

HAWASSA UNIVERSITY



COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

**APPLICATION OF TIME SERIES ANALYSIS FOR MODELLING AND
PREDICTING MONTHLY AVERAGE TEMPERATURE IN GLOBAL
AND EAST AFRICA**

MSc. THESIS

NEBIYU MOHAMMED WOYESO

ADVISOR: CHERU ATSMEGIORGIS (PhD)

HAWASSA, ETHIOPIA

MAY, 2024

**APPLICATION OF TIME SERIES ANALYSIS FOR MODELLING AND PREDICTING
MONTHLY AVERAGE TEMPERATURE IN GLOBAL AND EAST AFRICA**

NEBIYU MOHAMMED WOYESO

ADVISOR: CHERU ATSMEGIORGIS (PhD)

**THESIS SUBMITTED TO DEPARTMENT OF STATISTICS, COLLEGE OF NATURAL
AND COMPUTATIONAL SCIENCES, SCHOOL OF GRADUATE STUDIES,
HAWASSA UNIVERSITY, HAWASSA, ETHIOPIA**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTERS OF SCIENCE IN STATISTICS (APPLIED STATISTICS)**

HAWASSA, ETHIOPIA

MAY, 2024

DECLARATION

I hereby declare that this MSc thesis is my original work and has not been presented for a degree in any other university, and all sources of material used for this thesis have been duly acknowledged.

Name: _____

Signature: _____

This MSc thesis has been submitted for examination with my approval as thesis advisor.

Name: _____

Signature: _____

Place and Date of Submission: _____

HAWASSA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
ADVISORS' APPROVAL SHEET

This is to certify that the thesis entitled “**APPLICATION OF TIME SERIES ANALYSIS FOR MODELLING AND PREDICTING MONTHLY AVERAGE TEMPERATURE IN GLOBAL AND EAST AFRICA**” submitted in partial fulfillment of the requirements for the degree of Master’s with specialization in Applied Statistics, the graduate program of the department of Statistics, and has been carried out by **Nebiyu Mohammed**, ID No: **GpApStR/0002/15** under my supervision. Therefore, I recommend that the student has fulfilled the requirements and hence hereby can submit the thesis to the department.

Name of Advisor

Signature

Date

Dr. Cheru Atsmegiorgis

HAWASSA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
EXAMINERS' APPROVAL SHEET

We, the undersigned, members of the Board of Examiners of the final open defense by **Nebiyu Mohammed Woyeso** have read and evaluated his thesis entitled “**APPLICATION OF TIME SERIES ANALYSIS FOR MODELLING AND PREDICTING MONTHLY AVERAGE TEMPERATURE IN GLOBAL AND EAST AFRICA**” and examined the candidate. This is, therefore, to certify that the thesis has been accepted in partial fulfillment of the requirements for the degree of Masters of Science in Applied Statistics.

_____	_____	_____
Name of the Chairperson	Signature	Date
<u>Dr. Cheru Atsmegiorgis</u>	_____	_____
Name of advisor	Signature	Date
_____	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of External Examiner	Signature	Date
_____	_____	_____
SGS approval	Signature	Date

Final approval and acceptance of the thesis is contingent upon the submission of the final copy of the thesis to the School of Graduate Studies (SGS) through the School Graduate Committee (SGC) of the candidate's department.

Stamp of SGS: _____ Date: _____

ACKNOWLEDGEMENTS

First and foremost, I offer my deepest gratitude and praise to God, the Almighty, for His abundant blessings throughout my research work, enabling me to reach this beautiful day marking the end of a long journey.

I am indebted to several individuals for their invaluable support and contributions to the completion of this research project. Among them, I would like to express my profound thanks to my advisor, Dr. Cheru Atsmegiorgis, for his keen interest in this research project. His vision, sincerity, and motivation have deeply inspired me. It has been a great privilege and honor to work and study under his guidance.

I am extremely grateful to my parents for their unwavering love, prayers, care, and sacrifices in educating and preparing me for my future.

Many thanks to Mr. Sergei Rogushin. He has been an exceptional sounding board and provider of constructive feedback. I truly appreciate every conversation we had and every piece of moral support I received from him.

Lastly, but not least, I would like to thank my friends for their unwavering support, encouragement, and valuable ideas throughout my thesis journey and in my future endeavors.

Table of Contents

DECLARATION.....	III
ACKNOWLEDGEMENTS	VI
Acronyms /Abbreviations	IX
List of Tables.....	XI
List of Figures	XII
Abstracts	XIV
1. Introduction.....	1
1.1 Background of the Study	1
1.2 Statement of the problem	3
1.3 Objectives of the Study	4
1.3.1 General Objective.....	4
1.3.2 Specific Objective	5
1.4 Significance of the Study	5
2. Literature Review.....	7
2.1 Overview of Time Series Analysis	7
2.2 Overview of Climate Change.....	8
2.2.1 Global Climate Change Impacts	8
2.2.2 Agriculture	8
2.2.3 Climate Change Impacts for the East Africa region.....	10
2.3 Application of Time Series Analysis in Climate Modelling	12
3. Methods and Materials.....	14
3.1 Study Area	14
3.2 Data Source	14
3.3 Time Series Analysis	15
3.4 Decomposition of Series	16
3.5 Stationarity	16
3.5.1 Examining Stationary of Time Series Data	17
3.6 Unit Root Tests.....	19
3.6.1 Augmented Dickey-Fuller (ADF) test.....	19
3.6.2 Phillips-Perron (PP) test.....	20
3.7 Transforming Nonstationary Data	20
3.8 Seasonal ARIMA (SARIMA) Process	21
3.9 Building ARIMA/SARIMA Models	21

3.9.1	Model Identification	22
3.9.2	Model Selection.....	23
3.9.3	Parameter Estimation	24
3.9.4	Diagnostic checking	24
3.9.5	Forecasting	26
3.10	Conditional Heteroscedastic Models	27
3.10.1	Volatility Testing.....	28
3.10.2	ARCH Model.....	28
3.10.3	The Generalized ARCH (GARCH) Models	28
4	Results and Discussion	31
4.1	Descriptive statistics.....	31
4.2	Test of Stationarity	44
4.3	Building SARIMA Model	48
4.3.1	Identifying parameters of SARIMA models	48
4.3.2	SARIMA Model Selection	49
4.3.3	Parameter Estimation	57
4.3.4	Diagnostic Checking	62
4.3.5	Forecasting	67
4.3.6	Forecasting Accuracy	70
4.4	Conditional Heteroscedasticity Models	72
4.4.1	Volatility Testing	72
4.4.2	Model Selection for Volatility	73
4.4.3	Parameter Estimation of EGARCH.....	77
4.4.4	Diagnostics Checking of EGARCH Model	81
4.5	Discussion	85
5.	Conclusion and Recommendation	89
5.1	Conclusion.....	89
5.2	Recommendations	90
	References.....	92
	Appendices.....	99
	Appendix A: Figures	99
	Appendix B: Tables.....	109

Acronyms /Abbreviations

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike's Information Criterion
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
DSC	Deterministic and Stochastic Combined
ECA	Economic Commission for Africa
EGARCH	Exponential GARCH
FAO	Food and Agriculture Organization
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GDP	Gross Domestic Product
GISS	Goddard Institute for Space Studies
GMT	Global Mean Temperature
IGARCH	Integrated GARCH
IPCC	Intergovernmental Panel on Climate Change
JB	Jarque-Bera
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MAT	Monthly Average Temperature
MSE	Mean Squared Errors
NASA	National Aeronautics and Space Administration
NGARCH	Non-Linear Asymmetric GARCH
PACF	Partial Autocorrelation Function
PP	Phillips-Perron
RMSE	Root Mean Square Error
SARIMA	Seasonal Autoregressive Integrated Moving Average

SSA	Sub-Saharan Africa
TGARCH	Threshold GARCH
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
UNICEF	United Nations International Children's Emergency Fund

List of Tables

Table 3.1 Behavior of ACF and PACF for Non-seasonal ARMA (p, q)	23
Table 3.2 Behavior of ACF and PACF for seasonal ARMA (P, Q).....	23
Table 4.1 Descriptive statistics of monthly average temperature	31
Table 4.2 Statistical tests of Stationarity for Regional Temperature.....	45
Table 4.3 Parameter estimation of SARIMA(4,0,0)(0,1,1,12) for Global MAT.....	58
Table 4.4 Parameter estimation of SARIMA(2,0,2)(3,1,1,12) for Ethiopia MAT	59
Table 4.5 Parameter estimation of SARIMA(2,0,3)(0,1,1,12) for Kenya MAT	60
Table 4.6 Parameter estimation of SARIMA(1,0,4)(0,1,1,12) for Uganda MAT	61
Table 4.7 Forecasting accuracy metrics	70
Table 4.8 ARCH effect test	73
Table 4.9 Model selection for ARCH model	74
Table 4.10 Model selection of GARCH model for Global and Ethiopia MAT	74
Table 4.11 Model selection of GARCH model for Kenya and Uganda MAT	75
Table 4.12 Model selection of EGARCH model for Global and Ethiopia MAT	75
Table 4.13 Model selection of EGARCH model for Kenya and Uganda MAT	76
Table 4.14 Overall volatility model selection	77
Table 4.15 Parameter Estimation of EGARCH(4,1) for Global MAT	78
Table 4.16 Parameter Estimation of EGARCH(3,1) for Ethiopia MAT	78
Table 4.17 Parameter Estimation of EGARCH(3,2) for Kenya MAT	79
Table 4.18 Parameter Estimation of EGARCH(4,1) for Uganda MAT	80
Table 0.1 SARIMA model selection for Global MAT	109
Table 0.2 SARIMA model selection for Ethiopia, Kenya, and Uganda MAT	111

List of Figures

Figure 3.1 Flowchart of the proposed methodology	30
Figure 4.1 Exploratory Data Analysis of Global Monthly Average Temperature.....	33
Figure 4.2 Exploratory Data Analysis of Ethiopia Monthly Average Temperature	35
Figure 4.3 Exploratory Data Analysis of Kenya Monthly Average Temperature	37
Figure 4.4 Exploratory Data Analysis of Uganda Monthly Average Temperature	39
Figure 4.5 Zoomed monthly average temperature	40
Figure 4.6 Yearly Average Temperature.....	41
Figure 4.7 Train-Test Split of Monthly Average Temperature	42
Figure 4.8 Test of Stationarity for Global MAT.....	44
Figure 4.9 Evaluation of selected models in validation set of Global MAT	51
Figure 4.10 Evaluation of selected models in validation set of Ethiopia MAT	52
Figure 4.11 Evaluation of selected models in validation set of Kenya MAT	53
Figure 4.12 Evaluation of selected models in validation set of Uganda MAT	54
Figure 4.13 Evaluation of selected models in the test set of Global MAT	54
Figure 4.14 Evaluation of selected models in the test set of Ethiopia MAT.....	55
Figure 4.15 Evaluation of selected models in the test set of Kenya MAT.....	56
Figure 4.16 Evaluation of selected models in the test set of Uganda MAT	57
Figure 4.17 Residual analysis of SARIMA(4,0,0)(0,1,1,12) for Global MAT	63
Figure 4.18 Residual analysis of SARIMA(2,0,2)(3,1,1,12) for Ethiopia MAT	64
Figure 4.19 Residual analysis of SARIMA(2,0,3)(0,1,1,12) for Kenya MAT	65
Figure 4.20 Residual analysis of SARIMA(1,0,4)(0,1,1,12) for Uganda MAT	66
Figure 4.21 Forecast of Global MAT	68
Figure 4.22 Forecast of Ethiopia MAT	69
Figure 4.23 Forecast of Kenya MAT	69
Figure 4.24 Forecast of Uganda MAT	70
Figure 4.25 Residual analysis of EGARCH(3,1) for Ethiopia MAT	81
Figure 4.26 Residual analysis of EGARCH(3,2) for Kenya MAT	83
Figure 4.27 Residual analysis of EGARCH(4,1) for Uganda MAT	84
Figure 0.1 Time Series Plot of Monthly Average Temperature Variation.....	99

Figure 0.2 Global MAT persistence model predictions	99
Figure 0.3 Ethiopia MAT persistence model predictions.....	100
Figure 0.4 Kenya MAT persistence model predictions.....	100
Figure 0.5 Uganda MAT persistence model predictions.....	101
Figure 0.6 Test of Stationarity for Ethiopia MAT	101
Figure 0.7 Test od Stationarity for Kenya MAT	102
Figure 0.8 Test of Stationarity for Uganda MAT	102
Figure 0.9 Seasonal decomposition of Global MAT.....	103
Figure 0.10 Seasonal decomposition of Ethiopia MAT	103
Figure 0.11 Seasonal decomposition of Kenya MAT	104
Figure 0.12 Seasonal decomposition of Uganda MAT	104
Figure 0.13 De-seasoned Global MAT	105
Figure 0.14 De-seasoned Ethiopia MAT.....	105
Figure 0.15 De-seasoned Kenya MAT.....	106
Figure 0.16 De-seasoned Uganda MAT.....	106
Figure 0.17 Bar-Chart of SARIMA models in validation and test set of Global MAT	107
Figure 0.18 Bar-Chart of SARIMA models in validation and test set of Ethiopia MAT.....	107
Figure 0.19 Bar-Chart of SARIMA models in validation and test set of Kenya MAT.....	108
Figure 0.20 Bar-Chart of SARIMA models in validation and test set of Uganda MAT.....	108

Abstracts

Global warming has garnered significant attention in recent years due to its profound implications for the environment, economy, and society. This study aims to model and forecast the monthly average temperature globally and in East Africa using time series analysis. Secondary data from the Berkeley method, spanning from 1850 to the present, were utilized. The average temperatures recorded were 14.103°C globally, 23.025°C in Ethiopia, 23.979°C in Kenya, and 23.019°C in Uganda. After applying the first seasonal differencing to achieve stationarity, statistical models such as SARIMA were employed. Based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), SARIMA(4,0,0)(0,1,1,12) for global temperatures, SARIMA(2,0,2)(3,1,1,12) for Ethiopia, SARIMA(2,0,3)(0,1,1) for Kenya, and SARIMA(1,0,4)(0,1,1,12) for Uganda were identified as the best-fit models. Forecasts for the next two decades indicate an increasing temperature trend across all regions. Model performance was evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), all of which indicated good forecasting accuracy. To capture temperature volatility, ARCH/GARCH models, including ARCH, GARCH, and EGARCH, were applied. The EGARCH models were found to be most effective, with EGARCH(4,1) for global temperatures, EGARCH(3,1) for Ethiopia, EGARCH(3,2) for Kenya, and EGARCH(4,1) for Uganda showing superior performance based on lower AIC and BIC values and higher log-likelihood values. The findings have significant implications for climate adaptation planning, particularly in regions like East Africa that are highly susceptible to climate change impacts. By providing reliable forecasting tools, this study supports efforts in agricultural planning, water resource management, energy demand forecasting, public health, and environmental impact assessment.

Keywords: Temperature, Time Series Analysis, SARIMA, East Africa

1. Introduction

1.1 Background of the Study

Human-induced climate change is causing dangerous and widespread destruction of nature, affecting the lives of billions of people around the world, despite efforts to reduce the risks. Scientists said in the latest Intergovernmental Panel on Climate Change (IPCC) report that the people and ecosystems least able to cope will be hit hardest (IPCC, 2023).

Since the development of the industrial revolution, human productive activities have increased rapidly, and the concentration of carbon dioxide and other greenhouse gases has increased rapidly. Earth's surface temperature is rising because so much heat is trapped in the atmosphere. Global warming will cause a series of extreme weather events that can not only destroy natural ecosystems but also seriously threaten sustainable economic development, normal production, and human life. Therefore, global warming is a widespread concern of the international community (Cheng 2023).

Climate change can have both natural and anthropogenic causes. In the past, temperature fluctuations of several degrees over decades or centuries were always caused by nature. Climate change is primarily caused by changes in greenhouse gas concentrations and solar radiation. However, global warming in recent decades has been mainly caused by humans. Typically, it takes thousands of years for the Earth's temperature to rise by 2 degrees. It took humanity less than a century to achieve such an increase. Even more rapid warming has been observed in the past two decades alone (WorldData.info 2023).

Global warming is a very important issue and has received a lot of attention in recent years, from environmental engineers to scientists and economists. If global temperatures continue to rise, it will have a major impact on the global economy. Furthermore, due to its economic implications, global warming is no longer an obscure technical problem, but is becoming a widespread societal and international regulatory concern (Kellogg 1987). Surface temperature is one of the most important environmental factors. Modeling surface temperature variations and producing reliable forecasts is the basis for sound environmental policy (Romilly 2005).

The temperatures we experience locally and over short periods of time can vary significantly due to predictable cyclical events (day and night, summer and winter) and unpredictable wind and precipitation patterns. However, Earth's temperature is determined primarily by how much energy

the planet receives from the sun and how much energy it radiates into space. Although the energy from the sun varies little from year to year, the amount of energy released by the Earth is closely related to the chemical composition of the atmosphere, especially the amount of heat-trapping greenhouse gases (Earth Observatory, 2022).

The world will inevitably face numerous climate changes over the next 20 years, with global warming reaching 1.5°C (2.7°F). Exceeding this level of warming, even temporarily, will lead to further severe impacts, some of which will be irreversible. Risks to society, including infrastructure and low-lying coastal communities, will increase (IPCC, 2023).

Africa has one of the lowest rates of emitting greenhouse gases that cause climate change, key developing sectors are already suffering from severe impacts such as biodiversity loss, water scarcity, reduced food production, and losses in food production. Suffering widespread loss and damage caused by anthropogenic climate change. Life expectancies will increase and economic growth will slow (high confidence). Limiting global warming to 1.5°C is expected to significantly reduce damage to Africa's economy, agriculture, human health, and ecosystems compared to higher levels of global warming (high confidence) (IPCC, 2023). Africa contributes insignificantly to climate change, with only about 2-3% of global emissions, it still stands out as the world's most vulnerable region. This vulnerability is due to the continent's low levels of socioeconomic growth (UNEP 2022).

The State of the Climate in Africa 2019 report highlights increasing climate change threats for human health, food and water security, and socio-economic development in Africa. It emphasizes the need for accurate data for adaptation planning and the importance of climate information in resilience building. The report also discusses rising temperatures, sea level rise, coastal erosion, extreme events, food security impacts, and health impacts in Africa (UNFCCC 2020).

The African Center for Climate Policy predicts that the gross domestic product of five African subregions will decline significantly as a result of rising global temperatures. In scenarios where global temperatures increase by 1°C to 4°C compared to pre-industrial levels, the continent's total GDP is expected to decline by 2.25% to 12.12%. West, Central and East Africa shows a greater negative impact than South Africa and North Africa (UNFCCC 2020). According to UN environment programme (2022) reports parts of East Africa, particularly Somalia, Djibouti,

Ethiopia and Kenya, are experiencing the driest conditions and hottest temperatures since satellite records began.

1.2 Statement of the problem

Climate change is a serious and growing threat to our well-being and the health of our planet. Our actions today will determine how people adapt and respond to growing climate risks. It will influence how nature responds (IPCC, 2023). The greenhouse effect today has an increasingly significant impact on our daily lives and the future of humanity. As pollution increased, temperatures around the world dramatically increase, causing disasters such as droughts, floods, and melting glaciers (Qiao 2023).

The Earth's temperature is not static, and understanding its fluctuations is crucial for a plethora of fields, from agriculture and resource management to human health and societal resilience. Decades of research have established the reality of global warming, by highlighting a rising average temperature (Chen, Jiang, and Li 2023; Naiqian, Yuxin, and Xuelian 2019)

Most empirical modeling of climate change time series uses the traditional method of autoregressive moving average (ARMA) modeling proposed by the Box Jenkins (1971) methodology to model and predict the mean behavior of time series has been focused on. Traditional models are based on the assumption that the series is stationary, that is, with a zero mean and either constant or seasonal variance, but do not take into account second moments or conditional variance. The use of linear time-series models has been broadly examined by various workers for temperature such as Cheng (2023) uses ARIMA for global temperature, Chen, Jiang, and Li (2023) also apply an ARIMA model on the global temperature and get ARIMA(2,1,1) is a best model, SARIMA modelling for precipitation (ARIMA(0,1,1)(0,1,1)₁₂) and temperature (ARIMA(0,1,0)(0,1,1)₁₂) are modeled as a best models in India by Dimri, Ahmad, and Sharif (2020), other study by Ye et al. (2013) specifies DSC(1, 1) + SARIMA(3, 1, 1)(0, 1, 1)₁₂ as best model for global monthly absolute temperature.

However, the above studies do not take into account higher-order moments such as dispersion, which are important for improving forecasting of complex climate change processes such as temperature. Higher moment orders of temperature (such as dispersion) can be successfully modeled using the ARCH/GARCH technique. However, the ARCH/GARCH model has not received much attention in the fields of environment. The literature shows that very few studies

have applied the hybrid autoregressive integrated moving average (ARIMA) method with GARCH to model temperature, such as (Romilly 2005).

Generally, motivation for this study stems from the limited research focus on East Africa, despite its vulnerability to climate variability, necessitating a dedicated investigation into temperature in this region. Additionally, there is a recognized need for long-term predictions, extending beyond the short-term scope of existing studies, to facilitate informed decision-making and planning in various sectors. The study was adopted an integrated approach, analyzing temperature globally and in East Africa, aiming to unveil common patterns and regional differences, thereby enhancing our understanding of temperature variability at different scales. Through the development of advanced time series methodologies, potentially incorporating GARCH models or other suitable approaches, the research seeks to capture the temporal dynamics of temperature, thereby enabling forecasting and analysis script.

Through this comprehensive approach, the research aims to contribute valuable insights to the field of climate science, bridging the existing gaps in modeling precision and providing decision-makers with more reliable tools for climate adaptation planning in both global and East African contexts. Thereby, this study as tried to respond the following fundamental research questions.

- What are the historical trends of monthly average temperature in global and east Africa?
- What is the appropriate volatility models to fit the average monthly temperature both in global and east Africa?
- What is the predicted value of average monthly temperature in the next decades both in global and east Africa?
- What common patterns and distinctive features of average monthly temperature globally and in East Africa?

1.3 Objectives of the Study

1.3.1 General Objective

The general objective of this study is to apply time series analysis techniques for modeling and predicting the monthly average temperature in both the Global and East Africa regions.

1.3.2 Specific Objective

The Specific objectives of this study is:

- To examine historical trends of average monthly temperature in Global and East Africa.
- To model the average monthly temperature in Global and East Africa and its volatility using the family of GARCH models.
- To predict average monthly temperature both in global and east Africa based on the developed models one-decade ahead.
- To investigate the common and distinct features of global and East Africa average monthly temperature.

1.4 Significance of the Study

Understanding and accurately predicting temperature dynamics is important for various sectors such as agriculture, water resources management, energy planning, and public health. This study on applying time series analysis for modeling and predicting monthly average temperature in global and East Africa holds significant importance due to the following reasons:

I. Climate Change Monitoring:

By analyzing historical temperature data and developing predictive models, we can contribute to the ongoing efforts of monitoring and understanding climate change patterns. This is crucial for informed decision making and policy formulation as it helps us assess long-term trends and variations in temperature.

II. Agricultural Planning:

Accurate temperature predictions are essential for effective agricultural planning and management. Farmers can optimize their crop selection, planting schedules, and irrigation strategies based on reliable temperature forecasts. Hence, this study provides valuable insights to agricultural stakeholders in mitigating risks associated with temperature fluctuations and improving crop yields.

III. Water Resources Management:

Hydrological cycle is the circulation of water among the land, atmosphere and oceans, and water availability depends on those temperature variations in these systems. For appropriate water management, there is need to know how temperature patterns that are relevant impact water

resources. This study will enable water resource manager's plan for droughts, floods and other related challenges. Additionally, it will contribute to sustainable water use and conservation.

IV. Energy Demand Forecasting:

The use of energy for heating and cooling purposes depends on the temperature variations around us. By predicting future climatic temperatures, planners in the energy sector can anticipate future energy needs thus optimizing generation and grid distribution for this purpose. The findings provide a basis for better planning and management of energy.

V. Public Health and Safety:

Public health conditions are greatly influenced by extreme temperature events such as heat waves or cold spells. Through accurate predictions on temperatures, health authorities as well as emergency response teams can be warned about likely risks due to extreme temperatures hence reducing vulnerability to them.

VI. Environmental Impact Assessment:

Temperature changes impact ecosystems and biodiversity. This investigation provides information on temperature variations that will help scientists and politicians in creating strategies for maintaining the diversity of life and mitigating ecological damage due to global warming.

VII. Regional Focus on East Africa:

It is against this backdrop that this study specifically focuses on East Africa because it is highly susceptible to climate change, with significant reliance on climate sensitive sectors such as agriculture and tourism. The findings of this study can directly inform regional policies and adaptation strategies to cope with changing temperature patterns.

In general, this study related to time series analysis of modeling and predicting monthly average temperatures has substantial implications for climate change monitoring, agricultural planning, water resource management, energy demand forecast, public health and safety, environmental impact assessment, regional development among others. It is through such studies that informed choices will be made leading to sustainable practices across different fields hence resilience building in the face of global warming challenges.

2. Literature Review

2.1 Overview of Time Series Analysis

A time series is a series of individual observations made over a specific period of time. In general, time series are considered to contain time information that cannot be reproduced (George E.P. Box, Jenkins, and Reinsel 2013). Data obtained from observations collected continuously over long periods of time are very common. In business, observe weekly interest rates, daily closing prices, monthly His price index, annual sales, etc. Meteorology observes daily maximum and minimum temperatures, annual precipitation and drought index, and hourly wind speed. Agriculture collects annual statistics on crop and livestock production, soil erosion, and export sales. In the life sciences, the electrical activity of the heart is observed at millisecond intervals. Ecology records the abundance of animal species. The list of fields in which time series are studied is almost endless. There are two general purposes for time series analysis. It is the understanding or modeling of the stochastic mechanisms that produce an observed sequence and the predicting or predicting the future value of a sequence based on the history of that sequence. There are probably other related series and factors as well (J.D. Cryer 1985).

Data is the foundation of time series analysis, and innovations in time series analysis are closely related to data collection, recording, and visualization. Around the 4th century BC, scientists used instruments to collect weather-related data over short periods of time, such as hourly or daily. Astronomers are a group of people who rely heavily on plotting objects, orbits, and measurements over time and are masters of time series. In particular, chronological collections of sunspots date back to 800 BC (Nielsen 2019).

A formal numerical approach to time series analysis was created by (Yule 1927), to consider the time series conduct of sunspots, as an application of autoregressive (AR) models. Wold (1938) created Wold decomposition and the most result was that a stationary handle can be expressed as a sum of a deterministic term and a stochastic term, where the last-mentioned term can be expressed as an infinite moving average, that's autoregressive moving average (ARMA) model. However, there was no likelihood function that empowers researchers to determine maximum likelihood estimation for the parameters. Until 1970, the ARMA models got to be popular as Box and Jenkins (1970) proposed the Box-Jenkins method for choosing the order of the models and evaluating the coefficients. The model selection criteria have been created (e.g. Akaike

1974, Hannan 1980), as well as pattern identification strategies (e.g. Gray, Kelley, and McIntire 1978) and an expanded sample auto-correlation function for order determination for both stationary and non-stationary series (Tsay and Tiao 1984)

2.2 Overview of Climate Change

Climate is sometimes confused with weather. But climate is different from weather because it is measured over a long period of time, while weather can change from day to day or from year to year. A region's climate includes average seasonal temperatures and precipitation as well as wind patterns. Different places have different climates. For example, deserts are called arid climates because little water falls as rain or snow during the year. Other climate types include hot and humid tropical climates and temperate climates with hot summers and cooler winters (National Geographic Society 2023).

Climate change is a long-term change in the typical temperature and weather conditions in a location. Climate change can affect a specific location or the entire planet. Climate change can make weather conditions less predictable. These unexpected weather conditions can make it difficult to maintain and grow crops in areas that depend on agriculture because of heat and rainfall. Expectations can no longer be relied upon. Climate change is also linked to other damaging weather events, such as more frequent and intense hurricanes, floods, downpours and winter storms (National Geographic Society 2023).

2.2.1 Global Climate Change Impacts

2.2.2 Agriculture

The impact of global climate change on agriculture is a critical concern due to its potential to disrupt food production and food security. Climate change affects agricultural productivity through changes in weather conditions, which directly influence crop growth, yield, and hydrological balances (Knox et al. 2012). Studies have shown that climate change has led to a decrease in global agricultural productivity growth, making agriculture more vulnerable to ongoing climate change (Ortiz-Bobea et al. 2021). The effects of climate change on agricultural systems are not limited to specific regions but have global implications, affecting relative competitiveness, international trade, global supply, and prices (J. S. Baker et al. 2018). Furthermore, climate change impacts on agriculture are projected to have negative effects on global agricultural production, with potential decreases ranging from 0.5% to 2.5% (Van Meijl et al. 2018).

Smallholder production systems in regions like South Asia are particularly vulnerable to the impacts of climate change on agriculture. Research has highlighted the need for adaptation options in these systems to mitigate the effects of climate change on crop production (Aryal et al. 2020). Additionally, the adverse effects of climate change on agriculture have been observed in various regions, such as Ethiopia, where changes in climate have led to decreased crop yields and altered livestock feed availability (Singh 2019).

Farmers' perceptions and indigenous strategies for climate change adaptation in agriculture are also crucial factors to consider. Studies have shown that farmers are aware of the effects of climate change on crop production and have been making adjustments to adapt their agricultural practices in response to these changes (Abid et al. 2015). Furthermore, access to agricultural information is essential for farmers to make informed decisions for adapting to climate change and mitigating its impacts on agriculture (Mwalukasa 2013).

The impact of global climate change on agriculture is multifaceted, affecting various aspects of agricultural systems, including productivity, adaptation strategies, and farmers' perceptions. Addressing these challenges requires a comprehensive understanding of the biophysical and socioeconomic impacts of climate change on agriculture, as well as the development and implementation of effective adaptation strategies to ensure global food security.

2.2.2.1 Health

The impact of global climate change on human health is a growing concern, with potential effects including higher rates of respiratory and heat-related illness, increased prevalence of vector-borne and waterborne diseases, food and water insecurity, malnutrition, and adverse impacts on mental health (Crowley 2016). Climate change policies and strategies have the potential to have significant and immediate beneficial effects on public health and health equity through improving social determinants of health across many sectors (Gould and Rudolph 2015). The impacts of climate change on health and wellbeing may occur directly through extreme weather events and indirectly through impacts on food security, air and water quality, and other environmental amenities (Boylan et al. 2018). There is growing evidence that climate change has caused a range of direct and indirect effects on global and human health in the past few decades (Shi 2018). Future climate change is projected to exacerbate various current health problems, including the increase in global temperature, which will likely have negative repercussions for public health, including

exacerbations of chronic diseases, mental illness, food shortages, malnutrition, human displacement, and increased transmission and susceptibility to infectious diseases (Ebi et al. 2009; Miller 2019). Furthermore, climate change has a direct influence on human health, especially where infectious diseases are involved (Al-Ghamdi, Ali khan, and Mahyoub 2009; Wang et al. 2015). Recent climate changes have had negative impacts on human health, and climate changes are expected to contribute to significant increases in morbidity and mortality associated with a range of health problems (Watkiss and Ebi 2022). Vector-borne diseases are likely to be affected by global warming, although other factors may play an important role, as evidenced by the significant decline in malaria incidence over the decades. recently in areas where climate change is having a profound and significant impact (Flahault, de Castaneda, and Bolon 2016).

2.2.3 Climate Change Impacts for the East Africa region

2.2.3.1 Agriculture and Food Security

East Africa has seen impressive economic growth over the past decade, but structural changes to the economy have been slower. Agricultural production still accounts for about 25% of regional GDP (more than 30% in Ethiopia, Somalia, Kenya and Burundi), and most agricultural systems are highly vulnerable to climate shocks, because agriculture relies rain dominates agricultural activities (FAO 2018).

As temperatures rise, heat stress will put pressure on all livelihood systems in the region. Rising temperatures have a dual impact on agriculture. First, high temperatures themselves can influence plants to reach their maximum temperatures, reducing yields. Second, high temperatures increase transpiration and reduce soil moisture, increasing the amount of water plants need (Richardson et al. 2022). The effects of high temperatures follow a similar pattern for livestock. Heat stress directly increases animal mortality, reduces milk production and reproduction, and increases water requirements. With lower seasonal yields, households will be forced to buy more food to meet food security needs. Because these shocks tend to affect agriculture across the region, price fluctuations are likely to force households to spend a larger portion of their income to meet their food needs.

In cotton, high temperatures reduce boll maturity, reduce boll size, reduce the number of seeds per boll, reduce boll length and reduce cotton yield. When temperatures exceed 35°C throughout the year, livelihoods based on mixed maize will be less productive than livelihoods based on cereal

root crops, as sorghum and millet are more resilient than before. projected temperature increases and drought periods over maize (Abera et al. 2018; Sultan et al. 2013). Coffee production is an important cash crop in the region for household income. The optimal growing temperature ranges from 18 to 23°C for Arabica varieties and 22 to 26°C for Robusta varieties (Adhikari, Nejadhashemi, and Woznicki 2015). Coffee plants are very sensitive to climate change, which can reduce the quantity and quality of coffee beans (Adane and Bewket 2021). Currently, the majority of coffee growing areas in Ethiopia are located in the southwest and southeast, with small production in the north. A study examining the relationship between temperature and coffee production in Ethiopia found that current coffee growing areas will no longer be suitable, but there will likely be a significant increase in suitable area. For coffee production in higher altitude areas, especially in southwestern Ethiopia (Moat et al. 2017). Banana, the most important food crop in Uganda and Rwanda and a key crop in highland perennial agricultural systems, is sensitive to prolonged exposure to extreme temperatures (above 35°C) and low soil moisture, but are resistant to increasing temperature-related stresses until 2050 (Thornton and Cramer 2012).

Pastoral livelihoods face heat stress related to climate change, with animal reproduction rates and milk production slowing or ceasing altogether. completely under higher temperatures. Livestock raised by pastoralists or agricultural producers will be affected by heat stress, especially livestock (USAID, 2019). Higher temperatures are also likely to reduce the reproductive rate of goats, although they are relatively better physiologically and behaviorally adapted to heat stress than sheep and cattle (Sejian et al. 2021). At the same time, rising temperatures and extended periods of drought can negatively impact pasture productivity and livestock drinking water supplies.

2.2.3.2 Water Resources and Water-Dependent Services

The East African region has no shortage of water, at least in terms of available water resources. This area averages about 1,400 $m^3/capita$ in 11 countries, with measurements that do not include large volumes of stored groundwater held in scarcity (Cuthbert et al. 2019; Taylor et al. 2013). With the exception of Sudan, all countries fall into the “low or no water stress” category used to monitor the Sustainable Development Goals. However, measurements at regional and national scales hide major problems of distribution and variability, as well as long-term challenges related to mobilizing and managing water for lives and livelihoods – challenges that could be exacerbated by climate change.

The combination of changes in river flows and rising temperatures is expected to have major negative impacts on freshwater ecosystems and water quality (Döll et al. 2014). Higher water temperatures promote algal blooms and increase risks from cyanotoxins and naturally occurring organic matter in water sources. Increased runoff results in more fertilizer, animal waste and particulate matter. Conversely, low flows reduce the river's ability to dilute, degrade, and remove pollution and sediment.

2.2.3.3 Health

Climate model projections show high confidence in the predicted 1-4°C average annual temperature increase in the region by the 2050s compared to the 1981-2010 baseline period. Climate projections show a strong and consistent signal of further warming, with the largest temperature increases occurring during the hottest times of the year. High temperatures have been a serious problem in the region, especially during the summer months, with 4,444 daily maximum temperatures exceeding 40°C between May and September. Human physiological limits in tolerating extreme temperatures, heat stress, and even the basic habitability of certain areas are also a factor (Raymond, Matthews, and Horton 2020).

Some of the most devastating impacts of climate change will affect human health, and many of these health burdens fall on children (UNICEF 2014). In the arid, drought-prone lands of East Africa, malnutrition poses serious health risks to children. Higher ambient temperatures in sub-Saharan Africa are associated with poorer nutritional status in the short and long term; In the short term, high temperatures can reduce a child's ability to absorb nutrients due to loss of appetite, increased dehydration or diarrhea, which can contribute to weight loss. In the long term, higher temperatures during the growing season can reduce total crop yields. A study that isolated these effects found significant effects of temperature on the weight-to-height ratio of rural children during the growing season, especially in corn and wheat production areas in SSA (R. E. Baker and Anttila-Hughes 2020).

2.3 Application of Time Series Analysis in Climate Modelling

Time series analysis is a powerful tool in climate modeling. It can be used to investigate the dynamic structure of climate by studying changes in precipitation and temperature over time (Dimri, Ahmad, and Sharif 2020). One such study used seasonal ARIMA approach to analyze monthly mean minimum and maximum temperatures and precipitation for the Bhagirathi river

basin in India. The study found that the most appropriate model for time series analysis of precipitation data was SARIMA (0,1,1) (0,1,1) 12 (with constant) and of temperature data was SARIMA (0,1,0) (0,1,1) 12 (with constant). The information of pattern and trends can assist as a prediction tool for development of better water management practices in the area.

Romilly (2005) used the univariate GMT time series to develop a forecast model for the short term (5 to 10 years). Statistical techniques used include seasonal and non-seasonal unit root tests, as well as ARIMA and GARCH models. The data set includes a monthly 1602 observations from 1870:1 to 2003:6 of global mean near-surface temperature, expressed as the difference (or anomaly) between the actual monthly temperature and the average temperature from 1961 to 1990. Temperature is measured in degrees Celsius and the average temperature from 1961 to 1990 is 14.08 °C. The SARMA–GARCH (1, 1)-M model gives the best RMSE = 0.111.

The monthly absolute surface temperature record modeled by Ye et al. (2013) used a deterministic and stochastic coupled (DSC) approach, in which deterministic processes are consists the variation pattern of the global trend and periodic fluctuations of the thermal signal degrees, involving polynomial functions and stochastic processes are used to consider for any remaining trends in the temperature signal, involving SARIMA models. Several DSC models have been built and tested. The best RMSE obtained was 0.1112. This $DSC (1, 1) + (3, 1, 1) \times (0, 1, 1)$ model is known to have the best overall performance, combining both in-sample and out-of-sample consistent forecasts. Monthly global temperature projections using the combined deterministic and stochastic model suggest that global temperatures will likely continue to increase at a rate twice the average over the past 150 years for the 2nd decade of the 21st century.

The study by Naiqian, Yuxin, and Xuelian (2019) examines the global temperature using ARMA model by aiming to show modeling techniques from a viewpoint of empirical time series analysis and forecast comparison. They use the land-surface temperature monthly mean series based on GISS, NASA dataset. The sample period is from January 1880 to December 2018. They develop different ARIMA models for monthly average temperature and select ARIMA(2,1,1) as a best fit for global land surface air temperature by using standard variance of each coefficient.

3. Methods and Materials

3.1 Study Area

The study area for this thesis encompasses a comprehensive exploration of the global climate system with a specific focus on East African countries. The research delves into the application of sophisticated time series modeling techniques to predict the average monthly temperature patterns. The global scope ensures a thorough understanding of overarching climatic trends, while the targeted East African countries allows for a nuanced analysis of regional variations.

East Africa, also known as the Horn of Africa, includes 11 countries (Sudan, South Sudan, Eritrea, Djibouti, Ethiopia, Somalia, Kenya, Uganda, Rwanda, Burundi, and Tanzania). In the tropical belt, East Africa stands out as a relatively dry region, despite its equatorial location. Between 12°S and 12°N, this is in fact the driest land (Yang et al. 2015). Green and Trewartha (1963) highlighted the region's "widespread rainfall deficit" as "the most dramatic climatic anomaly in all of Africa".

East Africa includes the hottest place on earth known as Dallol in the Afar Depression (130 m below sea level), with an average annual temperature of 34.6°C (1960-66; Pedgley 1967)). The lowest average annual temperature is -7.1°C at the summit of Mount Kilimanjaro (Thompson et al. 2002). In most of East Africa, frosts occur sporadically at altitudes of about 2,200-2,400 m, although above 3,500 m they recur almost daily.

Based on the availability of time and data, three East African countries were selected for this study: Ethiopia, Kenya, and Uganda.

3.2 Data Source

The monthly average temperature data for the global scale, as well as for Ethiopia, Kenya, and Uganda, were sourced from Berkeley¹ Earth, an organization renowned for its commitment to environmental data science and analysis. They provided monthly global average temperature records spanning from the early 18th century to the present. For this study, the specific periods covered are 1850 to 2023 for global data, 1880 to 2020 for Ethiopia, and 1892 to 2020 for both Kenya and Uganda. The raw data, initially presented as anomalies, was preprocessed according to the website's data description for different months. Following this preprocessing, the monthly

¹ <https://berkeleyearth.org/data/>

average data was converted to actual temperature values by referencing the absolute temperature records from January 1951 to December 1980.

3.3 Time Series Analysis

A time series is a sequence of data points collected or recorded over time. The data in a time series is typically ordered chronologically (Chatfield 2005). Time series data often exhibits certain patterns and structures, which can be decomposed into different components. The main components of a time series (Box and Jenkins 1970) :

1. **Trend:** The trend component represents the long-term movement or directionality in the time series. It indicates whether the values are increasing, decreasing, or remaining relatively constant over time. Trends can be linear or nonlinear.
2. **Seasonality:** Seasonality refers to the repeating patterns or fluctuations in the data that occur at regular intervals. These intervals might correspond to days, months, seasons, or other periodic time frames. Seasonal patterns are often influenced by factors like weather, holidays, or business cycles.
3. **Cycle:** The cycle component represents longer-term undulating patterns that are not strictly tied to fixed periods like seasonality. Cycles are typically associated with economic or business cycles and have a duration longer than the seasonality period.
4. **Irregular or Residual:** The irregular component, also known as the residual or noise, represents the random fluctuations or variability in the time series that cannot be attributed to the trend, seasonality, or cycle. It includes the effects of unpredictable events, measurement errors, and other unexplained factors.

