Time Series Analysis of Average Temperatures in Armagh: Modeling and Forecasting for Climate Trends

*Part A: Time Series Analysis

Saheli Dutta
x21246513student@ncirl.ie
National College of Ireland
Masters in Data Analytics

***Abstract*—This research study presents a time series analysis of average temperatures in Armagh, focusing on modeling and forecasting climate trends. Two datasets are used: monthly average temperatures and yearly average temperatures from 1844 to 2004. The monthly data shows stationarity, while the yearly data displays a trend. Three categories of models are considered: exponential smoothing, ARIMA/SARIMA, and simple time series models. The models are evaluated using diagnostic tests, and the data up to 2003 are used to forecast temperatures for 2004. The forecasts are compared with the actual 2004 data. Optimum models are selected based on performance. Different modeling approaches are required for monthly and yearly data. The selected models accurately captured patterns and provided reliable forecasts. This study contributes to understanding climate trends in Armagh, aiding informed decision-making and proactive measures to address climate risks.**

***Index Terms*—Climate trends, Time series analysis, Descriptive Statistics, Naive Method, Exponential smoothing, ARIMA models, SARIMA models, Stationarity, Forecasting, and Model evaluation.**

I. Introduction

This report presents a comprehensive analysis of average temperatures in Armagh using time series analysis techniques. The dataset consists of monthly average temperatures from January 1844 to December 2004 and yearly average temperatures from 1844 to 2004. These datasets provide insights into long-term trends and seasonal variations in Armagh's average temperatures.

Distinct characteristics are observed in the two datasets. The monthly data exhibits stationarity indicating consistent statistical properties over time and prominent seasonality. Conversely, the yearly data displays a noticeable trend.

Various time series models are employed to capture patterns and forecast future temperature trends. For the monthly data, methods such as seasonal naive, Holt-Winters' seasonal exponential smoothing, and SARIMA models are utilized. These models consider seasonality and provide short-term temperature variations. For the yearly data with a trend, models such as naive, exponential smoothing, and ARIMA models are applied. These models aim to capture underlying trends and forecast long-term temperature trends.

Diagnostic tests and visualizations are conducted to assess the models' performance and adequacy. These include the Augmented Dickey-Fuller (ADF) test, autocorrelation function (ACF), and partial autocorrelation function (PACF) plots, as well as the Ljung-Box test. Additionally, metrics such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), Theil's U, Akaike information criterion (AIC), and Bayesian information criterion (BIC) are used to evaluate the models' performance.

The analysis is performed using the R programming language, leveraging its time series analysis libraries and functions. The objective of this study is to provide accurate and reliable forecasts for average temperatures in Armagh based on historical data.

II. Related Work

In the study "A Complete Tutorial on Time Series Modeling in R" by Tavish Srivastava, the author discusses the importance of time in business and the methods of prediction and forecasting. The study focuses on time series modeling as a powerful tool for analyzing time-based data and deriving hidden insights for informed decision-making.

Time series modeling is particularly useful when dealing with serially correlated data, commonly encountered in various business domains. The study serves as a guide, providing various levels of time series modeling and introducing the associated techniques. It aims to bridge the knowledge gap and equip analysts with the necessary skills to effectively utilize time series modeling in their analyses [1].

In the context of the present research on average temperatures in Armagh, time series modeling techniques discussed in the tutorial, such as exponential smoothing, ARIMA, and seasonal methods are applied to effectively model and forecast the temperature patterns. These techniques take into account the temporal nature of the data and can capture seasonality, trends, and other relevant components.

III. Methodology

A. Data Sets and Characteristics

The dataset for monthly data consists of 1932 observations of one variable. The data is stored in a data frame format. Each observation represents the average temperature for a specific month. The values range is in the sequential order of the months. For example, the first value corresponds to January, the second value to February, and so on (**Fig 1**).

The yearly dataset consists of 161 observations of one variable.

	x
1	8.5
2	8.3
3	9.7
4	8.9
5	8.5
6	8.7

Fig. 1: Monthly Time Series Dataset

Similar to the monthly data, the data is stored in a data frame format. Each observation represents the average temperature for a specific year. For example, the first value corresponds to the temperature for the first year, the second value to the temperature for the second year, and so on (**Fig 2**).

	x
1	4.5
2	2.4
3	4.8
4	9.1
5	10.9
6	12.9

Fig. 2: Yearly Time Series Dataset

Both datasets are valuable resources for analyzing and modeling average temperatures in Armagh. The monthly data provides a higher level of granularity, allowing for the examination of temperature variations within individual months. On the other hand, the yearly data provides a broader perspective, enabling the identification of long-term trends in average temperatures.

B. Data Exploration and Descriptive Statistics

This section aims to perform a descriptive analysis to gain insights into the variables present in the dataset.

1) The time series analysis is performed on the monthly temperature data using the time series object. The resulting time series object spans from January 1844 to December 2004, with a frequency of 12 (representing monthly observations). The class "ts", confirms its representation as a time series (**Fig 3**).
   Moving on to the summary statistics, the minimum, 1st quartile, median, mean, 3rd quartile, and maximum values of the monthly temperature data is calculated (**Fig 4**).
   The summary reveals the following characteristics:
   - The median value of 8.2 represents the middle value in the sorted distribution of the monthly temperature data in Armagh. It indicates that approximately 50% of the observed temperatures fall below 8.2, while the other 50% fall above this value.
   - The mean temperature of 8.501 represents the average temperature calculated over the observed period.

