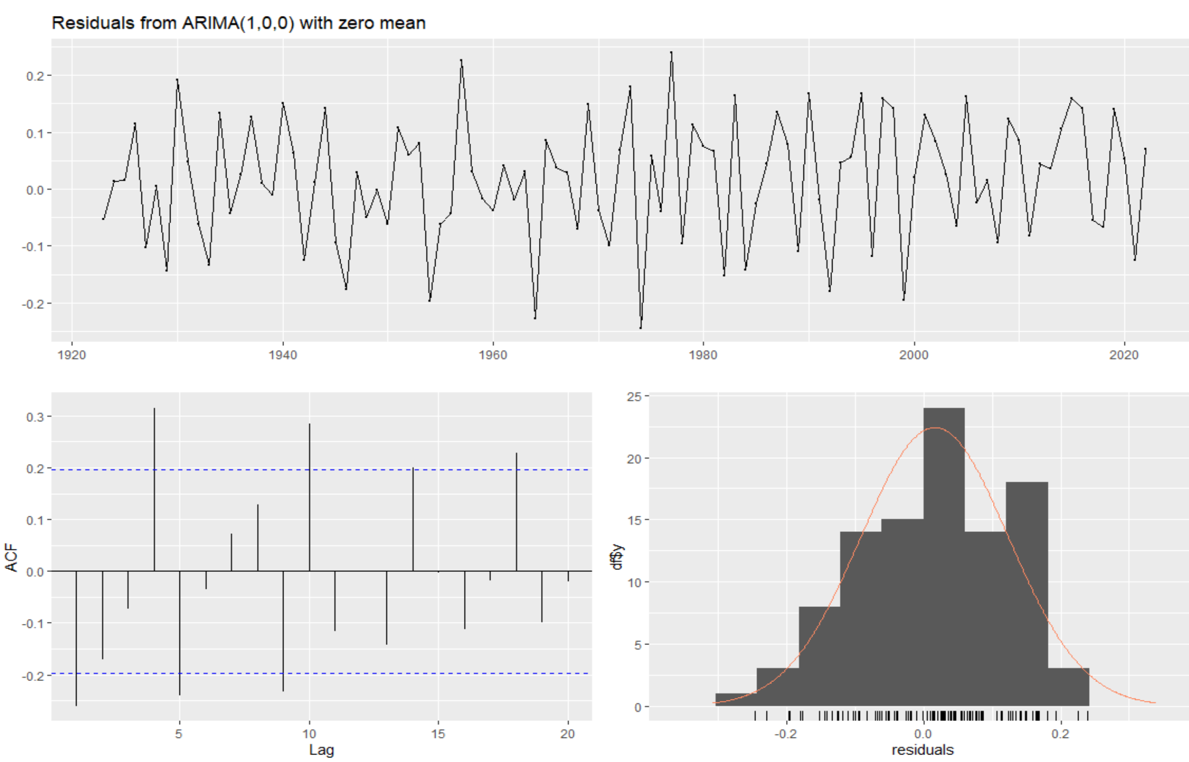
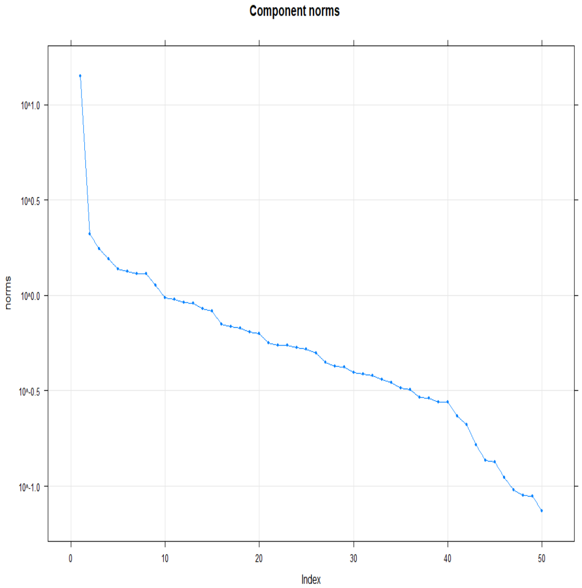
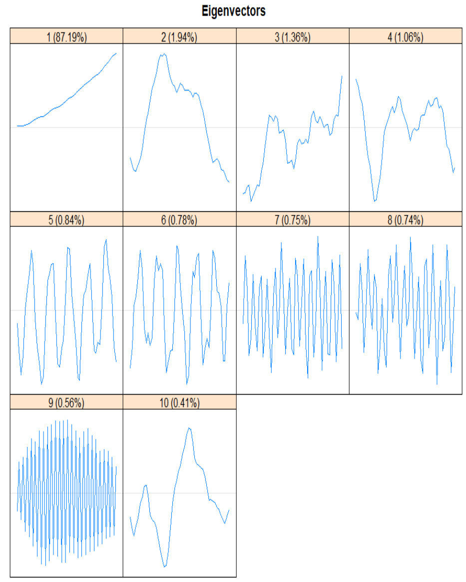