Time series analysis is a fascinating approach to unraveling the hidden insights within a sequence of data points gathered over a specific span of time. Unlike random data recording, analysts meticulously capture data points at regular intervals throughout a defined period. However, it's important to note that time series analysis goes far beyond the mere act of data collection.

What sets time series data apart from other types of data analysis is its ability to reveal how variables evolve over time. By considering time as a critical factor, we gain a profound understanding of how the data adjusts itself across different points as well as the final outcomes.

Time series analysis offers a remarkable additional layer of information by introducing a well-defined order of dependencies between various data points.

Conducting a thorough time series analysis requires a substantial number of data points to guarantee consistency and reliability. An extensive dataset is crucial as it ensures the presence of a representative sample size, enabling us to cut through the noise present in the data. Moreover, a sizable dataset helps identify trends or patterns that emerge, ensuring they are not mere outliers and considering any seasonal variances.

3.4 Decomposition of Series

Decomposition is an important technique for time series analysis because it gives us brief information of the data and thus helps us to better analyse the data (Chourasia 2020). Decomposition is a process that can split the data into three components: trend, seasonal variety, and random error (irregular). Based on the dominance of trend and seasonal effects in time series models, two models that are based on these components are additive decomposition show (Equation 3.1) and multiplicative decomposition model (Equation 3.2), defined below, respectively.

$$Y_t = T_t + S_t + R_t \quad 3.1$$

$$Y_t = T_t \times S_t \times R_t \quad 3.2$$

Where Y_t is a monthly average temperature, T_t is the trend, S_t is the seasonal effects, and R_t is an error term. The random error, R_t , is a sequence of correlated variables with mean zero in general. The main difference between these two models is, for the multiplicative decomposition model the seasonal effect tends to increase as the trend increases, whereas additive decomposition model does not.

3.5 Stationarity

When it comes to analyzing time series data, one crucial concept to understand is stationarity. Stationarity refers to the statistical properties of a time series dataset remaining constant over time. In simpler terms, a time series is considered stationary if its statistical characteristics, such as mean, variance, and autocorrelation, do not show significant changes over different time periods (Hyndman and Athanasopoulos 2021).

The importance of stationarity in time series analysis cannot be overstated. Many time series models and techniques assume or work best with stationary data. Therefore, ensuring stationarity is crucial for accurate analysis and reliable predictions (Hamilton 1993). There are two main types of stationarity: strict stationarity and weak stationarity.

A time series is said to be strictly stationary if the joint distribution of any set of time points remains the same for all subsets of those time points. Specifically, a process $\{Y_t\}$ a monthly average temperature is strictly stationary if the joint distribution of $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ is the same as the joint distribution of $Y_{t_1-k}, Y_{t_2-k}, \dots, Y_{t_n-k}$ for all choices of time points $t_1, t_2, \dots, t_n$ and at all choices of lag k . mathematically,

$$F(Y_{t_1-k}, Y_{t_2-k}, \dots, Y_{t_n-k}) = F(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n})$$

In particular, $E[Y_t]$ is constant and the auto-covariance $cov(Y_t, Y_s)$ only depends on lag $k = |t - s|$. A time series variable $\{Y_t\}$ is weakly stationary if

- ❖ $E(Y_t^2) < \infty$ for all t
- ❖ $E(Y_t) = \mu$ for all t
- ❖ $cov(Y_t, Y_s)$ only depends on lag $t - s$ for all t and s

The condition $E(Y_t^2) < \infty$ for all t is the condition for variance to exist. Weak stationarity does not imply strict stationarity, and strict stationarity does not necessarily imply weak stationarity. As the condition $E(Y_t^2) < \infty$ for all t is not assumed in the definition of strict stationarity. That is, if a monthly average temperature $\{Y_t\}$ is strictly stationary and has finite second moment ($E(Y_t^2) < \infty$ for all t), then $\{Y_t\}$ is a weakly stationary process (Chatfield 2005; Hamilton 1993).

3.5.1 Examining Stationary of Time Series Data

Assessing the stationarity of a time series is essential in the field of data analysis. Stationarity refers to the property of a series to have constant statistical properties over time, such as mean and variance. There are various methods to test for stationarity.

A. Visual Inspection

Visual inspections offer a simple yet effective way to identify the stationarity of time series data. By plotting the data and visually inspecting it, we can uncover trends, seasonality, and underlying patterns.

- ❖ A time series plot is a line graph that illustrates how a variable change over time.
- ❖ For a stationary time series, the plot will display a flat line with no discernible pattern, indicating a constant mean and variance.

B. Autocorrelation and Partial Autocorrelation Plots

Autocorrelation and partial autocorrelation plots are commonly used tools to assess the stationarity of time series data. These plots allow us to measure the correlation between a variable and its past values, providing valuable insights into the series' behavior.

- ❖ Autocorrelation measures the correlation between a variable and its lagged values.
- ❖ Partial autocorrelation measures the correlation between a variable and its past values after controlling for intermediate values.

In a stationary time series, the correlation between a variable and its past values should decay rapidly as the lag increases. Conversely, in a non-stationary time series, the correlation may persist over time, resulting in a slow decay of the correlation coefficients.

To determine the stationarity of a time series using autocorrelation and partial autocorrelation plots:

1. Look for a rapid decay in the magnitude of the correlation coefficients in the autocorrelation function (ACF) plot.
2. In the partial autocorrelation function (PACF) plot, observe a sharp drop-off after a certain lag.

If these conditions are met, it suggests that the time series is stationary.

C. The Unit Root Test

The presence of a unit root characterizes a non-stationary time series, where the data-generating process has a root close to one. It indicates a trend component that follows a stochastic process rather than a deterministic one. To identify the presence of a unit root, we employ unit root tests,

such as the Augmented Dickey-Fuller (ADF) (Dickey and Fuller 1979) and Phillips-Perron (PP) (Phillips and Xiao 1998) tests.

- ❖ The ADF test estimates an autoregressive model and tests whether the coefficient on the lagged dependent variable is statistically significant.
- ❖ The PP test, considered more robust in certain cases, also examines the presence of a unit root in a time series.

3.6 Unit Root Tests

Unit root tests can be used to determine whether a non-stationary process should be detrended or differencing to render the stationarity. Unit root tests are used for checking the presence of stochastic trend in univariate time series. If the series show evidence to reject the hypothesis of presence of unit root, then the two series is stationary. Three-unit root tests are discussed, the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test. The ADF and the PP test a null hypothesis of a unit root, which is non-stationary in variance. It is a good idea to use unit root tests with different null hypotheses, as we cannot accept the null hypothesis even, we have no evidence to reject the null hypothesis when p-value is greater than the expected significance level.

3.6.1 Augmented Dickey-Fuller (ADF) test

Dickey and Fuller (1979) invented the Dickey-Fuller (DF) test, to test the unit root against the null hypothesis of the presence of a unit root($\rho - 1 = 0$), in the series $y_t - y_{t-1} = (\rho - 1)y_{t-1} + u_t$. However, since it only considers $AR(1)$ models, it has extremely low power at the test for unit root. Therefore, an Augmented Dickey-Fuller (ADF) test was proposed, which allows for a more general procedure, i.e. a higher order autoregressive procedure(Dickey and Fuller 1979). The ADF test model is as follows:

$$\Delta y_t = \beta y_{t-1} + \sum_{k=1}^K \alpha_k \Delta y_{t-k} + w_t \quad 3.3$$

where $\Delta = 1 - B$ is the first difference operator, K is the lag order of the autoregressive process, and β and α_k are constants for $k = 1, \dots, K$. The summation term in the equation above is to correct for the presence of serial correlation.

The null hypothesis of the ADF test is the presence of unit root, that is $H_0: \beta = 0$ against the alternative which is $H_a: \beta < 0$. We compare a t-statistic with the values of this Dickey-Fuller

distribution. If the t-statistic less than the critical value, then we have evidence to reject the null hypothesis of presence of unit root.

3.6.2 Phillips-Perron (PP) test

There is another unit root test based on the null hypothesis of the presence of a unit root, which is the Phillips-Perron (PP) test, developed by Phillips and Xiao (1998). The main difference between ADF and PP tests is how they handle serial correlations and error heteroskedasticity. Specifically, the ADF test uses parametric autoregression to estimate the ARMA error structure in the regression test, while the PP test uses a corrected t-test form to correct for the presence of serial correlation and heteroskedasticity in residual term (Zivot and Wang 2006). Model is:

$$y_t = c + \delta t + \alpha y_{t-1} + \epsilon_t \quad 3.4$$

Where y_t is monthly average temperature, δ and α is the coefficient of trend and autoregressive, and ϵ is the error term. The null hypothesis of this test is $\alpha = 1$, compared to the alternative hypothesis $\alpha < 1$. The advantage of the PP test over the ADF test is that the PP test is robust to general forms of heteroskedasticity in the error class and we do not need to specify a lag length for the regression test. If the t-statistic is less than the critical value then we have evidence to reject the null hypothesis of the presence of a unit root.

3.7 Transforming Nonstationary Data

Once we've identified that our data is nonstationary, we need to transform it in order to induce stationarity. A common method used for this purpose is differencing the data. Differencing involves subtracting the value at the previous time step from the current value. Differencing can help stabilize the mean of a time series by eliminating changes in the level of the time series and thus eliminating (or reducing) trends and seasonality (Hamilton 1993). Mathematically it is written as

$$\nabla Y_t = (1 - B)Y_t = Y_t - Y_{t-1} \quad 3.5$$

Where ∇ is a Differencing notation, B is Backward shift Operator, which is a very useful notation in time series analysis, especially the ARIMA model. A single backshift operator represents a time series by a single shift(lag), and Y_t a monthly average temperature which is our variable of interest.

In addition to differencing, seasonal differencing can also be performed. This involves subtracting the data from the previous season from the current data point. Seasonal differencing helps remove any seasonality in the data and further induces stationarity (Hyndman and Athanasopoulos 2021). The **seasonal difference** of period s for the series $\{Y_t\}$ is denoted $\nabla_s Y_t$ and is defined as

$$\nabla_s Y_t = (1 - B^s)Y_t = Y_t - Y_{t-s} \quad 3.6$$

Apart from differencing, there are other transformations that can be applied to make the data stationary. These include logarithmic transformation, power transformation, and smoothing techniques. Each of these methods serves its purpose in removing specific patterns from the data (Montgomery, Jennings, and Kulahci 2015).

Once we have successfully made our data stationary, we can move on to the next phase: fitting a model and making forecasts. With stationarity achieved, we can have confidence in the reliability of our models and the accuracy of our predictions.

3.8 Seasonal ARIMA (SARIMA) Process

Autoregressive Integrated Moving Average, or ARIMA, is a broadly utilized method for univariate time series data estimating. Even though the method can handle data with a trend, it does not support time series with a seasonal component. An extension to the Autoregressive Integrated Moving Average that supports the direct modeling of the seasonal component of the series is called Seasonal Autoregressive Integrated Moving Average (SARIMA) (Box and Jenkins 1970). A SARIMA model with non-seasonal terms of order (p, d, q) and seasonal terms of order (P, D, Q) is abbreviated as $SARIMA(p, d, q) \times (P, D, Q)_s$ model (Wei 2006) and Mathematically the model may be written as

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D Y_t = \theta(B)\Theta(B^s)\varepsilon_t \quad 3.7$$

where Φ, Θ denote polynomials in B^s of order P, Q , respectively.

3.9 Building ARIMA/SARIMA Models

There were some fundamental steps to fitting ARIMA/Seasonal ARIMA models to time series data. These steps include plotting the data, possibly transform the data, identifying the fitting

model, parameter estimation, model diagnosis and estimating. The first Box-Jenkins modeling method includes an iterative three-stage process of model choice, parameter estimation and diagnostic checking. But, further clarifications of the method by Briggs et al. (1999) often include a preliminary stage of data preparation and a final step of estimating and assessment of the forecast performance measure.

3.9.1 Model Identification

The primary step in ARIMA modeling is the identification of the appropriate model. ARIMA methods are mainly based on the value of the autocorrelation function (ACF) and the partial autocorrelation function (PACF), which are critical to check diagnostic of the series.

A comparison of the sample ACF and PACF to those of various hypothetical ARIMA forms may propose several plausible models. A common stationarity inducing transformation is to take logarithms and first difference of the series. Once we have achieved stationarity, the next step is identifying the p and q orders of the non-seasonal ARIMA model and P and Q order of SARIMA.

3.9.1.1 Autocorrelation Functions

Autocorrelation function is a very valuable in order to identify a time series model. It measures the direct dependence or the extent of the autocovariance of Y_t and Y_{t-k} to the variance of a dependent variable Y_t

$$ACF(k) = \rho_k = \frac{cov(Y_t, Y_{t-k})}{Var(Y_t)} \quad 3.8$$

The order of moving average was determined by the number of significant auto correlation.

3.9.1.2 Partial autocorrelation function (PACF)

PACF is the simple correlation between Y_t and Y_{t-k} minus the part explained by the intervening lags

$$\rho_k^* = corr[Y_t - E^*(Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}), Y_{t-k}] \quad 3.9$$

where $E^*(Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1})$ is the minimum mean-squared error predictor of Y_t by $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$.

When the stationarity condition of the monthly average temperature will be fulfilled, the order of the model which AR, MA, seasonal AR and MA terms were determined with the assistance of the ACF and the PACF plot of the stationary series (G. E. P. Box and Jenkins 1970). The ACF gives information about the inner correlation between observations in a time series at different distances apart, usually expressed as a function of the time slack between observations. These ACF and PACF plots propose the model should be built and the researcher will look at both seasonal and non-seasonal lags. Usually the ACF and the PACF has spikes at lag k and cuts off after lag k at the non-seasonal level. Moreover, the ACF and the PACF has spikes at lag ks and cuts off after lag ks at the seasonal(s) level. The number of significant spikes recommends the order of the model. Table 3.1 Behavior of ACF and PACF for Non-seasonal ARMA (p, q) and Table 3.2 below portrays the behavior of the ACF and PACF for both seasonal and the non-seasonal series.

Table 3.1 Behavior of ACF and PACF for Non-seasonal ARMA (p, q)

	AP(p)	MA(q)	ARMA(p, q)
ACF	Tails off at lag k	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off at lags k	Tails off

Table 3.2 Behavior of ACF and PACF for seasonal ARMA (P, Q)

	AP(P) s	MA(Q) s	ARMA(P, Q) s
ACF	Tails off at lag ks ,	Cuts off after lag Qs	Tails off at lag ks
PACF	Cuts off after lag Ps	Tails off at lags ks	Tails off at lag ks

where k is the lags, $k=1, 2, 3, \dots$ and s is a seasonal period

3.9.2 Model Selection

A model selection criterion is utilized to decide the approximating model from a set of evaluated models. In time series analysis there may be several satisfactory models that can fit a given data set. The well-known model selection criteria such as the Akaike data basis (AIC) proposed by (Akaike 1974) and the Schwartz Bayesian information criterion (SBIC) proposed by (Schwarz 1978) will be utilized to identify the ideal lag specification for the model.

3.9.2.1 Akaike's Information Criterion (AIC)

(Akaike 1974) proposed the goodness of fit for ARMA (p, q) by adjusting the error of the fit against the number of parameters within the model. It is characterized by:

$$AIC = -2\log L_k + 2k \quad 3.10$$

where p and q are the parameters of non-seasonal AR and MA respectively, L_k is the maximized log-likelihood and $k = 1 + p + q$, is the number of parameters within the model. For SARIMA models, the AIC is calculated similarly, but with extra parameters for the seasonal AR and MA components: where $k = 1 + p + q + P + Q$. The foremost suitable model is the one that minimizes the AIC, and there is no requirement for the models to be nested.

3.9.2.2 Bayesian Information Criterion (BIC)

Schwarz G (1978), introduced the BIC and is defined

$$BIC = -2\log L_k + k\log T \quad 3.11$$

where T is the number of observations in the time series. Diverse simulation studies have tended to identify that BIC does well at getting the proper order in large samples, while AIC tends to be predominant in smaller samples where the relative number of parameters is large (McQuarrie and Tsai 1998) as pointed out by Shumway and Stoffer (2016).

3.9.3 Parameter Estimation

Parameter estimation of ARIMA/SARIMA model will be accomplished by utilizing maximum likelihood estimation. At this stage precise estimates of the coefficients of the chosen model will be assessed. If the provisional model has significant parameters, whose values lie inside the bounds of stationary and inevitability and was not highly correlated, then the researcher will continue to the final stage, diagnostic checking. If not, the researcher will return to the identification stage and define an alternate model based on the information gained at the estimation stage.

3.9.4 Diagnostic checking

Once a model has been identified and the parameters estimated, diagnostic checks are then applied to the fitted model. These include tests of serial correlation and normality of the residuals

3.9.4.1 Ljung-Box test of serial correlation

Before an estimated model is utilized for statistical inference, the residuals must be examined for the presence of serial correlation. The Q-statistic is frequently utilized as a test of whether the series is white noise (Ljung and Box 1978). The Q-statistic at lag m could be a test statistic for the null hypothesis that there's no autocorrelation up to order m and is computed as:

$$Q = T(T + 2) \sum_{j=m}^m \frac{\hat{\rho}_j}{T - j} \quad 3.12$$

where $\hat{\rho}_j$ is the j^{th} lag autocorrelation and T is the number of observations. If the series is not based upon results of ARIMA/SARIMA estimation, then under the null hypothesis, Q is asymptotically distributed as chi-square with degrees of freedom equal to $(m-k)$, where k denotes the number of parameters.

3.9.4.2 Jarque-Bera Test of Normality

One of the foremost commonly applied tests for normality is the Jarque-Bera (JB) test. JB utilized for testing whether the series is normally distributed. It measures the contrast within the skewness and kurtosis of the series with those from the normal distribution (Bera and Jarque 1981). The test statistic is always nonnegative. If it is distant from zero, it signals the data don't have a normal distribution. The statistic JB is computed as:

$$JB = \frac{n}{6} \left(S^2 + \frac{1}{4}(K - 3)^2 \right) \quad 3.13$$

where n is the number of observations, S is the sample skewness, and K is the sample kurtosis:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\left(\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right)^{3/2}},$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^4}{\left(\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right)^2},$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are estimates of third and fourth central moments, respectively, $\bar{y}$ is the sample mean, and $\hat{\sigma}^2$ is the estimate of the second central moment, the variance. Under the null hypothesis

of a normal distribution, the JB statistic asymptotically has a chi-squared distribution with two degrees of freedom. The reported Probability is the probability that a JB statistic exceeds (in absolute value) the observed value under the null hypothesis - a small probability value leads to the rejection of the null hypothesis of a normal distribution.

3.9.5 Forecasting

The final step in time series modeling is forecasting. It is the method of estimation in unknown circumstances, which is commonly utilized in discussion of time-series data. There are two kinds of forecasts: sample period forecasts and post-sample period forecasts. The former one is used to create confidence within the model and the last mentioned to create genuine desired forecasts. In forecasting, the objective is to predict future values of a time series $y_{t+m} = 1, 2, 3, \dots$, based on the data collected to the present $y = y_t, y_{(t-1)}, y_{(t-2)}, \dots, y_1$. All through this section, we'll assume y_t is stationary and the model parameters are known (Chatfield 2005; Hamilton 1993).

3.9.5.1 Forecasting Accuracy Measures

The accuracy of the forecasting method is determined by analyzing forecast errors. The forecasting performance of the estimators is evaluated based on the difference between forecast and actual. The smaller the difference between the predicted value and the actual value of the dependent variable, the better the forecasting performance of the estimator (Tsay 2010). The forecasting performance of the system in the sample must be evaluated using standard statistical tools such as mean square error, mean absolute error, mean absolute percentage error.

Mean Squared Errors (MSE)

Mean squared error is a measure of accuracy calculated by squaring the individual error for each element in a data set and then finding the mean or median of the sum of these squares.

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad 3.14$$

Where MSE is mean squared error, n is time periods and e^2 is forecast error.

Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error is the mean of the sum of all percentage errors for a given data set, taken without regard to sign to avoid the problem of positive values and negative values cancel each other out.

$$MAPE = \frac{1}{n} \sum_{t=1}^n |P_t| \quad 3.15$$

Where P_t is percentage error given by $100e_t/y_t$, y_t is monthly average temperature for time period t , MAPE is mean absolute percentage error and n is time periods.

Root Mean Square Error

The RMSE is one of the most widely used measures of forecast accuracy.

$$RMSE = \sqrt{\frac{1}{n} (F_t - Y_t)^2} \quad 3.16$$

where RMSE is root mean squared error, n is time periods, F_t is the forecasted value and Y_t is the actual value.

3.10 Conditional Heteroscedastic Models

Box-Jenkins convective time series models rely on the assumption of homoscedasticity, which implies that the conditional variance of the residuals extracted from the model mean always remains constant. However, in the practical application of time series models, the conditional variance of the error term changes over time, indicating the existence of heteroskedasticity in the data. To capture these effects, Engle (1982) proposed the autoregressive conditional heteroskedasticity (ARCH) model, which states that the variance of the errors relative to the model mean at time t depends on previous information or shocks. Bollerslev (1986) generalized the ARCH model to GARCH, which is sophisticated and allows for a more flexible lag structure in the conditional variance.

3.10.1 Volatility Testing

In heteroskedasticity data modeling, volatility is an important aspect. If clustering of variation is present in the data, ARCH or GARCH family models can be used. Volatility clustering can be defined as large changes tend to entail large changes in either sign, and small changes tend to entail small changes in either sign or sign other brand.

3.10.2 ARCH Model

The first model to provide a systematic framework for modeling volatility was Engle (1982) ARCH model. The basic idea of ARCH models is that the shock is serially uncorrelated but dependent and this dependence structure can be described by a simple quadratic function of its lagged values. More precisely, the ARCH (p) model is specified as:

$$\sigma_{i,t}^2 = \alpha_0 + \sum_{k=1}^p \alpha_k \varepsilon_{i-k,t}^2 \quad 3.17$$

where α_0 is a constant term and α_p is the coefficient of the ARCH term.

Since negative variance is meaningless, we impose non-negative constraints, i.e. $\alpha_0 > 0$ and $\alpha_k \geq 0, k = 1, 2, \dots, p$. The value p is chosen so that the residual of the conditional variance equation is white noise. The process is covariant stationary if and only if $\sum_{k=1}^p \alpha_k < 1$.

However, (Tsay 2010) identified limitations of the ARCH model in that it requires a large number of parameters to estimate the volatility of temperature series and does not provide insight into the impact of past volatility on current volatility levels.

3.10.3 The Generalized ARCH (GARCH) Models

Due to the short time frame and easier implementation of the model, in this thesis the researcher will only use a subset of the models used in (Hansen and Lunde, 2005). The researcher will use the following models in the study: simple GARCH model, exponential GARCH model (EGARCH).

Bollerslev (1986) developed a more flexible generalization of ARCH, called the generalized autoregressive conditional heteroskedasticity model (GARCH). The GARCH model is a mechanism that adds lagged conditional variances to explain future conditional variance and also allows the models to be parameterized in greater detail. This greatly reduces the number of parameters that need to be estimated.

The generalized GARCH formula (p, q) , in which the current conditional variance is parameterized to depend on the bias of the squared error and the bias q of the conditional variance, is specified as follows:

$$\sigma_{i,t}^2 = \alpha_0 + \sum_{k=1}^p \alpha_k \varepsilon_{i-k}^2 + \sum_{j=1}^q \beta_j \sigma_{i-j}^2 \quad 3.18$$

where σ_t^2 is the conditional variance of the time series at time t , α_0 is the constant term representing the long-term average volatility, and α_k , and β_j are the non-negative parameters representing the short- and long-term effects.

In the context of climate change, we can consider disturbances to the Earth's ecosystem, such as temperature changes, as causing further temperature disturbances that take some time to stabilize to relatively “normal” conditions. Therefore, some summers may be much hotter and longer than usual, or other phenomena such as rainfall and floods may be more intense and last longer. Then, the heteroscedasticity or unequal variance of the relevant variable may represent an Autoregressive structure, such that its conditional variance depends on past variance.

Figure 3.1 depicts the flowchart of the proposed methodology to be followed for analysis of monthly average temperature.

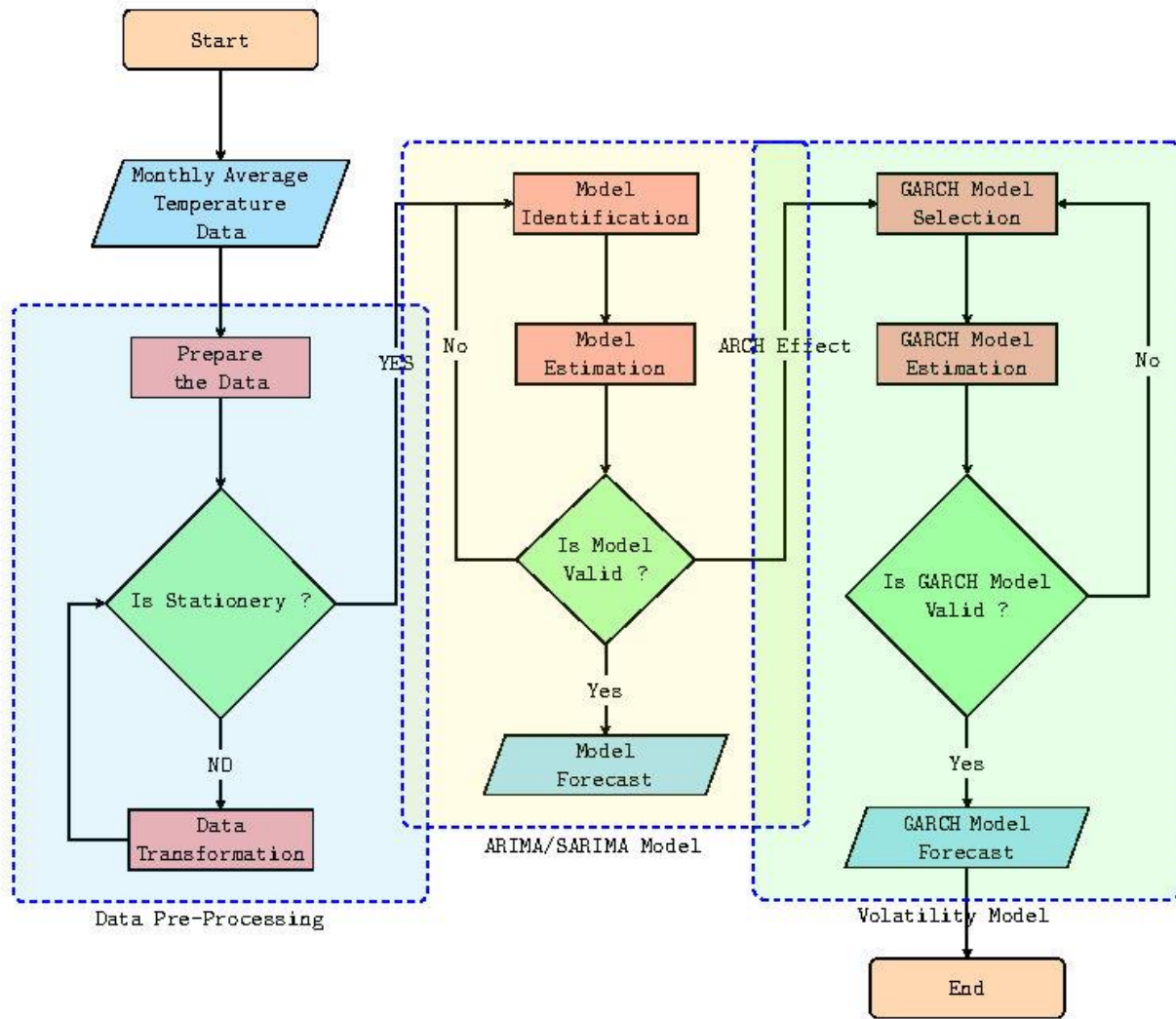


Figure 3.1 Flowchart of the proposed methodology

4 Results and Discussion

4.1 Descriptive statistics

Descriptive statistics provide a summary of the distribution, central tendency, and variability of a dataset. In this case, the dataset comprises monthly average temperatures for four different regions: Global, Ethiopia, Kenya, and Uganda.

Table 4.1 Descriptive statistics of monthly average temperature

	GMAT	EMAT	KMAT	UMAT
Count	2089	1679	1548	1548
Mean	14.103	23.025	23.979	23.019
Std	1.401	1.106	1.084	0.898
Min	11.202	20.144	21.337	21.016
50%	14.112	22.968	23.944	22.930
Max	17.213	26.961	27.659	26.468

The count represents the number of observations in each dataset. For the global dataset (GMAT), there are 2089 observations, while for Ethiopia (EMAT), Kenya (KMAT), and Uganda (UMAT), there are 1679, 1548, and 1548 observations respectively.

The mean (or average) temperature gives a measure of central tendency for each dataset. The global monthly average temperature (GMAT) is approximately 14.103 degrees Celsius. In contrast, the mean temperatures for Ethiopia (EMAT), Kenya (KMAT), and Uganda (UMAT) are 23.025°C, 23.979°C, and 23.019°C respectively. The mean temperature for GMAT (14.103°C) is significantly lower compared to EMAT (23.025°C), KMAT (23.979°C), and UMAT (23.019°C). This suggests that, on average, the global temperature is much cooler than the temperatures observed in Ethiopia, Kenya, and Uganda.

The standard deviation provides a measure of the dispersion or spread of the data points around the mean. A higher standard deviation indicates greater variability. For the global dataset, the standard deviation is approximately 1.401°C. Meanwhile, the standard deviations for Ethiopia, Kenya, and Uganda are 1.106°C, 1.084°C, and 0.898°C respectively, indicating relatively lower variability in temperature compared to the global dataset. GMAT also exhibits a higher standard deviation (1.401°C) compared to EMAT (1.106°C), KMAT (1.084°C), and UMAT (0.898°C). This indicates that the global temperature variability is greater than that of the individual countries,

implying more fluctuation in global temperatures. Also Ethiopia exhibits the second higher variability as compared with other East African countries.

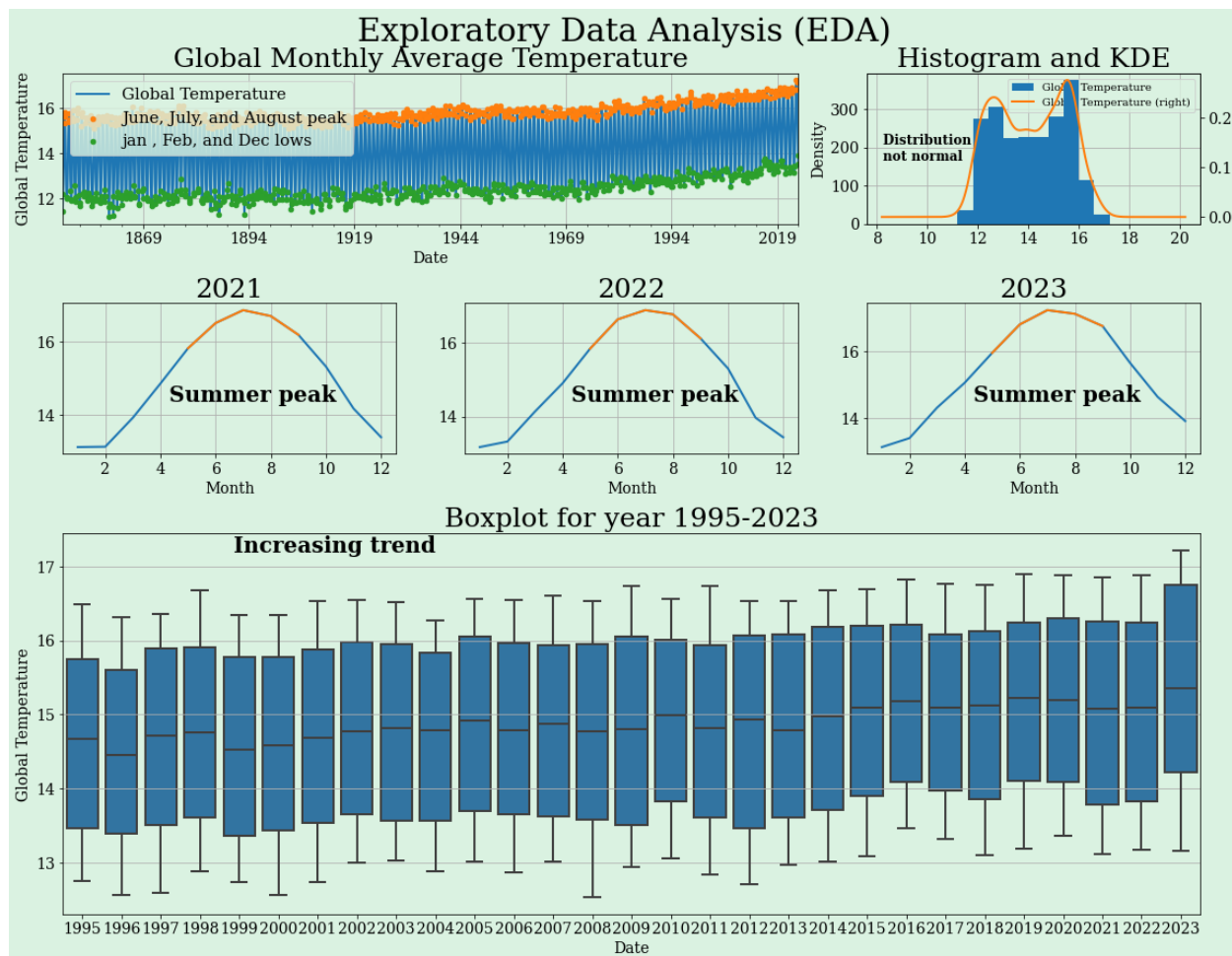
The minimum temperature recorded in each dataset. The lowest recorded temperature globally (GMAT) is 11.202°C, while for Ethiopia (EMAT), Kenya (KMAT), and Uganda (UMAT), the minimum temperatures are 20.144°C, 21.337°C, and 21.016°C respectively. This suggests that the coldest temperatures occur at the global level but are relatively warmer in the individual countries.

The median (50th Percentile or Q2) temperature represents the midpoint of the dataset, separating the higher half from the lower half. For GMAT, EMAT, KMAT, and UMAT, the median temperatures are 14.112°C, 22.968°C, 23.944°C, and 22.930°C respectively. This further supports the observation that global temperatures are generally cooler than those in the individual countries.

The maximum temperature recorded in each dataset. The highest recorded temperature globally (GMAT) is 17.213°C, while for Ethiopia (EMAT), Kenya (KMAT), and Uganda (UMAT), the maximum temperatures are 26.961°C, 27.659°C, and 26.468°C respectively. This indicates that while the global dataset includes warmer temperatures, they are not as extreme as those observed in the individual countries.

The descriptive statistics reveal that global monthly average temperatures tend to be cooler and more variable compared to those observed in Ethiopia, Kenya, and Uganda. While the global dataset includes a wide range of temperatures, they generally fall within a lower temperature range compared to the individual countries, which experience warmer temperatures with less variability.

Figure 4.1 Exploratory Data Analysis of Global Monthly Average Temperature



The top left panel is a Global Monthly Average Temperature Line Plot: This plot illustrates the global monthly average temperature from 1850 to 2023. It provides a comprehensive view of temperature fluctuations, seasonal patterns, and long-term trends. The line plot allows us to observe how temperatures have evolved over time, highlighting periods of warming, cooling, and stability. Furthermore, the orange dots denote the peak(warm) temperature months, comprising June, July, and August, while the green dots signify the low-temperature(cool) months, including January, February, and December. The plot shows a clear upward trend in global temperatures over time.

The top right panel is a Histogram and KDE Plot: This graph presents a histogram and Kernel Density Estimate (KDE) of the global temperature distribution. The histogram shows the

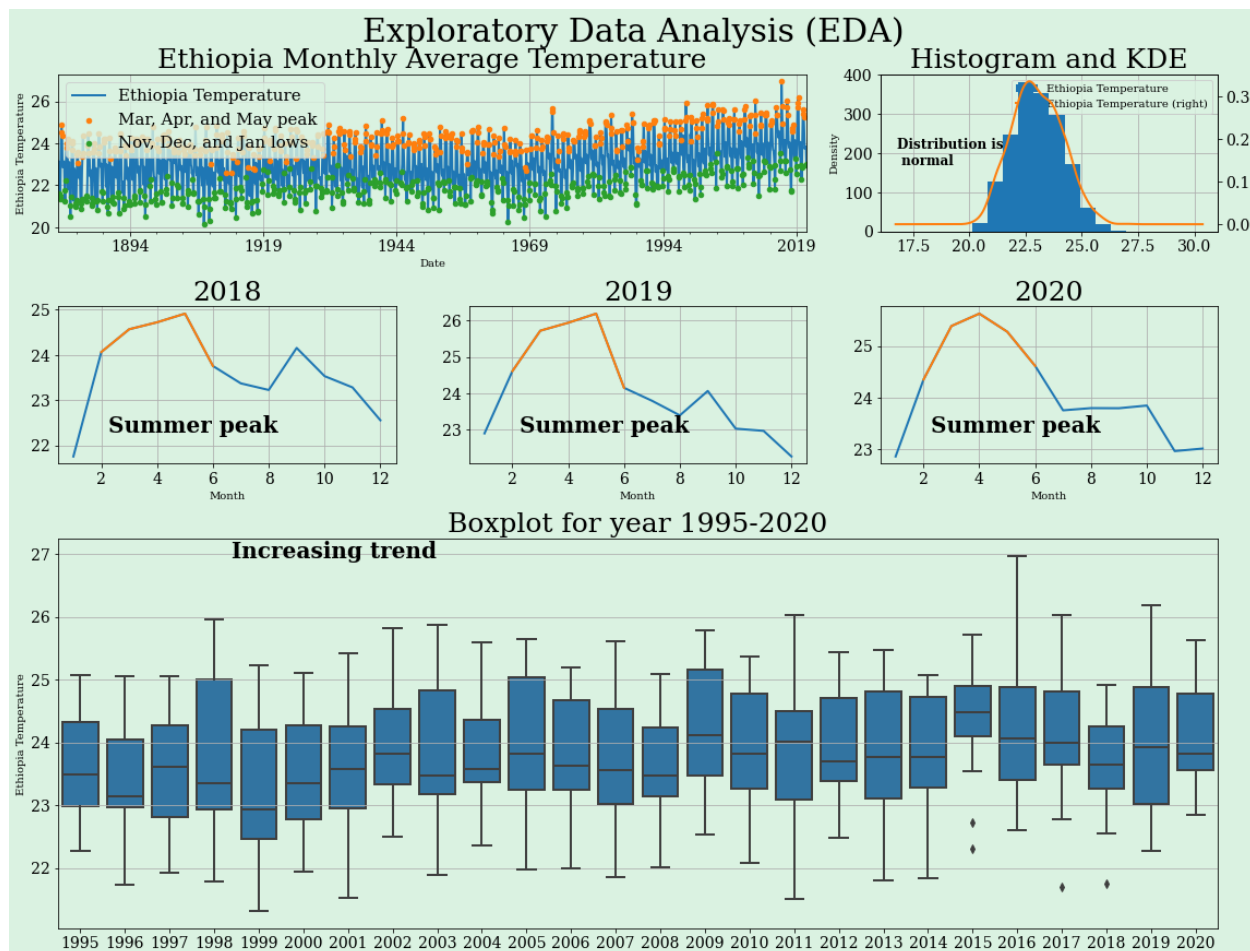
frequency distribution of temperature values, while the KDE smooths out the distribution curve. From this plot, we can infer that the distribution of global temperatures is non-normal, indicating potential skewness or kurtosis in the data.

The middle row panel is a Line Plots for Specific Years (2021-2023): These line plots focus on the years 2021, 2022, and 2023, offering a detailed examination of temperature variations. The blue lines depict temperature trends, while the orange lines highlight seasonal peaks, providing insights into temperature fluctuations within each year. These plots enable a closer analysis of temperature dynamics over shorter time periods. It's noteworthy that the seasonal peaks for global temperatures occur in the months of May, June, July, August, and September.

Boxplot for Years 1995-2023: The boxplot summarizes the distribution of global temperatures from 1995 to 2023. It provides a visual representation of temperature variability, including measures such as median, quartiles, and outliers. The increasing trend in the boxplot suggests that global temperatures have been gradually rising over the past few decades, indicating a long-term warming trend. This plot helps to identify patterns and anomalies in temperature distributions over time.

Overall, the above figure suggests that global average temperatures have been increasing over time. The histogram and KDE plot show that the distribution of temperatures is not normal, and the boxplot shows that the median temperature has been increasing. The time series plot shows the long-term trend in global temperatures.