	Jan	Feb	Mar	Apr	May	Jun
1844	4.5	2.4	4.8	9.1	10.9	12.9
Class	ts	ts	ts	ts	ts	ts
Freq	12	12	12	12	12	12

Start	1844,1
End	2004,12

Fig. 3: Monthly Dataset Description

	Values
Min.	-0.900
1st Qu.	5.300
Median	8.200
Mean	8.501
3rd Qu.	12.100
Max.	17.200

Fig. 4: Monthly Dataset Summary

2) The 'start(ts_data)' function indicates that the yearly time series data starts in 1844 and ends in 2004, with a frequency of 1, indicating that the observations are collected on an annual basis (**Fig 5**).

	Start	End	Frequency	Class
	1844	2004	1	Time Series

Yearly Summary:

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	6.700	8.200	8.500	8.489	8.800	9.700

Fig. 5: Yearly Dataset Summary

The summary statistics for the yearly temperature data in Armagh reveal the following characteristics:

- The minimum temperature recorded is 6.7, the median is 8.5, and the maximum temperature is 9.7.
- The 1st quartile (25th percentile) represents the temperature value below which 25% of the yearly temperature data in Armagh falls, while the 3rd quartile (75th percentile) represents the temperature value below which 75% of the yearly temperature data falls.

C. Data Visualization

1) In the analysis of monthly average temperature data in Armagh, several visualizations are created to explore the characteristics of the time series.
   - First, a time plot of the monthly average temperature is generated, providing a visual representation of the data over time. The plot allows for a visual assessment of any trends, patterns, or outliers in the data (**Fig 6**).

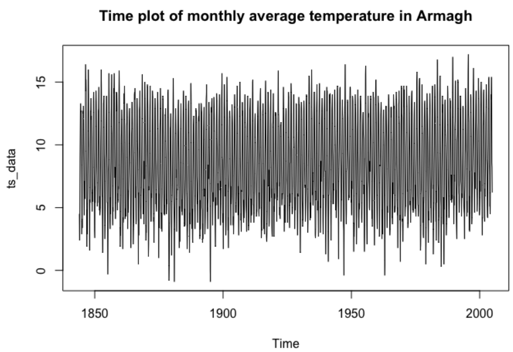

Fig. 6: Time plot of monthly average temperature in Armagh

   The time plot of the monthly average temperature reveals a consistent range of values between 5 and 15 on the y-axis, suggesting no discernible trends or patterns in the data.
   - Next, seasonal patterns are examined using two different visualizations [2]. The seasonal plot displayed the seasonal effects in the data, highlighting recurring patterns or fluctuations across the months (**Fig 7**).

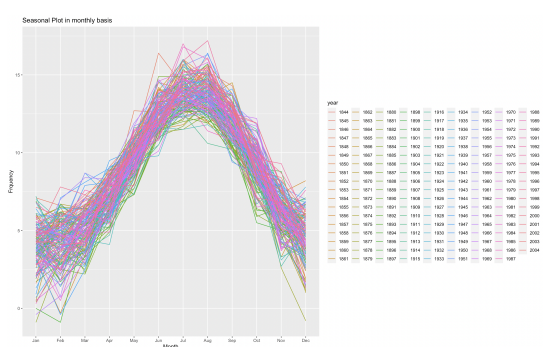

Fig. 7: Seasonal Plot

   Additionally, the subseries plot provides insights into the distribution of observations within each season (**Fig 8**).
   Both the seasonal plot and the subseries plot reveal consistent temperature rises from May to September, indicating a recurring seasonal pattern. Furthermore, a peak in temperatures is observed specifically during the months of July and August.

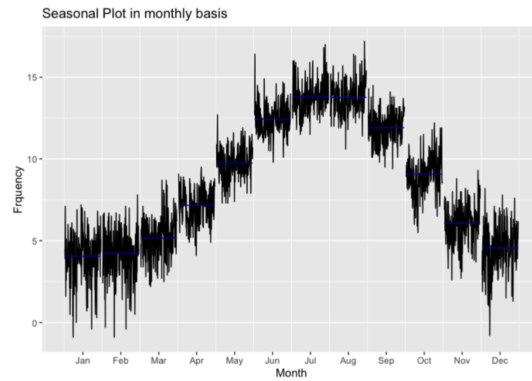

Fig. 8: Seasonal Subseries Plot

   - To further investigate both the seasonal effects and potential outliers within the data, a boxplot is constructed to illustrate the variation in temperature across the 12 months.

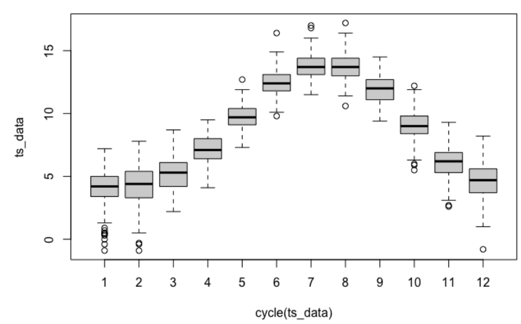

Fig. 9: Box Plot of Monthly Data

   The boxplot reveals the same outcome as the season plot (**Fig 9**).
   The dataset ts_data does not have any values that fall outside of the "whiskers" of the boxplot. This means that there are no outliers detected based on the criteria used by the boxplot.stats() function [3].