Figure 6: ARMA model residual diagnostic

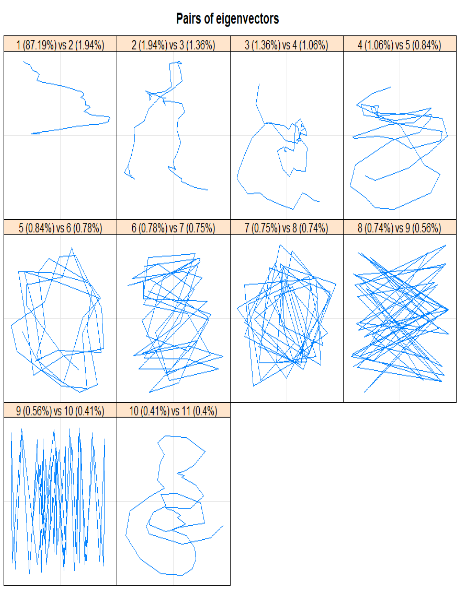


(a) SSA eigenvalues

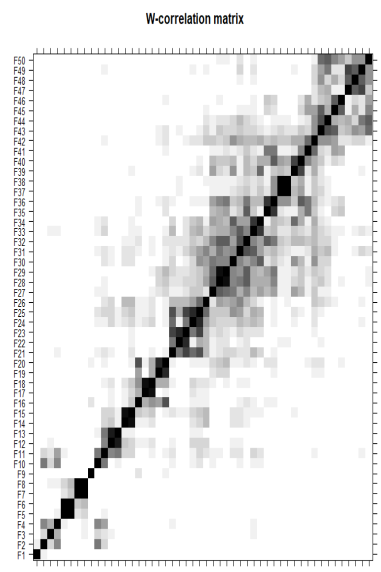


(b) SSA eigenvectors

Figure 7: SSA components



(a) SSA pairs of eigenvectors



(b) SSA W-correlation matrix

Figure 8: SSA components