Figure 4.2 Exploratory Data Analysis of Ethiopia Monthly Average Temperature



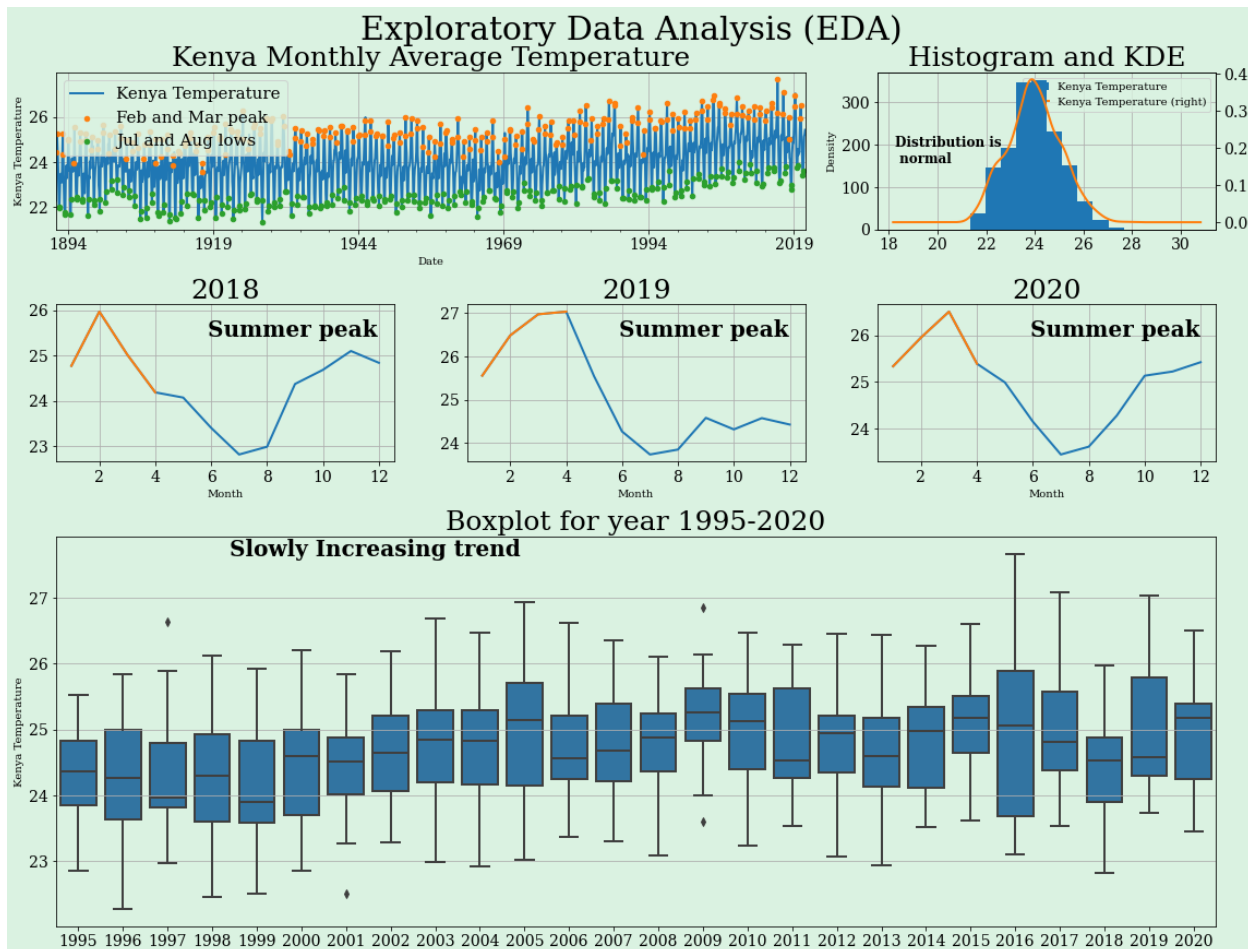
The top left panel showcases a Line Plot illustrating Ethiopia's Monthly Average Temperature. This plot provides a comprehensive view of the monthly average temperature trends across the year 1850-2020. The blue line represents the overall temperature trend over time, capturing any significant long-term changes or patterns. Furthermore, the orange dots denote the peak temperature months, comprising March, April, and May, while the green dots signify the low-temperature months, including November, December, and January. By presenting this data, the plot offers valuable insights into the temperature fluctuations observed throughout the years, thereby aiding in the identification of notable trends and irregularities. The line in the plot shows a clear upward trend, indicating that the average monthly temperature in Ethiopia has been increasing over time.

The top right panel is a Histogram and KDE (Kernel Density Estimation) Plot: The histogram visualizes the distribution of Ethiopia's monthly average temperatures, providing insights into the frequency of different temperature ranges. The KDE plot overlays a smooth curve on the histogram, offering a more continuous representation of the data distribution. The shape of the KDE plot suggests that the distribution of temperatures is not symmetrical, but leans slightly towards warmer temperatures. This indicates that there may be more months with warmer temperatures than cooler temperatures in the data set.

The middle row panel shows a Line Plots for Specific Years (2018-2020): These line plots focus on the years 2018, 2019, and 2020, offering a detailed examination of temperature variations within each year. The blue lines depict temperature trends, while the orange lines highlight seasonal peaks, providing insights into temperature fluctuations throughout the year. The seasonal peak months are February, March, April, May, and June. These plots allow for a closer analysis of temperature dynamics over shorter time periods, revealing any seasonal patterns or anomalies.

The bottom panel is a Boxplot for the Years 1995-2020: The boxplot summarizes the distribution of Ethiopia's monthly average temperatures from 1995 to 2020. It displays key statistical measures such as the median, quartiles, and outliers, providing a concise summary of the temperature data over the specified period. This visualization helps identify any trends or changes in temperature distribution over time, such as shifts in the median or the presence of outliers.

Figure 4.3 Exploratory Data Analysis of Kenya Monthly Average Temperature



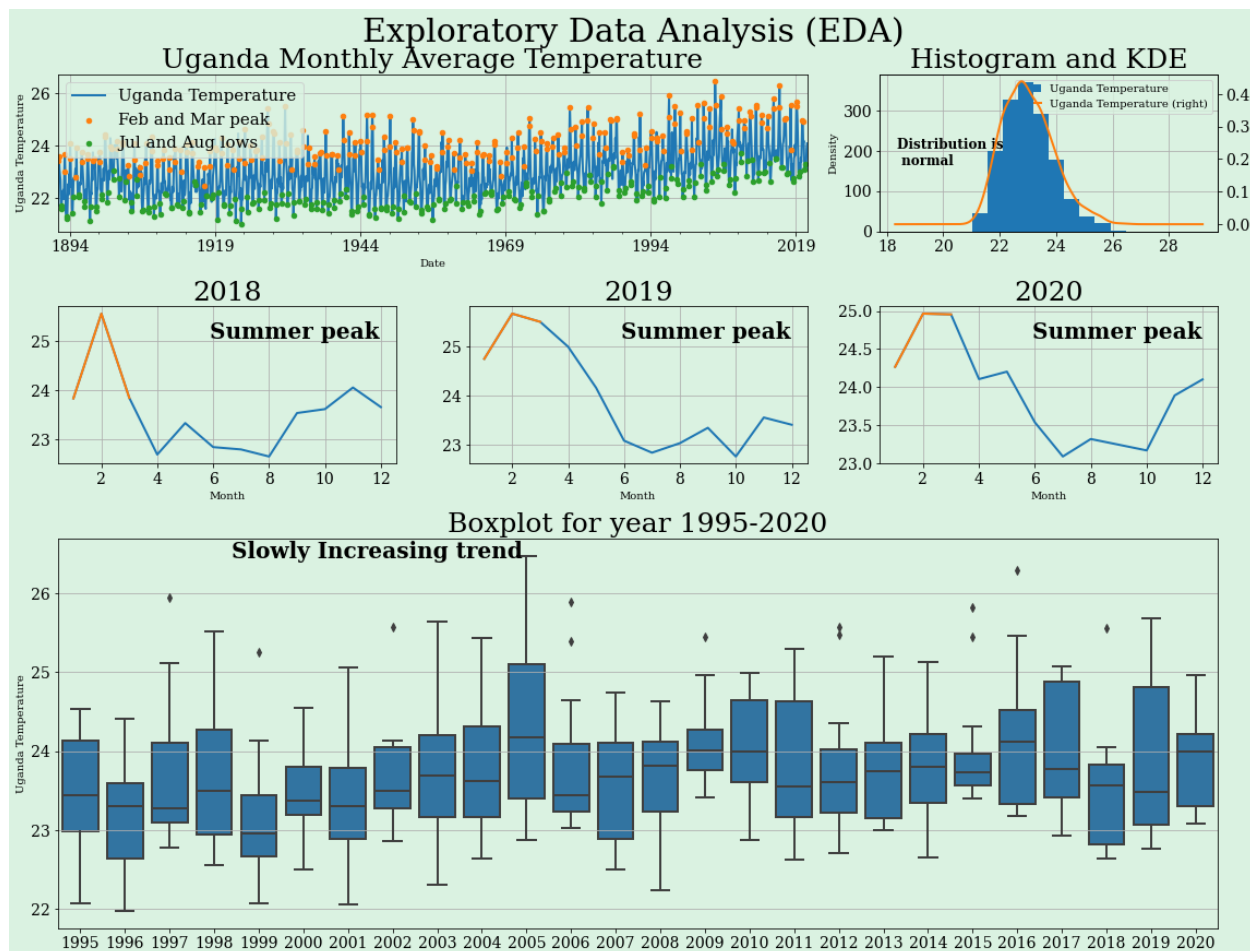
Main Line Plot: The primary focus of this plot is to visualize Kenya's Monthly Average Temperature over the duration 1850-2020. The blue line represents the temperature trend over time, offering insights into long-term increasing changes or patterns. Additionally, scatter points in orange highlight peak temperatures observed in February and March, while green points denote lower temperatures experienced in July and August. These annotations allow for a deeper understanding of seasonal temperature variations and their impacts.

Histogram and KDE Plot: This subplot provides a statistical overview of the temperature distribution in Kenya. The histogram illustrates the frequency distribution of temperature values, while the KDE plot presents a smoothed estimate of the probability density function. The KDE plot itself appears somewhat skewed to the right, with a longer tail towards warmer temperatures. This suggests that there may be slightly more months with warmer temperatures than cooler temperatures in the data set, although the overall distribution is close to normal.

Subplot Line Graphs: These line graphs depict temperature trends for specific years (2018, 2019, and 2020), allowing for a closer examination of temperature fluctuations over shorter time periods. The blue lines represent temperature trends, while the orange lines highlight seasonal peaks observed within each year. These plots offer valuable insights into temperature dynamics, revealing any noticeable trends or irregularities over the selected years.

Boxplot: This plot is a boxplot that summarizes the spread of monthly average temperatures in Kenya for each year between 1995 and 2020. The box in the middle of each vertical line represents the middle 50% of the data points for that year. The line in the middle of the box is the median temperature, and the ends of the lines (whiskers) extend to data points that are within 1.5 times the interquartile range (IQR) from the quartile. Points outside the whiskers are considered outliers. This is reflected in the fact that the median temperature (the line in the middle of each box) appears to be increasing over time. For example, the median temperature in 2020 is likely higher than the median temperature in 1995.

Figure 4.4 Exploratory Data Analysis of Uganda Monthly Average Temperature



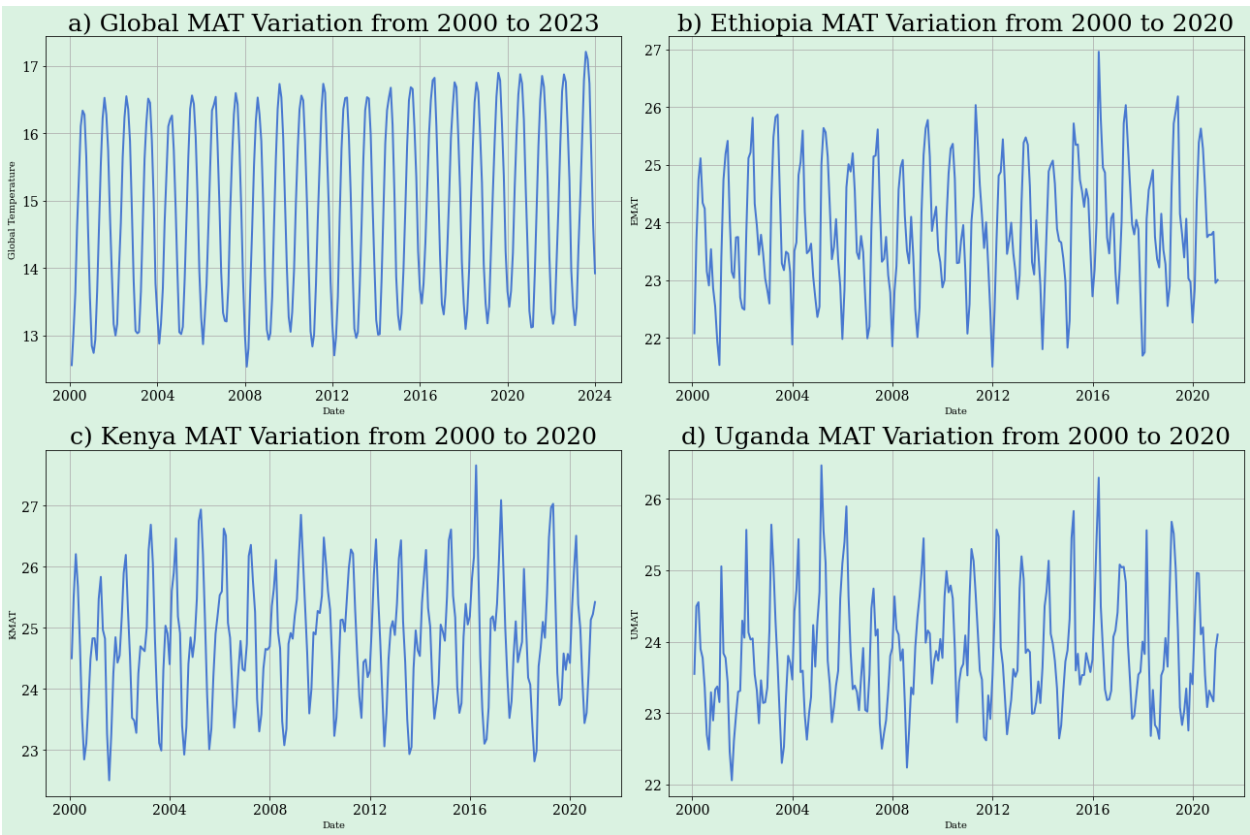
Main Line Plot: The primary visualization in this graph showcases Uganda's Monthly Average Temperature over the entire dataset. The blue line illustrates the temperature trend over time, capturing any notable long-term changes or patterns. Additionally, scatter points in orange denote peak temperatures observed in February and March, while green points represent lower temperatures experienced in July and August. These annotations help to identify seasonal temperature variations and their impacts on Uganda's climate. The line graph displays a clear upward trend, visually confirming a rise in Uganda's average monthly temperature over time.

Histogram and KDE Plot: This subplot provides a statistical overview of the temperature distribution in Uganda. The histogram shows the data binned into frequency counts, while the KDE plot is a smoothed representation of the distribution. The smooth curve of the KDE plot suggests a relatively normal distribution of temperatures, though a slight rightward skew might indicate a tendency for warmer temperatures compared to cooler ones.

Subplot Line Graphs: These line graphs depict temperature trends for specific years (2018, 2019, and 2020), enabling a detailed examination of temperature fluctuations over shorter time periods. The blue lines represent temperature trends, while the orange lines highlight seasonal peaks observed within each year. These plots offer valuable insights into temperature dynamics, revealing any noticeable trends or irregularities over the selected years.

Boxplot: The boxplot summarizes the temperature distribution from 1995 to 2020, providing a visual representation of the central tendency, dispersion, and skewness of the data. The plot reveals a slowly increasing trend in temperatures over the years, suggesting potential long-term climate changes or environmental influences. This information is vital for understanding the overall trajectory of temperature variations in Uganda and can inform further research and policy decisions.

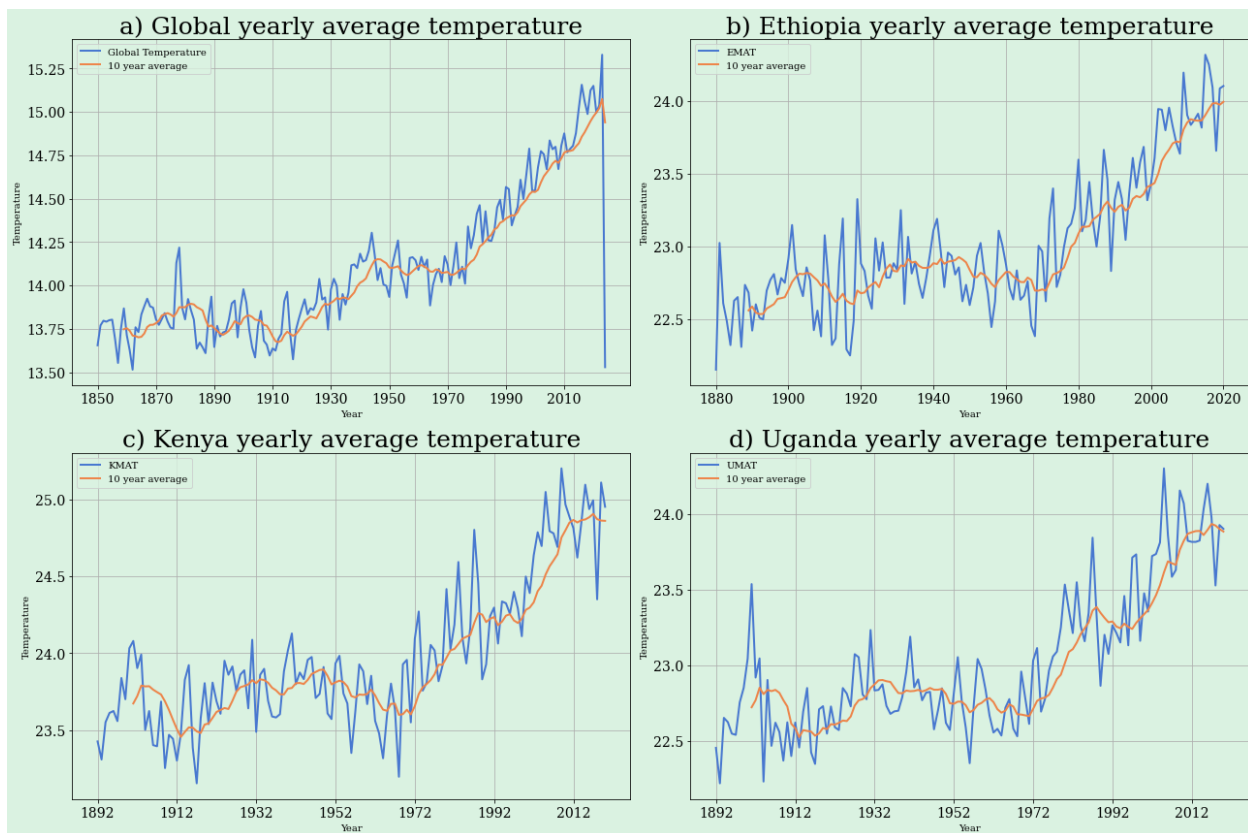
Figure 4.5 Zoomed monthly average temperature



These plots provide a detailed examination of monthly average temperature variations for four different regions from 2000 to 2023. Each subplot zooms in on the temperature data within the

specified timeframe, allowing for a closer inspection of the temperature patterns and trends. From a visual perspective, the plots reveal fluctuations and trends in temperature over the selected period, highlighting the temporal variability of temperature dynamics. By focusing on a narrower timeframe, these zoomed-in plots offer insights into short-term temperature variations and potential climatic changes occurring within the analyzed years. Such detailed visualizations enable a deeper understanding of the temperature patterns and facilitate the identification of potential anomalies or trends that may influence regional climate conditions.

Figure 4.6 Yearly Average Temperature



The plots depict the yearly average temperature trends for four different countries: global, Ethiopia, Kenya, and Uganda. Each plot showcases the fluctuation of yearly temperature averages over time, alongside a 10-year moving average trend line for smoother visualization. The visual examination of these plots reveals evident seasonality in the temperature data, with recurring patterns indicating seasonal fluctuations. Additionally, a discernible trend is observed in each country's temperature data, suggesting long-term changes in temperature patterns. These visual insights contribute valuable information to our understanding of historical temperature variations

and provide a basis for further analysis and research into the underlying climate dynamics in these regions.

Figure 4.7 Train-Test Split of Monthly Average Temperature

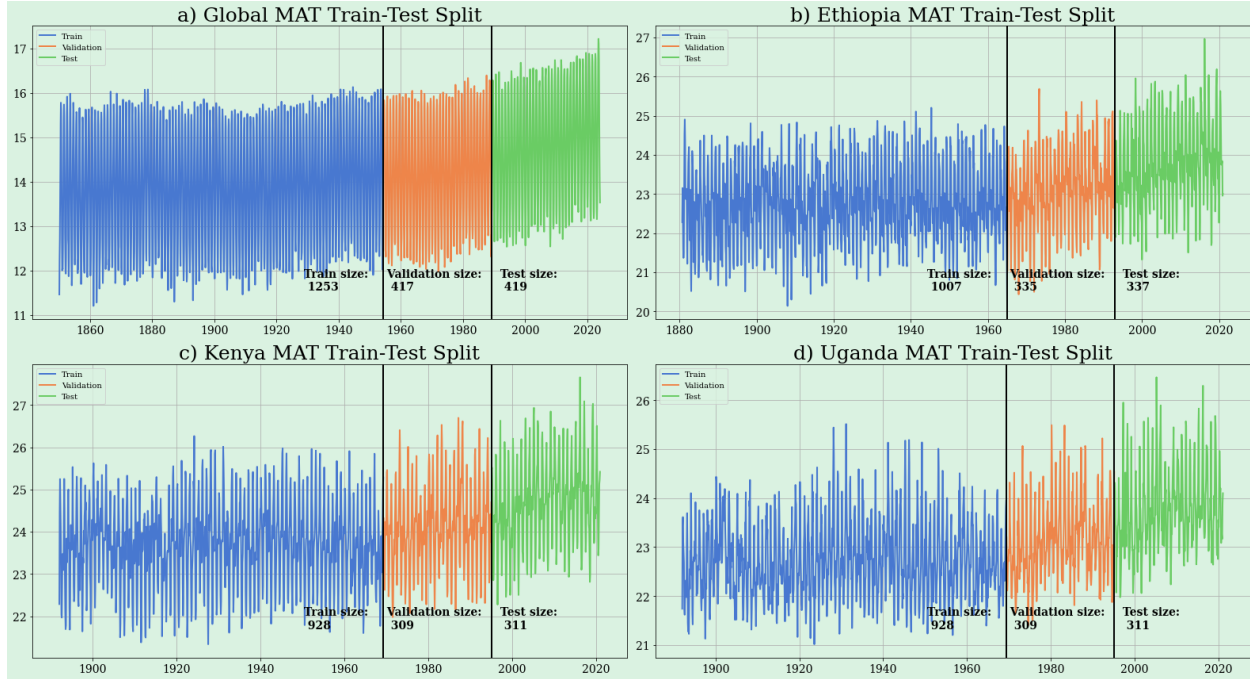


Figure 4.7a) Global MAT Train-Test Split: This plot illustrates the division of the global Monthly Average Temperature (MAT) dataset into training, validation, and test sets for modeling purposes. The training set (in blue) encompasses approximately 60% about 1253 of the total data, providing a substantial portion for model training. The validation set (in orange) comprises around 20% about 417 of the data, serving as a separate subset for fine-tuning the model's hyperparameters and assessing its performance during training. Finally, the test set (in green) includes the remaining data, around 20%(419), and is reserved for evaluating the model's performance after training and validation.

Figure 4.7b) Ethiopia MAT Train-Test Split: Similar to the first plot, this graph demonstrates the partitioning of the Ethiopia Monthly Average Temperature (MAT) dataset into training, validation, and test sets. The blue line represents the training data, consisting of approximately 60%(1007) of the total observations, utilized for training the model. The orange line indicates the validation set, constituting around 20%(335) of the data, used for optimizing model parameters and assessing its

performance. Lastly, the green line represents the test set, comprising the remaining 20%(337) of the data, reserved for evaluating the model's performance on unseen data.

Figure 4.7c) Kenya MAT Train-Test Split: This plot depicts the split of the Kenya Monthly Average Temperature (MAT) dataset into training, validation, and test sets. The blue line corresponds to the training set, which encompasses roughly 60%(928) of the total data, utilized for training the predictive model. The orange line represents the validation set, constituting approximately 20%(309) of the data, employed for tuning the model's parameters and assessing its performance. Lastly, the green line denotes the test set, comprising the remaining 20%(311) of the data, used for evaluating the model's predictive accuracy on unseen data.

Figure 4.7d) Uganda MAT Train-Test Split: Similar to the previous plots, this graph showcases the division of the Uganda Monthly Average Temperature (MAT) dataset into training, validation, and test sets. The blue line indicates the training set, which contains approximately 60%(928) of the total observations, employed for training the predictive model. The orange line represents the validation set, constituting around 20%(309) of the data, utilized for optimizing the model's hyperparameters and evaluating its performance. Lastly, the green line depicts the test set, comprising the remaining 20%(311) of the data, reserved for assessing the model's predictive accuracy on unseen data.

The baseline model, established through the persistence method, serves as a fundamental benchmark against which the performance of more sophisticated forecasting models can be evaluated. By simply predicting future values based on the previous time step's value, this baseline model provides a straightforward reference point for comparison.

As illustrated in Figure 0.2, Figure 0.3, Figure 0.4, Figure 0.5 for each region, including Global, Ethiopia, Kenya, and Uganda, the baseline validation scores were calculated to be 0.697, 0.75, 0.706, and 0.593, respectively. These scores represent the root mean square error (RMSE) between the actual and predicted values of the validation datasets.

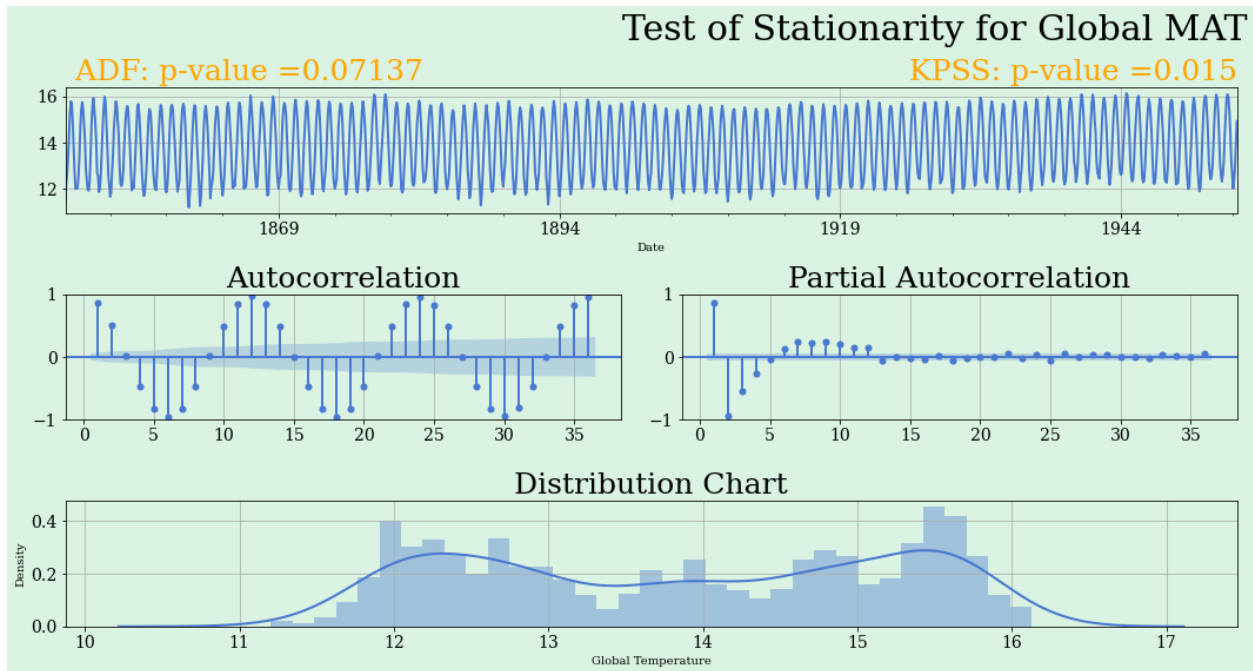
The resulting baseline DataFrame encapsulates these baseline scores, offering a clear and concise summary for comparison with future forecasting models, such as SARIMA. This comparison enables researchers to assess whether more complex models yield substantial improvements in predictive accuracy over the simple baseline approach.

Ultimately, the establishment of a baseline forecast is a critical step in forecasting analysis, providing researchers with a foundational reference point for evaluating the performance and efficacy of their models. By leveraging this baseline, researchers can make informed decisions about the adoption and refinement of forecasting methodologies, thereby enhancing the reliability and accuracy of future predictions.

4.2 Test of Stationarity

The test of stationarity is a crucial aspect of time series analysis. To assess whether a time series dataset exhibits stationarity, a widely used statistical test called the Augmented Dickey-Fuller (ADF) test is employed. This test evaluates the presence of unit roots in the data, which are indicative of non-stationarity. The null hypothesis of the ADF test is that the time series has a unit root, implying it is non-stationary. Conversely, the alternative hypothesis suggests stationarity. By conducting this test, I aim to determine the underlying nature of the time series data, which is essential for the subsequent modeling and forecasting tasks in my research.

Figure 4.8 Test of Stationarity for Global MAT



The above Figure 4.8 were used for examining the Global Monthly Average Temperature (MAT) time series data. The time series plot visually represents the temperature fluctuations over time, allowing for observations of any discernible trends, seasonality, or irregular patterns. There is a

clear seasonality was seen from the plot. Additionally, autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were utilized to assess the correlation structure within the data. Peaks or patterns in the ACF plot which indicate seasonal behavior, while the PACF plot aids in identifying significant autocorrelation lags for modeling purposes. Furthermore, a distribution plot illustrates the underlying distribution of Global temperature values which is not normal.

Statistical tests for stationarity, namely the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests, were conducted. The ADF test evaluates the presence of a unit root in the data, with a higher test statistic compared to critical values indicating non-stationarity. Conversely, the KPSS test examines for a deterministic trend, where a p-value less than a chosen significance level suggests non-stationarity. In this analysis, the ADF test statistic exceeds the critical value at the 5% significance level, indicating non-stationarity. Similarly, the KPSS p-value is less than 0.05, supporting the rejection of the null hypothesis of stationarity. Consequently, the time series exhibits non-stationary behavior, requiring further analysis and modeling techniques, such as differencing or appropriate time series models, to capture and address the underlying patterns in the data.

Table 4.2 Statistical tests of Stationarity for Regional Temperature

Statistical Tests	Ethiopia MAT	Kenya MAT	Uganda MAT
ADF	0.000	0.000	0.000
KPSS	0.010	0.10	0.10

The statistical tests conducted on the monthly average temperature data for Ethiopia, Kenya, and Uganda reveal compelling insights into the stationarity of each respective time series. The Augmented Dickey-Fuller (ADF) test results indicate remarkably low p-values of 0.000 for all three countries, suggesting strong evidence against the presence of a unit root in the data. This outcome implies that the temperature series for Ethiopia, Kenya, and Uganda are likely stationary, as the ADF test statistic falls below the critical values at conventional significance levels.

Additionally, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test results further corroborate the stationarity findings, albeit with some nuances. While Ethiopia's temperature series demonstrates a KPSS p-value of 0.010, which is below the typical significance level of 0.05, indicating strong evidence against non-stationarity, both Kenya and Uganda exhibit slightly higher KPSS p-values

of 0.10. Although these values are marginally above the conventional threshold, they still suggest the presence of stationarity in the temperature data for Kenya and Uganda, albeit with slightly less certainty compared to Ethiopia.

Overall, the combined results of the ADF and KPSS tests provide robust evidence supporting the stationarity of the monthly average temperature series for Ethiopia, Kenya, and Uganda. These findings are crucial for subsequent time series analyses and modeling efforts, as they affirm the suitability of employing various statistical techniques and models that assume stationarity in the data.

Despite the statistically significant results indicating stationarity in the monthly average temperature (MAT) time series for Ethiopia, Kenya, and Uganda, the presence of clear seasonality evident in the time series plots and autocorrelation function (ACF) plots warrants further consideration. While the stationarity tests provide valuable insights into the overall behavior of the data, they may not fully capture the cyclic patterns and fluctuations inherent in temperature series, especially those influenced by seasonal variations.

The discrepancy between the stationarity tests and the observed seasonality could be attributed to the specific characteristics of temperature data. Temperature series often exhibit cyclical patterns due to seasonal variations driven by factors such as changes in weather patterns, solar radiation, and geographical location. These cyclical patterns can lead to significant autocorrelation at seasonal lags, resulting in apparent non-stationarity in the ACF plots.

In interpreting these results, it is essential to recognize that stationarity tests primarily assess the presence of a unit root or deterministic trend in the data, rather than capturing cyclicity or seasonality explicitly. Therefore, while the ADF and KPSS tests provide valuable initial assessments of stationarity, they may not fully capture the complexities of temperature time series data, particularly the seasonal components.

In light of these findings, researcher were exercise caution when interpreting stationarity test results in the context of temperature data and consider complementary analyses, such as seasonal decomposition techniques or time series modeling approaches that explicitly account for seasonality. Additionally, acknowledging the presence of seasonality in the data underscores the

importance of selecting appropriate modeling techniques that can effectively capture and account for these cyclic patterns in forecasting and analysis.

The seasonal decomposition graph provides valuable insights into the underlying patterns within the dataset. In this analysis, the researcher examines the seasonal decomposition graphs for the monthly average temperatures (MAT) of four regions: Ethiopia, Kenya, Uganda, and the global temperature as presented in Figure 0.9, Figure 0.10, Figure 0.11, and Figure 0.12 respectively.

Upon initial observation, we notice a prominent increasing trend in all four datasets, particularly evident from around 1960 onwards. This upward trend suggests a consistent rise in temperatures over time across these regions, reflecting a potential long-term climate change phenomenon.

When we delve deeper into the seasonal component of the decomposition, we observe distinct seasonality patterns embedded within the data. This seasonality manifests as recurring fluctuations that follow a consistent pattern over the course of each year. These fluctuations could be attributed to various factors such as seasonal variations in weather patterns, annual cycles, or other environmental influences.

Specifically, in the seasonal component of the decomposition graphs, we can discern periodic peaks and troughs that align with the changing seasons. For instance, we might observe peaks during warmer months and troughs during colder months, reflecting the cyclical nature of temperature changes throughout the year.

By analyzing the seasonal decomposition graphs for all four regions, we gain a comprehensive understanding of the underlying temporal patterns in the MAT datasets. This analysis aids in identifying both long-term trends and recurring seasonal variations.

The application of seasonal differencing as a remedial action for addressing non-stationarity in the Global MAT and accounting for seasonality in the MAT of Ethiopia, Kenya, and Uganda has yielded notable improvements in the stationarity tests. With the ADF p-values registering as 0.00 and KPSS p-values at 0.1 for all four regions, the results suggest that seasonal differencing has effectively mitigated the non-stationarity concerns while preserving the seasonal patterns inherent in the data.

The attainment of a significant ADF p-value of 0.00 indicates a rejection of the null hypothesis of a unit root, providing evidence in favor of stationarity in the differenced series. Similarly, the KPSS p-value of 0.1 suggests that the differenced series does not exhibit significant departure from stationarity under the null hypothesis. These results align with the objective of seasonal differencing, which aims to remove the seasonal trends and cyclic patterns while retaining the underlying stationarity in the data.

Furthermore, the consistency of these findings across all four regions Global MAT, Ethiopia, Kenya, and Uganda underscores the effectiveness of seasonal differencing as a robust approach for addressing non-stationarity and seasonality in temperature time series data. By incorporating seasonal adjustments into the analysis, researchers can obtain more reliable and interpretable results, enabling a better understanding of the underlying dynamics and facilitating more accurate forecasting and modeling efforts.

4.3 Building SARIMA Model

4.3.1 Identifying parameters of SARIMA models

In the process of identifying the parameters of the SARIMA (Seasonal Autoregressive Integrated Moving Average) model, crucial insights were gleaned through the analysis of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. Given the seasonal nature of the time series under consideration, characterized by patterns recurring at 12-month intervals, both the non-seasonal and seasonal components of the SARIMA model were meticulously scrutinized.

The middle row of Figure 0.13 presents a comprehensive examination of the ACF and PACF plots for the deseasoned data pertaining to the Global monthly average temperature. Notably, the ACF plot for the non-seasonal component showcases a distinctive pattern of tailing off, indicative of an Autoregressive (AR) model. By analyzing the PACF plot, it was discerned that the plot becomes truncated after the fourth lag, signaling the presence of an AR(4) model.

In addition to the non-seasonal component, attention was directed towards identifying the seasonal aspects of the time series. The ACF plot for the seasonal component exhibits a termination after the 12th lag, suggesting the inclusion of a Moving Average (MA) component, possibly MA(1), within the model. Moreover, the PACF plot for the seasonal component truncates after the 36th lag, hinting at the presence of a seasonal Autoregressive (AR) component, specifically SAR(3).

In Figure 0.14, Figure 0.15, Figure 0.16 the middle row displays the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for the deseasoned data corresponding to the Ethiopia, Kenya, and Uganda monthly average temperature respectively. Through a precise examination of these plots, insights into both the non-seasonal and seasonal components of the SARIMA (Seasonal Autoregressive Integrated Moving Average) model were gleaned.

The ACF plot for the non-seasonal component exhibits a distinct cutoff after the fourth lag, suggestive of a Moving Average (MA) component of order 4, denoted as MA(4). Conversely, the PACF plot for the non-seasonal component reveals a truncation after the second lag, indicating the presence of an Autoregressive (AR) component of order 2, signified as AR(2).

Turning attention to the seasonal component, the ACF plot displays a cessation after the 12th lag, suggesting the inclusion of a Moving Average (MA) component, plausibly MA(1), within the model. Similarly, the PACF plot for the seasonal component truncates after the 36th lag, implying the presence of a seasonal Autoregressive (AR) component, denoted as SAR(3).

This thorough analysis of the ACF and PACF plots provides valuable insights into the parameterization of the SARIMA model for the Ethiopia, Kenya, and Uganda monthly average temperature dataset. By discerning the appropriate orders for both non-seasonal and seasonal components, this methodology facilitates the development of robust forecasting models essential for climate analysis and prediction in the region.

4.3.2 SARIMA Model Selection

In time series analysis, selecting the appropriate model from a range of candidates is essential for accurate forecasting. The Akaike Information Criterion (AIC) and the Schwarz Bayesian Information Criterion (SBIC) are widely used model selection criteria that balance model fit and complexity. These criteria penalize models with excessive parameters, favoring parsimony without sacrificing predictive accuracy.

For Global MAT, the researcher considered several SARIMA models with different lag specifications. The researcher limits the maximum number of non-seasonal autoregressive (AR) parameters (p) to four (4) and the seasonal AR parameters (P) to three (3). The non-seasonal moving average (MA) parameter (q) were set to 0, and the maximum seasonal MA parameter (Q) value was 1.

By fitting each SARIMA model to the dataset and calculating the AIC and SBIC values, we can compare their performance. The model with the lowest AIC or SBIC value indicates the best trade-off between goodness of fit and model complexity. This approach allows us to identify the ideal lag specification that captures the underlying dynamics of the time series data while avoiding overfitting.

By utilizing these model selection criteria, we can ensure that the chosen SARIMA model provides reliable forecasts while remaining interpretable and robust to potential fluctuations in the data.

Based on the AIC and BIC values, SARIMA(4,0,0)(0,1,1,12) is identified as the model with the lowest AIC and BIC values in Appendix B: Tables

Table 0.1 SARIMA model selection for Global MAT, suggesting it as the best-fit model for the given data. Following closely, SARIMA (4,0,0)(2,1,1,12) exhibits decent AIC and BIC values, positioning it as the second-best candidate among the evaluated models. These results indicate that SARIMA(4,0,0)(0,1,1,12) and SARIMA(4,0,0)(2,1,1,12) are promising models for forecasting the observed Global monthly average temperature, providing valuable insights into future trends and patterns.

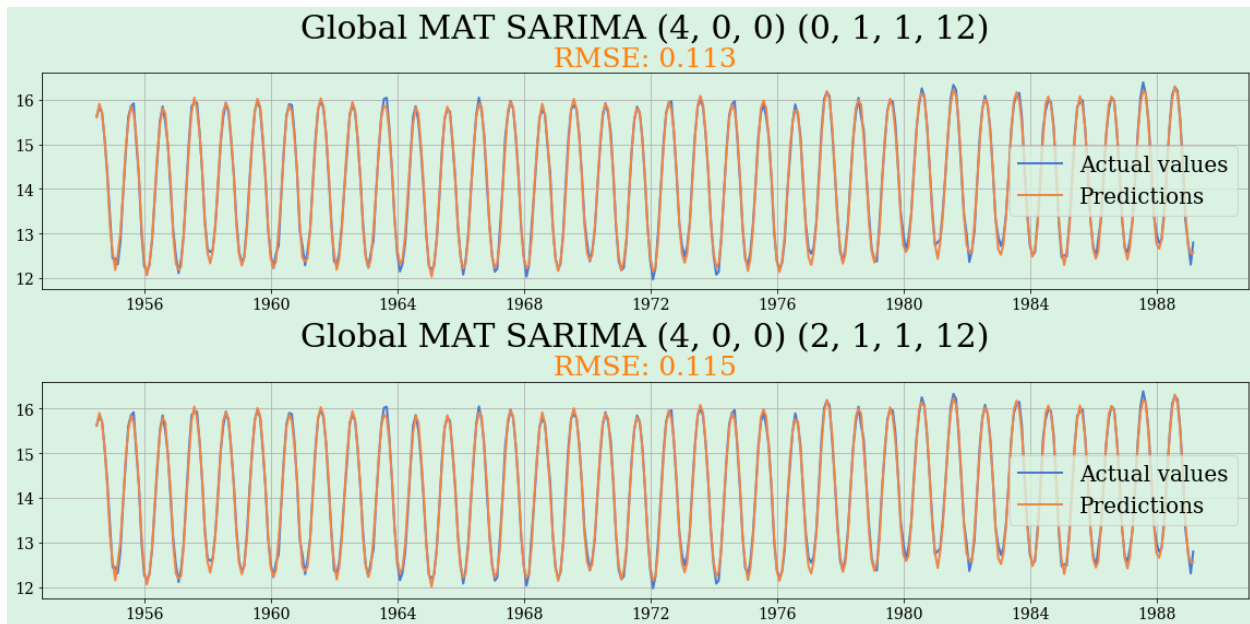
Table 0.2 presents the outcomes of the grid search, unveiling 120 candidate models for monthly average temperature across three distinct regions: Ethiopia, Kenya, and Uganda. For Ethiopia, SARIMA(0,0,4)(3,1,1,12) emerged as the optimal model, boasting the lowest AIC of 1365.62 and BIC of 1412.36. Following closely, SARIMA(2,0,2)(3,1,1,12) secured the second position with an AIC of 1366.19 and BIC of 1412.93. In the case of Kenya, SARIMA(2,0,3)(0,1,1,12) stood out as the most suitable model, backed by an AIC of 698.973 and BIC of 734.748. Another competitive contender was SARIMA(2,0,3)(1,1,1,12) with an AIC of 700.057 and BIC of 740.943. For Uganda, SARIMA(1,0,4)(0,1,1,12) claimed the top spot, supported by an AIC of 1130.724 and BIC of 1166.499, while SARIMA(2,0,0)(0,1,1,12) followed closely as the runner-up with an AIC of 1131.424 and BIC of 1151.867. These findings underscore the effectiveness of SARIMA models in capturing the complex temporal patterns exhibited by monthly average temperatures across diverse geographical regions.