   - To decompose the time series into its underlying components, a decomposition plot is generated [4]. The plot displays the trend, seasonal, and remainder elements, allowing for a better understanding of the individual components' contributions to the overall series (**Fig 10**).
   The decomposition plot reveals an additive seasonality pattern and a slight presence of trend in the time series. To further assess the trend's stationarity, an Augmented Dickey-Fuller (ADF) test will be conducted.

   - Lastly, the correlogram, consisting of the autocorrelation function (ACF) and partial autocorrelation

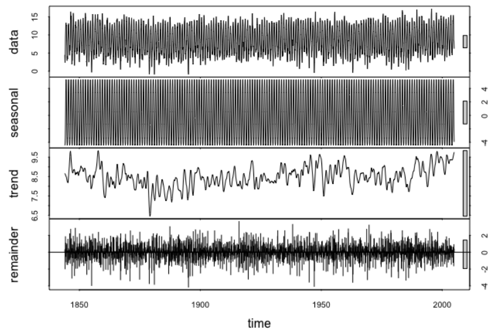

Fig. 10: Decomposed Data

function (PACF) plots [5], is utilized to assess the presence of autocorrelation in the time series (**Fig 11**).

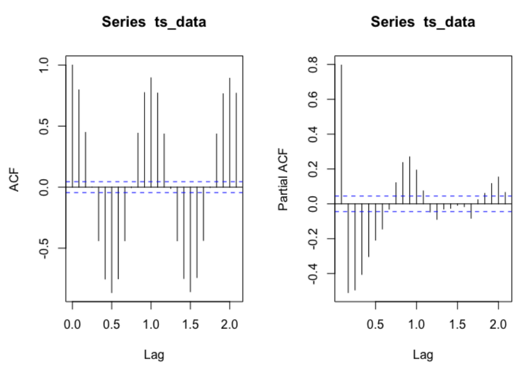

Fig. 11: Correlogram of Monthly Data

The autocorrelation plot shows strong positive correlations at lag 0 and lags 1 and 2, indicating a potential seasonal pattern in the data. The partial autocorrelation plot displays a significant positive partial autocorrelation at lag 0, implying a direct relationship between consecutive observations. The negative partial autocorrelation suggests the impact of the immediate previous observation.

2) In the analysis of yearly average temperature data in Armagh, various visualizations are generated to investigate the properties of the time series.
   - A time plot is created to visualize the yearly average temperature data over time (**Fig 12**). Clearly, the data is not stationary and exhibit some pattern in the plot.
     To make the data stationary, differencing is applied with the help of ndiffs() function. The differenced yearly average temperature time series is shown in a time plot, which helps in identifying any remaining patterns or trends (**Fig 13**).
     By differencing the data, the mean is adjusted to zero, helping to remove any persistent trends and

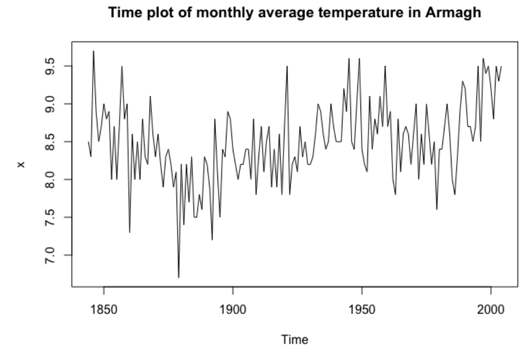

Fig. 12: Time plot of yearly average temperature in Armagh

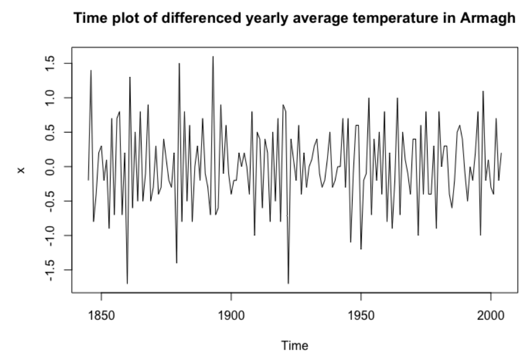

Fig. 13: Time plot of differenced yearly average temperature in Armagh

making the data trend stationary.
   - The correlogram, consisting of the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, is utilized to examine the presence of autocorrelation in the differenced time series (**Fig 14**).

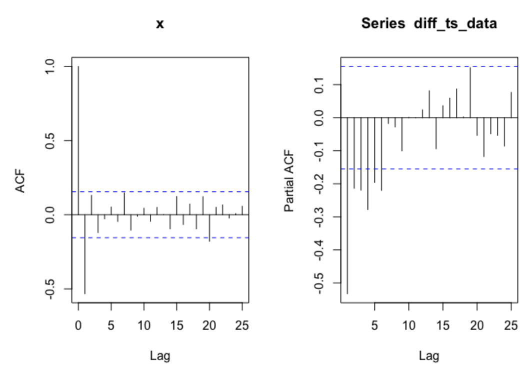

Fig. 14: Correlogram of Yearly Data

In the ACF, no prominent lags are evident, as most correlation values fall within the range of -0.2 to 0.2. Similarly, in the PACF, a negative correlation

is observed at lag 1, indicating a direct influence of the previous observation. However, there are no significant or prominent lags beyond lag 1 in the PACF, indicating no distinct pattern in the partial autocorrelations.
Overall, the correlogram does not show any significant or consistent autocorrelation patterns, suggesting the absence of strong dependencies or trends in the differenced time series data.