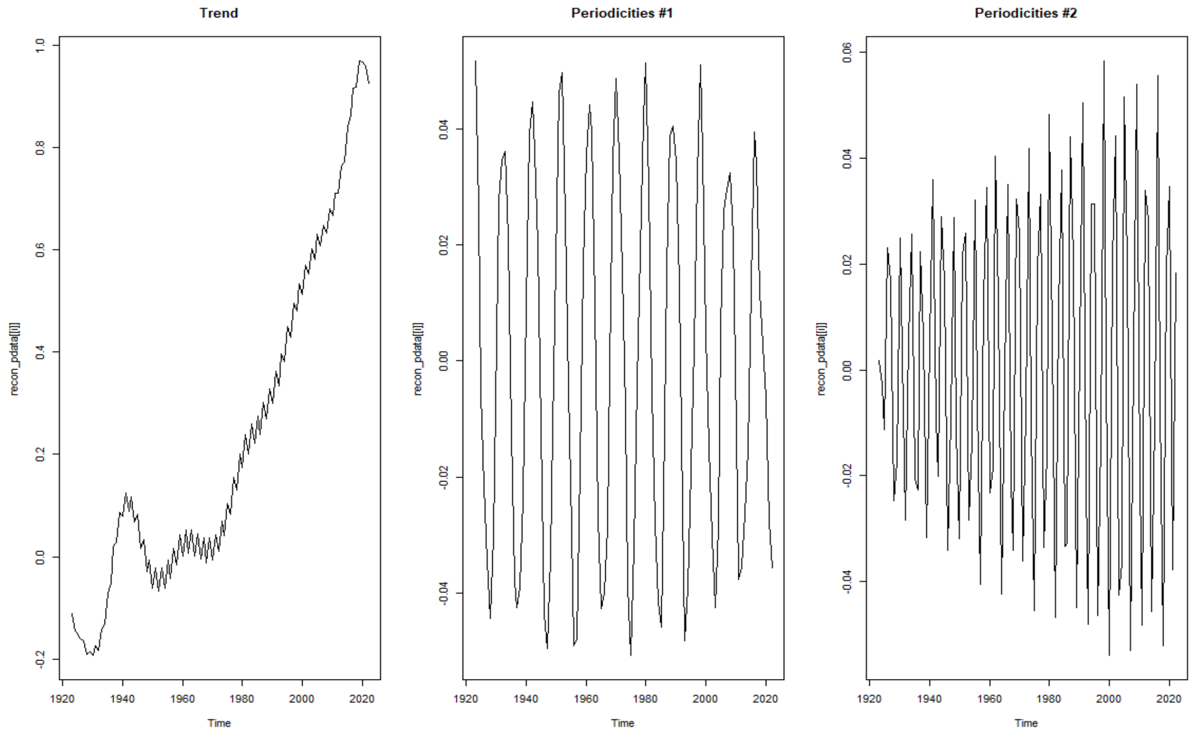


Figure 9: SSA reconstructed components

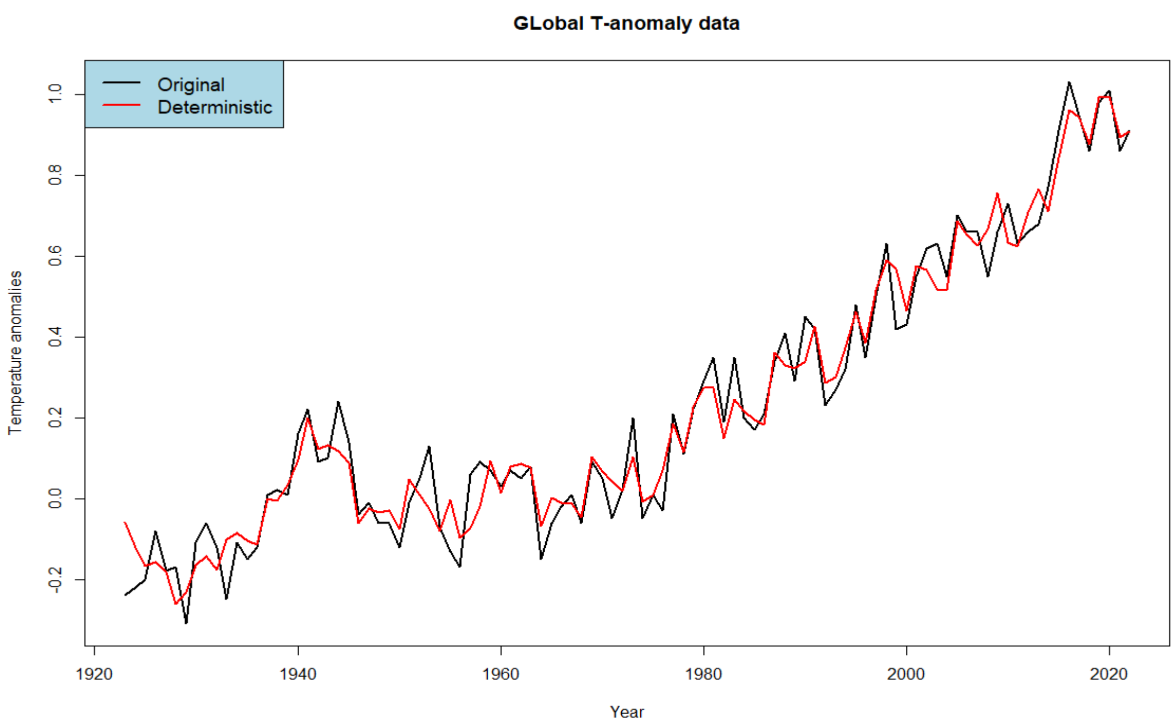


Figure 10: SSA original time series with deterministic pattern

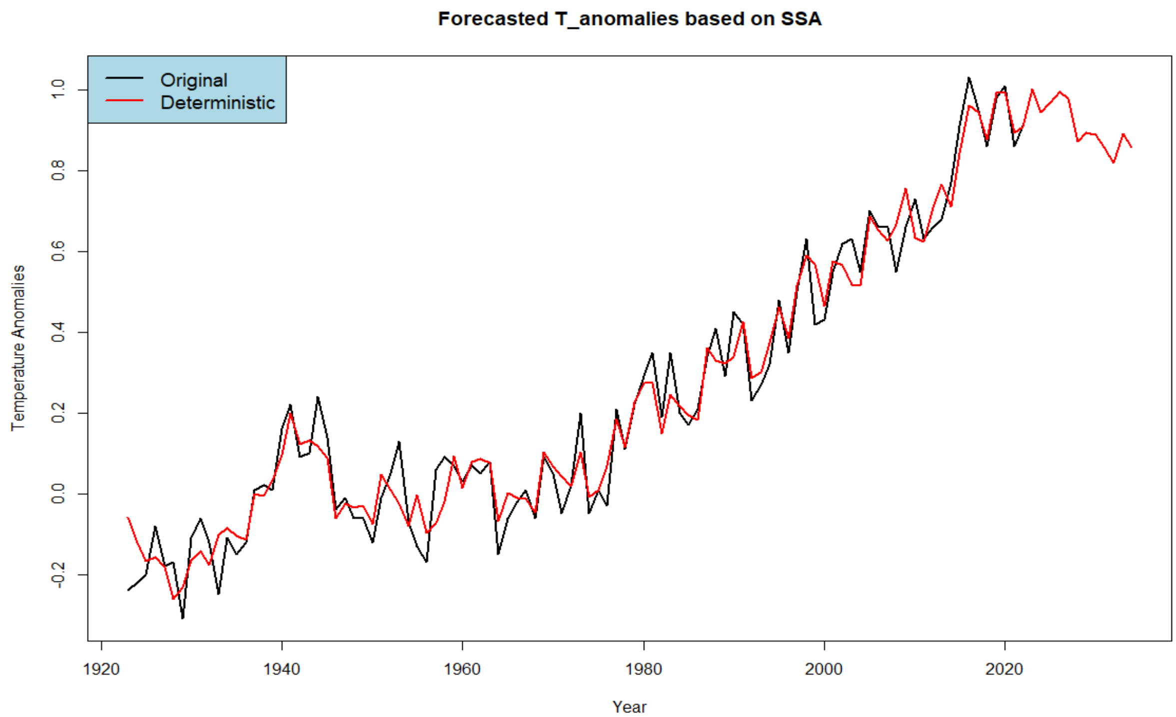


Figure 11: Forecasted temperature anomalies using SSA

```
> dm_test =dm.test(res_ssa,arma_res, alternative = "two.sided", h=12)
> dm_test

        Diebold-Mariano Test

data:  res_ssaarma_res
DM = -4.3257, Forecast horizon = 12, Loss function power = 2, p-value = 3.634e-05
alternative hypothesis: two.sided
```

Figure 12: forecasting performances evaluation