After conducting a grid search, we identified the two best models for our Global monthly average temperature forecasting task. These models were SARIMA(4, 0, 0)(0, 1, 1, 12) and SARIMA(4,

0, 0)(2, 1, 1, 12). We then evaluated both models using the validation set. Our evaluation metric was the root mean square error (RMSE), with lower values indicating better performance.

Upon testing, we found that the SARIMA(4, 0, 0)(0, 1, 1, 12) model slightly outperformed the SARIMA(4, 0, 0)(2, 1, 1, 12) model on the validation set as presented below in Figure 4.9. This means that the first model provided more accurate predictions compared to the second model, as it achieved a slightly lower RMSE. Therefore, based on our evaluation criteria, we consider the SARIMA(4, 0, 0)(0, 1, 1, 12) model to be the better choice for forecasting in this scenario.

Figure 4.9 Evaluation of selected models in validation set of Global MAT



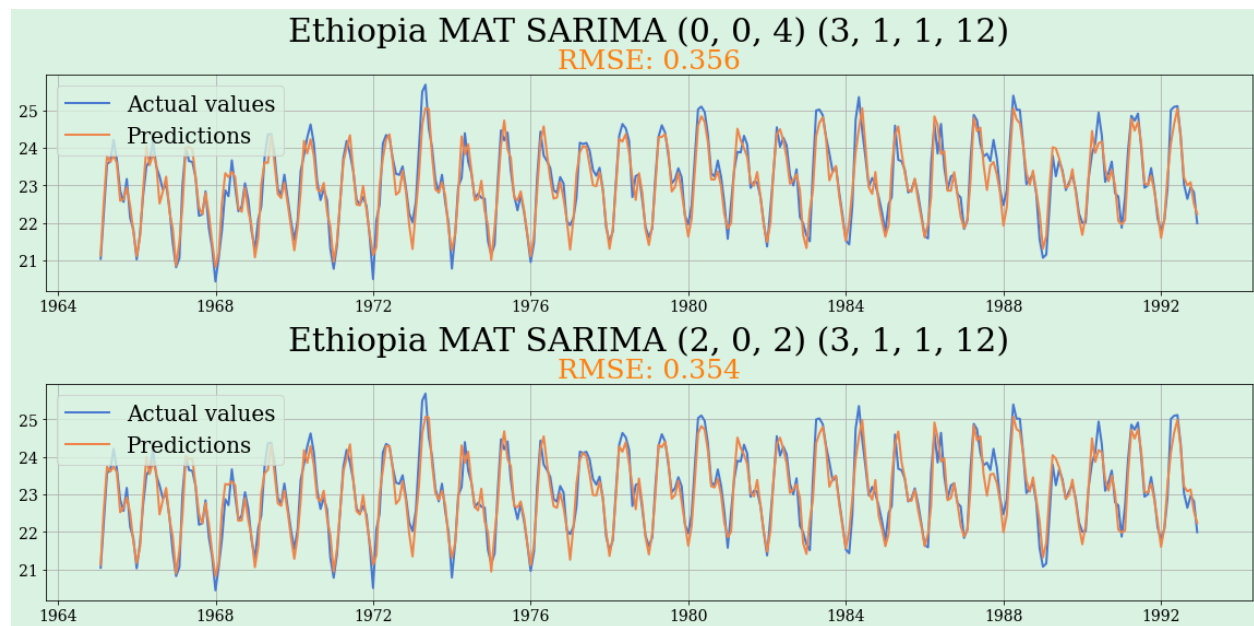
In our evaluation of forecasting models for Ethiopia's Monthly Average Temperature (MAT), we selected the two best-performing models identified through grid search for final assessment. These models, namely SARIMA(2, 0, 2)(3, 1, 1, 12) and SARIMA(0, 0, 4)(3, 1, 3, 12), were subjected to validation to gauge their predictive accuracy.

Upon examining their performance on the validation set, it was observed that the SARIMA(2, 0, 2)(3, 1, 1, 12) model slightly outperformed the SARIMA(0, 0, 4)(3, 1, 3, 12) model as shown in Figure 4.10. This conclusion was drawn based on comparative analysis of their respective validation scores, with the former exhibiting a marginally better predictive capability.

The evaluation process involved generating predictions for Ethiopia's MAT using both models and assessing their alignment with actual validation data. The resulting plots depicted the predicted values alongside the observed values, facilitating a visual comparison of their accuracy.

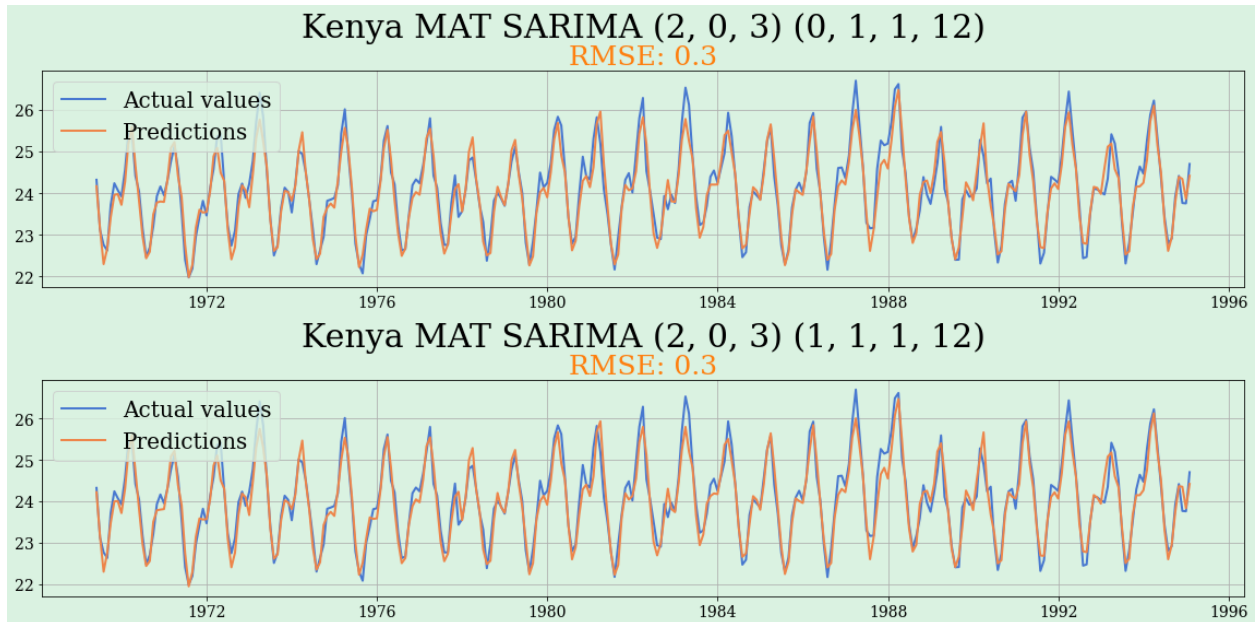
These findings contribute to our understanding of forecasting methodologies for Ethiopia's MAT and provide insights into the efficacy of SARIMA models in capturing its temporal patterns. The superiority of the SARIMA(2, 0, 2)(3, 1, 1, 12) model underscores its potential utility in practical applications requiring accurate temperature forecasts for Ethiopia.

Figure 4.10 Evaluation of selected models in validation set of Ethiopia MAT



The plotted Figure 4.11 showcases the predictive performance of two SARIMA models, which were identified as the best models for Kenya's Monthly Average Temperature (MAT) dataset based on grid search optimization. The SARIMA(2, 0, 3)(0, 1, 1, 12) and SARIMA(2, 0, 3)(1, 1, 1, 12) models were selected for final evaluation due to their superior performance on the validation set, as evidenced by their respective RMSE values of 0.3. These models were chosen after a thorough grid search process aimed at identifying the most suitable parameter configurations. The figure provides a visual representation of the actual temperature values compared to the predictions generated by each SARIMA model. Such evaluation is crucial for assessing the models' efficacy in capturing the underlying patterns and variations in Kenya's monthly temperature data.

Figure 4.11 Evaluation of selected models in validation set of Kenya MAT



The Figure 4.12 displayed provides an evaluation of two SARIMA models chosen for forecasting Uganda's Monthly Average Temperature (MAT) dataset. After an extensive grid search, the SARIMA(1, 0, 4)(0, 1, 1, 12) and SARIMA(2, 0, 0)(0, 1, 1, 12) models emerged as the primary contenders, demonstrating promising predictive abilities. With RMSE values of 0.369 and 0.367 respectively on the validation set, these models showcase their proficiency in capturing the complex temperature variations inherent in Uganda's climate. By comparing the observed temperature data with the model forecasts, researchers gain valuable insights into the accuracy and reliability of the models.

Figure 4.12 Evaluation of selected models in validation set of Uganda MAT

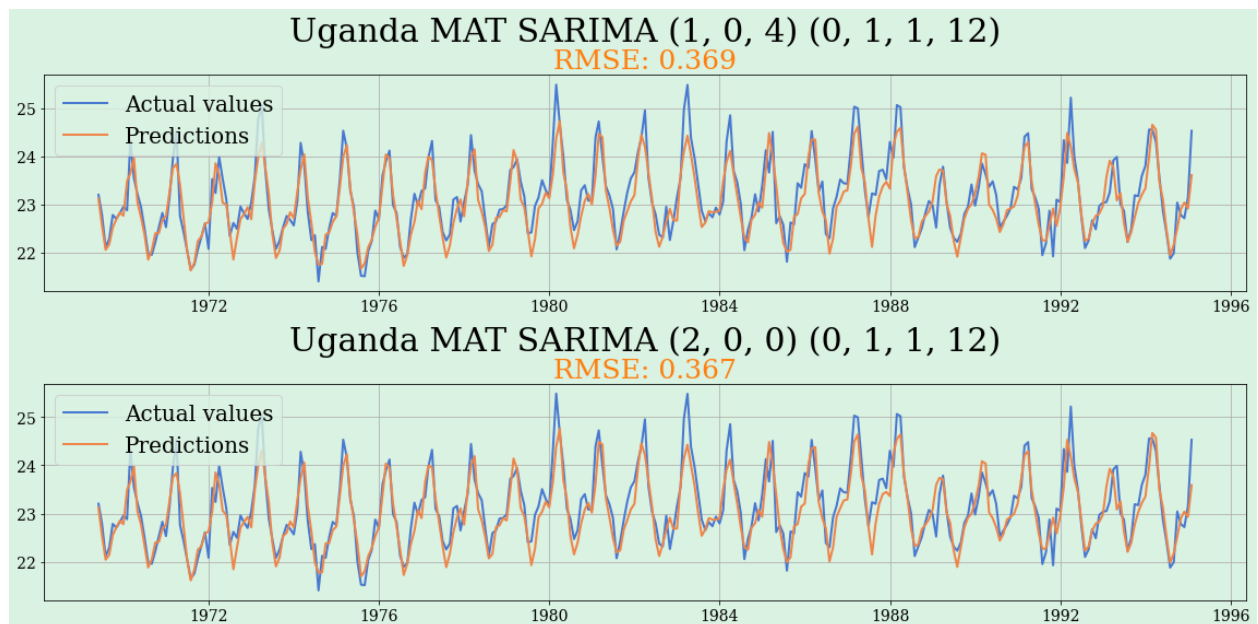
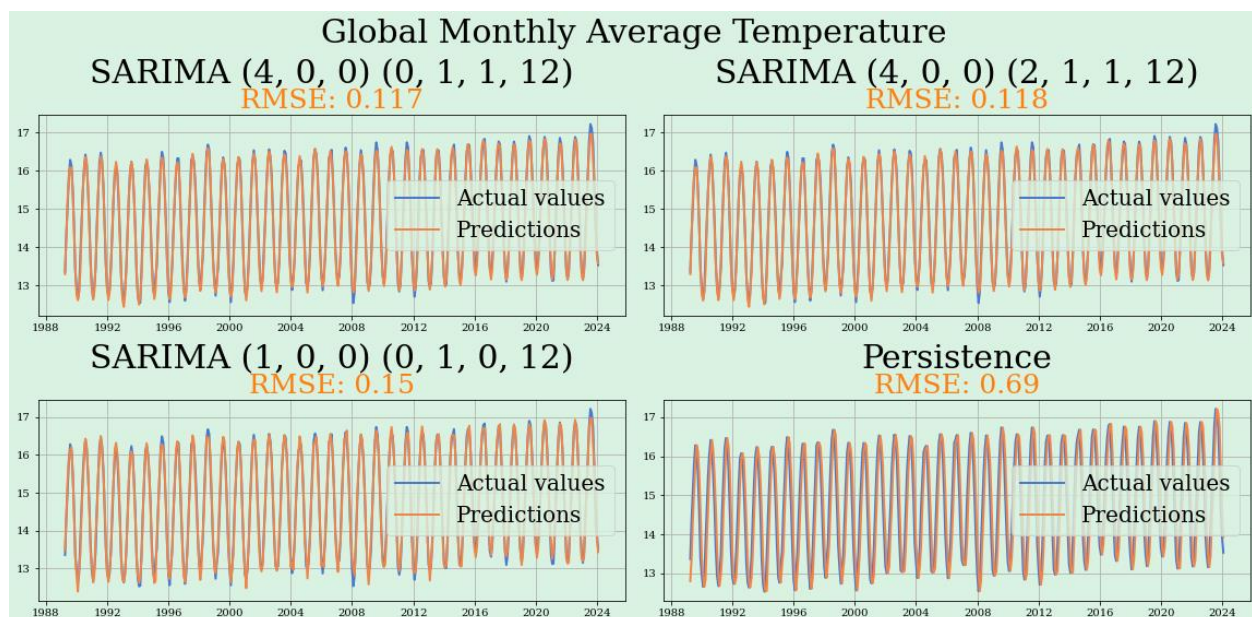


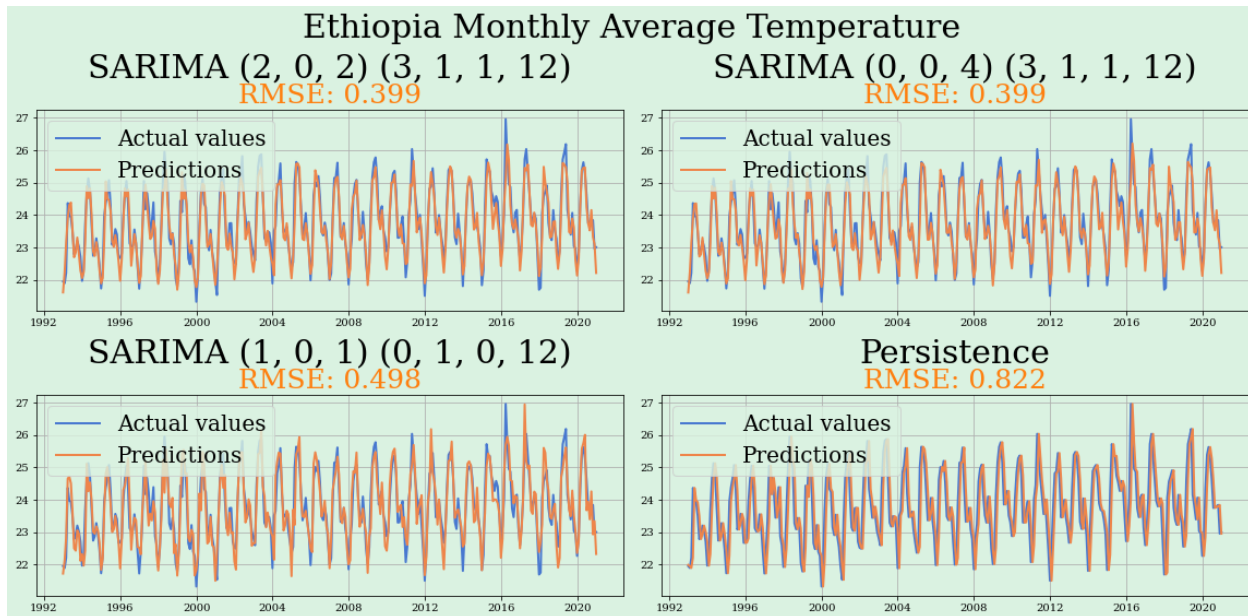
Figure 4.13 Evaluation of selected models in the test set of Global MAT



The plotted Figure 4.13 above illustrates the test scores of four models, including the two best models identified through grid search, a simple model, and the persistence model, applied to the test set of the Global monthly average temperature. The models are evaluated based on their predictive performance on unseen data.

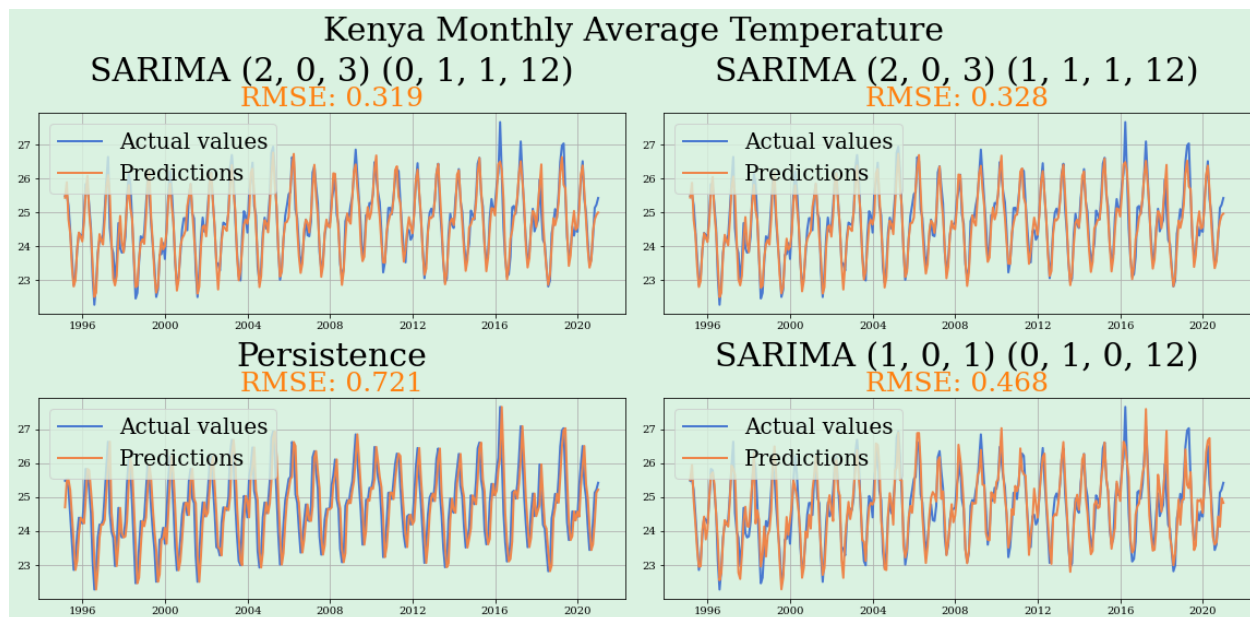
Upon examination of the test scores, it is evident that the SARIMA(4, 0, 0) (0, 1, 1, 12) model outperforms the other models, including the simple model and the persistence model, based on both test and validation scores. This conclusion is drawn from the comparison of the test scores associated with each model, indicating the superior predictive accuracy of the SARIMA(4, 0, 0) (0, 1, 1, 12) model in forecasting global monthly average temperatures.

Figure 4.14 Evaluation of selected models in the test set of Ethiopia MAT



The visual representation in Figure 4.14 juxtaposes the predictive capabilities of four distinct models applied to forecast Ethiopia's monthly average temperature, employing the Root Mean Square Error (RMSE) metric as the yardstick for evaluation. Notably, among the models scrutinized, SARIMA(2,0,2)(3,1,1,12) and SARIMA(0,0,4)(3,1,1,12) emerge as the top performers, both yielding an identical RMSE score of 0.399. This parity in predictive accuracy underscores the efficacy of these models in capturing the intricate temperature patterns inherent in Ethiopia's climate. The figure provides a succinct yet insightful glimpse into the comparative performance of these models.

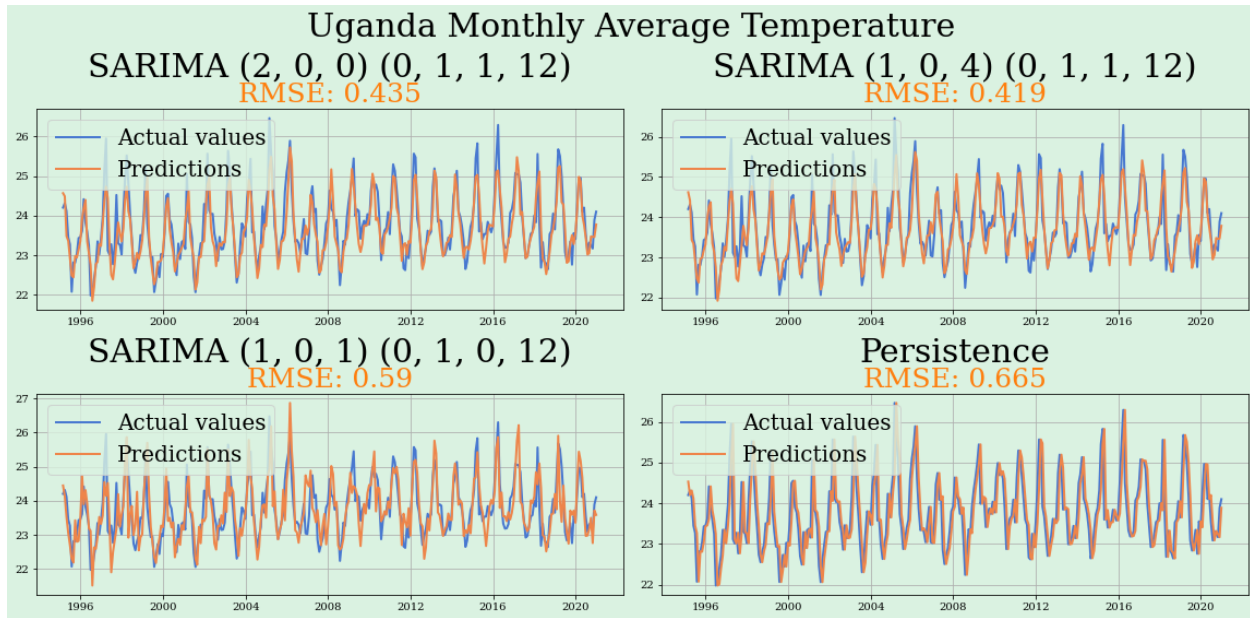
Figure 4.15 Evaluation of selected models in the test set of Kenya MAT



The above Figure 4.15 presents a straightforward comparison of four distinct models applied to forecast Kenya's monthly average temperature, utilizing the Root Mean Square Error (RMSE) as the metric for assessment. Notably, SARIMA(2,0,3)(0,1,1,12) emerges as the frontrunner among the models evaluated, boasting the lowest RMSE score of 0.319. This clear-cut depiction underscores the superior predictive accuracy of SARIMA(2,0,3)(0,1,1,12) in capturing Kenya's temperature dynamics, making it the preferred choice for forecasting purposes.

The illustration Figure 4.16 Evaluation of selected models in the test set of Uganda MAT below offers a straightforward comparison of four models applied to forecast Uganda's monthly average temperature, focusing on the Root Mean Square Error (RMSE) metric for evaluation. Among these models, SARIMA(1,0,4)(0,1,1,12) stands out as the top performer, boasting the lowest RMSE score of 0.419. This clear presentation underscores SARIMA(1,0,4)(0,1,1,12)'s superior predictive capability in capturing Uganda's temperature trends accurately. As such, it emerges as the preferred choice for forecasting Uganda's monthly average temperature due to its notable performance in the test set.

Figure 4.16 Evaluation of selected models in the test set of Uganda MAT



The selection process for the best model across four different regions was precisely conducted, considering both validation and test scores, as depicted in the bar plot. In Figure 0.17, SARIMA(4,0,0)(0,1,1,12) emerges as the optimal choice for global Monthly Average Temperature (MAT), exhibiting impressive validation and test scores of 0.11 and 0.12, respectively. Transitioning to Figure 0.18, SARIMA(2,0,2)(3,1,1,12) stands out as the most suitable model for Ethiopia MAT, showcasing a validation score of 0.35 and a test score of 0.4. Within Figure 0.19, the models for Kenya MAT are scrutinized, where SARIMA(2,0,3)(0,1,1,12) demonstrates commendable performance with a validation score of 0.3 and a test score of 0.32, positioning it as the optimal model. Lastly, in Figure 0.20, SARIMA(1,0,4)(0,1,1,12) emerges as the preferred choice for Uganda MAT, characterized by its lowest validation score of 0.37 and a corresponding test score of 0.42. This systematic approach ensures the selection of robust models tailored to the unique temperature patterns of each region.

4.3.3 Parameter Estimation

The parameter estimation process plays a crucial role in fitting the SARIMA model to the observed data. Through maximum likelihood estimation (MLE), the objective is to determine the parameter values that maximize the likelihood function of the model given the observed data. Essentially, MLE seeks to find the parameter values that make the observed data most probable under the assumed SARIMA model. By maximizing the likelihood function, the model parameters are

optimized to provide the best possible fit to the data. This iterative process involves evaluating different parameter combinations and selecting the ones that result in the highest likelihood of observing the given data. Ultimately, MLE enables us to obtain parameter estimates that accurately capture the underlying dynamics of the time series data, facilitating reliable forecasting and analysis.

Table 4.3 Parameter estimation of SARIMA(4,0,0)(0,1,1,12) for Global MAT

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.5233	0.020	25.922	0.000	0.484	0.563
ar.L2	0.1215	0.025	4.845	0.000	0.072	0.171
ar.L3	0.0249	0.028	0.897	0.370	-0.030	0.079
ar.L4	0.1479	0.025	5.865	0.000	0.098	0.197
ma.S.L12	-0.9289	0.011	-85.462	0.000	-0.950	-0.908
sigma2	0.0167	0.000	37.319	0.000	0.016	0.018

The parameter estimation for the SARIMA(4,0,0)(0,1,1,12) model, as detailed in Table 4.3, provides insights into the model's fit for the time series data of global monthly average temperatures.

The autoregressive parameters (ar) show significant coefficients for three out of four lags. The first lag (ar.L1) has a coefficient of 0.5233 with a very small standard error (0.020) and a highly significant z-value of 25.922 ($p < 0.001$). This indicates a strong positive influence of the first lag on the current value. The second lag (ar.L2) has a coefficient of 0.1215 (standard error of 0.025, z-value of 4.845, $p < 0.001$), also suggesting a significant but smaller positive effect. The third lag (ar.L3) is not statistically significant (coefficient of 0.0249, $p = 0.370$), indicating it does not contribute meaningfully to the model. The fourth lag (ar.L4) has a coefficient of 0.1479 (standard error of 0.025, z-value of 5.865, $p < 0.001$), showing a significant positive effect.

The seasonal moving average parameter (ma.S.L12) has a coefficient of -0.9289 with a very small standard error (0.011) and a highly significant negative z-value of -85.462 ($p < 0.001$). This indicates a strong seasonal component, with the twelfth lag having a substantial negative influence on the current value, which is typical in seasonal time series data.

The variance of the error term (sigma2) is estimated at 0.0167 with a highly significant z-value of 37.319 ($p < 0.001$), confirming that the model captures the variability in the data well.

The SARIMA(4,0,0) (0,1,1,12) model can be written as:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$$

$$y_t = 0.5233y_{t-1} + 0.1215y_{t-2} + 0.0249y_{t-3} + 0.1479y_{t-4} + \varepsilon_t - 0.9289\varepsilon_{t-12}$$

Overall, the SARIMA(4,0,0)(0,1,1,12) model's parameters indicate a strong autoregressive structure, significant seasonal effects, and a well-fitted model for predicting global monthly average temperatures.

Table 4.4 Parameter estimation of SARIMA(2,0,2)(3,1,1,12) for Ethiopia MAT

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	1.0394	0.206	5.038	0.000	0.635	1.444
ar.L2	-0.3597	0.129	-2.782	0.005	-0.613	-0.106
ma.L1	-0.4596	0.208	-2.211	0.027	-0.867	-0.052
ma.L2	0.1915	0.040	4.744	0.000	0.112	0.271
ar.S.L12	0.1309	0.026	4.944	0.000	0.079	0.183
ar.S.L24	0.1112	0.027	4.082	0.000	0.058	0.165
ar.S.L36	0.0844	0.029	2.906	0.004	0.027	0.141
ma.S.L12	-0.9306	0.016	-58.086	0.000	-0.962	-0.899
sigma2	0.1587	0.005	29.405	0.000	0.148	0.169

The parameter estimation for the SARIMA(2,0,2)(3,1,1,12) model for Ethiopia MAT, presented in Table 4.4, reveals key insights into the dynamics of Ethiopia's monthly average temperature data.

The autoregressive (AR) parameters for the non-seasonal component indicate significant relationships at both the first and second lags. The first lag (ar.L1) has a coefficient of 1.0394 with a standard error of 0.206, resulting in a z-value of 5.038 ($p < 0.001$). This suggests a strong positive influence of the first lag on the current value. In contrast, the second lag (ar.L2) shows a negative coefficient of -0.3597 (standard error of 0.129, z-value of -2.782, $p = 0.005$), indicating a significant negative influence.

The moving average (MA) parameters for the non-seasonal component are also significant. The first lag (ma.L1) has a coefficient of -0.4596 with a standard error of 0.208, and a z-value of -2.211 ($p = 0.027$), indicating a significant negative effect. The second lag (ma.L2) has a coefficient of 0.1915 with a standard error of 0.040, and a z-value of 4.744 ($p < 0.001$), indicating a significant positive effect.

The seasonal autoregressive (AR) parameters at lags 12, 24, and 36 months are all positive and significant, suggesting a substantial seasonal influence. Specifically, ar.S.L12 has a coefficient of

0.1309 (standard error of 0.026, z-value of 4.944, $p < 0.001$), ar.S.L24 has a coefficient of 0.1112 (standard error of 0.027, z-value of 4.082, $p < 0.001$), and ar.S.L36 has a coefficient of 0.0844 (standard error of 0.029, z-value of 2.906, $p = 0.004$).

The seasonal moving average parameter (ma.S.L12) is strongly negative with a coefficient of -0.9306 (standard error of 0.016, z-value of -58.086, $p < 0.001$), indicating a significant and substantial negative seasonal effect.

The variance of the error term (sigma2) is estimated at 0.1587 with a highly significant z-value of 29.405 ($p < 0.001$), suggesting that the model adequately captures the variability in the temperature data.

The SARIMA(4,0,0) (0,1,1,12) model can be written as:

$$y_t = 1.0394y_{t-1} - 0.3597y_{t-2} - 0.4596\varepsilon_{t-1} + 0.1915\varepsilon_{t-2} + 0.1309y_{t-12} + 0.1112y_{t-24} + 0.0844y_{t-36} - 0.9306\varepsilon_{t-12} + \varepsilon_t$$

Overall, the SARIMA(2,0,2)(3,1,1,12) model's parameters for Ethiopia MAT indicate a well-fitted model with significant autoregressive and moving average components, both at the non-seasonal and seasonal levels, effectively capturing the temperature dynamics for accurate forecasting.

Table 4.5 Parameter estimation of SARIMA(2,0,3)(0,1,1,12) for Kenya MAT

	coef	std err	z	P> z	0.025	0.975
ar.L1	1.5583	0.093	16.724	0.000	1.376	1.741
ar.L2	-0.5687	0.086	-6.587	0.000	-0.738	-0.399
ma.L1	-0.9902	0.097	-10.169	0.000	-1.181	-0.799
ma.L2	0.1073	0.051	2.112	0.035	0.008	0.207
ma.L3	-0.0593	0.042	-1.416	0.157	-0.141	0.023
ma.S.L12	-0.9331	0.013	-72.289	0.000	-0.958	-0.908
sigma2	0.1007	0.004	26.809	0.000	0.093	0.108

Table 4.5 provides the parameter estimates for the SARIMA(2,0,3)(0,1,1,12) model applied to the monthly average temperature data for Kenya. This model includes autoregressive (AR) and moving average (MA) components along with a seasonal MA component.

The SARIMA(2,0,3) (0,1,1,12) model can be written as:

$$y_t = 1.5583y_{t-1} - 0.5687y_{t-2} - 0.9902\varepsilon_{t-1} + 0.1073\varepsilon_{t-2} - 0.0593\varepsilon_{t-3} - 0.9331\varepsilon_{t-12} + \varepsilon_t$$

The non-seasonal AR parameters show significant effects at both the first and second lags. The coefficient for the first lag (ar.L1) is 1.5583 with a standard error of 0.093, yielding a z-value of 16.724 ($p < 0.001$). This indicates a strong positive influence of the first lag on the current temperature value. The second lag (ar.L2) has a negative coefficient of -0.5687 (standard error of 0.086, z-value of -6.587, $p < 0.001$), suggesting a significant negative influence on the current temperature.

The non-seasonal MA parameters include three lags. The first lag (ma.L1) has a highly significant negative coefficient of -0.9902 with a standard error of 0.097 (z-value of -10.169, $p < 0.001$). The second lag (ma.L2) is positive and significant with a coefficient of 0.1073 (standard error of 0.051, z-value of 2.112, $p = 0.035$). The third lag (ma.L3) has a negative coefficient of -0.0593 with a standard error of 0.042, but it is not statistically significant (z-value of -1.416, $p = 0.157$).

The seasonal MA component at lag 12 (ma.S.L12) is highly significant and negative, with a coefficient of -0.9331 (standard error of 0.013, z-value of -72.289, $p < 0.001$). This indicates a strong seasonal effect with a 12-month lag.

The variance of the error term (sigma2) is estimated at 0.1007 with a z-value of 26.809 ($p < 0.001$), signifying a well-fitted model capturing the variability in the temperature data.

In summary, the SARIMA(2,0,3)(0,1,1,12) model for Kenya effectively captures the key dynamics of the monthly average temperature data, with significant AR and MA components both at non-seasonal and seasonal levels. This model can be utilized to make reliable forecasts and understand the underlying temperature patterns in Kenya.

Table 4.6 Parameter estimation of SARIMA(1,0,4)(0,1,1,12) for Uganda MAT

	coef	std err	Z	P> z	0.025	0.975
ar.L1	0.9869	0.008	119.668	0.000	0.971	1.003
ma.L1	-0.5528	0.028	-20.021	0.000	-0.607	-0.499
ma.L2	-0.0919	0.031	-3.005	0.003	-0.152	-0.032
ma.L3	-0.1466	0.032	-4.600	0.000	-0.209	-0.084
ma.L4	-0.0963	0.029	-3.348	0.001	-0.153	-0.040
ma.S.L12	-0.9724	0.011	-90.902	0.000	-0.993	-0.951
sigma2	0.1418	0.005	27.074	0.000	0.132	0.152

Table 4.6 provides the parameter estimates for the SARIMA(1,0,4)(0,1,1,12) model applied to the monthly average temperature data for Uganda. This model incorporates autoregressive (AR) and moving average (MA) components along with a seasonal MA component.

The non-seasonal AR parameter at lag 1 (ar.L1) is highly significant, with a coefficient of 0.9869 and a standard error of 0.008, yielding a z-value of 119.668 ($p < 0.001$). This indicates a strong positive influence of the first lag on the current temperature value.

The non-seasonal MA parameters include four lags. The first lag (ma.L1) has a significant negative coefficient of -0.5528 with a standard error of 0.028 (z-value of -20.021, $p < 0.001$). The second lag (ma.L2) is also significant with a negative coefficient of -0.0919 (standard error of 0.031, z-value of -3.005, $p = 0.003$). The third lag (ma.L3) is significant with a negative coefficient of -0.1466 (standard error of 0.032, z-value of -4.600, $p < 0.001$). The fourth lag (ma.L4) has a negative coefficient of -0.0963, which is also significant (standard error of 0.029, z-value of -3.348, $p = 0.001$).

The seasonal MA component at lag 12 (ma.S.L12) is highly significant and negative, with a coefficient of -0.9724 and a standard error of 0.011 (z-value of -90.902, $p < 0.001$). This indicates a strong seasonal effect with a 12-month lag.

The variance of the error term (sigma2) is estimated at 0.1418 with a z-value of 27.074 ($p < 0.001$), indicating a well-fitted model that captures the variability in the temperature data.

The SARIMA(1,0,4) (0,1,1,12) model can be written as:

$$y_t = 0.9869y_{t-1} - 0.5528\varepsilon_{t-1} - 0.0919\varepsilon_{t-2} - 0.1466\varepsilon_{t-3} - 0.0963\varepsilon_{t-4} - 0.9724\varepsilon_{t-12} + \varepsilon_t$$

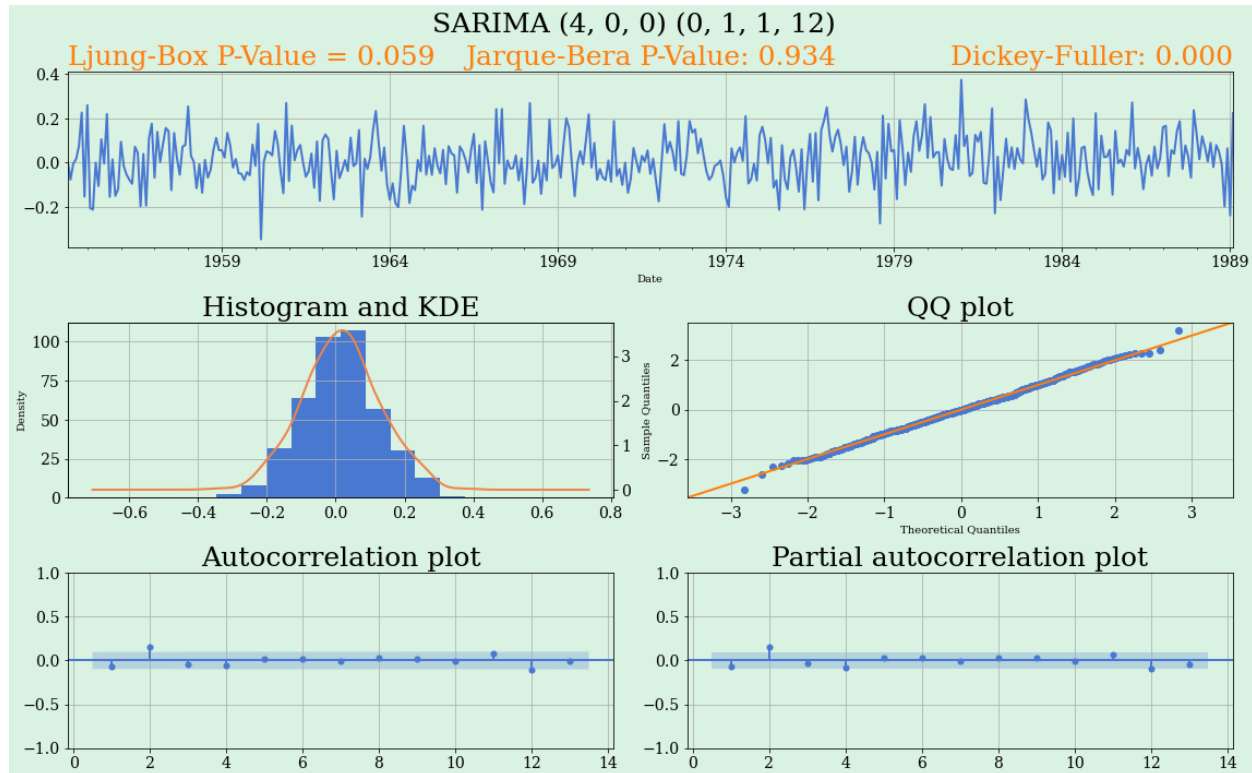
In summary, the SARIMA(1,0,4)(0,1,1,12) model for Uganda effectively captures the key dynamics of the monthly average temperature data, with significant AR and MA components both at non-seasonal and seasonal levels. This model provides reliable forecasts and a comprehensive understanding of the temperature patterns in Uganda.

4.3.4 Diagnostic Checking

The diagnostic analysis of the SARIMA(4,0,0)(0,1,1) [12] model aimed to evaluate its goodness of fit and identify any potential model deficiencies. Two primary diagnostic tests were conducted:

the Ljung-Box test for serial correlation of residuals and the Jarque-Bera test for normality of residuals.