D. Time Series Properties

- For the yearly average temperature data in Armagh, the Augmented Dickey-Fuller (ADF) test is conducted to examine the stationarity of the data [6]. The test yielded a Dickey-Fuller statistic of -2.861 with a p-value of 0.2172. The p-value suggests that there is insufficient evidence to reject the null hypothesis of non-stationarity.
Similarly, for the differenced yearly average temperature data, the ADF test is performed. The test resulted in a very small p-value of 0.01. The p-value indicates strong evidence to reject the null hypothesis of non-stationarity in favor of the alternative hypothesis of stationarity (**Fig 15**). The mean seasonal cycle is 1, indicating a consistent pattern observed across the years. The lack of significant variability in the seasonal component suggests a consistent annual temperature pattern in the region.

Test	Dickey-Fuller	Lag Order	p-value	Stationary?
Original Data	-2.861	5	0.2172	No
Differenced	-9.8954	5	0.01	Yes

	Value
Mean seasonal cycle	1
Standard deviation of seasonal cycle	0

Fig. 15: Stationary Yearly Data

For the monthly average temperature data in Armagh, the ADF test is conducted. The test produced a Dickey-Fuller statistic of -7.6152 with a p-value of 0.01. The p-value suggests significant evidence to reject the null hypothesis of non-stationarity in favor of the alternative hypothesis of stationarity (**Fig 16**).

Test	Dickey-Fuller	Lag Order	p-value	Stationary?
Original Data	-7.6152	12	0.01	Yes

	Mean	Standard Deviation
Monthly Data	6.5	3.452946

Fig. 16: Stationary Monthly Data

The mean seasonal cycle for the monthly data is computed as 6.5, indicating an average fluctuation around this value over the course of a year. The standard deviation of the seasonal cycle is calculated as 3.452946, suggesting the degree of variability in the monthly average temperature within each season.

IV. Evaluation Metrics

The evaluation metrics used to assess the accuracy of the generated forecasts against the actual data for the year 2004 include ME (Mean Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), MPE (Mean Percentage Error), MAPE (Mean Absolute Percentage Error), MASE (Mean Absolute Scaled Error), ACF1 (Autocorrelation of Residuals at Lag 1), and Theil's U statistic [7].

V. Model Performance - Monthly Data

1) The Seasonal Naive method is applied to forecast the monthly average temperature in Armagh. The dataset is split into a training set (until December 2003) and a test set (starting from January 2004). The Residual sd value of 1.705 represents the standard deviation of the forecast errors. A higher value indicates larger deviations from the actual values, suggesting a relatively higher level of uncertainty in the forecasts (**Fig 17**) [8].

```
Forecasts:
         Point Forecast     Lo 80     Hi 80      Lo 95     Hi 95
Jan 2004            4.3  2.114926  6.485074  0.9582173  7.641783
Feb 2004            4.2  2.014926  6.385074  0.8582173  7.541783
Mar 2004            6.4  4.214926  8.585074  3.0582173  9.741783
Apr 2004            8.8  6.614926 10.985074  5.4582173 12.141783
May 2004           10.0  7.814926 12.185074  6.6582173 13.341783
Jun 2004           13.2 11.014926 15.385074  9.8582173 16.541783
Jul 2004           15.2 13.014926 17.385074 11.8582173 18.541783
Aug 2004           15.4 13.214926 17.585074 12.0582173 18.741783
Sep 2004           12.9 10.714926 15.085074  9.5582173 16.241783
Oct 2004            8.5  6.314926 10.685074  5.1582173 11.841783
Nov 2004            7.6  5.414926  9.785074  4.2582173 10.941783
Dec 2004            5.1  2.914926  7.285074  1.7582173  8.441783
```

Fig. 17: Point Forcast Value for Seasonal Naive

The Ljung-Box test is conducted to check the residuals of the Seasonal Naive method, and the obtained p-value is less than 0.05, indicating significant evidence of autocorrelation in the residuals (**Fig 18**) [9].

```
	Ljung-Box test

data:  Residuals from Seasonal naive method
Q* = 674.17, df = 24, p-value < 2.2e-16

Model df: 0.   Total lags used: 24
```

Fig. 18: Ljung Box Test for Seasonal Naive

The Seasonal Naive method performs reasonably well in forecasting the monthly average temperature in Armagh. The MPE (Mean Percentage Error) value of 3.507% indicates that, on average, the forecasts have a percentage error of approximately 3.507%. The MASE (Mean Absolute Scaled Error) value of 0.437 suggests that the Seasonal Naive method captures the variability of the data reasonably well when compared to a naive benchmark model.
The ACF1 (Autocorrelation of Residuals at Lag 1) value of -0.031 indicates minimal correlation in the residuals at Lag 1.

Theil's U statistic of 0.301 represents moderate forecasting performance. A value closer to 0 indicates better forecasting performance. (**Fig 19**).

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	0.005	1.705	1.334	-Inf	Inf	1	0.198	NA
Test set	0.200	0.729	0.583	3.507	7.01	0.437	-0.031	0.301

Fig. 19: Evaluation Metrics for Seasonal Naive

A plot of the forecasts and the actual data for the test period visually depicts the performance of the Seasonal Naive method (**Fig 20**).

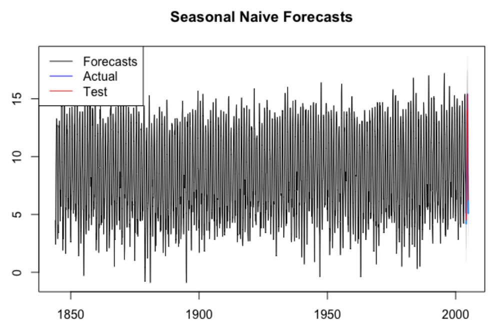

Fig. 20: Forecast Data

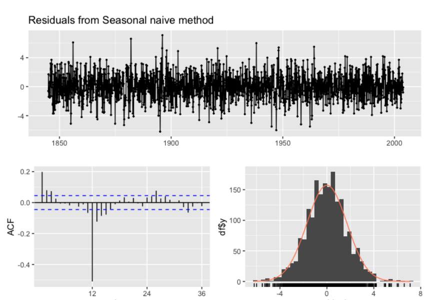

Fig. 21: Residuals of Seasonal Naive

Furthermore, the residuals of the Seasonal Naive method are checked for white noise, normal distribution, and autocorrelation (at Lag 12 autocorrelation present) (**Fig 21**).
However, there is room for improvement in terms of reducing the autocorrelation in the residuals and enhancing the accuracy of the forecasts.

2) Holt's method with additive seasonality is applied to forecast the monthly average temperature in Armagh. The Residual Standard Error (RSE) of 1.250518 indicates the average magnitude of the forecast errors.
The forecasts generated by the HoltWinters model for the test data along with 80% and 95% confidence intervals. The forecasts indicate the expected values for

```
Forecasts:
         Point Forecast     Lo 80     Hi 80     Lo 95     Hi 95
Jan 2004       4.755004  3.152846  6.357161  2.304715  7.205292
Feb 2004       4.935544  3.330648  6.540441  2.481068  7.390021
Mar 2004       6.253573  4.645899  7.861246  3.794848  8.712297
Apr 2004       7.862041  6.251552  9.472531  5.399010 10.325073
May 2004      10.526687  8.913342 12.140032  8.059288 12.994085
Jun 2004      12.918052 11.301812 14.534292 10.446226 15.389877
Jul 2004      14.778224 13.159050 16.397398 12.301911 17.254537
Aug 2004      14.797802 13.175654 16.419949 12.316941 17.278662
Sep 2004      12.784633 11.159472 14.409794 10.299163 15.270102
Oct 2004      10.062583  8.434368 11.690797  7.572444 12.552722
Nov 2004       7.168671  5.537364  8.799979  4.673802  9.663541
Dec 2004       5.299075  3.664634  6.933516  2.799414  7.798736
```

Fig. 22: Point Forcast Data

each month in 2004 (**Fig 22**).
The p-value obtained from the Ljung-Box test is highly significant, suggesting evidence of autocorrelation (**Fig 23**).

```
        Ljung-Box test

data:  Residuals from HoltWinters
Q* = 97.756, df = 24, p-value = 7.233e-11

Model df: 0.   Total lags used: 24
```

Fig. 23: Ljung Box Test

When evaluating the model's performance against the actual data for the year 2004, Lower MAE and RMSE value is achieved. The MASE value is indicating that the model performs reasonably well in capturing the data variability compared to the naive benchmark. The ACF1 value suggests a moderate level of correlation in the residuals at lag 1. Theil's U statistic of 0.326 indicates moderate forecasting performance (**Fig 24**).

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	0.041	1.251	0.975	-Inf	Inf	0.731	0.162	NA
Test set	0.155	0.798	0.676	1.612	7.647	0.507	-0.398	0.326

Fig. 24: Evaluation Metrics of Holt's Method

The plot of the forecasts visually compares the HoltWinters forecasts (black line) with the actual test data (red line) (**Fig 25**).

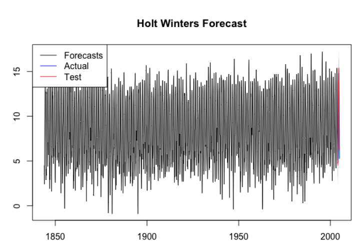

Fig. 25: Forecast Data

The residuals of Holt's method with the additive seasonality model are evaluated for white noise, normal distribution, and autocorrelation (**Fig 26**).

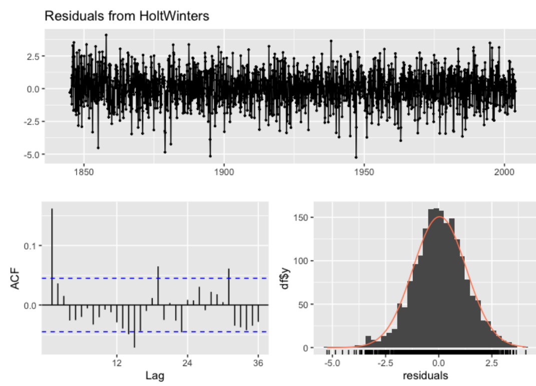

Fig. 26: Residuals of Holt's Method

Holt's method with additive seasonality provides reasonably accurate forecasts for the monthly average temperature in Armagh. However, there is room for improvement as the model should account for the significant autocorrelation present in the residuals.