Figure 4.17 Residual analysis of SARIMA(4,0,0)(0,1,1,12) for Global MAT

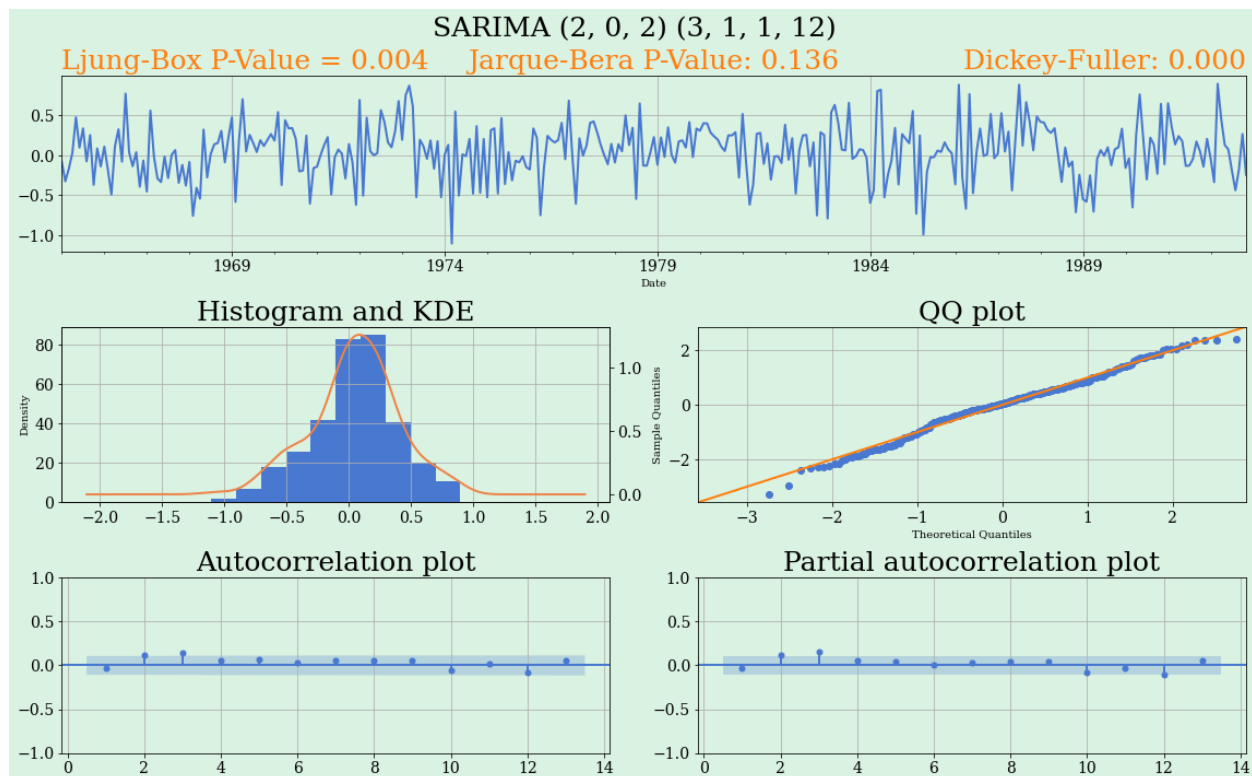


Upon examining the results, several key observations were made. The Ljung-Box test yielded a p-value of 0.059, indicating that there is no significant serial correlation present in the residuals up to lag 20. The Jarque-Bera test's p-value of 0.934 indeed suggests that there is no significant evidence to reject the null hypothesis of normality for the residuals. Therefore, it indicates that the residuals can reasonably be considered to follow a normal distribution. Additionally, the Fuller test for stationarity yielded a p-value of 0.000, indicating strong evidence against the null hypothesis of non-stationarity.

Further examination of the residuals' distribution through histogram and kernel density estimation (KDE) plots indicated a close adherence to a normal distribution, bolstering confidence in the model's appropriateness. Moreover, the quantile-quantile (qq) plot revealed that the majority of residuals aligned closely with the diagonal reference line, suggesting normality. Although two outliers were observed at the extremes of the qq plot, the overall pattern remained consistent with normality.

Finally, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were examined to detect any significant residual correlations. Most lag values fell within the expected confidence intervals, indicating no substantial autocorrelation. However, a slight deviation was observed at lag 2, where the correlation exceeded the expected bounds in both the ACF and PACF plots. While this suggests a potential area for improvement, the overall diagnostic analysis indicates that the SARIMA(4,0,0)(0,1,1,12) model performs satisfactorily, with minor discrepancies noted in the residual behavior at lag 2.

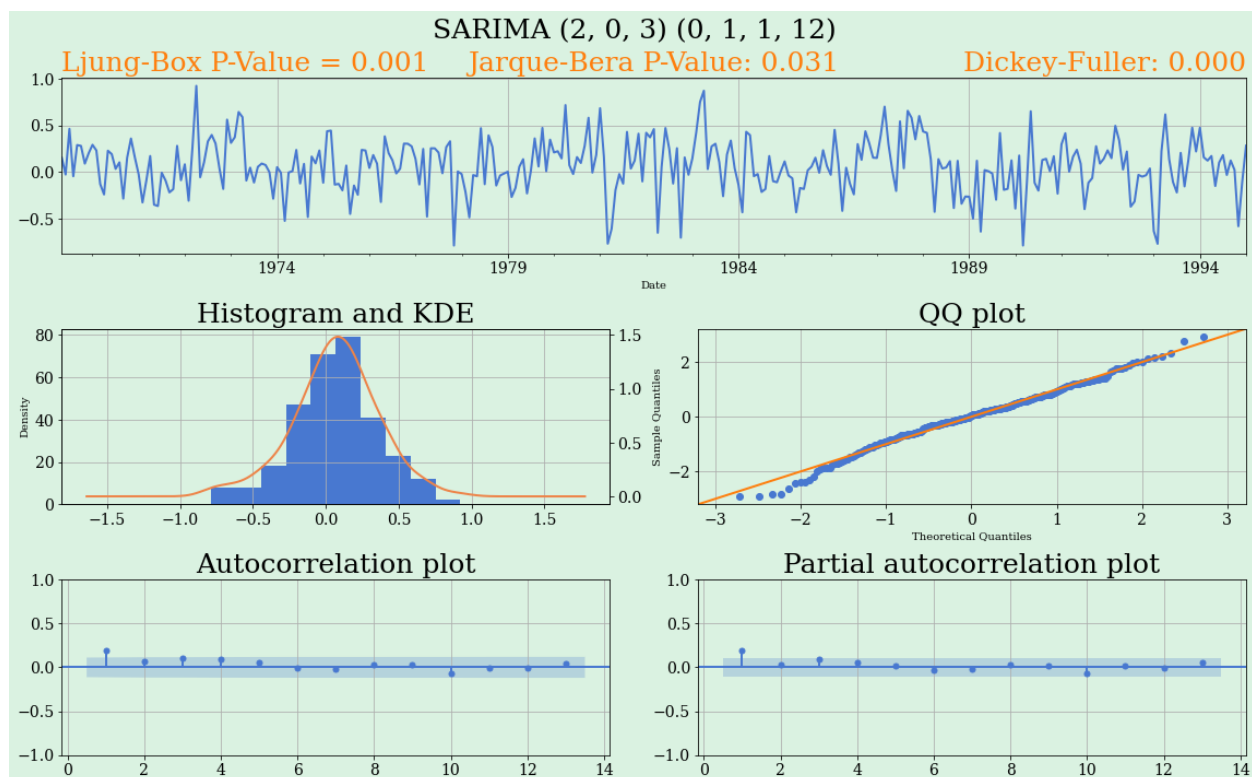
Figure 4.18 Residual analysis of SARIMA(2,0,2)(3,1,1,12) for Ethiopia MAT



The diagnostic evaluation of the SARIMA(2,0,2)(3,1,1) [12] models aimed to assess their adequacy in fitting the Ethiopia monthly average temperature and uncover any potential issues. Two crucial diagnostic tests, focusing on the serial correlation and normality of residuals, were conducted to gauge the model's performance. The analysis revealed insightful findings: the Fuller test yielded a p-value of 0.000, indicating strong evidence against the null hypothesis of non-stationarity, affirming the stationarity of residuals. Meanwhile, the Jarque-Bera test yielded a p-value of 0.136, suggesting that the residuals follows Gaussian distribution. Additionally, the Ljung-Box test produced a p-value of 0.004, indicating significant evidence against the null hypothesis

of no autocorrelation, suggesting the presence of residual correlation at certain lags. However, visual inspection of the histogram, KDE plot, QQ plot, and autocorrelation plots revealed promising outcomes. The histogram and KDE plot showcased a distribution closely resembling normality, while the QQ plot exhibited alignment with the diagonal line, indicating conformity to normal distribution assumptions. Furthermore, the ACF and PACF plots demonstrated that most residual lags fell within acceptable bounds, albeit with minor deviations. Overall, while minor issues such as residual correlation and slight departure from normality were observed, the SARIMA(2,0,2)(3,1,1) [12] models exhibited satisfactory performance in capturing the Ethiopia monthly average temperature temporal dynamics.

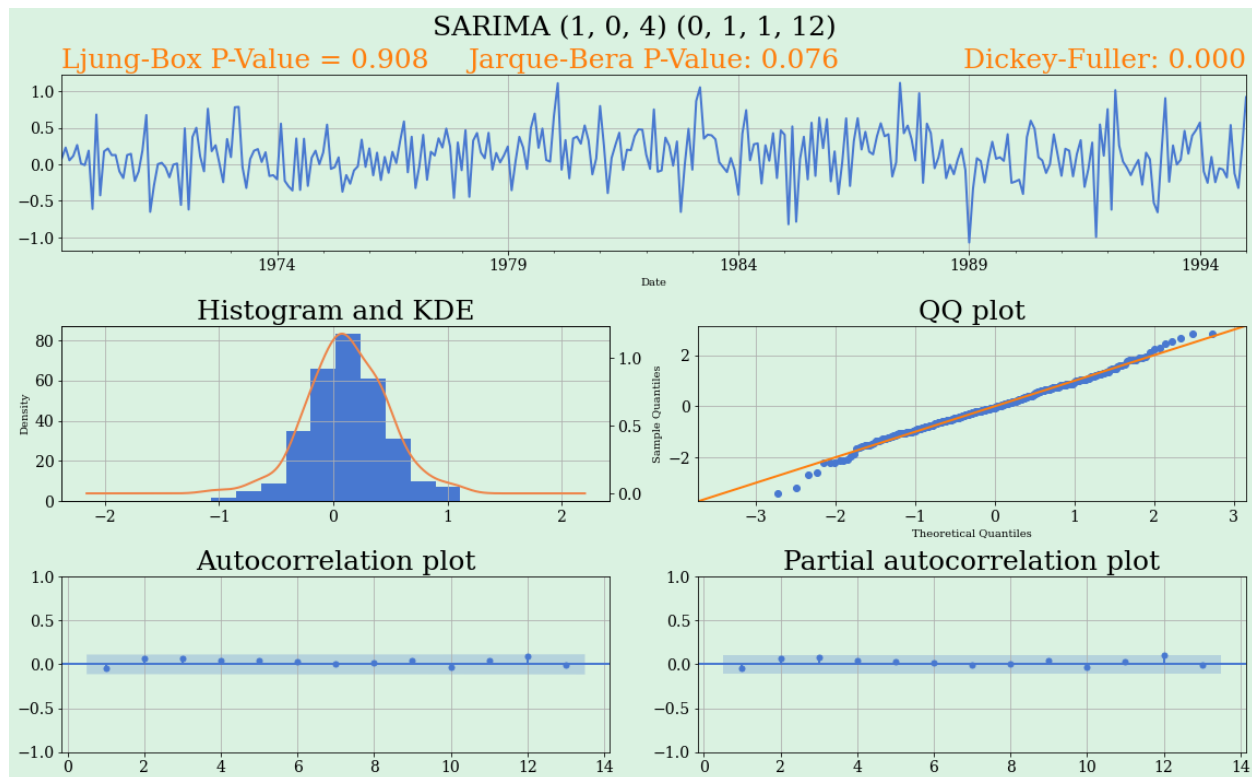
Figure 4.19 Residual analysis of SARIMA(2,0,3)(0,1,1,12) for Kenya MAT



The diagnostic assessment conducted on the chosen SARIMA(2,0,3)(0,1,1) [12] models for the monthly average temperature in Kenya served as a pivotal step in evaluating the model's adequacy and identifying potential limitations. By subjecting the residuals to rigorous scrutiny, key statistical tests were employed to scrutinize their behavior. The results revealed significant outcomes, with the Fuller test yielding a p-value of 0.000, indicating strong evidence against the null hypothesis of non-stationarity. Similarly, the Jarque-Bera test yielded a p-value of 0.031, implying some

deviation from normality, albeit not statistically significant. Additionally, the Ljung-Box test produced a p-value of 0.001, indicative of no significant autocorrelation in the residuals. Visual inspection further reinforced these findings, with the histogram and Kernel Density Estimation (KDE) exhibiting a symmetrical distribution resembling normality. Furthermore, the Quantile-Quantile (QQ) plot demonstrated a linear relationship between the observed and theoretical quantiles, reinforcing the normality assumption. However, slight discrepancies were noted in the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, particularly with lag 1 showing a slight departure from the expected range. Despite this, overall diagnostic evaluation suggests that the SARIMA model adequately captures the temporal dynamics of Kenya's monthly temperature averages, albeit with minor areas for improvement.

Figure 4.20 Residual analysis of SARIMA(1,0,4)(0,1,1,12) for Uganda MAT



The diagnostic evaluation conducted on the chosen SARIMA(1,0,4)(0,1,1) [12] models for Uganda's monthly average temperature played a crucial role in scrutinizing the model's adequacy and uncovering potential shortcomings. Through comprehensive statistical tests, the behavior of residuals was precisely examined to glean insights into the model's performance. Notably, the Fuller test yielded a strikingly low p-value of 0.000, indicative of strong evidence against the null

hypothesis of non-stationarity, affirming the model's ability to capture temporal patterns effectively. Additionally, the Jarque-Bera test returned a p-value of 0.076, suggesting no significant departure from normality, albeit with some marginal deviation. Moreover, the Ljung-Box test produced a p-value of 0.908, indicating no discernible autocorrelation in the residuals. Visually, the diagnostic plots further bolstered these findings, with the histogram and Kernel Density Estimation (KDE) exhibiting a symmetric distribution resembling normality. Furthermore, the Quantile-Quantile (QQ) plot demonstrated a close alignment between observed and theoretical quantiles, supporting the normality assumption. Notably, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots revealed lags well within the expected intervals, underscoring the model's robustness in capturing Uganda's monthly temperature variations accurately. Overall, the diagnostic assessment portrays the SARIMA model as a reliable tool for forecasting Uganda's temperature dynamics, with minimal discernible issues warranting attention.

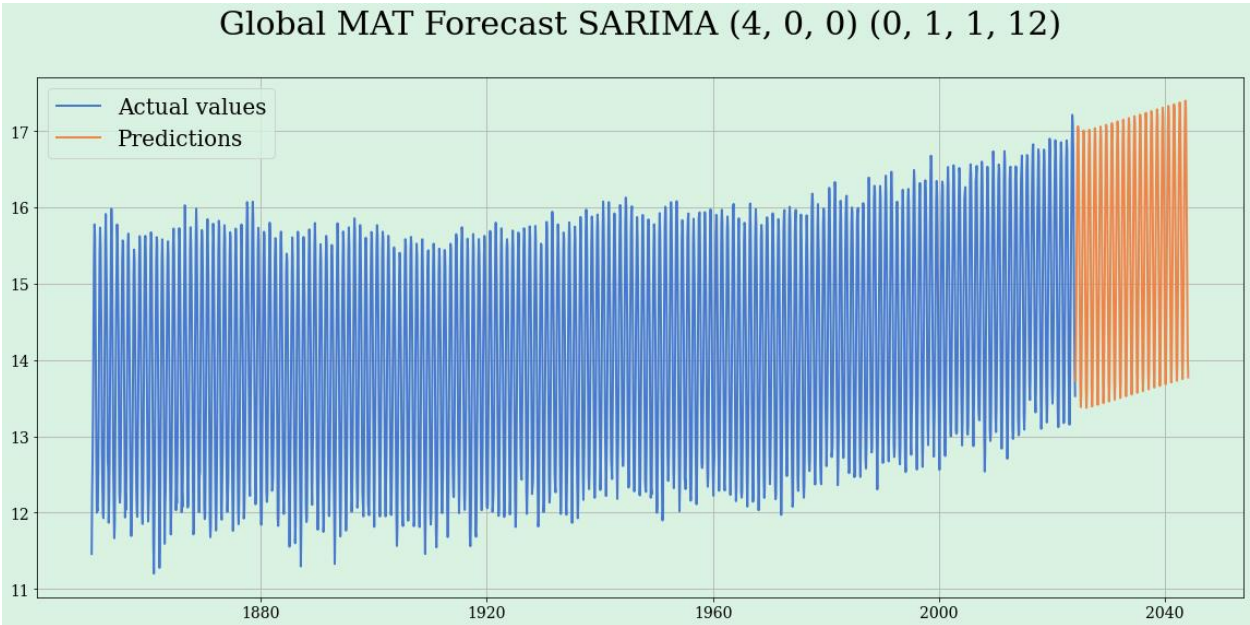
4.3.5 Forecasting

The culmination of time series modeling lies in the crucial step of forecasting. This process involves estimating values in unknown circumstances, a fundamental technique extensively employed in the analysis of time-series data. Upon completing the rigorous stages of parameter estimation and diagnostic checking, the carefully selected SARIMA model emerges as a potent tool for forecasting future trends. Forecasts are pivotal in both sample and post-sample periods, each serving distinct purposes. Sample period forecasts are instrumental in instilling confidence in the model's efficacy, while post-sample forecasts yield tangible predictions for desired outcomes. In our analysis, we focus on predicting the monthly average temperature for the next two decades. The sample period forecasts, as detailed in Section 4.3.2 on selecting SARIMA models, showcase the model's adeptness in predicting both validation and test sets across all four regions. Subsequently, we present the post-sample period forecasts, offering insights into anticipated trends beyond the sampled data timeframe.

Figure 4.21 illustrates the forecast of the global monthly average temperature for the forthcoming two decades, revealing a discernible upward trend. The forecast highlights key metrics: the maximum temperature of 17.39696042 degrees Celsius is projected for 7/31/2043, while the minimum temperature of 13.37399342 degrees Celsius is anticipated for 1/31/2026. On average,

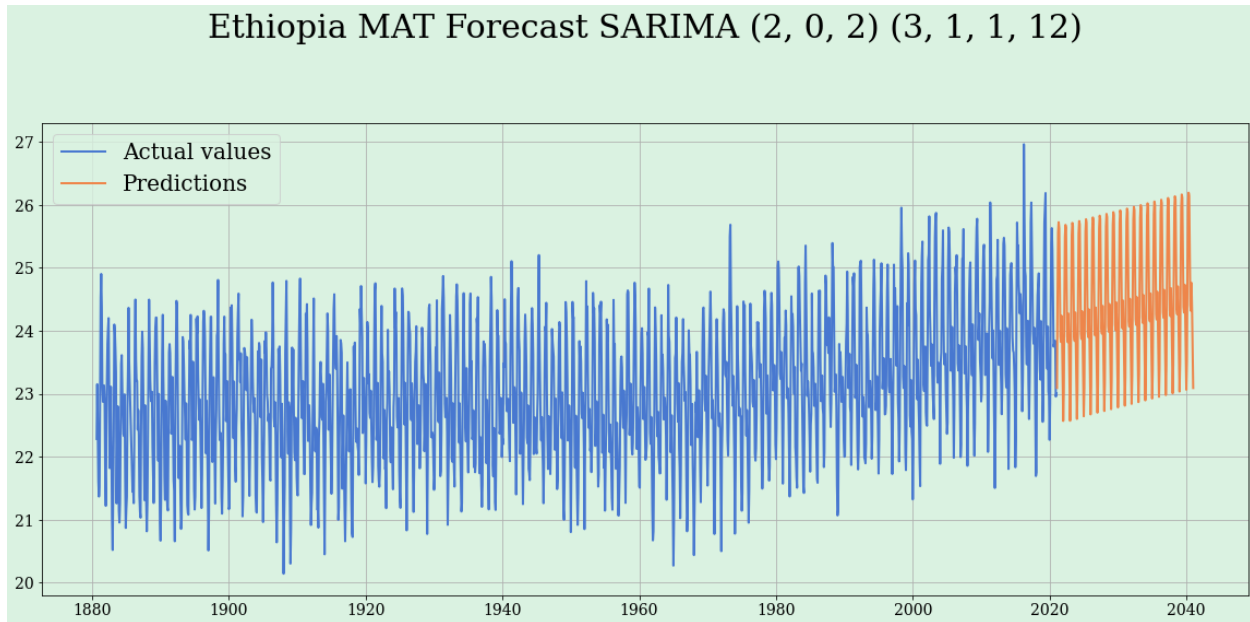
a temperature of 15.41471211 degrees Celsius is forecasted over the next two decades. These insights provide valuable foresight into the anticipated temperature patterns.

Figure 4.21 Forecast of Global MAT



In Figure 4.22 below, we embark on a journey into the predictive realm of Ethiopia's monthly average temperature over the next two decades. As the plot unfolds before us, it unveils a compelling narrative of an upward trajectory in temperature values, hinting at a future marked by rising thermometric indices. With each passing month, we witness the temperature scale new heights, reaching a maximum of 26.192230 degrees Celsius, while also observing dips to a minimum of 22.570154 degrees Celsius. Amidst this fluctuating landscape, the average temperature over the next two decades settles at a moderate 24.461581 degrees Celsius, embodying the essence of equilibrium amidst variability. Through this lens, we glean insights into the intricate dance of temperature dynamics, painting a vivid picture of what the future may hold for Ethiopia's climatic tapestry.

Figure 4.22 Forecast of Ethiopia MAT



In our study of Kenya's monthly temperatures, Figure 4.23 paints a vivid picture of what's to come over the next two decades. It suggests things are heating up, with temperatures expected to reach a peak of 27.25°C. But there are cooler spells too, with lows dipping to 23.52°C. On average, expect temperatures to settle around 25.35°C. This journey through Kenya's climate offers valuable insights into its future path.

Figure 4.23 Forecast of Kenya MAT

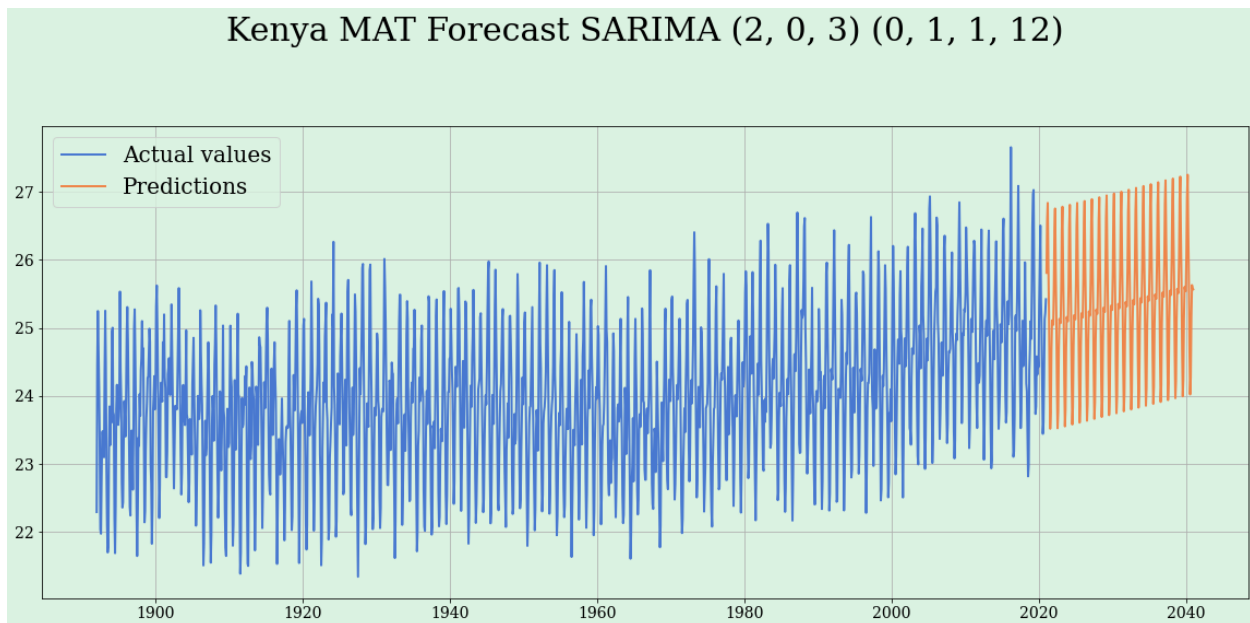
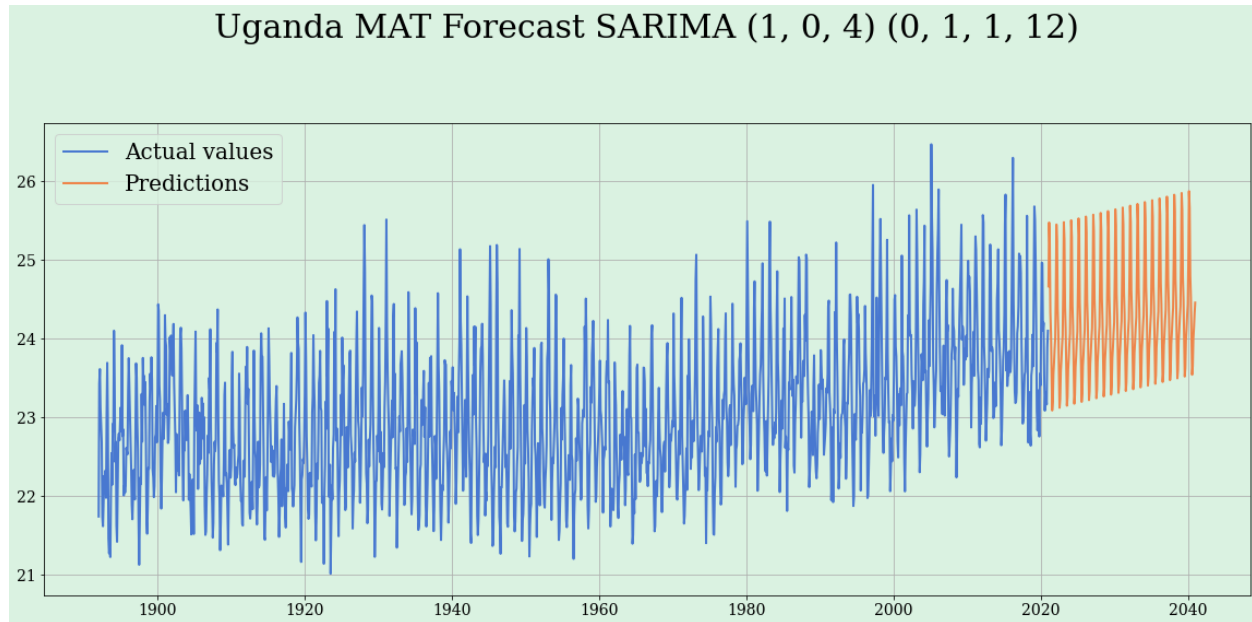


Figure 4.24 Forecast of Uganda MAT



In our analysis of Uganda's monthly temperatures, we anticipate a range of climatic conditions over the next two decades. The forecast suggests that temperatures could climb to a high of 25.87°C, while also experiencing cooler periods with lows around 23.09°C. On average, temperatures are projected to hover around 24.28°C. This forecast provides valuable insights into the future climate trends of Uganda.

4.3.6 Forecasting Accuracy

The forecasting accuracy of the SARIMA models for global and East African monthly average temperatures is evaluated using three standard metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics provide a comprehensive assessment of the models' predictive performance by measuring the average magnitude of errors, the square root of the average squared errors, and the average percentage errors, respectively.

Table 4.7 Forecasting accuracy metrics

Regions	Model	Metrics		
		MAE	RAMSE	MAPE
Global	SARIMA(4,0,0)(0,1,1,12)	0.0939	0.1169	0.6442
Ethiopia	SARIMA(2,0,2)(3,1,1,12)	0.3043	0.3988	1.2733

Kenya	SARIMA(2,0,3)(0,1,1,12)	0.2417	0.3188	0.9733
Uganda	SARIMA(1,0,4)(0,1,1,12)	0.3360	0.4352	1.4016

The SARIMA model for global temperatures demonstrates high forecasting accuracy, as indicated by the low MAE, RMSE, and MAPE values. The MAE of 0.0939 suggests that, on average, the forecasted temperatures deviate from the observed values by approximately 0.0939°C. The RMSE of 0.1169 further corroborates the model's accuracy, indicating that the typical deviation of the forecast errors is around 0.1169°C. The MAPE of 0.6442% reflects that the model's forecasted temperatures are, on average, within 0.6442% of the actual values, highlighting the model's robustness in capturing global temperature trends and variability.

The forecasting accuracy for Ethiopia's temperatures, while reasonably good, is lower compared to the global model. The MAE of 0.3043 indicates an average deviation of about 0.3043°C between the forecasted and observed temperatures. The RMSE of 0.3988 suggests a typical error magnitude of 0.3988°C, slightly higher than the MAE, indicating the presence of some larger forecast errors. The MAPE of 1.2733% shows that the forecasted temperatures are, on average, within 1.2733% of the actual values. This indicates a satisfactory level of accuracy, but also suggests that the model could benefit from further refinement to better capture the underlying temperature dynamics in Ethiopia.

The SARIMA model for Kenya exhibits relatively strong forecasting performance. The MAE of 0.2417 indicates that the forecasted temperatures deviate from the observed values by approximately 0.2417°C on average. The RMSE of 0.3188 suggests that the typical forecast error magnitude is around 0.3188°C. The MAPE of 0.9733% implies that the forecasted temperatures are, on average, within 0.9733% of the actual values. These metrics reflect a good model fit and accurate temperature predictions for Kenya, demonstrating the model's ability to effectively capture the seasonal and trend components of the temperature series.

The forecasting accuracy for Uganda's temperatures is slightly lower compared to Kenya and Ethiopia. The MAE of 0.3360 indicates an average forecast error of 0.3360°C. The RMSE of 0.4352 suggests that the typical error magnitude is about 0.4352°C, which is higher than both the MAE and RMSE for Kenya and Ethiopia. The MAPE of 1.4016% shows that the forecasted temperatures are, on average, within 1.4016% of the actual values. These results suggest that while

the model provides a reasonable forecast, there is room for improvement in capturing the unique temperature dynamics of Uganda.

When comparing the forecasting accuracy across the regions, the global SARIMA model performs the best, with the lowest MAE, RMSE, and MAPE values. This high accuracy is likely due to the larger dataset and more pronounced seasonal and trend components in global temperatures, which the model can effectively capture. Among the East African countries, Kenya's SARIMA model shows the highest accuracy, followed by Ethiopia and Uganda. The differences in forecasting performance can be attributed to regional climatic variations, data quality, and the inherent complexity of local temperature patterns.

The SARIMA models demonstrate varying degrees of forecasting accuracy for global and regional temperatures. The global model achieves the highest accuracy, highlighting its robustness in capturing large-scale temperature trends. For East African countries, the models perform reasonably well, with Kenya's model showing the best performance among the three. However, the higher error metrics for Ethiopia and Uganda suggest that additional refinements, such as incorporating exogenous variables or more sophisticated modeling techniques, could further enhance forecasting accuracy. These insights provide valuable guidance for improving temperature prediction models and developing more effective climate adaptation strategies for these regions.

4.4 Conditional Heteroscedasticity Models

4.4.1 Volatility Testing

In the realm of time series analysis, a critical aspect of model assessment involves probing the residuals for potential heteroskedasticity, particularly within the framework of SARIMA modeling. Heteroskedasticity, denoting the variance of residuals exhibiting systematic changes over time, poses a significant concern as it can compromise the reliability of forecasting outcomes. To scrutinize this phenomenon, researchers often turn to Autoregressive Conditional Heteroskedasticity (ARCH) effect tests. ARCH models postulate that the variance of residuals is contingent upon their past values, offering insights into temporal variations in volatility. The ARCH-LM test serves as a valuable diagnostic tool in this context, probing whether the squared residuals display serial correlation—a hallmark of heteroskedasticity. By leveraging the Lagrange Multiplier (LM) test statistic, researchers can gauge the statistical significance of ARCH effects,

thereby informing model refinement endeavors aimed at bolstering the forecasting accuracy and robustness of SARIMA models.

Table 4.8 ARCH effect test

Global		Ethiopia		Kenya		Uganda	
LM-Statistic	P-Value	LM-Statistic	P-Value	LM-Statistic	P-Value	LM-Statistic	P-Value
396.9247	0.000	506.0192	0.0000	464.4682	0.0000	469.4020	0.0000

The ARCH LM test results for the selected SARIMA models in each region provide crucial insights into the behavior of the residuals. For the global model, an LM-Statistic of 396.9247 with a p-value of 0.000 indicates strong evidence against the null hypothesis of no ARCH effects, suggesting significant heteroskedasticity in the residuals. Similar patterns are observed in the regional models for Ethiopia, Kenya, and Uganda, with LM-Statistics of 506.0192, 464.4682, and 469.4020, respectively, all accompanied by p-values of 0.0000. These results confirm the presence of ARCH effects, indicating that the residuals' variance changes over time. This heteroskedasticity implies that simple models assuming constant variance are inadequate for accurate forecasting. Therefore, incorporating ARCH or GARCH models to address this variability is essential for improving the robustness and precision of temperature forecasts across these regions.

4.4.2 Model Selection for Volatility

In the pursuit of understanding the drivers of temperature volatility, an exploration into the lag parameters suitable for ARCH and GARCH family models was initiated subsequent to the identification of ARCH effects. While the GARCH(1,1) model is often favored for its simplicity in data fitting, this study opted for a broader consideration of higher-order models. Despite the widespread preference for GARCH(1,1), higher-order models extending up to three months were included in the analysis to comprehensively assess volatility dynamics. This deliberation resulted in the selection of 36 models for each of the ARCH, GARCH, and EGARCH categories, predicated on criteria such as AIC, BIC, and adherence to normal distribution assumptions. Notably, although discernible differences existed in AIC, BIC, and log-likelihood values among the models, their overall proximity underscored the complexity of model selection and the nuanced nature of temperature volatility dynamics.

Table 4.9 Model selection for ARCH model

ARCH		Global		Ethiopia		Kenya		Uganda	
	p	aic	bic	aic	bic	aic	bic	aic	bic
0	1	1519.072	1527.148	1494.035	1501.675	954.2079	961.6874	952.4116	959.8912
1	2	1521.072	1533.186	1496.035	1507.495	956.2079	967.4272	954.4116	965.631
2	3	1523.072	1539.224	1498.035	1513.315	958.2079	973.167	956.4116	971.3708
3	4	1525.072	1545.262	1500.035	1519.135	960.2079	978.9068	958.4116	977.1106

Among the ARCH models evaluated, ARCH(1) emerged as the best model across all four regions, demonstrating the lowest AIC and BIC values. Specifically, for the global temperature series, ARCH(1) exhibited an AIC of 1519.072 and a BIC of 1527.148. For Ethiopia, the AIC and BIC values were 1494.035 and 1501.675, respectively. Similarly, the ARCH(1) model for Kenya had an AIC of 954.2079 and a BIC of 961.6874, while for Uganda, the model recorded an AIC of 952.4116 and a BIC of 959.8912. These results indicate that ARCH(1) provides the most parsimonious fit among the models considered, effectively capturing the volatility dynamics in the temperature data across these regions.

Table 4.10 Model selection of GARCH model for Global and Ethiopia MAT

GARCH			Global		Ethiopia			
	p	q	aic	bic	p	q	aic	bic
0	1	1	1521.072	1533.186	1	1	1496.035	1507.495
1	1	2	1523.072	1539.224	1	2	1498.035	1513.315
2	2	1	1523.072	1539.224	2	1	1498.035	1513.315
3	2	2	1525.072	1545.261	1	3	1500.035	1519.135
4	3	1	1525.072	1545.261	2	2	1500.035	1519.135
5	1	3	1525.072	1545.261	3	1	1500.035	1519.135
6	1	4	1527.072	1551.299	3	2	1502.035	1524.956
7	3	2	1527.072	1551.299	1	4	1502.035	1524.956
8	2	3	1527.072	1551.299	2	3	1502.035	1524.956
9	4	1	1527.072	1551.299	2	4	1504.035	1530.776
10	4	2	1529.072	1557.337	3	3	1504.035	1530.776
11	3	3	1529.072	1557.337	4	2	1504.035	1530.776
12	2	4	1529.072	1557.337	4	3	1506.035	1536.596
13	4	3	1531.072	1563.375	3	4	1506.035	1536.596
14	3	4	1531.072	1563.375	4	4	1508.035	1542.416
15	4	4	1533.072	1569.413	4	1	1557.82	1580.74

Table 4.11 Model selection of GARCH model for Kenya and Uganda MAT

GARCH			Kenya		Uganda			
	p	q	aic	bic	p	q	aic	bic
0	1	1	956.2079	967.4272	1	1	954.4116	965.631
1	1	2	958.2079	973.167	1	2	956.4116	971.3708
2	2	1	958.2079	973.167	2	1	956.4116	971.3708
3	1	3	960.2079	978.9068	2	2	958.4116	977.1106
4	3	1	960.2079	978.9068	1	3	958.4116	977.1106
5	2	2	960.2079	978.9068	3	1	958.4116	977.1106
6	2	3	962.2079	984.6466	3	2	960.4116	982.8504
7	1	4	962.2079	984.6466	2	3	960.4116	982.8504
8	3	2	962.2079	984.6466	4	1	960.4116	982.8504
9	2	4	964.2079	990.3864	1	4	960.4116	982.8504
10	4	2	964.2079	990.3864	4	2	962.4116	988.5902
11	3	3	964.2079	990.3864	3	3	962.4116	988.5902
12	4	3	966.2079	996.1262	2	4	962.4116	988.5902
13	3	4	966.2079	996.1262	4	3	964.4116	994.33
14	4	4	968.2079	1001.866	3	4	964.4116	994.33
15	4	1	1031.95	1054.389	4	4	966.4116	1000.07

As shown in Table 4.10 and Table 4.11, GARCH(1,1) is identified as the best model with the lowest AIC and BIC values across all regions. For the global temperature series, GARCH(1,1) recorded an AIC of 1521.072 and a BIC of 1533.186. In Ethiopia, the model achieved an AIC of 1496.035 and a BIC of 1507.495. Similarly, for Kenya, the AIC and BIC values were 956.2079 and 967.4272, respectively, while for Uganda, the values were 954.4116 and 965.631. These results highlight that GARCH(1,1) effectively models the temperature volatility with the most parsimonious fit across the regions considered.

Table 4.12 Model selection of EGARCH model for Global and Ethiopia MAT

EGARCH			Global		Ethiopia			
	p	q	aic	bic	p	q	aic	bic
0	4	1	1447.202	1471.429	3	1	1485.514	1504.614
1	4	2	1449.202	1477.467	4	1	1485.543	1508.463
2	4	3	1451.202	1483.505	3	2	1487.514	1510.435
3	4	4	1453.202	1489.542	4	2	1487.543	1514.283
4	3	1	1496.185	1516.375	3	4	1491.514	1522.075
5	3	2	1498.185	1522.413	2	1	1491.665	1506.945
6	3	3	1500.185	1528.45	1	1	1492.099	1503.559
7	3	4	1502.185	1534.488	2	2	1493.665	1512.765
8	2	1	1502.314	1518.466	1	2	1494.099	1509.379

9	2	2	1504.314	1524.504	3	3	1494.532	1521.273
10	2	3	1506.314	1530.541	4	4	1495.292	1529.673
11	2	4	1508.314	1536.579	2	3	1495.665	1518.585
12	1	1	1509.732	1521.846	1	3	1496.099	1515.2
13	1	2	1511.732	1527.884	4	3	1496.246	1526.807
14	1	3	1513.732	1533.921	2	4	1497.665	1524.405
15	1	4	1515.732	1539.959	1	4	1498.099	1521.02

Table 4.13 Model selection of EGARCH model for Kenya and Uganda MAT

EGARCH			Kenya		Uganda			
	p	q	aic	bic	p	q	aic	bic
0	3	2	939.7578	962.1965	4	1	954.4837	976.9225
1	3	1	940.2385	958.9374	1	1	955.6215	966.8409
2	4	2	942.0157	968.1942	4	2	956.07	982.2486
3	4	1	942.2086	964.6474	2	1	957.3316	972.2908
4	1	1	943.2102	954.4296	1	2	957.6215	972.5807
5	4	3	944.1943	974.1126	3	1	957.9753	976.6743
6	2	1	944.8909	959.8501	4	3	958.3567	988.2751
7	1	2	945.2102	960.1694	2	2	959.3316	978.0305
8	3	3	946.8701	973.0486	1	3	959.6215	978.3205
9	2	2	946.8909	965.5899	3	2	959.9753	982.4141
10	1	3	947.2102	965.9092	2	3	961.3316	983.7703
11	4	4	947.2439	980.902	1	4	961.6215	984.0603
12	3	4	948.8701	978.7884	3	3	961.9753	988.1539
13	2	3	948.8909	971.3297	2	4	963.3316	989.5101
14	1	4	949.2102	971.6489	3	4	963.9753	993.8937
15	2	4	950.8909	977.0695	4	4	965.3013	998.9595

From Table 4.12 and Table 4.13, it is evident that EGARCH models provide the best fit across all regions with the lowest AIC and BIC values. For the global dataset, EGARCH(4,1) emerges as the optimal model with an AIC of 1447.202 and a BIC of 1471.429. In Ethiopia, EGARCH(3,1) is identified as the best model, recording an AIC of 1485.514 and a BIC of 1504.614. Similarly, for Kenya, EGARCH(3,2) is the most suitable model with an AIC of 939.7578 and a BIC of 962.1965. Lastly, for Uganda, EGARCH(4,1) is the preferred model, achieving an AIC of 954.4837 and a BIC of 976.9225. These results underscore the effectiveness of EGARCH models in capturing the volatility dynamics of temperature data across the different regions.

Table 4.14 Overall volatility model selection

Global				Kenya			
Model	Aic	bic	LL	Model	Aic	bic	LL
ARCH(1)	1519.07	1527.15	-757.536	ARCH(1)	954.207	961.687	-475.104
GARCH(1,1)	1521.07	1533.19	-757.536	GARCH(1,1)	956.207	967.427	-475.104
EGARCH(4,1)	1447.20	1471.43	-717.601	EGARCH(3,2)	939.757	962.196	-463.879
Ethiopia				Uganda			
Model	Aic	bic	LL	Model	Aic	bic	LL
ARCH(1)	1494.03	1501.67	-745.018	ARCH(1)	952.411	959.891	-474.206
GARCH(1,1)	1496.03	1507.49	-745.018	GARCH(1,1)	954.411	965.631	-474.206
EGARCH(3,1)	1485.51	1504.61	-737.757	EGARCH(4,1)	954.483	976.922	-471.242

In the overall model selection process among the three categories (ARCH, GARCH, and EGARCH) presented in Table 4.14, EGARCH models emerged as the most parsimonious volatility models for addressing the conditional heteroscedasticity of residuals from the SARIMA model. For the global dataset, EGARCH(4,1) was identified as the best fit, with the lowest AIC and BIC values and the highest log-likelihood value. Similarly, for Ethiopia, EGARCH(3,1) was the optimal model, while for Kenya, EGARCH(3,2) proved to be the most suitable. In Uganda, EGARCH(4,1) demonstrated the best performance. These models were selected based on their superior fit, evidenced by the lowest AIC and BIC values and the highest log-likelihood values, highlighting the effectiveness of EGARCH models in capturing the volatility dynamics across all regions.

4.4.3 Parameter Estimation of EGARCH

This pivotal step involves determining the optimal values for the model's parameters, which are essential for accurately characterizing the conditional volatility dynamics of temperature data. The estimation procedure encompasses fitting the EGARCH model to the observed data and iteratively adjusting the parameters to minimize the discrepancy between the model's predictions and the actual observations. Through this iterative optimization process, the EGARCH model seeks to capture the intricate patterns and nuances inherent in temperature volatility, thereby facilitating more precise forecasting and informed decision-making.

Table 4.15 Parameter Estimation of EGARCH(4,1) for Global MAT

	coef	std err	t	P> t	95.0% Conf. Int.
omega	0.2781	2.661e-02	10.451	1.456e-25	[0.226, 0.330]
alpha[1]	1.7176	6.067e-02	28.312	2.465e-176	[1.599, 1.837]
alpha[2]	-0.4454	8.254e-02	-5.396	6.832e-08	[-0.607, -0.284]
alpha[3]	-0.9135	5.996e-02	-15.236	2.047e-52	[-1.031, -0.796]
alpha[4]	-0.8869	7.143e-02	-12.417	2.117e-35	[-1.027, -0.747]
beta[1]	0.6103	3.340e-02	18.273	1.365e-74	[0.545, 0.676]

The parameter estimation for the EGARCH(4,1) model applied to the global monthly average temperature data reveals significant characteristics of the volatility structure. The omega (ω) parameter is 0.2781, with a highly significant t-value of 10.451 ($p < 0.001$), indicating a substantial constant variance component within the model.

The alpha parameters, which capture the influence of past shocks on current volatility, demonstrate varying effects. Alpha[1] (α_1) has a coefficient of 1.7176 and is highly significant ($p < 0.001$), suggesting a strong positive impact of past shocks on current volatility. Alpha[2] (α_2), with a coefficient of -0.4454, is significantly negative ($p < 0.001$), indicating a reduction in the impact of past shocks. Alpha[3] (α_3) and Alpha[4] (α_4) have coefficients of -0.9135 and -0.8869, respectively, both highly significant ($p < 0.001$), further suggesting a notable decrease in volatility influenced by past shocks.

The beta parameter (β_1) is 0.6103 with a highly significant t-value of 18.273 ($p < 0.001$), indicating moderate persistence in volatility, meaning that past volatility significantly affects current volatility levels but to a lesser extent compared to α_1 's impact.

Overall, the EGARCH(4,1) model effectively captures the intricate volatility dynamics in global monthly average temperatures. The model shows significant and varying influences of past shocks, coupled with moderate persistence in volatility, making it a robust tool for modeling and forecasting the conditional heteroscedasticity in the temperature data.

Table 4.16 Parameter Estimation of EGARCH(3,1) for Ethiopia MAT

	coef	std err	t	P> t	95.0% Conf. Int.
omega	0.6258	9.601e-02	6.518	7.114e-11	[0.438, 0.814]
alpha[1]	1.5470	9.680e-02	15.981	1.738e-57	[1.357, 1.737]
alpha[2]	-0.6802	7.936e-02	-8.571	1.026e-17	[-0.836, -0.525]
alpha[3]	-0.7078	6.706e-02	-10.555	4.802e-26	[-0.839, -0.576]

beta[1]	0.5813	5.487e-02	10.593	3.209e-26	[0.474, 0.689]
---------	--------	-----------	--------	-----------	-----------------

The parameter estimation results for the EGARCH(3,1) model fitted to Ethiopia's monthly average temperature data reveal significant insights into the volatility dynamics. The omega (ω) parameter, with a coefficient of 0.6258 and a highly significant p-value (7.114e-11), indicates a substantial constant variance component in the model. The alpha parameters (α_1 , α_2 , α_3) provide a nuanced understanding of the impact of past shocks on current volatility. Alpha[1] (α_1) has a coefficient of 1.5470, demonstrating a strong influence of past shocks on current volatility, supported by its extremely significant p-value (1.738e-57). Conversely, the negative coefficients of alpha[2] (α_2) at -0.6802 and alpha[3] (α_3) at -0.7078, both highly significant ($p < 0.001$), suggest a moderating effect, indicating that while past shocks do influence current volatility, their impact diminishes over time. The beta parameter (β_1), with a coefficient of 0.5813 and a highly significant p-value (3.209e-26), reflects the persistence in volatility, signifying that current volatility is strongly influenced by past volatility. Overall, these parameter estimates underscore the robustness of the EGARCH(3,1) model in capturing the conditional heteroscedasticity in Ethiopia's monthly average temperature data, highlighting both the persistence of volatility and the complex interplay of past shocks.

Table 4.17 Parameter Estimation of EGARCH(3,2) for Kenya MAT

	coef	std err	t	P> t	95.0% Conf. Int.
omega	0.0223	7.532e-06	2961.308	0.000	[2.229e-02,2.232e-02]
alpha[1]	1.1237	9.493e-03	118.376	0.000	[1.105, 1.142]
alpha[2]	-0.1912	7.419e-03	-25.765	2.164e-146	[-0.206, -0.177]
alpha[3]	-1.2062	1.207e-02	-99.943	0.000	[-1.230, -1.183]
beta[1]	0.0938	5.137e-05	1826.624	0.000	[9.373e-02,9.393e-02]
beta[2]	0.8627	3.185e-12	2.708e+11	0.000	[0.863, 0.863]

The parameter estimation results for the EGARCH(3,2) model applied to Kenya's monthly average temperature data provide detailed insights into the volatility structure. The omega (ω) parameter has a coefficient of 0.0223, which is extremely significant ($p < 0.001$), indicating a substantial constant variance component in the model. The alpha parameters reveal the effects of past shocks on current volatility. Alpha[1] (α_1) is significantly positive with a coefficient of 1.1237, suggesting a strong positive influence of past shocks on current volatility. Alpha[2] (α_2), with a negative

coefficient of -0.1912, indicates a reduction in the effect of shocks over time, while $\alpha[3]$ (α_3) at -1.2062 shows a further strong negative impact, all of which are highly significant ($p < 0.001$). The beta parameters reflect the persistence in volatility. Beta[1] (β_1) has a coefficient of 0.0938, indicating that current volatility is influenced by past volatility, though less strongly compared to other parameters, and is highly significant ($p < 0.001$). Beta[2] (β_2), with a coefficient of 0.8627 and an extremely significant p-value ($p < 0.001$), shows a strong persistence in volatility, suggesting that past volatility has a long-lasting impact. Overall, these estimates underscore the complexity and persistence in Kenya's monthly average temperature volatility, validating the EGARCH(3,2) model as a robust tool for capturing these dynamics.

Table 4.18 Parameter Estimation of EGARCH(4,1) for Uganda MAT

	coef	std err	t	P> t	95.0% Conf. Int.
omega	0.0121	5.660e-05	214.491	0.000	[1.203e-02,1.225e-02]
alpha[1]	1.0227	5.468e-03	187.035	0.000	[1.012, 1.033]
alpha[2]	-0.6460	3.816e-03	-169.269	0.000	[-0.653, -0.639]
alpha[3]	-0.2638	9.375e-02	-2.814	4.900e-03	[-0.448,-8.003e-02]
alpha[4]	-0.2395	9.356e-02	-2.560	1.047e-02	[-0.423,-5.614e-02]
beta[1]	0.9863	8.774e-14	1.124e+13	0.000	[0.986, 0.986]

The parameter estimation for the EGARCH(4,1) model applied to Uganda's monthly average temperature data provides significant insights into the volatility dynamics. The omega (ω) parameter is 0.0121, highly significant ($p < 0.001$), indicating a consistent variance component within the model. The alpha parameters, representing the impact of past shocks on current volatility, show diverse effects. Alpha[1] (α_1) is significantly positive with a coefficient of 1.0227, suggesting a strong positive influence of past shocks on current volatility. Alpha[2] (α_2), with a negative coefficient of -0.6460, indicates a substantial reduction in the impact of past shocks. Both α_1 and α_2 are highly significant ($p < 0.001$).

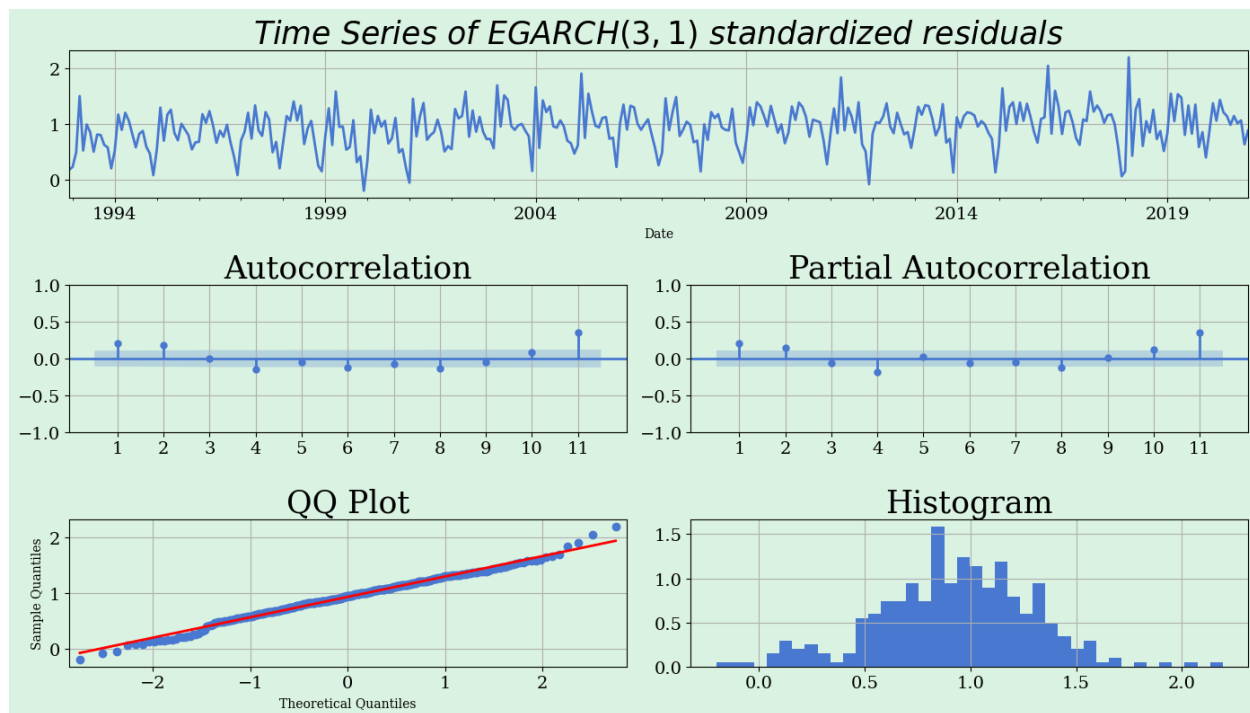
Alpha[3] (α_3), with a coefficient of -0.2638, and Alpha[4] (α_4), with a coefficient of -0.2395, also negatively influence volatility, albeit to a lesser extent compared to α_2 . These coefficients are significant with p-values of 0.0049 and 0.0105, respectively, suggesting a diminishing impact of past shocks over time. The beta parameter (β_1) is 0.9863, with an extremely high significance ($p < 0.001$), indicating strong persistence in volatility, meaning that past volatility has a lasting influence on current volatility levels.

Overall, the results highlight the EGARCH(4,1) model's capacity to capture the complex volatility patterns in Uganda's monthly average temperature, demonstrating significant and varying influences of past shocks and high persistence in volatility. This model effectively characterizes the conditional heteroscedasticity observed in the temperature data.

4.4.4 Diagnostics Checking of EGARCH Model

The EGARCH(3,1) model was applied to the residuals of the SARIMA model fitted to the monthly average temperature data of Ethiopia. The following diagnostic plots and their interpretations are provided to assess the adequacy and performance of the EGARCH model.

Figure 4.25 Residual analysis of EGARCH(3,1) for Ethiopia MAT



The time series plot of the residuals from the EGARCH(3,1) model reveals that the residuals fluctuate around a constant trend close to zero, with no apparent systematic patterns over time. This behavior indicates that the model has successfully captured the conditional heteroskedasticity present in the data, leaving residuals that are approximately white noise. The constant trend suggests that the model is adequately removing the volatility clustering typically present in temperature data, leaving residuals that are independent and identically distributed (i.i.d.).

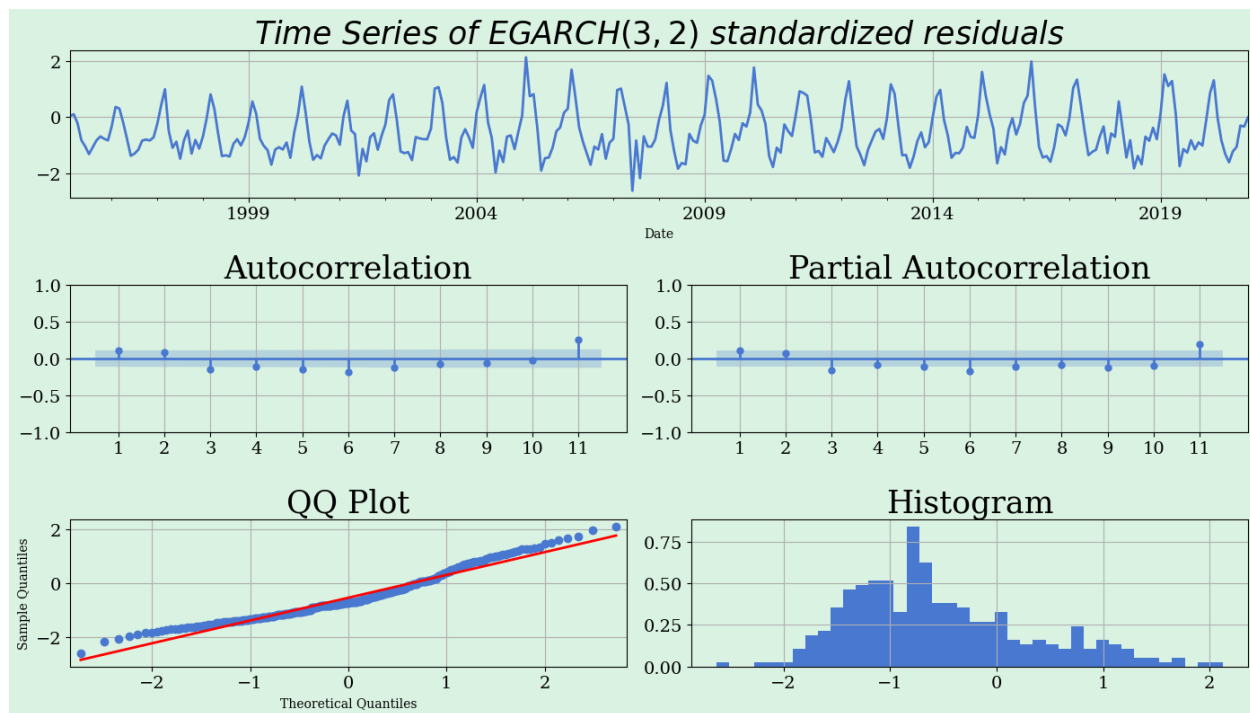
The ACF and PACF plots of the residuals provide further evidence on the adequacy of the EGARCH(3,1) model. In these plots, while some lags fall outside the confidence intervals, the majority are within the bounds. The presence of a few significant lags suggests that there may still be some minor autocorrelation left in the residuals. However, the overall pattern indicates that most of the autocorrelation has been effectively removed by the EGARCH model. This residual autocorrelation may be attributed to either model specification issues or intrinsic properties of the temperature time series that are not fully captured by the model.

The Q-Q plot is used to assess the normality of the residuals. In this case, the Q-Q plot shows that almost all the dots are aligned along the diagonal line, suggesting that the residuals of the EGARCH(3,1) model follow a normal distribution. This alignment indicates that the assumption of normality is reasonably satisfied, which is important for the validity of subsequent statistical inferences and forecasts based on the model. Deviations from the diagonal, if any, are minimal and do not significantly impact the overall conclusion of normality.

The histogram of the residuals further corroborates the findings from the Q-Q plot. The shape of the histogram closely resembles a normal distribution, with a bell-shaped curve centered around zero. This visual confirmation supports the notion that the residuals are approximately normally distributed. The absence of significant skewness or kurtosis in the histogram implies that the EGARCH(3,1) model has adequately captured the underlying volatility dynamics of the temperature series.

Based on these diagnostic plots, the EGARCH(3,1) model appears to be a good fit for the residuals of the SARIMA model applied to Ethiopia's monthly average temperature data. The constant trend in the residual time series plot, the alignment along the diagonal in the Q-Q plot, and the normal distribution-like histogram all indicate that the model has effectively captured the conditional variance in the data. The minor autocorrelation observed in the ACF and PACF plots suggests that there may be some room for model refinement, but it does not significantly detract from the overall adequacy of the model.

Figure 4.26 Residual analysis of EGARCH(3,2) for Kenya MAT



The EGARCH(3,2) model was used to analyze the residuals from the SARIMA model fitted to Kenya's monthly average temperature data. The following diagnostics provide an assessment of the model's performance:

The time series plot of the residuals from the EGARCH(3,2) model shows that the residuals fluctuate around a constant trend close to zero. This indicates that the model effectively captures the conditional volatility in the temperature data, leaving residuals that appear to be stationary and mean-reverting.

The ACF and PACF plots of the residuals show that while some lags fall outside the confidence intervals, most are within the bounds. This suggests that there is minimal remaining autocorrelation in the residuals, indicating that the EGARCH model has successfully modeled most of the volatility in the series. The few significant lags may indicate minor remaining patterns that could be explored further.

The Q-Q plot displays that almost all the dots are aligned along the diagonal line, suggesting that the residuals follow a normal distribution. This alignment indicates that the assumption of

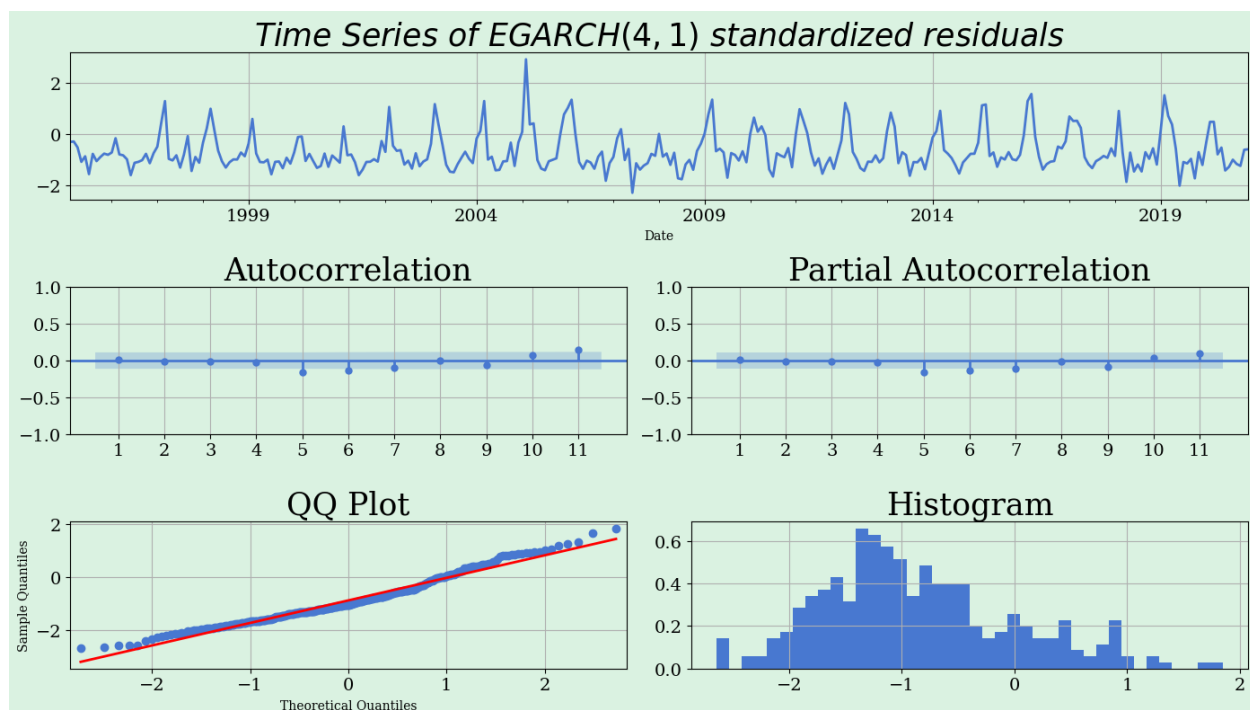
normality is largely satisfied, which is crucial for making valid statistical inferences based on the model.

The histogram of the residuals resembles a normal distribution, with a bell-shaped curve centered around zero. This further supports the Q-Q plot's indication that the residuals are approximately normally distributed. The lack of significant skewness or kurtosis confirms the adequacy of the EGARCH model in capturing the volatility of the temperature data.

Overall, the EGARCH(3,2) model diagnostics indicate that it is well-suited for modeling the volatility in Kenya's monthly average temperature data. The residuals exhibit a constant trend around zero, minimal autocorrelation, and approximate normality. These results suggest that the EGARCH model effectively addresses the heteroskedasticity in the temperature series, providing a reliable foundation for accurate temperature forecasting.

The diagnostics of the EGARCH(4,1) model for Uganda's monthly average temperature reveal several important characteristics about the residuals, which indicate how well the model fits the data and captures the volatility.

Figure 4.27 Residual analysis of EGARCH(4,1) for Uganda MAT



The time series plot of the residuals from the EGARCH(4,1) model shows that the residuals fluctuate around a constant trend between -1 and 0. This pattern suggests that the model effectively captures the mean structure of the temperature series, leaving the residuals centered around a relatively stable value. The lack of any apparent trend in the residuals indicates that the model has successfully removed any systematic components from the data.

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for the residuals show that most of the lags fall within the confidence intervals. This indicates that there is no significant autocorrelation left in the residuals, implying that the model has effectively captured the temporal dependencies in the data. The few lags that fall outside the intervals are not necessarily problematic, as they may represent minor deviations that do not significantly impact the overall model performance.

The Quantile-Quantile (QQ) plot of the residuals shows that almost all the dots are aligned along the diagonal line. This alignment suggests that the residuals follow a normal distribution quite closely, which is a desirable property in model diagnostics. A normal distribution of residuals indicates that the model's assumptions about the error terms are valid, further validating the model's fit.

The histogram of the residuals also resembles a normal distribution, reinforcing the findings from the QQ plot. This normality in the residuals is indicative of a well-specified model, as it suggests that the residuals do not exhibit significant skewness or kurtosis.

Overall, the diagnostics of the EGARCH(4,1) model for Uganda's monthly average temperature suggest that the model is well-specified and effectively captures the volatility in the temperature series. The residuals are centered around a constant trend, exhibit minimal autocorrelation, and follow a normal distribution. These characteristics indicate that the model is robust and reliable for forecasting purposes. This high-quality fit is crucial for accurate prediction and analysis of temperature patterns, providing a strong foundation for climate-related decision-making and policy development.

4.5 Discussion

The application of time series analysis to model and predict monthly average temperatures in global and East African regions has yielded significant insights. For global temperatures, the

SARIMA(4,0,0)(0,1,1,12) model demonstrated high forecasting accuracy, with MAE, RMSE, and MAPE values indicating strong predictive performance. In the East African region, the SARIMA models for Ethiopia, Kenya, and Uganda also provided reasonably good forecasts, though with varying degrees of accuracy. The SARIMA(2,0,2)(3,1,1,12) model for Ethiopia, SARIMA(2,0,3)(0,1,1,12) for Kenya, and SARIMA(1,0,4)(0,1,1,12) for Uganda showed that Kenya's model performed the best among the three, followed by Ethiopia and Uganda.

The findings of this study are consistent with previous research on global and regional temperature trends and variability. The upward trend in global temperatures observed in this study aligns with the extensive body of literature, including reports by the Intergovernmental Panel on Climate Change (IPCC), which attribute this rise to anthropogenic influences, particularly greenhouse gas emissions (IPCC, 2021).

The finding of this study on Ethiopia's monthly average temperature reveals clear seasonality, making the SARIMA model a suitable choice. This finding is consistent with the work of Hailemeskel Abebe (2020), who identified the SARIMA(2,0,1)(2,0,1,12) model for the monthly average temperature in case of Ambo area, Ethiopia. This reinforces the robustness of the SARIMA model in capturing the seasonal components of temperature data in Ethiopia.

Ethiopia's monthly average temperatures reveals an upward trajectory, with a forecasted maximum of 26.19°C and an average of 24.46°C. This is consistent with other studies on Ethiopia's climate, which have also documented warming trends. A study by Gebrechorkos, Hülsmann, and Bernhofer (2019) found that average temperatures in Ethiopia have increased by 0.37°C per decade since the 1980s, with the trend expected to continue in the coming decades.

The forecasted temperature range for Kenya, with a maximum of 27.25°C and an average of 25.35°C, corresponds well with other research on Kenya's climate. A study by Omondi et al. (2012) reported that average temperatures in Kenya have increased by 1.0°C since the 1960s, and future projections indicate further warming in the coming decades.

Uganda's monthly temperatures shows a forecasted maximum of 25.87°C and an average of 24.28°C. These findings are aligned with existing studies on Uganda's climate. For instance, a report by the Uganda National Meteorological Authority (2019) noted that average temperatures

in Uganda have risen by 1.3°C since the 1960s, and projections suggest continued warming in the future.

For global temperatures, the SARIMA(4,0,0)(0,1,1,12) model demonstrated high forecasting accuracy, with MAE, RMSE, and MAPE values of 0.0939, 0.1169, and 0.6442, respectively. These metrics indicate strong predictive performance, reinforcing the model's reliability for long-term temperature forecasting. The forecasted maximum global temperature of 17.397°C over the next two decades aligns with Cheng (2023), who used an XGBoost-LSTM model to predict that global temperatures in 2050 would not reach 20.00°C. This consistency across different modeling approaches highlights the robustness of the SARIMA model.

By applying the ARCH effect test on the residuals of the SARIMA model, the test results indicated the existence of an ARCH effect in the model. Consequently, volatility models such as SARIMA-EGARCH were employed, with SARIMA-EGARCH proving to be the best fit. This finding is consistent with the research by Iqelan (2016), who identified the AR(1)-GARCH(1,1) model as the best fit for Monthly Temperature Data of Jerusalem/Palestine, ensuring the existence of the ARCH effect. Similarly, Romilly (2005) found that the SARMA–GARCH(1,1)-M model was the best fit for Global temperature, reinforcing the application of GARCH models in temperature forecasting.

In the East African context, the warming trends detected in Ethiopia, Kenya, and Uganda corroborate the findings of Nicholson (2014) and Tierney, Ummenhofer, and DeMenocal (2015), who documented similar rising temperature patterns in the region. These studies also emphasize the significant impact of climate change on regional temperature dynamics, reinforcing the evidence presented in this thesis.

The greater variability observed in global temperatures compared to regional temperatures aligns with the findings of Hawkins and Sutton (2009), who attribute this phenomenon to the diverse climatic conditions across different parts of the world. The aggregation of these varying conditions leads to a more variable global temperature record, which is consistent with the variability patterns identified in this study.

Furthermore, the seasonal patterns observed, such as global temperature peaks during the Northern Hemisphere summer months, are well-documented in climatological literature. The seasonal

cycles in Ethiopia, Kenya, and Uganda identified in this study are consistent with the findings of Camberlin and Philippon (2002), who detailed the climatic seasonality in these regions. This alignment with previous studies not only validates the models used in this research but also underscores the reliability of the findings.

5. Conclusion and Recommendation

5.1. Conclusion

This study sought to address the question: How can we model and forecast temperature variability globally and in East Africa? The descriptive statistics reveal that the maximum temperatures observed during the historical period were 17.213°C globally and 26.961°C, 27.659°C, and 26.468°C for Ethiopia, Kenya, and Uganda, respectively. These findings indicate that East African countries experience higher temperature extremes compared to the global average, though the overall temperature variability within these regions is relatively small.

The stationarity of the monthly average temperature was assessed using the Augmented Dickey-Fuller (ADF) and KPSS tests. The ADF test results suggested global temperatures were non-stationary, while East African temperatures were stationary. Contrarily, the KPSS test indicated non-stationarity for all temperature series. Seasonality was evident in the data, and after applying the first seasonal differencing, the temperature series for all regions became stationary.

To construct the SARIMA model, we identified parameters using ACF and PACF plots and selected the optimal models based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The selected models were SARIMA(4,0,0)(0,1,1,12) for global temperatures, SARIMA(2,0,2)(3,1,1,12) for Ethiopia, SARIMA(2,0,2)(0,1,1,12) for Kenya, and SARIMA(1,0,1)(0,1,1,12) for Uganda. Model diagnostics validated these choices, indicating their suitability for forecasting.

The SARIMA models forecasted a significant increasing trend in monthly average temperatures over the next two decades across all regions. In the forecasted two decades, the monthly average temperature is projected to reach a maximum of 17.397°C globally, 26.192°C for Ethiopia, 27.25°C for Kenya, and 25.87°C for Uganda, with Kenya experiencing the highest temperatures among the regions studied. Forecast accuracy metrics confirmed the models' robustness.

To account for temperature volatility, conditional heteroscedasticity models such as ARCH, GARCH, and EGARCH were employed. Based on AIC, BIC, and Log-Likelihood values, EGARCH models were determined to best capture temperature volatility. Specifically, EGARCH(4,1) for global temperatures, EGARCH(3,1) for Ethiopia, EGARCH(3,2) for Kenya,

and EGARCH(4,1) for Uganda were the most suitable. Model diagnostics further supported the effectiveness of these models in capturing temperature volatility.

The findings suggest that advanced volatility models, particularly EGARCH, significantly enhance the accuracy of temperature forecasts. The robust forecasting accuracy of the SARIMA models, especially when enhanced with the EGARCH approach for capturing volatility, has significant implications for climate modeling and prediction. Accurate temperature forecasts are crucial for developing effective climate adaptation and mitigation strategies, particularly in regions vulnerable to climate variability and change, such as East Africa.

Future research should focus on incorporating more diverse climatic factors and exploring the integration of these models with other predictive techniques to further refine temperature forecasts. Additionally, enhancing data collection efforts in regions with sparse historical data will improve model accuracy and reliability.

Overall, this study contributes to the field of climate science by offering a robust framework for modeling and forecasting temperature variability, which is crucial for anticipating and mitigating the impacts of climate change.

5.2. Recommendations

Based on the results and discussion of this study, several recommendations can be made to enhance climate data analysis and its applications:

1. **Integrate Climate Models in Policy Making:** Policymakers should integrate findings from advanced climate models into their decision-making processes. Enhanced temperature forecasts can support more informed decisions in sectors such as agriculture, water resources, and energy management. For example, agricultural planners can use these forecasts to optimize planting schedules and improve crop yields.
2. **Focus on Regional Specificity:** Given the variability in model performance across different regions, it is crucial to tailor climate models to the specific characteristics of each region. The study showed that different configurations of EGARCH and SARIMA models were optimal for the regions studied, highlighting the need for region-specific approaches in climate modeling.

3. Incorporate Exogenous Variables: Future research should explore incorporating exogenous variables, such as sea surface temperatures and atmospheric circulation patterns, to further enhance model accuracy. Additionally, investigating the potential impacts of extreme weather events and long-term climate change on temperature patterns could provide deeper insights and more comprehensive predictive models.

4. Advance Modeling Techniques: Researchers should explore the use of neural network models, such as LSTM (Long Short-Term Memory) and RNN (Recurrent Neural Network) models, which can capture the complex dynamics of temperature better than traditional SARIMA-GARCH models. These advanced techniques can offer improved forecasting capabilities and more robust analysis of temperature data.

By implementing these recommendations, stakeholders can better anticipate and mitigate the impacts of climate change, thereby enhancing resilience and sustainability in affected regions..

References

- Abera, Kidist, Olivier Crespo, Jemal Seid, and Fasil Mequanent. 2018. "Simulating the Impact of Climate Change on Maize Production in Ethiopia, East Africa." *Environmental Systems Research* 7(1). doi:10.1186/s40068-018-0107-z.
- Abid, M., J. Scheffran, U. A. Schneider, and M. Ashfaq. 2015. "Farmers' Perceptions of and Adaptation Strategies to Climate Change and Their Determinants: The Case of Punjab Province, Pakistan." *Earth System Dynamics* 6(1). doi:10.5194/esd-6-225-2015.
- Adane, Asnake, and Woldeamlak Bewket. 2021. "Effects of Quality Coffee Production on Smallholders' Adaptation to Climate Change in Yirgacheffe, Southern Ethiopia." *International Journal of Climate Change Strategies and Management* 13(4–5). doi:10.1108/IJCCSM-01-2021-0002.
- Adhikari, Umesh, A. Pouyan Nejadhashemi, and Sean A. Woznicki. 2015. "Climate Change and Eastern Africa: A Review of Impact on Major Crops." *Food and Energy Security* 4(2). doi:10.1002/fes3.61.
- Akaike, Hirotugu. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19(6). doi:10.1109/TAC.1974.1100705.
- Al-Ghamdi, Khalid, Masroor Ali khan, and Jazem Mahyoub. 2009. "Role of Climatic Factors in the Seasonal Abundance of *Aedes Aegypti* L. and Dengue Fever Cases in Jeddah Province of Saudi Arabia." *Current World Environment* 4(2). doi:10.12944/cwe.4.2.07.
- Aryal, Jeetendra Prakash, Tek B. Sapkota, Ritika Khurana, Arun Khatri-Chhetri, Dil Bahadur Rahut, and M. L. Jat. 2020. "Climate Change and Agriculture in South Asia: Adaptation Options in Smallholder Production Systems." *Environment, Development and Sustainability* 22(6). doi:10.1007/s10668-019-00414-4.
- Baker, Justin S., Petr Havlík, Robert Beach, David Leclère, Erwin Schmid, Hugo Valin, Jefferson Cole, et al. 2018. "Evaluating the Effects of Climate Change on US Agricultural Systems: Sensitivity to Regional Impact and Trade Expansion Scenarios." *Environmental Research Letters* 13(6). doi:10.1088/1748-9326/aac1c2.
- Baker, Rachel E., and Jesse Anttila-Hughes. 2020. "Characterizing the Contribution of High Temperatures to Child Undernourishment in Sub-Saharan Africa." *Scientific Reports* 10(1). doi:10.1038/s41598-020-74942-9.
- Bartholomew, D. J., G. E. P. Box, and G. M. Jenkins. 1971. "Time Series Analysis Forecasting and Control." *Operational Research Quarterly (1970-1977)* 22(2). doi:10.2307/3008255.
- Bera, Anil K., and Carlos M. Jarque. 1981. "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. Monte Carlo Evidence." *Economics Letters* 7(4). doi:10.1016/0165-1765(81)90035-5.

- Bollerslev, Tim. 1986. 31 *Journal of Econometrics* *GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY*.
- Box, G. E. P., and G. M. Jenkins. 1970. "Time Series Analysis Forecasting and Control." *Operational Research Quarterly (1970-1977)* 22(2). doi:10.2307/3008255.
- Box, George E.P., Gwilym M. Jenkins, and Gregory C. Reinsel. 2013. Time Series Analysis: Forecasting and Control: Fourth Edition *Time Series Analysis: Forecasting and Control: Fourth Edition*. doi:10.1002/9781118619193.
- Boylan, Sinead, Kathleen Beyer, David Schlosberg, Anastasia Mortimer, Neil Hime, Benjamin Scalley, Robyn Alders, Carlos Corvalan, and Anthony Capon. 2018. "A Conceptual Framework for Climate Change, Health and Wellbeing in NSW, Australia." *Public Health Research and Practice* 28(4). doi:10.17061/phrp2841826.
- Briggs, William M., Spyros Makridakis, Steven C. Wheelwright, Rob J. Hyndman, and Francis X. Diebold. 1999. "Forecasting: Methods and Applications." *Journal of the American Statistical Association* 94(445). doi:10.2307/2669717.
- Camberlin, Pierre, and N. Philippon. 2002. "The East African March-May Rainy Season: Associated Atmospheric Dynamics and Predictability over the 1968-97 Period." *Journal of Climate* 15(9). doi:10.1175/1520-0442(2002)015<1002:TEAMMR>2.0.CO;2.
- Chatfield, Chris. 2005. "Time-Series Forecasting." *Significance* 2(3). doi:10.1111/j.1740-9713.2005.00117.x.
- Chen, Yujia, Ting Jiang, and Siben Li. 2023. "Global Temperature Prediction Based on Time Series and Global Climate." *Highlights in Science, Engineering and Technology* 44. doi:10.54097/hset.v44i.7336.
- Cheng, Xinyi. 2023. "Global Temperature Study Based on ARIMA and XGBoost-LSTM Models." *Frontiers in Economics and Management* 4(4): 224–34. doi:10.6981/FEM.202304_4(4).0028.
- Chourasia, Abhilasha. 2020. "Decomposition in Time Series Data." <https://medium.com/analytics-vidhya/decomposition-in-time-series-data-b20764946d63> (January 23, 2024).
- Crowley, Ryan A. 2016. "Climate Change and Health: A Position Paper of the American College of Physicians." *Annals of Internal Medicine* 164(9). doi:10.7326/M15-2766.
- Cuthbert, Mark O., Richard G. Taylor, Guillaume Favreau, Martin C. Todd, Mohammad Shamsudduha, Karen G. Villholth, Alan M. MacDonald, et al. 2019. "Observed Controls on Resilience of Groundwater to Climate Variability in Sub-Saharan Africa." *Nature* 572(7768). doi:10.1038/s41586-019-1441-7.
- Dickey, David A., and Wayne A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series With a Unit Root." *Journal of the American Statistical Association* 74(366). doi:10.2307/2286348.

- Dimri, Tripti, Shamshad Ahmad, and Mohammad Sharif. 2020. "Time Series Analysis of Climate Variables Using Seasonal ARIMA Approach." *Journal of Earth System Science* 129(1). doi:10.1007/s12040-020-01408-x.
- Döll, P., B. Jiménez-Cisneros, T. Oki, N. W. Arnell, G. Benito, J. G. Cogley, T. Jiang, et al. 2014. "Integrating Risks of Climate Change into Water Management." *Hydrological Sciences Journal* 60(1). doi:10.1080/02626667.2014.967250.
- Earth Observatory. 2022. "World of Change: Global Temperatures." <https://earthobservatory.nasa.gov/world-of-change/decadaltemp.php> (December 23, 2023).
- Ebi, Kristie L., John Balbus, Patricia L. Kinney, Erin Lipp, David Mills, Marie S. O'Neill, and Mark L. Wilson. 2009. "U.S. Funding Is Insufficient to Address the Human Health Impacts of and Public Health Responses to Climate Variability and Change." *Environmental Health Perspectives* 117(6). doi:10.1289/ehp.0800088.
- Engle, Robert F. 1982. *50 Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. <https://www.jstor.org/stable/1912773>.
- FAO. 2018. "AQUASTAT Core Database. Food and Agriculture Organization of the United Nations." <https://www.fao.org/aquastat/en/databases/maindatabase> (January 31, 2024).
- Flahault, Antoine, Rafael Ruiz de Castaneda, and Isabelle Bolon. 2016. "Climate Change and Infectious Diseases." *Public Health Reviews* 37(1). doi:10.1186/s40985-016-0035-2.
- Gebrechorkos, Solomon H., Stephan Hülsmann, and Christian Bernhofer. 2019. "Changes in Temperature and Precipitation Extremes in Ethiopia, Kenya, and Tanzania." *International Journal of Climatology* 39(1). doi:10.1002/joc.5777.
- Gould, Solange, and Linda Rudolph. 2015. "Challenges and Opportunities for Advancing Work on Climate Change and Public Health." *International Journal of Environmental Research and Public Health* 12(12). doi:10.3390/ijerph121215010.
- Gray, H. L., G. D. Kelley, and D. D. McIntire. 1978. "A New Approach to Arma Modeling." *Communications in Statistics - Simulation and Computation* 7(1). doi:10.1080/03610917808812057.
- Green, F. H. W., and Glenn T. Trewartha. 1963. "The Earth's Problem Climates." *The Journal of Animal Ecology* 32(2). doi:10.2307/2545.
- Hailemeskel Abebe, Teshome. 2020. "Time Series Analysis of Monthly Average Temperature and Rainfall Using Seasonal ARIMA Model (in Case of Ambo Area, Ethiopia)." *International Journal of Theoretical and Applied Mathematics* 6(5). doi:10.11648/j.ijtam.20200605.13.
- Hamilton. 1993. "Time Series Analysis / James D. Hamilton." *Time series analysis*.