3) The best model for forecasting the data is identified as SARIMA(2,0,2)(1,1,1)[12]. Additionally, alternative ARIMA models with different parameter configurations are tested, including ARIMA(2,1,0)(1,0,1)[12], ARIMA(1,1,1)(2,0,0), and ARIMA(1,1,1)(1,0,0). However, the SARIMA(2,0,2)(1,1,1)[12] model outperformed these alternatives based on the evaluation metrics and achieved a lower AIC value, indicating its superior fit to the data (**Fig 27**).

```
ARIMA(2,0,2)(1,1,1)[12]

Coefficients:
         ar1      ar2      ma1     ma2      sar1     sma1
      1.3267  -0.3287  -1.1169  0.1313  -0.0106  -0.9777
s.e.  0.1179   0.1175   0.1239  0.1210   0.0238   0.0074

sigma^2 = 1.428:  log likelihood = -3062.05
AIC=6138.11   AICc=6138.17   BIC=6176.98
```

Fig. 27: Best Sarima Model

The SARIMA(2,0,2)(1,1,1)[12] model also produces point forecasts with corresponding lower and upper bounds at the 80% and 95% confidence levels (**Fig 28**).

Forecasts:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2004	4.560434	3.029053	6.091816	2.218388	6.902480
Feb 2004	4.702597	3.137880	6.267315	2.309568	7.095626
Mar 2004	5.938963	4.369354	7.508572	3.538453	8.339472
Apr 2004	7.723171	6.152466	9.293875	5.320985	10.125356
May 2004	10.354576	8.783433	11.925719	7.951720	12.757432
Jun 2004	12.917692	11.346263	14.489121	10.514399	15.320985
Jul 2004	14.502324	12.930654	16.073995	12.098662	16.905987
Aug 2004	14.443730	12.871833	16.015627	12.039721	16.847739
Sep 2004	12.542178	10.970060	14.114296	10.137831	14.946525
Oct 2004	9.875781	8.303444	11.448117	7.471099	12.280462
Nov 2004	6.764779	5.192225	8.337332	4.359766	9.169791
Dec 2004	5.185063	3.612294	6.757831	2.779721	7.590404

Fig. 28: Point Forecast Data

When evaluating the forecasts against the actual data for the year 2004, the model provides a mean error (ME) of 0.374, an RMSE of 0.861, and an MAE of 0.771. The MASE value is 0.578, indicating that the model performs reasonably well in capturing the data variability compared to the naive benchmark (**Fig 29**).

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	0.012	1.189	0.927	-Inf	Inf	0.695	0.000	NA
Test set	0.374	0.861	0.771	4.316	8.645	0.578	-0.429	0.364

Fig. 29: Evaluation Matrics

The plot of the forecasts visually compares the SARIMA forecasts (black line) with the actual test data (red line) (**Fig 30**).

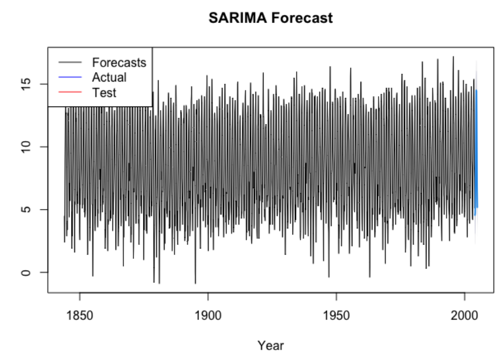

Fig. 30: SARIMA Forecast

The residuals of the SARIMA model are checked for autocorrelation using the Ljung-Box test, which yields a p-value of 0.1792, suggesting no significant autocorrelation in the residuals (**Fig 31**).

```
	Ljung-Box test

data:  Residuals from ARIMA(2,0,2)(1,1,1)[12]
Q* = 23.302, df = 18, p-value = 0.1792

Model df: 6.   Total lags used: 24
```

Fig. 31: Ljung Box Test

In addition, the residuals of the SARIMA model are checked for white noise and normal distribution. The residual standard error (RSE) for the SARIMA model is 1.189311 (**Fig 32**).

In conclusion, the SARIMA(2,0,2)(1,1,1)[12] model provides accurate and reliable forecasts for the specified time series data. The model exhibits satisfactory performance, as evidenced by low error measures. The residuals of the model demonstrate no significant autocorrelation, indicating that the model captures the underlying patterns in the data effectively.

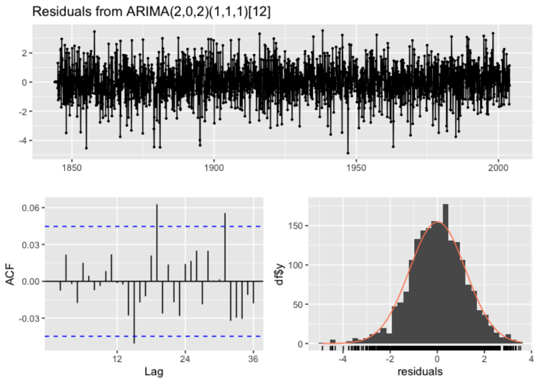

Fig. 32: Residuals of the Sarima Model

VI. Model Performance - Yearly Data

1) The data is split into training and test sets, with the training set consisting of data up to the end of 2003 and the test set starting from 2004.
   A Naive method is applied to forecast the data. The model yielded a residual standard deviation of 1.07. The forecasts for the test data showed a point forecast of -0.2 for 2004 (**Fig 33**).

```
Forecasts:
     Point Forecast       Lo 80    Hi 80     Lo 95    Hi 95
2004           -0.2 -1.571206 1.171206 -2.297079 1.897079
```

Fig. 33: Forecast Data

The accuracy evaluation on the test set indicates a mean absolute error (MAE) of 0.400 and a mean absolute percentage error (MAPE) of 200% (**Fig 34**).