- Hannan, E. J. 1980. "The Estimation of the Order of an ARMA Process." *The Annals of Statistics* 8(5). doi:10.1214/aos/1176345144.
- Hansen, Peter R., and Asger Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics* 20(7). doi:10.1002/jae.800.
- Hawkins, Ed, and Rowan Sutton. 2009. "The Potential to Narrow Uncertainty in Regional Climate Predictions." *Bulletin of the American Meteorological Society* 90(8). doi:10.1175/2009BAMS2607.1.
- Hyndman, R.J., and G. Athanasopoulos. 2021. OTexts: Melbourne, Australia. OTexts.com/fpp3 *Forecasting: Principles and Practice (3rd Ed)*.
- Intergovernmental Panel on Climate Change (IPCC). 2023. Climate Change 2022 – Impacts, Adaptation and Vulnerability *Climate Change 2022 – Impacts, Adaptation and Vulnerability*. doi:10.1017/9781009325844.
- Iqelan, Bisher Mamoun. 2016. "Time Series Modelling of Monthly Temperature Data of Jerusalem/Palestine." *MATEMATIKA*. doi:10.11113/matematika.v31.n2.773.
- J.D. Cryer, K. Chan. 1985. 20 European Journal of Operational Research *Time Series Analysis with Application in R*.
- Kellogg, William W. 1987. "Mankind's Impact on Climate: The Evolution of an Awareness." *Climatic Change* 10(2). doi:10.1007/BF00140251.
- Knox, Jerry, Tim Hess, Andre Daccache, and Tim Wheeler. 2012. "Climate Change Impacts on Crop Productivity in Africa and South Asia." *Environmental Research Letters* 7(3). doi:10.1088/1748-9326/7/3/034032.
- Ljung, G. M., and G. E.P. Box. 1978. "On a Measure of Lack of Fit in Time Series Models." *Biometrika* 65(2). doi:10.1093/biomet/65.2.297.
- McQuarrie, Allan D R, and Chih-Ling Tsai. 1998. Regression and Time Series Model Selection *Regression and Time Series Model Selection*. doi:10.1142/3573.
- Van Meijl, Hans, Petr Havlik, Hermann Lotze-Campen, Elke Stehfest, Peter Witzke, Ignacio Pérez Domínguez, Benjamin Leon Bodirsky, et al. 2018. "Comparing Impacts of Climate Change and Mitigation on Global Agriculture by 2050." *Environmental Research Letters* 13(6). doi:10.1088/1748-9326/aabdc4.
- Miller, Alanna. 2019. "Climate Change and the Future of Public Health." *Health Science Inquiry* 10(1). doi:10.29173/hsi30.
- Moat, Justin, Jenny Williams, Susana Baena, Timothy Wilkinson, Tadesse W. Gole, Zeleke K. Challa, Sebsebe Demissew, and Aaron P. Davis. 2017. "Resilience Potential of the Ethiopian Coffee Sector under Climate Change." *Nature Plants* 3. doi:10.1038/nplants.2017.81.

- Montgomery, Douglas C, Cheryl L Jennings, and Murat Kulahci. 2015. *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons.
- Mwalukasa, Nicholas. 2013. "Agricultural Information Sources Used for Climate Change Adaptation in Tanzania." *Library Review* 62(4). doi:10.1108/LR-12-2011-0096.
- Naiqian, Guan, Guan Yuxin, and Sun Xuelian. 2019. "Global Temperature Forecast Based on ARIMA Model." In *Proceedings - 2019 4th International Conference on Communication and Information Systems, ICCIS 2019*, doi:10.1109/ICCIS49662.2019.00026.
- National Geographic Society. 2023. "Climate Change." <https://education.nationalgeographic.org/resource/climate-change/> (January 28, 2024).
- Nicholson, Sharon E. 2014. "A Detailed Look at the Recent Drought Situation in the Greater Horn of Africa." *Journal of Arid Environments* 103. doi:10.1016/j.jaridenv.2013.12.003.
- Nielsen, Aileen. 2019. 21 O'Reilly *Practical Time Series: Prediction with Statistics & Machine Learning*.
- Omondi, P., J. L. Awange, L. A. Ogallo, R. A. Okoola, and E. Forootan. 2012. "Decadal Rainfall Variability Modes in Observed Rainfall Records over East Africa and Their Relations to Historical Sea Surface Temperature Changes." *Journal of Hydrology* 464–465. doi:10.1016/j.jhydrol.2012.07.003.
- Ortiz-Bobea, Ariel, Toby R. Ault, Carlos M. Carrillo, Robert G. Chambers, and David B. Lobell. 2021. "Anthropogenic Climate Change Has Slowed Global Agricultural Productivity Growth." *Nature Climate Change* 11(4). doi:10.1038/s41558-021-01000-1.
- Pedgley, D.E. 1967. "Air Temperature at Dallol, Ethiopia." *Meteorological Magazine*: 265–71.
- Phillips, Peter C.B., and Zhijie Xiao. 1998. "A Primer on Unit Root Testing." *Journal of Economic Surveys* 12(5). doi:10.1111/1467-6419.00064.
- Qiao, Liuyiyan. 2023. "Time-Series Global Temperature Prediction and Visualization." *Highlights in Science, Engineering and Technology* 39. doi:10.54097/hset.v39i.6709.
- Raymond, Colin, Tom Matthews, and Radley M. Horton. 2020. "The Emergence of Heat and Humidity Too Severe for Human Tolerance." *Science Advances* 6(19). doi:10.1126/sciadv.aaw1838.
- Richardson, K, R Calow, F Pichon, S New, and R Osborne. 2022. *Climate Risk Report for the East Africa Region*. Met Office, ODI, FCDO. www.metoffice.gov.uk.
- Romilly, Peter. 2005. "Time Series Modelling of Global Mean Temperature for Managerial Decision-Making." *Journal of Environmental Management* 76(1). doi:10.1016/j.jenvman.2005.01.008.
- Schwarz G. 1978. "Estimating the Dimension of a Model." *The annals of statistics* 6(2).

- Sejian, Veerasamy, Mullakkalparambil V. Silpa, Mini R. Reshma Nair, Chinnasamy Devaraj, Govindan Krishnan, Madijagan Bagath, Surinder S. Chauhan, et al. 2021. "Heat Stress and Goat Welfare: Adaptation and Production Considerations." *Animals* 11(4). doi:10.3390/ani11041021.
- Shi, Zhiying. 2018. "Impact of Climate Change on the Global Environment and Associated Human Health." *OALib* 05(10). doi:10.4236/oalib.1104934.
- Shumway, Robert H, and David S Stoffer. 2016. Springer Texts in Statistics EZ Edition *Time Series Analysis and Its Applications With R Examples EZ Edition*.
- Singh, S.N. 2019. "Climate Change and Agriculture in Ethiopia: A Case Study of Mettu Woreda." *SocioEconomic Challenges* 3(3). doi:10.21272/sec.3(3).61-79.2019.
- Sultan, B., P. Roudier, P. Quirion, A. Alhassane, B. Muller, M. Dingkuhn, P. Ciais, et al. 2013. "Assessing Climate Change Impacts on Sorghum and Millet Yields in the Sudanian and Sahelian Savannas of West Africa." *Environmental Research Letters* 8(1). doi:10.1088/1748-9326/8/1/014040.
- Taylor, Richard G., Bridget Scanlon, Petra Döll, Matt Rodell, Rens Van Beek, Yoshihide Wada, Laurent Longuevergne, et al. 2013. "Ground Water and Climate Change." *Nature Climate Change* 3(4). doi:10.1038/nclimate1744.
- Thompson, Lonnie G., Ellen Mosley-Thompson, Mary E. Davis, Keith A. Henderson, Henry H. Brecher, Victor S. Zagorodnov, Tracy A. Mashiotta, et al. 2002. "Kilimanjaro Ice Core Records: Evidence of Holocene Climate Change in Tropical Africa." *Science* 298(5593). doi:10.1126/science.1073198.
- Thornton, Philip, and Laura Cramer. 2012. *Impacts of Climate Change on the Agricultural and Aquatic Systems and Natural Resources within the CGIAR's Mandate*. www.ccafs.cgiar.org.
- Tierney, Jessica E., Caroline C. Ummenhofer, and Peter B. DeMenocal. 2015. "Past and Future Rainfall in the Horn of Africa." *Science Advances* 1(9). doi:10.1126/sciadv.1500682.
- Tsay, Ruey S. 2010. Analysis of Financial Time Series *Analysis of Financial Time Series*. doi:10.1002/9780470644560.
- Tsay, Ruey S., and George C. Tiao. 1984. "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models." *Journal of the American Statistical Association* 79(385). doi:10.1080/01621459.1984.10477068.
- Uganda National Meteorological Authority. 2019. "National Climate Change Policy." *Ministry of Water and Environment*.
- UNEP. 2022. "On Verge of Record Drought, East Africa Grapples with New Climate Normal." *UN environment programme* . <https://www.unep.org/news-and->

stories/story/verge-record-drought-east-africa-grapples-new-climate-normal (December 25, 2023).

UNFCCC. 2020. “Climate Change Is an Increasing Threat to Africa | UNFCCC.” *United Nations Climate Change*. <https://unfccc.int/news/climate-change-is-an-increasing-threat-to-africa> (December 24, 2023).

UNICEF. 2014. *Generation 2030 Africa : Child Demographics in Africa*. Division of Data, Research and Policy.

Wang, Yunjing, Yuhan Rao, Xiaoxu Wu, Hainan Zhao, and Jin Chen. 2015. “A Method for Screening Climate Change-Sensitive Infectious Diseases.” *International Journal of Environmental Research and Public Health* 12(1). doi:10.3390/ijerph120100767.

Watkiss, Paul, and Kristie L. Ebi. 2022. “A Lack of Climate Finance Is Harming Population Health.” *The BMJ* 376. doi:10.1136/bmj.o313.

Wei, William W.S. 2006. “Time Series Analysis _ Univariate and Multivariate Methods (2nd Edition).” *New introduction to Multiple Time Series Analysis*.

Wold, H. 1938. “A Study in the Analysis of Stationary Time Series.” Doctoral dissertation. Almqvist & Wiksell.

WorldData.info. 2023. “Climate Change and Global Warming.” *WorldData.info*.

Yang, Wenchang, Richard Seager, Mark A. Cane, and Bradfield Lyon. 2015. “The Annual Cycle of East African Precipitation.” *Journal of Climate* 28(6). doi:10.1175/JCLI-D-14-00484.1.

Ye, Liming, Guixia Yang, Eric Van Ranst, and Huajun Tang. 2013. “Time-Series Modeling and Prediction of Global Monthly Absolute Temperature for Environmental Decision Making.” *Advances in Atmospheric Sciences* 30(2). doi:10.1007/s00376-012-1252-3.

Yule, George Udny. 1927. “VII. On a Method of Investigating Periodicities Disturbed Series, with Special Reference to Wolfer’s Sunspot Numbers.” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 226(636–646). doi:10.1098/rsta.1927.0007.

Zivot, Eric, and Jiahui Wang. 2006. Modeling Financial Time Series with S-PLUS®. *Modeling Financial Time Series with S-PLUS®*. doi:10.1007/978-0-387-32348-0.

Appendices

Appendix A: Figures

Figure 0.1 Time Series Plot of Monthly Average Temperature Variation

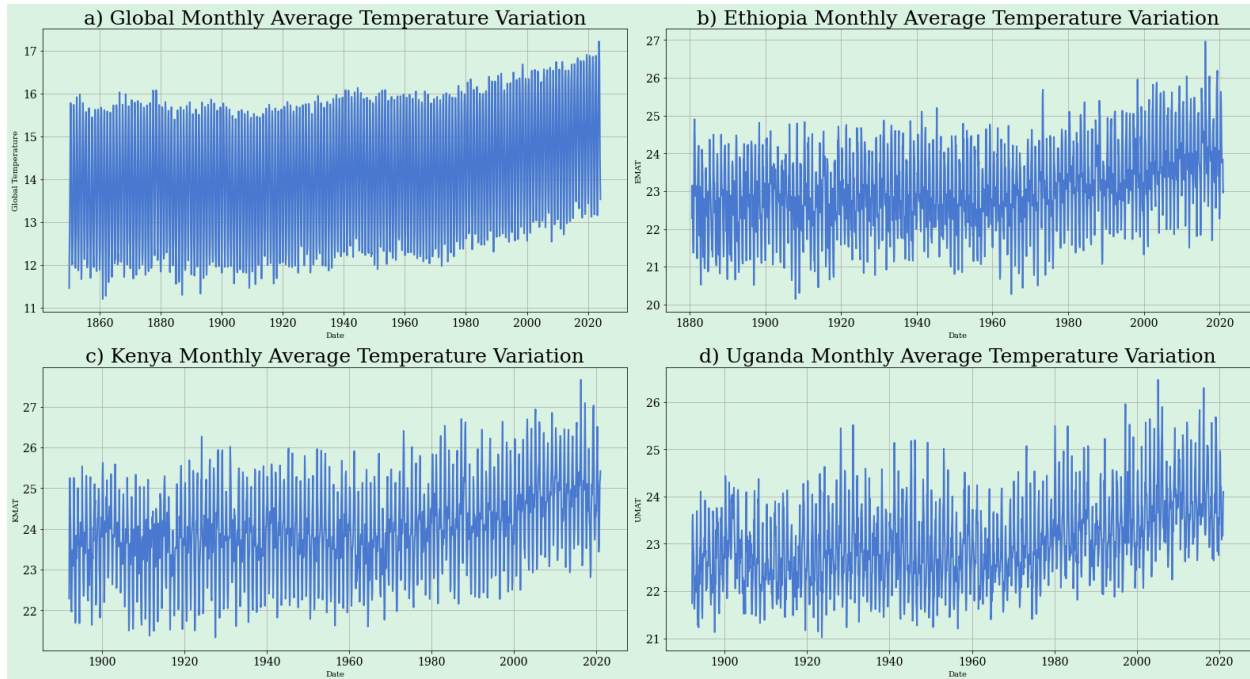


Figure 0.2 Global MAT persistence model predictions

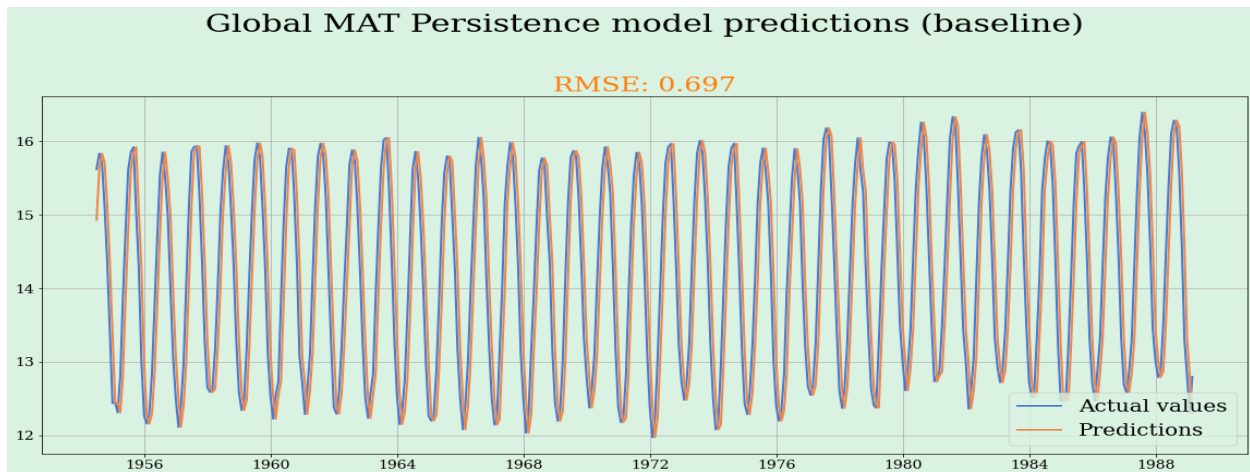


Figure 0.3 Ethiopia MAT persistence model predictions

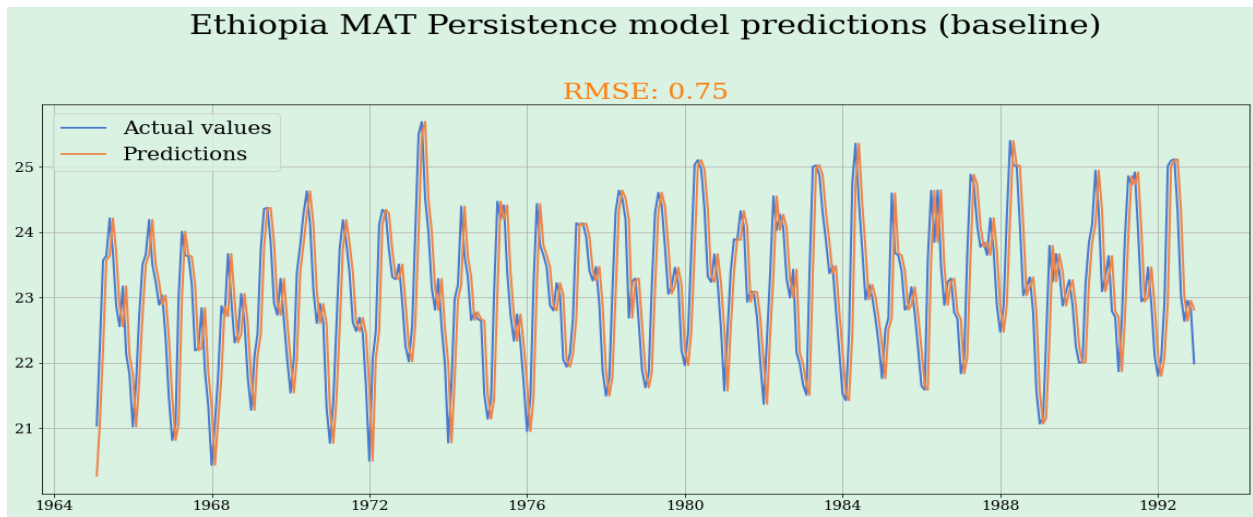


Figure 0.4 Kenya MAT persistence model predictions

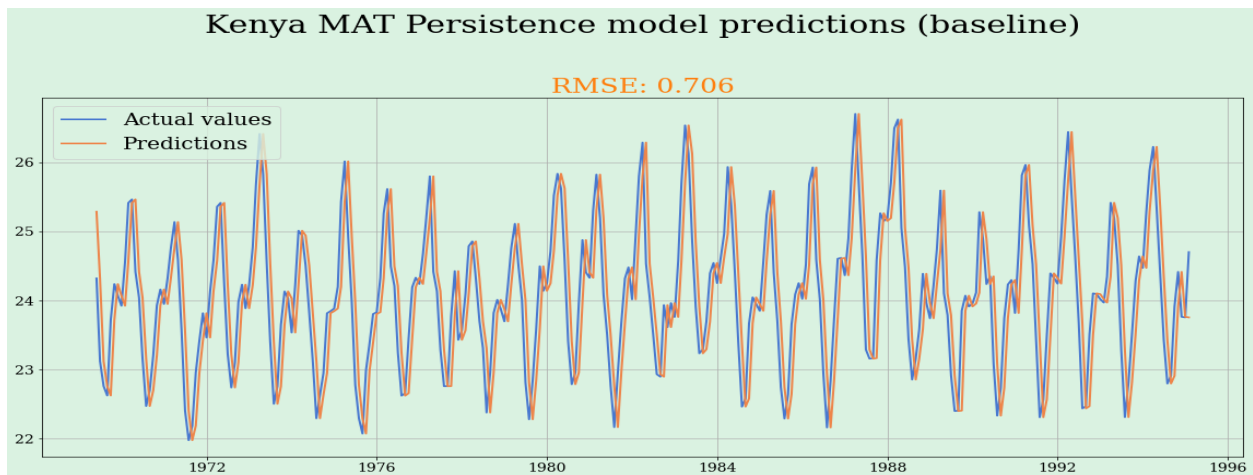


Figure 0.5 Uganda MAT persistence model predictions

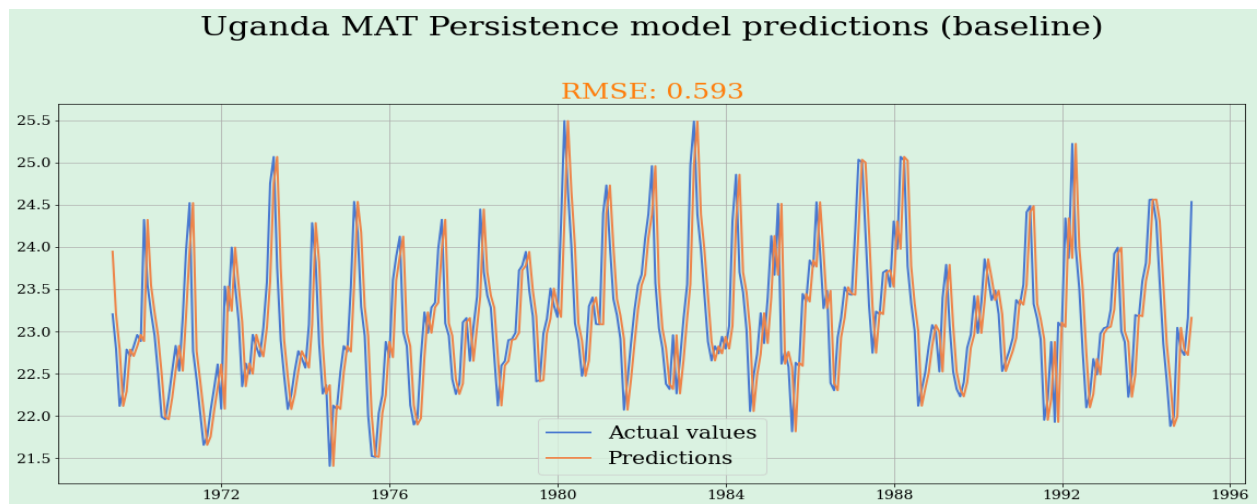


Figure 0.6 Test of Stationarity for Ethiopia MAT

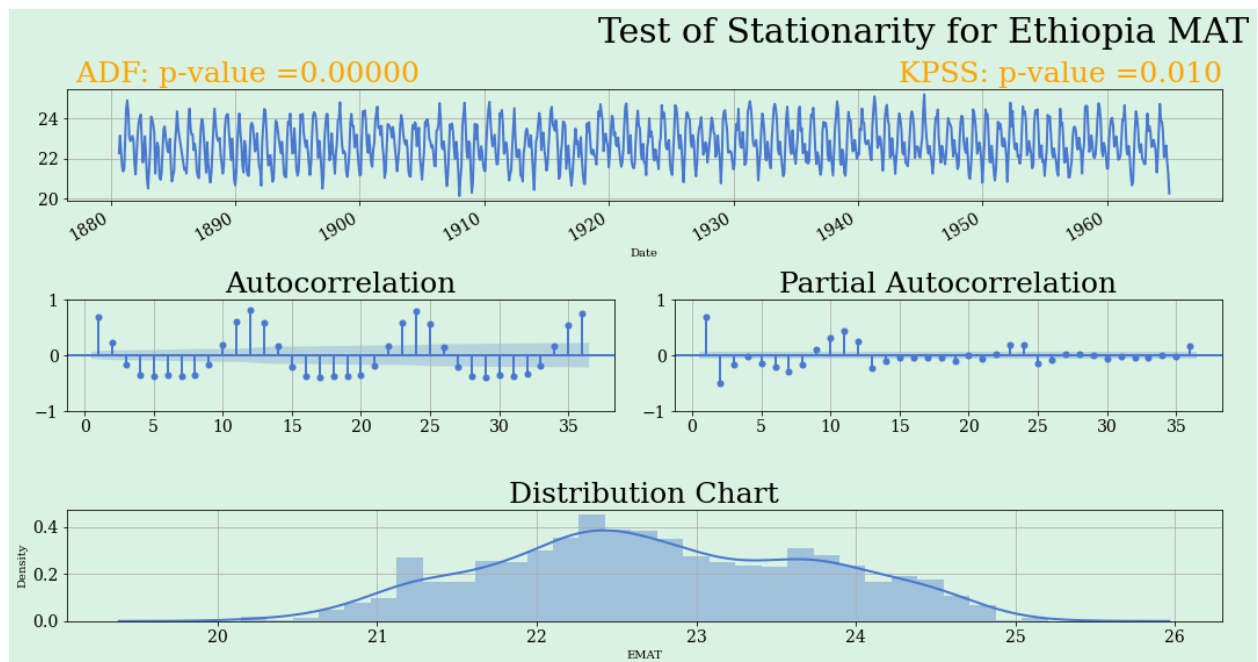


Figure 0.7 Test of Stationarity for Kenya MAT

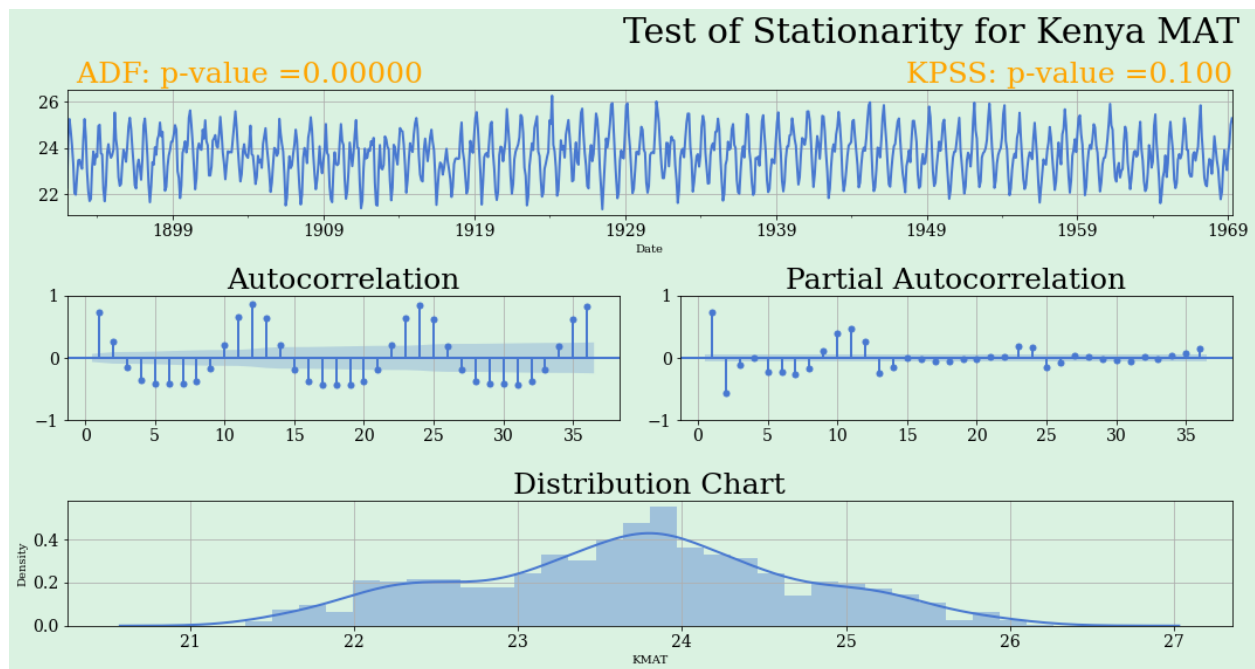


Figure 0.8 Test of Stationarity for Uganda MAT

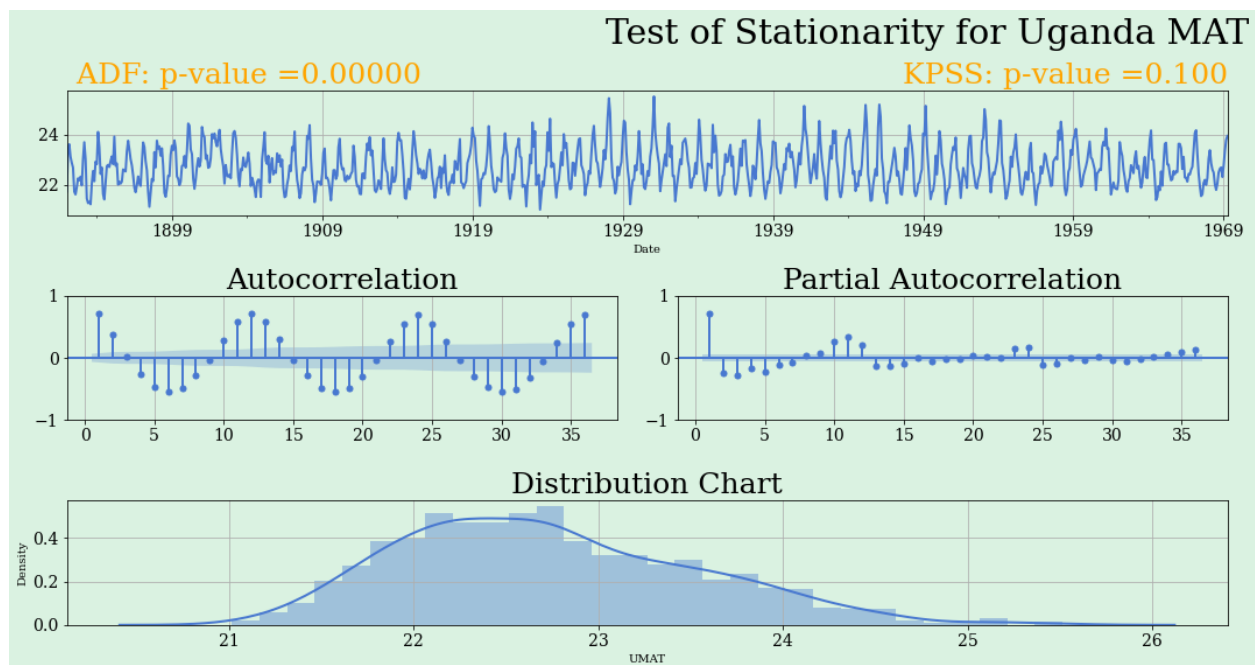


Figure 0.9 Seasonal decomposition of Global MAT

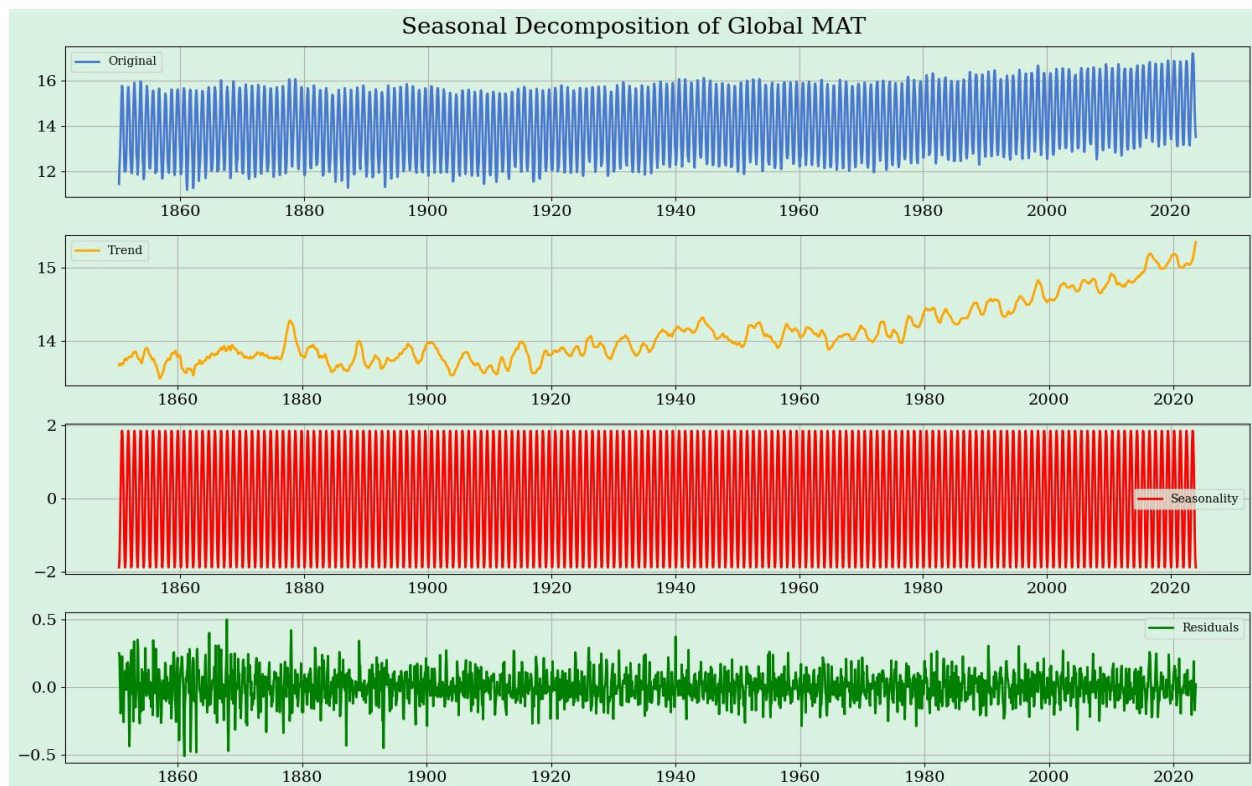


Figure 0.10 Seasonal decomposition of Ethiopia MAT

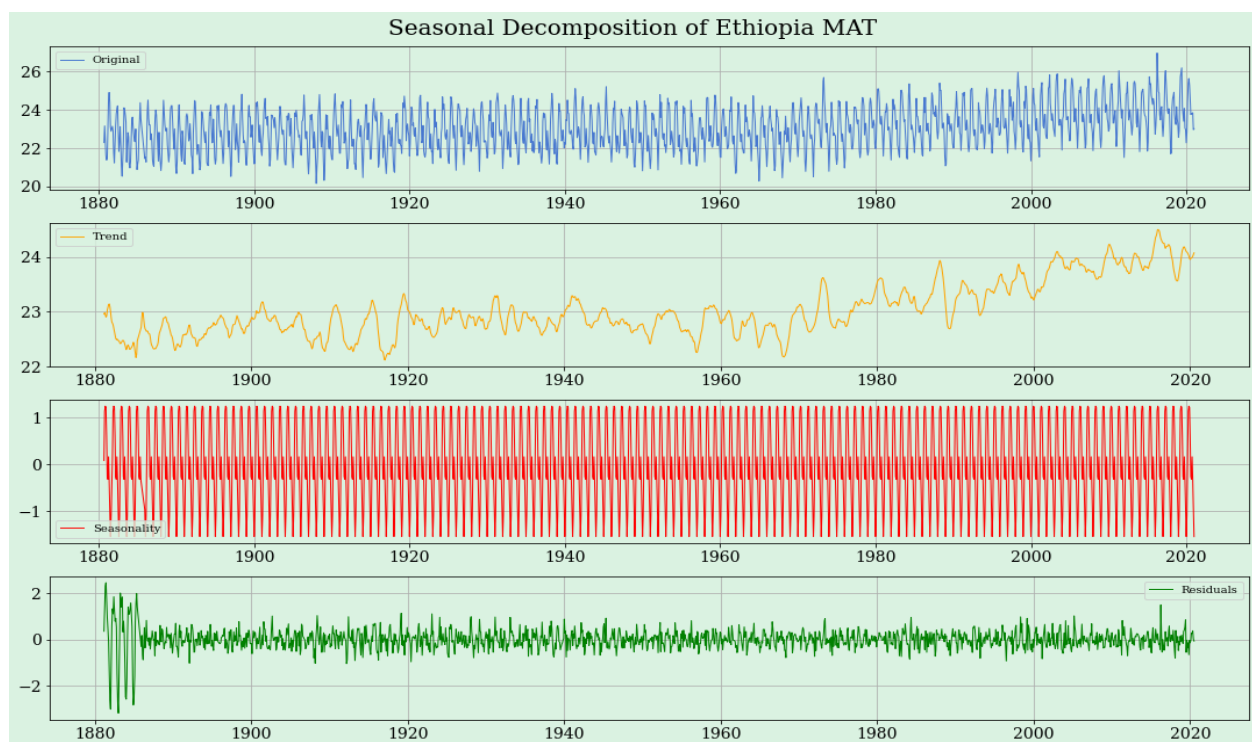


Figure 0.11 Seasonal decomposition of Kenya MAT

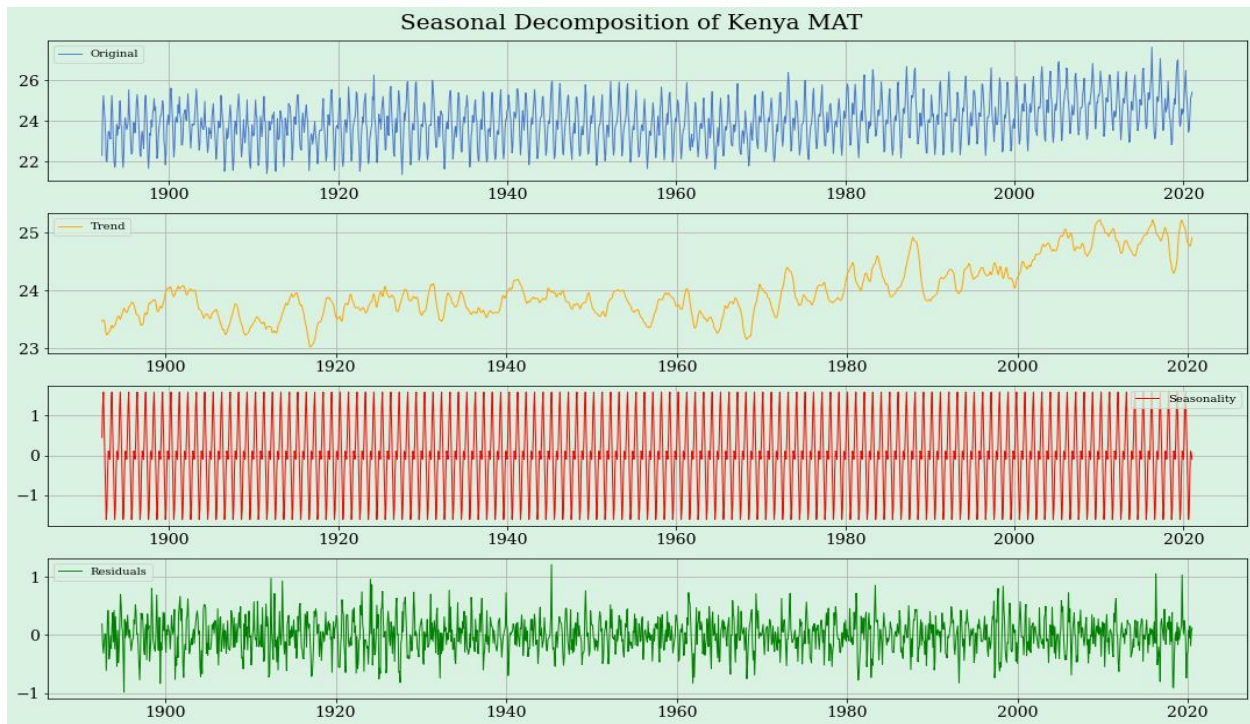


Figure 0.12 Seasonal decomposition of Uganda MAT

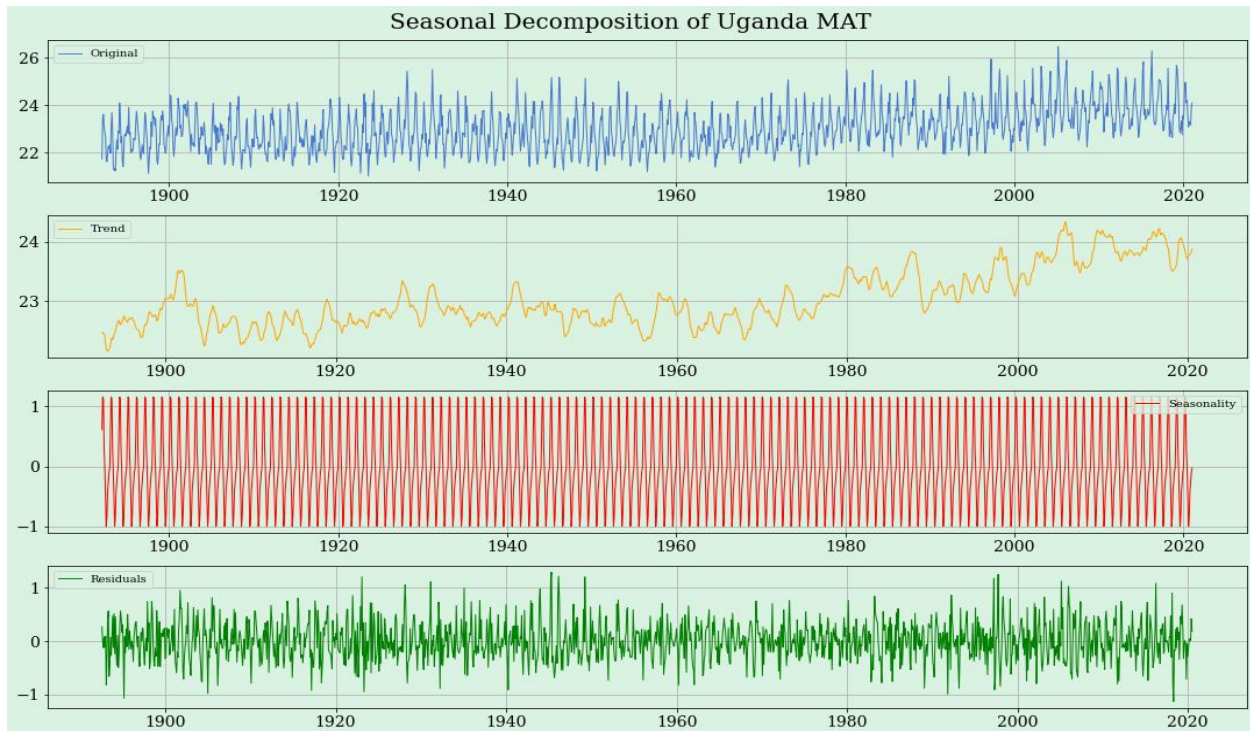


Figure 0.13 De-seasoned Global MAT

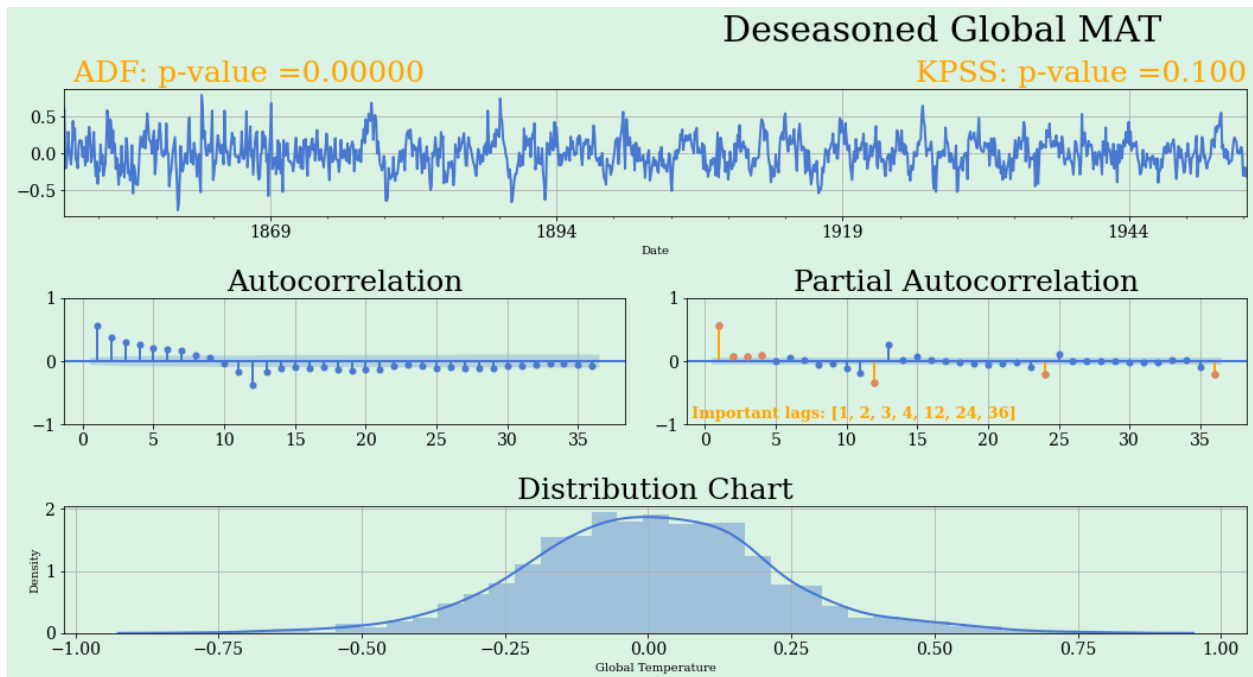


Figure 0.14 De-seasoned Ethiopia MAT

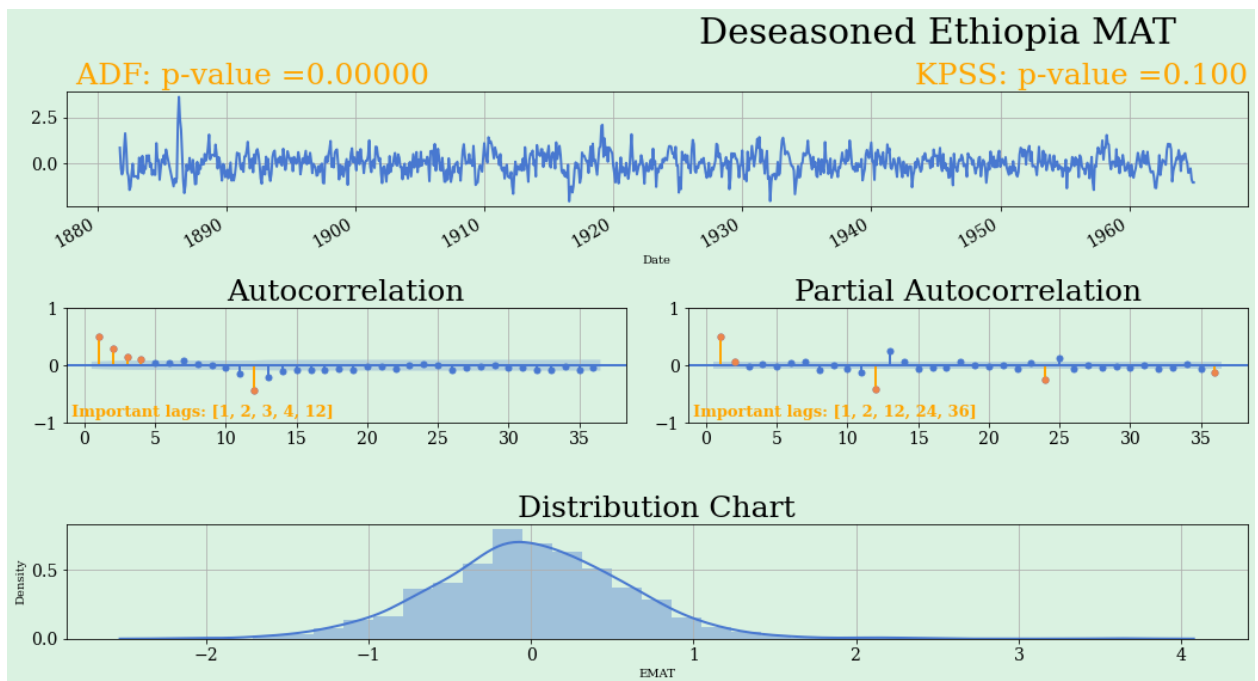


Figure 0.15 De-seasoned Kenya MAT

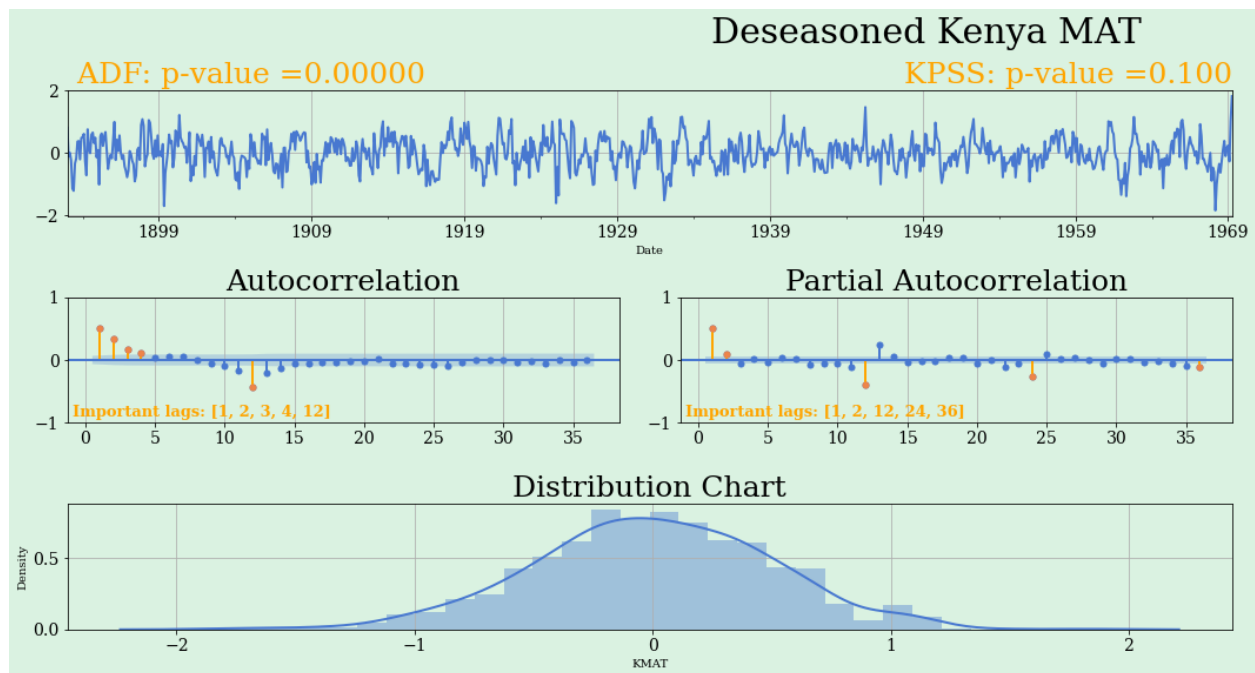


Figure 0.16 De-seasoned Uganda MAT

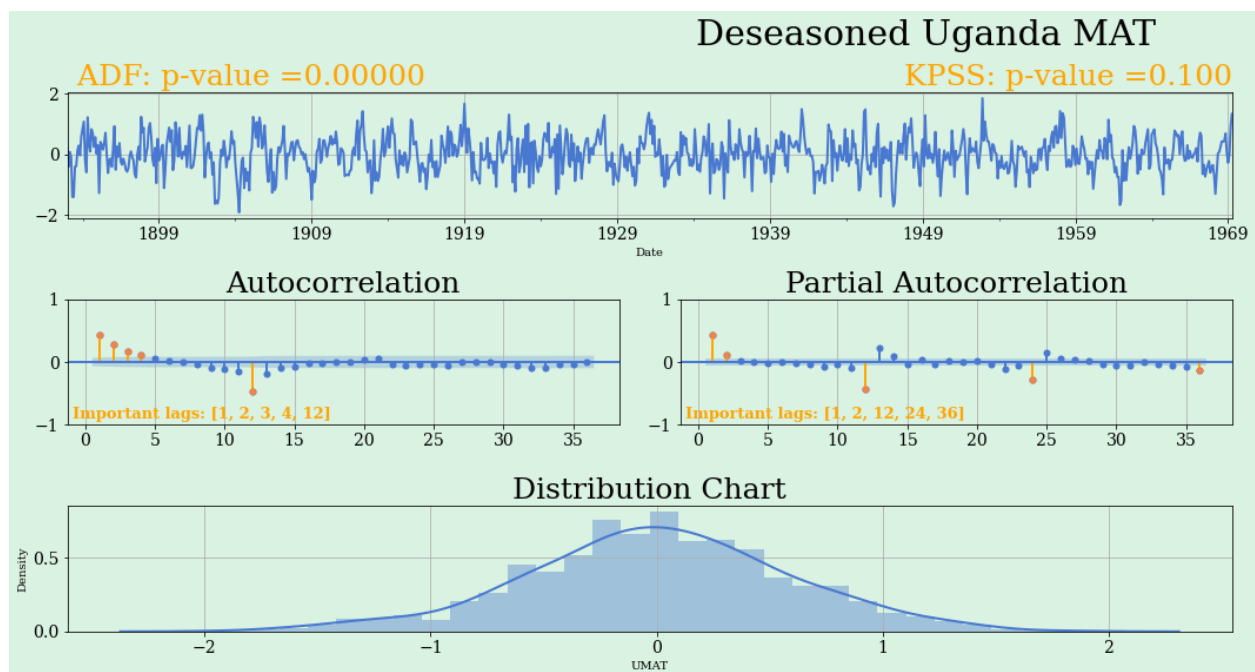


Figure 0.17 Bar-Chart of SARIMA models in validation and test set of Global MAT

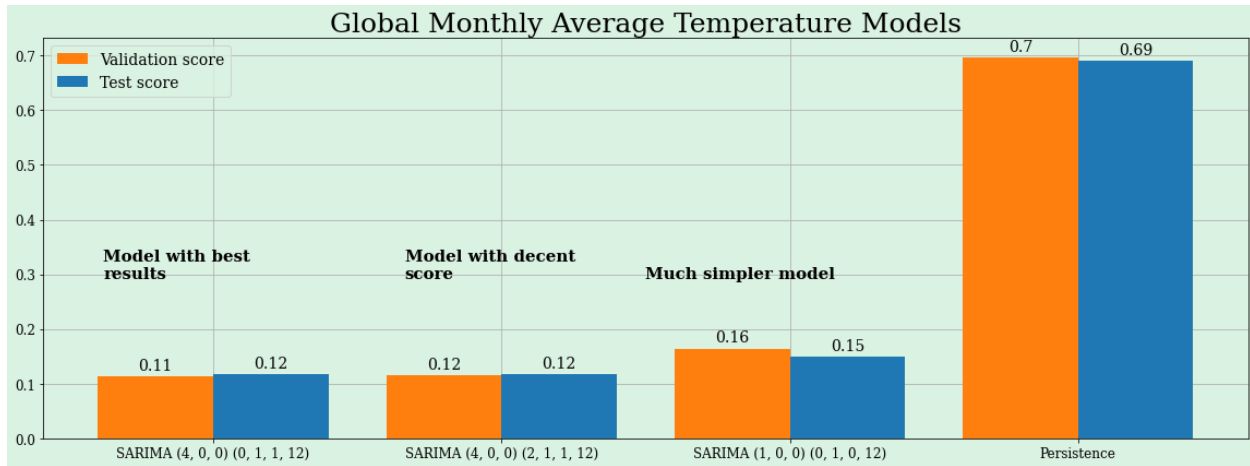


Figure 0.18 Bar-Chart of SARIMA models in validation and test set of Ethiopia MAT

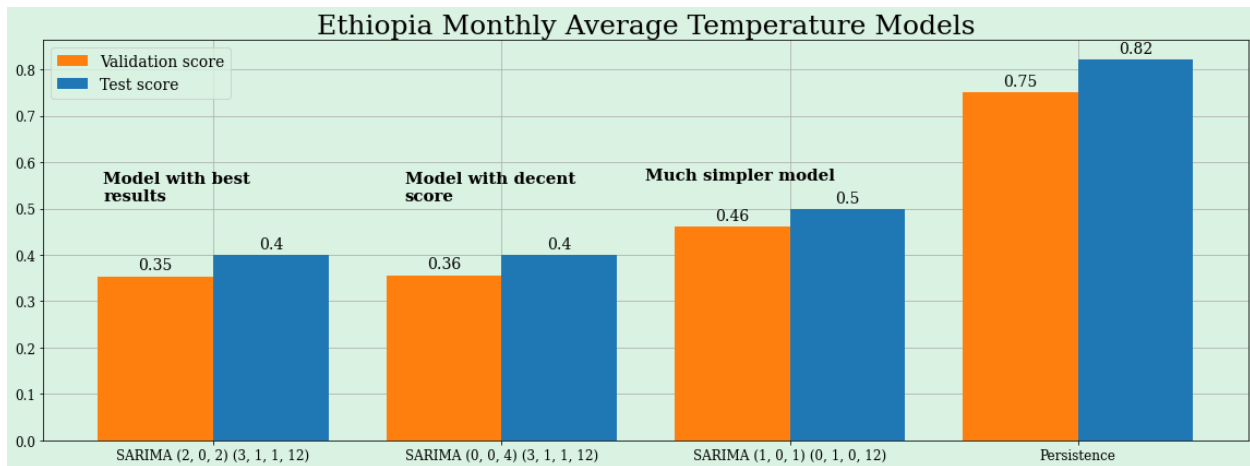


Figure 0.19 Bar-Chart of SARIMA models in validation and test set of Kenya MAT

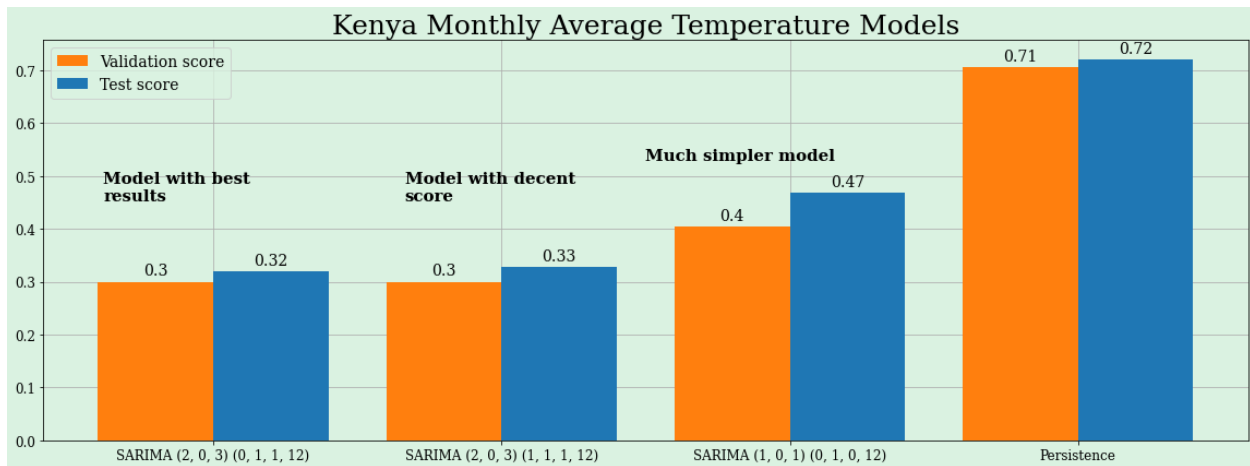
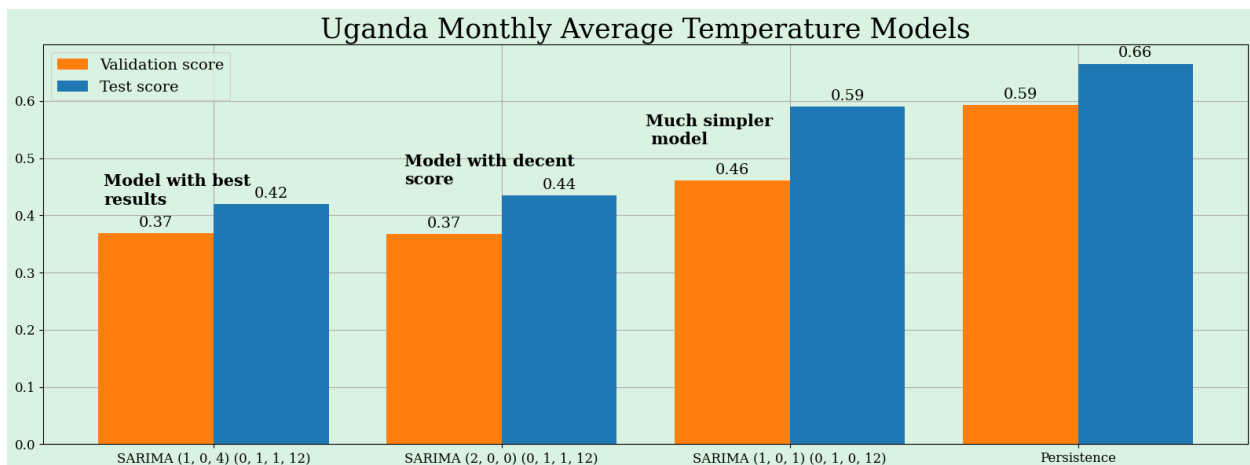


Figure 0.20 Bar-Chart of SARIMA models in validation and test set of Uganda MAT



Appendix B: Tables

Table 0.1 SARIMA model selection for Global MAT

	Parameters (p, q, P, Q)	aic	bic
0	(4, 0, 0, 1)	-2043.967308	-2011.487104
1	(4, 0, 2, 1)	-2043.656237	-2000.349298
2	(4, 0, 1, 1)	-2043.385849	-2005.492278
3	(4, 0, 3, 1)	-2042.674112	-1993.953806
4	(3, 0, 2, 1)	-2010.715495	-1972.821924
5	(3, 0, 1, 1)	-2010.529272	-1978.049068
6	(3, 0, 3, 1)	-2010.455295	-1967.148357
7	(3, 0, 0, 1)	-2010.412619	-1983.345782
8	(2, 0, 1, 1)	-1995.674178	-1968.607342
9	(2, 0, 0, 1)	-1995.180538	-1973.527068
10	(2, 0, 3, 1)	-1995.079811	-1957.186240
11	(2, 0, 2, 1)	-1994.910628	-1962.430424
12	(1, 0, 1, 1)	-1936.369425	-1914.715956
13	(1, 0, 2, 1)	-1935.941916	-1908.875079
14	(1, 0, 3, 1)	-1935.495415	-1903.015211
15	(1, 0, 0, 1)	-1934.652128	-1918.412026
16	(4, 0, 3, 0)	-1806.651268	-1763.344329
17	(3, 0, 3, 0)	-1779.136782	-1741.243210
18	(2, 0, 3, 0)	-1769.626500	-1737.146296

19	(1, 0, 3, 0)	-1729.555071	-1702.488234
20	(4, 0, 2, 0)	-1709.464130	-1671.570559
21	(3, 0, 2, 0)	-1683.354611	-1650.874407
22	(2, 0, 2, 0)	-1669.849214	-1642.782378
23	(1, 0, 2, 0)	-1632.574632	-1610.921163
24	(4, 0, 1, 0)	-1570.850201	-1538.369997
25	(3, 0, 1, 0)	-1551.200444	-1524.133607
26	(2, 0, 1, 0)	-1537.257337	-1515.603868
27	(1, 0, 1, 0)	-1502.048694	-1485.808592
28	(4, 0, 0, 0)	-1122.519753	-1095.452916
29	(3, 0, 0, 0)	-1112.718089	-1091.064619
30	(2, 0, 0, 0)	-1110.407041	-1094.166939
31	(1, 0, 0, 0)	-1084.637410	-1073.810676
32	(0, 0, 1, 1)	-1043.037134	-1026.797032
33	(0, 0, 3, 1)	-1042.595175	-1015.528339
34	(0, 0, 2, 1)	-1041.196135	-1019.542665
35	(0, 0, 0, 1)	-1021.485313	-1010.658579
36	(0, 0, 3, 0)	-931.369594	-909.716124
37	(0, 0, 2, 0)	-830.962170	-814.722068
38	(0, 0, 1, 0)	-718.811317	-707.984583
39	(0, 0, 0, 0)	-423.536104	-418.122737

Table 0.2 SARIMA model selection for Ethiopia, Kenya, and Uganda MAT

Ethiopia			Kenya			Uganda		
Parameters	AIC	BIC	Parameters	AIC	BIC	Parameters	AIC	BIC
(0, 4, 3, 1)	1365.62	1412.36	(2, 3, 0, 1)	698.973	734.748	(1, 4, 0, 1)	1130.724	1166.499
(2, 2, 3, 1)	1366.19	1412.93	(2, 3, 1, 1)	700.057	740.943	(2, 0, 0, 1)	1131.424	1151.867
(1, 4, 3, 1)	1366.76	1418.69	(1, 3, 1, 1)	700.417	736.192	(1, 4, 1, 1)	1131.977	1172.862
(1, 3, 3, 1)	1367.42	1414.16	(1, 3, 0, 1)	701.189	731.853	(2, 0, 1, 1)	1132.439	1157.993
(1, 2, 3, 1)	1367.76	1409.3	(1, 4, 1, 1)	702.094	742.979	(2, 2, 0, 1)	1132.567	1163.231
(2, 4, 3, 1)	1368.61	1425.73	(2, 4, 0, 1)	702.156	743.042	(1, 1, 0, 1)	1132.682	1153.125
(2, 2, 2, 1)	1370.01	1411.56	(2, 0, 1, 1)	702.326	727.879	(2, 1, 0, 1)	1133.07	1158.624
(0, 4, 2, 1)	1370.56	1412.11	(1, 3, 2, 1)	702.362	743.248	(1, 2, 0, 1)	1133.443	1158.997
(2, 2, 0, 1)	1371.23	1402.38	(1, 1, 1, 1)	702.486	728.04	(1, 1, 1, 1)	1133.579	1159.133
(2, 3, 3, 1)	1371.45	1423.38	(1, 3, 3, 1)	702.524	748.52	(1, 4, 2, 1)	1133.717	1179.713
(1, 0, 3, 1)	1371.69	1402.85	(1, 4, 0, 1)	702.612	738.387	(2, 2, 1, 1)	1133.861	1169.636
(2, 1, 3, 1)	1371.76	1413.3	(2, 3, 2, 1)	703.388	749.385	(2, 1, 1, 1)	1134.048	1164.712
(0, 4, 0, 1)	1371.81	1402.97	(1, 4, 2, 1)	704.05	750.046	(1, 3, 0, 1)	1134.217	1164.881
(1, 4, 2, 1)	1371.95	1418.69	(2, 0, 0, 1)	704.094	724.537	(2, 0, 2, 1)	1134.272	1164.937
(2, 0, 3, 1)	1371.95	1408.3	(2, 1, 1, 1)	704.167	734.831	(1, 2, 1, 1)	1134.388	1165.052
(2, 2, 1, 1)	1372.07	1408.42	(2, 0, 3, 1)	704.185	739.96	(2, 3, 0, 1)	1134.519	1170.294
(1, 1, 3, 1)	1372.3	1408.65	(1, 1, 0, 1)	704.208	724.65	(1, 4, 3, 1)	1134.917	1186.024
(0, 4, 1, 1)	1372.84	1409.19	(1, 4, 3, 1)	704.211	755.318	(2, 0, 3, 1)	1135.119	1170.894
(1, 3, 2, 1)	1373.03	1414.58	(2, 0, 2, 1)	704.281	734.945	(1, 1, 2, 1)	1135.218	1165.882
(1, 4, 0, 1)	1373.36	1409.71	(1, 1, 3, 1)	704.291	740.066	(1, 3, 1, 1)	1135.306	1171.081
(1, 2, 2, 1)	1373.41	1409.76	(2, 3, 3, 1)	704.339	755.446	(2, 2, 2, 1)	1135.626	1176.512
(2, 4, 2, 1)	1373.71	1425.64	(2, 4, 1, 1)	704.411	750.407	(2, 3, 1, 1)	1135.796	1176.682
(1, 4, 1, 1)	1374.39	1415.93	(1, 1, 2, 1)	704.458	735.123	(2, 1, 2, 1)	1135.804	1171.579
(1, 3, 0, 1)	1374.46	1405.62	(1, 2, 1, 1)	704.472	735.136	(1, 1, 3, 1)	1135.922	1171.697
(1, 2, 0, 1)	1374.63	1400.6	(2, 2, 1, 1)	704.996	740.771	(1, 2, 2, 1)	1136.11	1171.885
(2, 4, 0, 1)	1375.02	1416.57	(2, 1, 0, 1)	705.949	731.502	(2, 2, 3, 1)	1136.276	1182.273
(1, 3, 1, 1)	1375.57	1411.92	(2, 1, 3, 1)	706.04	746.926	(2, 4, 0, 1)	1136.517	1177.403
(1, 2, 1, 1)	1375.79	1406.95	(2, 1, 2, 1)	706.076	741.851	(2, 1, 3, 1)	1136.599	1177.485
(2, 4, 1, 1)	1376.05	1422.79	(1, 2, 0, 1)	706.203	731.757	(1, 2, 3, 1)	1136.873	1177.759
(2, 3, 2, 1)	1376.71	1423.45	(1, 2, 3, 1)	706.284	747.17	(1, 3, 2, 1)	1137.087	1177.972
(2, 3, 0, 1)	1376.76	1413.11	(2, 4, 2, 1)	706.358	757.465	(2, 3, 2, 1)	1137.561	1183.557
(2, 3, 1, 1)	1378	1419.55	(1, 2, 2, 1)	706.441	742.216	(2, 4, 1, 1)	1137.778	1183.775

(2, 1, 2, 1)	1378.76	1415.11	(2, 4, 3, 1)	706.524	762.742	(1, 3, 3, 1)	1137.976	1183.972
(2, 1, 0, 1)	1379.29	1405.26	(2, 2, 0, 1)	706.769	737.433	(2, 3, 3, 1)	1138.246	1189.353
(1, 0, 0, 1)	1380.03	1395.61	(2, 2, 3, 1)	708.289	754.285	(2, 4, 2, 1)	1139.538	1190.645
(2, 0, 2, 1)	1380.06	1411.21	(2, 2, 2, 1)	708.458	749.344	(2, 4, 3, 1)	1140.155	1196.373
(2, 0, 0, 1)	1380.06	1400.83	(1, 0, 1, 1)	715.767	736.21	(0, 4, 0, 1)	1149.517	1180.181
(1, 1, 0, 1)	1380.49	1401.26	(1, 0, 3, 1)	716.466	747.13	(0, 4, 1, 1)	1150.654	1186.429
(2, 1, 1, 1)	1380.64	1411.8	(1, 0, 0, 1)	717.357	732.689	(0, 4, 2, 1)	1152.545	1193.431
(1, 1, 2, 1)	1380.67	1411.82	(1, 0, 2, 1)	717.656	743.21	(0, 4, 3, 1)	1153.15	1199.146
(1, 0, 2, 1)	1380.7	1406.66	(0, 4, 1, 1)	723.477	759.252	(1, 0, 0, 1)	1173.144	1188.476
(1, 0, 1, 1)	1381.43	1402.2	(0, 4, 3, 1)	724.563	770.559	(1, 0, 1, 1)	1173.545	1193.988
(2, 0, 1, 1)	1381.49	1407.45	(0, 4, 2, 1)	725.393	766.278	(1, 0, 2, 1)	1175.413	1200.967
(1, 1, 1, 1)	1381.91	1407.88	(0, 4, 0, 1)	725.613	756.278	(1, 0, 3, 1)	1175.806	1206.47
(0, 3, 3, 1)	1382.3	1423.84	(0, 3, 1, 1)	754.038	784.702	(0, 3, 0, 1)	1177.636	1203.189
(0, 3, 0, 1)	1387.17	1413.14	(0, 3, 3, 1)	755.508	796.394	(0, 3, 1, 1)	1177.686	1208.351
(0, 3, 2, 1)	1387.29	1423.64	(0, 3, 2, 1)	756.034	791.809	(0, 3, 2, 1)	1179.656	1215.431
(0, 3, 1, 1)	1388.31	1419.47	(0, 3, 0, 1)	756.062	781.616	(0, 3, 3, 1)	1180.67	1221.556
(0, 2, 3, 1)	1422.39	1458.74	(0, 2, 1, 1)	782.846	808.399	(0, 2, 0, 1)	1204.138	1224.581
(0, 2, 0, 1)	1423.97	1444.74	(0, 2, 3, 1)	783.961	819.736	(0, 2, 1, 1)	1204.634	1230.187
(0, 2, 1, 1)	1425.11	1451.07	(0, 2, 0, 1)	784.537	804.98	(0, 2, 2, 1)	1206.62	1237.284
(0, 2, 2, 1)	1426.32	1457.48	(0, 2, 2, 1)	784.846	815.51	(0, 2, 3, 1)	1207.452	1243.227
(1, 2, 3, 0)	1437.23	1473.58	(2, 1, 3, 0)	847.526	883.301	(2, 0, 3, 0)	1292.337	1323.001
(2, 0, 3, 0)	1437.59	1468.74	(2, 0, 3, 0)	847.981	878.645	(2, 2, 3, 0)	1293.52	1334.405
(1, 1, 3, 0)	1438.12	1469.28	(1, 3, 3, 0)	848.943	889.829	(2, 1, 3, 0)	1294.329	1330.104
(0, 4, 3, 0)	1438.29	1479.83	(2, 2, 3, 0)	849.169	890.054	(1, 2, 3, 0)	1294.873	1330.648
(2, 1, 3, 0)	1438.3	1474.65	(1, 1, 3, 0)	849.199	879.864	(2, 3, 3, 0)	1295.491	1341.487
(2, 2, 3, 0)	1438.48	1480.02	(2, 3, 3, 0)	849.734	895.73	(1, 4, 3, 0)	1295.527	1341.523
(1, 3, 3, 0)	1438.61	1480.15	(1, 2, 3, 0)	850.127	885.902	(1, 1, 3, 0)	1295.786	1326.45
(1, 4, 3, 0)	1439.23	1485.97	(1, 4, 3, 0)	850.511	896.508	(2, 4, 3, 0)	1296.489	1347.596
(1, 0, 3, 0)	1439.51	1465.48	(2, 4, 3, 0)	851.728	902.835	(1, 3, 3, 0)	1296.666	1337.552
(2, 4, 3, 0)	1441.22	1493.15	(0, 4, 3, 0)	857.81	898.696	(0, 4, 3, 0)	1301.267	1342.153
(2, 3, 3, 0)	1441.22	1487.96	(1, 0, 3, 0)	860.306	885.859	(0, 1, 1, 1)	1316.384	1336.827
(0, 3, 3, 0)	1452.08	1488.43	(0, 3, 3, 0)	882.983	918.758	(0, 1, 2, 1)	1317.522	1343.076
(2, 2, 2, 0)	1473.7	1510.05	(0, 2, 3, 0)	903.552	934.217	(0, 1, 3, 1)	1317.688	1348.352
(0, 2, 3, 0)	1484.25	1515.41	(2, 1, 2, 0)	913.943	944.607	(0, 1, 0, 1)	1318.013	1333.345
(2, 0, 2, 0)	1487.11	1513.07	(2, 0, 2, 0)	914.372	939.926	(0, 3, 3, 0)	1323.493	1359.268

(1, 1, 2, 0)	1487.29	1513.25	(1, 1, 2, 0)	914.919	940.473	(1, 0, 3, 0)	1330.646	1356.2
(1, 2, 2, 0)	1488.96	1520.12	(2, 2, 2, 0)	915.785	951.56	(0, 2, 3, 0)	1342.104	1372.768
(2, 1, 2, 0)	1489.02	1520.18	(1, 2, 2, 0)	916.257	946.921	(2, 0, 2, 0)	1359.201	1384.754
(1, 0, 2, 0)	1489.33	1510.11	(1, 3, 2, 0)	916.438	952.213	(2, 2, 2, 0)	1360.169	1395.944
(1, 3, 2, 0)	1490.88	1527.23	(2, 3, 2, 0)	916.679	957.565	(2, 1, 2, 0)	1361.201	1391.865
(2, 3, 2, 0)	1492.45	1534	(1, 4, 2, 0)	917.704	958.589	(1, 2, 2, 0)	1361.584	1392.249
(1, 4, 2, 0)	1492.76	1534.3	(2, 4, 2, 0)	918.49	964.486	(1, 1, 2, 0)	1361.691	1387.244
(0, 4, 2, 0)	1493.63	1529.98	(1, 0, 2, 0)	920.346	940.789	(1, 4, 2, 0)	1361.919	1402.805
(2, 4, 2, 0)	1494.88	1541.61	(0, 1, 3, 1)	931.233	961.897	(2, 3, 2, 0)	1362.166	1403.052
(0, 3, 2, 0)	1507.29	1538.45	(0, 1, 1, 1)	933.964	954.407	(2, 4, 2, 0)	1362.613	1408.61
(0, 2, 2, 0)	1541.4	1567.36	(0, 1, 2, 1)	934.068	959.622	(1, 3, 2, 0)	1363.486	1399.261
(0, 1, 3, 1)	1556.07	1587.23	(0, 1, 0, 1)	936.484	951.816	(0, 4, 2, 0)	1366.61	1402.385
(0, 1, 0, 1)	1559.46	1575.03	(0, 3, 2, 0)	948.909	979.573	(0, 3, 2, 0)	1388.837	1419.501
(0, 1, 1, 1)	1561.01	1581.78	(0, 2, 2, 0)	971.057	996.611	(1, 0, 2, 0)	1390.245	1410.688
(0, 1, 2, 1)	1562.21	1588.17	(0, 4, 2, 0)	1016.09	1051.87	(0, 2, 2, 0)	1407.094	1432.647
(2, 2, 1, 0)	1595.73	1626.89	(0, 1, 3, 0)	1042	1067.56	(0, 1, 3, 0)	1441.006	1466.559
(0, 1, 3, 0)	1602.42	1628.39	(2, 1, 1, 0)	1056.81	1082.37	(0, 1, 2, 0)	1503.248	1523.69
(2, 0, 1, 0)	1618.55	1639.32	(2, 2, 1, 0)	1058.81	1089.48	(2, 0, 1, 0)	1512.691	1533.134
(2, 1, 1, 0)	1619.97	1645.93	(1, 3, 1, 0)	1059.81	1090.47	(2, 1, 1, 0)	1513.665	1539.218
(1, 2, 1, 0)	1620.1	1646.06	(2, 0, 1, 0)	1059.93	1080.38	(1, 2, 1, 0)	1514.909	1540.463
(1, 1, 1, 0)	1620.2	1640.97	(2, 3, 1, 0)	1060.12	1095.89	(2, 2, 1, 0)	1514.924	1545.588
(1, 3, 1, 0)	1621.81	1652.97	(1, 2, 1, 0)	1060.66	1086.21	(1, 3, 1, 0)	1516.42	1547.084
(2, 3, 1, 0)	1622.8	1659.15	(1, 1, 1, 0)	1061.43	1081.87	(2, 3, 1, 0)	1516.859	1552.634
(1, 4, 1, 0)	1622.83	1659.18	(1, 4, 1, 0)	1061.48	1097.25	(1, 4, 1, 0)	1517.344	1553.119
(2, 4, 1, 0)	1624.36	1665.9	(2, 4, 1, 0)	1062.12	1103	(1, 1, 1, 0)	1517.805	1538.248
(0, 4, 1, 0)	1627.36	1658.51	(0, 4, 1, 0)	1067.22	1097.88	(2, 4, 1, 0)	1518.439	1559.324
(1, 0, 1, 0)	1634.29	1649.87	(1, 0, 1, 0)	1069.61	1084.94	(0, 4, 1, 0)	1519.431	1550.096
(0, 3, 1, 0)	1643.61	1669.58	(0, 3, 1, 0)	1094.13	1119.69	(0, 3, 1, 0)	1534.095	1559.648
(0, 1, 2, 0)	1647.57	1668.34	(0, 1, 2, 0)	1106.47	1126.92	(1, 0, 1, 0)	1544.18	1559.512
(0, 2, 1, 0)	1682.31	1703.08	(0, 2, 1, 0)	1119.43	1139.87	(0, 2, 1, 0)	1552.579	1573.022
(0, 1, 1, 0)	1810.13	1825.71	(0, 1, 1, 0)	1279.24	1294.58	(0, 0, 2, 1)	1556.186	1576.629
(2, 3, 0, 0)	1913.09	1944.25	(0, 0, 2, 1)	1299.33	1319.77	(0, 0, 1, 1)	1556.521	1571.853
(2, 2, 0, 0)	1919.72	1945.68	(0, 0, 3, 1)	1299.98	1325.54	(0, 0, 3, 1)	1557.417	1582.971
(0, 0, 0, 1)	1950.6	1960.99	(0, 0, 1, 1)	1307.9	1323.23	(0, 0, 0, 1)	1558.97	1569.191
(0, 0, 1, 1)	1951.08	1966.66	(0, 0, 0, 1)	1312.68	1322.9	(0, 0, 3, 0)	1647.746	1668.189

(0, 0, 2, 1)	1951.46	1972.23	(2, 2, 0, 0)	1321.22	1346.78	(0, 1, 1, 0)	1657.313	1672.645
(0, 0, 3, 1)	1953.45	1979.41	(2, 1, 0, 0)	1358.83	1379.27	(0, 0, 2, 0)	1724.38	1739.712
(2, 0, 0, 0)	1967.56	1983.14	(2, 0, 0, 0)	1359.09	1374.43	(2, 0, 0, 0)	1824.59	1839.922
(1, 1, 0, 0)	1968.07	1983.64	(1, 1, 0, 0)	1360.6	1375.94	(1, 1, 0, 0)	1825.984	1841.316
(2, 1, 0, 0)	1969.53	1990.3	(1, 2, 0, 0)	1360.66	1381.1	(1, 2, 0, 0)	1826.42	1846.863
(1, 2, 0, 0)	1969.7	1990.47	(1, 3, 0, 0)	1361.85	1387.4	(2, 1, 0, 0)	1826.429	1846.871
(1, 3, 0, 0)	1971.17	1997.13	(2, 3, 0, 0)	1362.27	1392.94	(1, 4, 0, 0)	1828.176	1858.841
(1, 4, 0, 0)	1973.14	2004.3	(1, 4, 0, 0)	1363.25	1393.91	(1, 3, 0, 0)	1828.418	1853.972
(2, 4, 0, 0)	1974.95	2011.3	(2, 4, 0, 0)	1363.77	1399.54	(2, 2, 0, 0)	1828.419	1853.972
(0, 4, 0, 0)	1975.99	2001.95	(0, 4, 0, 0)	1365.74	1391.29	(2, 3, 0, 0)	1830.42	1861.084
(1, 0, 0, 0)	1979.64	1990.02	(1, 0, 0, 0)	1371.17	1381.39	(2, 4, 0, 0)	1831.838	1867.613
(0, 0, 3, 0)	1979.71	2000.48	(0, 0, 3, 0)	1375.95	1396.39	(0, 4, 0, 0)	1833.617	1859.17
(0, 3, 0, 0)	1990.4	2011.17	(0, 3, 0, 0)	1387.24	1407.69	(0, 3, 0, 0)	1842.374	1862.816
(0, 2, 0, 0)	2015.33	2030.91	(0, 2, 0, 0)	1411.57	1426.9	(1, 0, 0, 0)	1849.988	1860.209
(0, 0, 2, 0)	2030.47	2046.05	(0, 0, 2, 0)	1468.68	1484.01	(0, 2, 0, 0)	1862.848	1878.18
(0, 1, 0, 0)	2118.57	2128.95	(0, 1, 0, 0)	1540.97	1551.19	(0, 0, 1, 0)	1874.53	1884.751
(0, 0, 1, 0)	2157.69	2168.07	(0, 0, 1, 0)	1650.44	1660.67	(0, 1, 0, 0)	1937.824	1948.045
(0, 0, 0, 0)	2430.78	2435.97	(0, 0, 0, 0)	1869.44	1874.55	(0, 0, 0, 0)	2134.66	2139.771