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.0	1.07	0.852	NaN	Inf	1.00	-0.713
Test set	0.4	0.40	0.400	200	200	0.47	NA

Fig. 34: Evaluation Matrics

```
        Ljung-Box test

data:  Residuals from Naive method
Q* = 105.23, df = 10, p-value < 2.2e-16
```

Fig. 35: Ljung-Box Test

The residuals of the Naive forecast model don't appear to be white noise according to the Ljung-Box test (**Fig 35 and 36**).
Based on the evaluation metrics, the Naive method performed poorly in forecasting the yearly data, as indicated by the high mean absolute percentage error (MAPE) of 200%. Additionally, the residuals of the Naive forecast model do not appear to exhibit the characteristics of white noise, as indicated by the results of the Ljung-Box test. These findings suggest that the Naive method is not suitable for accurately predicting the yearly data.

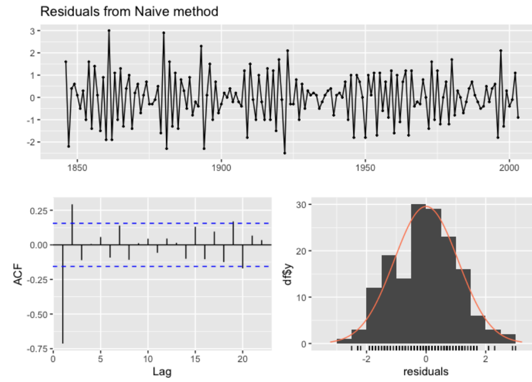

Fig. 36: Residuals From the Naive Model

2) The Exponential Smoothing (ETS) model, specifically the ETS(A, N, N) variant (Additive error, no trend, and no seasonality components) is applied to make forecasts for the test data.
   The model produced a point forecast of 0.005 with a residual standard error (RSE) of 0.61 (**Fig 37**). The

```
Forecasts:
     Point Forecast      Lo 80     Hi 80     Lo 95    Hi 95
2004    0.005019504 -0.7810119 0.7910509 -1.197112 1.207151
```

Fig. 37: Point Forecast Data

Akaike Information Criterion (AIC) value for the ETS model is 654.4951. The accuracy evaluation on the test set revealed a mean absolute error (MAE) of 0.195 and a mean absolute percentage error (MAPE) of 97.49%. The negative autocorrelation function (ACF1) value of -0.532 suggests a moderate inverse relationship between consecutive residuals ETS model (**Fig 38**).

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.000	0.609	0.494	-Inf	Inf	0.579	-0.532
Test set	0.195	0.195	0.195	97.49	97.49	0.229	NA

Fig. 38: Evaluation Metrics

The plot displayed the forecasts along with the actual test data for comparison (**Fig 39**).

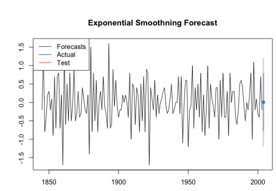

Fig. 39: Forecast Data

A residual plot is generated for the Exponential Smoothing (ETS) model, indicating whether the residuals exhibit white noise characteristics and follow a normal distribution (**Fig 40 and 41**).

```
        Ljung-Box test

data:  Residuals from ETS(A,N,N)
Q* = 57.314, df = 10, p-value = 1.164e-08

Model df: 0.   Total lags used: 10
```

Fig. 40: Ljung Box Test

Based on the Ljung-Box test results for the residuals of the ETS(A, N, N) model, the p-value is significantly low, indicating that the residuals are not consistent with white noise.

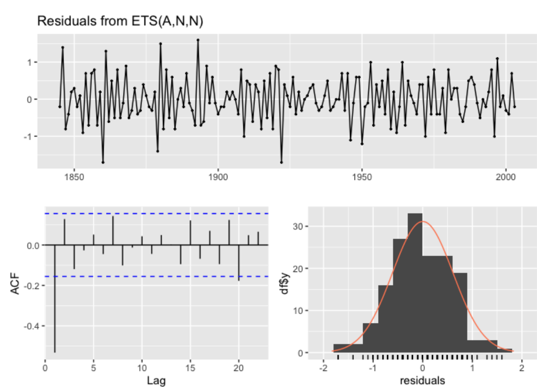

Fig. 41: Residual From Exponential Smoothing Model

This suggests that the ETS model may not adequately capture all the underlying patterns and information in the data.

3) Multiple models are evaluated for forecasting, including the ARIMA(1,1,1) and ARIMA(1,1,2) models. However, after considering the evaluation metrics and AIC value, the auto. arima function determined that the ARIMA(0,0,1) model yielded the best performance. The model yielded a point forecast of -0.145 and a residual standard deviation (RSE) of 0.46 (**Fig 42**).

```
Forecasts:
     Point Forecast      Lo 80     Hi 80    Lo 95     Hi 95
2004     -0.1454509 -0.7395974 0.4486956 -1.05412 0.7632177
```

Fig. 42: Point Forcast Data

The Akaike Information Criterion (AIC) value for the ARIMA model was 210.93. The accuracy evaluation on the test set indicated a mean absolute error (MAE) of 0.345 and a mean absolute percentage error (MAPE) of 172.725% (**Fig 43**).

```
                ME  RMSE   MAE     MPE    MAPE  MASE   ACF1
Training set 0.018 0.462 0.353     NaN     Inf 0.414 -0.032
Test set     0.345 0.345 0.345 172.725 172.725 0.406     NA
```

Fig. 43: Evaluation Matrics

The Ljung-Box test for the residuals showed a p-value of 0.3546, suggesting that the residuals are consistent with white noise (**Fig 44**).

```
        Ljung-Box test

data:  Residuals from ARIMA(0,0,1) with zero mean
Q* = 9.9498, df = 9, p-value = 0.3546

Model df: 1.   Total lags used: 10
```

Fig. 44: Ljung Box Test

The Ljung-Box test for the residuals showed a p-value of 0.3546, suggesting that the residuals are consistent with white noise (**Fig 45**).

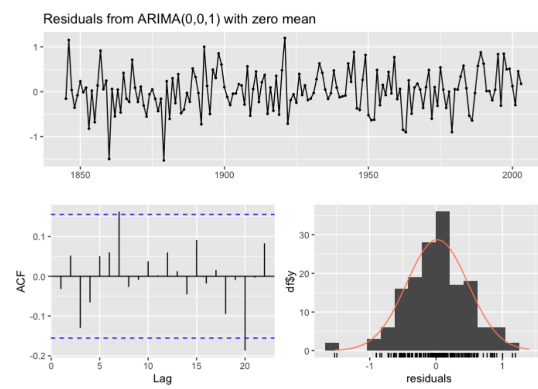

Fig. 45: Residual From ARIMA Model

The plot displayed the forecasts along with the actual test data for comparison (**Fig 46**).

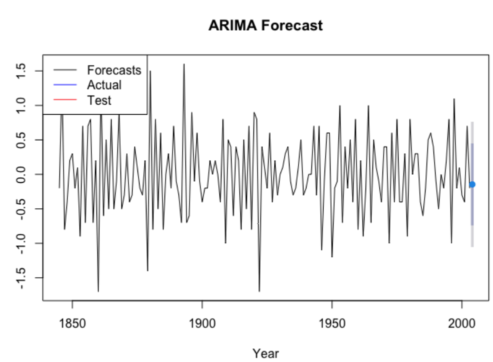

Fig. 46: ARIMA Forecast

Based on the evaluation of the test data for yearly forecasts, the ARIMA(0,0,1) model outperforms the

Exponential Smoothing (ETS) and Naive models. The ARIMA model exhibits the lowest mean absolute error (MAE) and root mean squared error (RMSE), indicating better accuracy in predicting the data. Furthermore, the residuals of the ARIMA model show lower autocorrelation (ACF1) compared to the ETS and Naive models. Therefore, the ARIMA(0,0,1) model is the preferred choice for yearly data forecasting based on these metrics.

VII. Conclusion

For Monthly data, Among the three forecasting methods, SARIMA, Holt-Winters, and Seasonal Naive, their performance is evaluated using various metrics on the test data. Holt-Winters demonstrated the best overall performance, with the lowest values for RMSE, MAE, MAPE, and MASE. It also had a reasonably good fit to the autocorrelation structure, as indicated by the near-zero negative autocorrelation function (ACF1) value. Theil's U value further supported the effectiveness of Holt-Winters in forecasting. However, SARIMA exhibits the lowest RSD (1.189311), indicating the smallest average forecast errors. To achieve accurate point forecasts Holt-Winters model is best and to capture the variability (complex patterns and autocorrelation) in the data SARIMA is the best option because of its lowest residual standard deviation.

In conclusion, when evaluating the yearly data, the Exponential Smoothing (ETS) model demonstrates better accuracy in terms of mean absolute error (MAE) and mean absolute percentage error (MAPE) for test sets. However, the ARIMA model exhibits superior overall forecast performance, considering the lower residual standard deviation (RSD) and significantly lower AIC value of 210.93. This indicates that the ARIMA model provides a better fit to the data and is more parsimonious. Therefore, based on the evaluation metrics, variability of residuals, and AIC value, the ARIMA model is the recommended choice for forecasting the yearly data.

References

[1] Tavish Srivastava, ”A Complete Tutorial on Time Series Modeling in R”, [Online]. Available: https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling//

[2] ”3 Ways to Visualize Time Series You May Not Know”, [Online]. Available: https://towardsdatascience.com/3-ways-to-visualize-time-series-you-may-not-know-c8572952ea9c/

[3] ”Stats and R” [Online]. Available: https://statsandr.com/blog/outliers-detection-in-r//

[4] ”Time Series Analysis in R - Decomposing Time Series” [Online]. Available: https://rpubs.com/davoodastaraky/TSA1/

[5] ”Interpreting ACF and PACF Plots for Time Series Forecasting” [Online]. Available: https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c: :text=Both

[6] ”Augmented Dickey-Fuller Test in R” [Online]. Available: https://www.r-bloggers.com/2022/06/augmented-dickey-fuller-test-in-r//

[7] ”Time Series Forecast Error Metrics You Should Know” [Online]. Available: https://towardsdatascience.com/time-series-forecast-error-metrics-you-should-know-cc88b8c67f27/

[8] Residual Standard Deviation: Definition, Formula, and Examples [Online]. Available: https://www.investopedia.com/terms/r/residual-standard-deviation.asp/

[9] ”Ljung-Box Test: Definition + Example” [Online]. Available: https://www.statology.org/ljung-box-test//