

HEDGING CROP YIELDS USING TEMPERATURE DERIVATIVES

MASTER OF SCIENCE IN MATHEMATICAL SCIENCES THESIS

PRINCE BLACKSON DENNIS CHIRWA

UNIVERSITY OF MALAWI

JUNE 2024



HEDGING CROP YIELDS USING TEMPERATURE DERIVATIVES

Master of Science in Mathematical Sciences Thesis

By

Prince Blackson Dennis Chirwa

Bachelor of Education (Sciences)-University of Livingstonia

Submitted to the Department of Mathematical Sciences, School of Applied and
Natural Science in Partial Fulfilment for the Award of Master of Science in
Mathematical Sciences

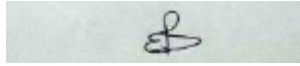
University of Malawi

June, 2024

DECLARATION

This thesis is my original work and has not been submitted to any other institution for similar purposes. Acknowledgements have been duly made where other people's work has been used. I bear the responsibility for the contents of this paper.

Signature:

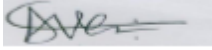
A rectangular box containing a handwritten signature in dark ink. The signature is stylized, with a large, looped initial 'P' followed by several smaller, less distinct characters.

Prince Blackson Dennis Chirwa

Date: 7th June, 2024

CERTIFICATE OF APPROVAL

The undersigned certifies that this thesis represents the student's work and effort and has been submitted with our approval.

Signature: 

Date: 7th June, 2024

Supervisor: Dr Nelson Christopher Dzupire

Signature: _____

Date: _____

Head of Department:

DEDICATION

I dedicate this work to my lovely wife, Alinafe Phiri-Chirwa and son, Dumisani Genius. Chirwa. They endured many days and nights of neglect just for the sake of this work.

ACKNOWLEDGEMENT

Firstly, I thank God, the Father Almighty, who provided light in times of darkness and thick clouds; He gave me hope when I had doubts and gave me strength, power and patience to complete this study. I exalt His name and adore Him.

As a continuation, I am indebted to express my sincere thanks to my supervisor, Dr Nelson Christopher Dzupire, for the support, guidance and words of encouragement he always provided to make this research project a successful one and for his constructive criticisms, his willingness and readiness to be consulted. When I started doubting my research success, he supported and encouraged me. He did not look at all kinds of mistakes as my downfall, but he looked at them and found the best way he could turn them into positive results.

I am also grateful to my Father, Mr Dennis J.S. Chirwa; Mother, Mrs Bridget M. Kachilika-Chirwa; brothers (Happy and Precious); sisters (Agnes and Queen); friend Happy Precious Nyasulu, and for their tireless support and encouragement they have given me throughout the process.

ABSTRACT

Agriculture production yield varies with weather changes. This causes farmers to incur losses. For instance, extreme temperature leads to low maize yield. This study describes incomplete temperature weather derivatives in agriculture markets and applies risk management hedging techniques. It focuses on hedging crop yield against extreme temperatures during irrigation farming, which is done without a greenhouse. This study primarily aims to hedge maize crop yields using temperature derivatives. This is achieved by (i) developing a daily average temperature stochastic model. (ii) Deriving statistical properties of the model based on the historical data of 31 years of our sample space (1990 – 2020 Kasungu District Temperature data). (iii) Pricing temperature derivatives to hedge maize crop yield. To achieve this, a stochastically Ornstein-Uhlenbeck process with the time-varying speed of reversion, seasonal mean, and local volatility that depends on the local average temperature was proposed. Based on the average temperature model, down and output, option pricing models for average temperature and growing degree day are presented. The study's findings suggest that the temperature will rise gradually but steadily. This scenario does not offer a positive outlook for agriculture production since a temperature rise can damage it. The premium for weather derivative options has been calculated as \$3.50 per GDD index contract. Farmers and agricultural stakeholders can hedge their crops against extreme temperature-related weather risks with these models. In line with Malawi's 2063 Millennium Development Goals (MDGs), this study acts as an eye opener for the government to put a policy on whether derivatives should be practised in our country hence, increase cash holding by improving the situation of the farmer and country.

TABLE OF CONTENTS

ABSTRACT	vi
LIST OF FIGURE	ix
LIST OF TABLES	x
LIST OF ACRONYMS	xi
CHAPTER ONE.....	1
INTRODUCTION	1
1.1 Background.....	1
1.2 Problem statement.....	5
1.3 Research objective	6
1.3.1 Main objective.....	6
1.3.2 Specific objective	6
1.4 Impact of the study	6
1.5 Limitation of the study	7
1.6 Structure of the study	7
CHAPTER TWO.....	8
LITERATURE REVIEW	8
2.1 Historical background of weather derivatives.....	8
2.2 Temperature process models.....	10
2.3. Pricing models of weather derivatives	13
CHAPTER THREE	16
METHODOLOGY	16
3.1 Daily average temperature data.....	16
3.2 Temperature indices.....	17
3.3 Daily average temperature Stochastic model	17
3.5 Pricing weather derivatives in an incomplete market.....	23
3.6 Temperature barrier option pricing	25
CHAPTER FOUR.....	29
RESULTS AND DISCUSSION	29
4.1 Descriptive statistics.....	29
4.2 Trend and seasonality	33
4.3 Estimation of mean reverting speed	35
Calculating the volatility of the temperature process.....	39
4.5 Hedging maize crop yield	42
CHAPTER FIVE.....	46
CONCLUSION AND RECOMMENDATION	46
5.1 Conclusion	46

5.2 Recommendation	47
5.3 Further studies	47
REFERENCES	48
MANUSCRIPT	53
APPENDIX	85

LIST OF FIGURE

Figure 4.1 shows a stochastic model of average daily temperature in Kasungu from 1990 to 2020.	30
Figure 4.2: Histogram of temperature differences of Kasungu district, Malawi.	31
Figure 4:4: Decomposed plot of the Average Temperature of Kasungu	38
Figure 4:5: shows the Auto Correlation Function (ACF) plot	38
Figure 4:6, a residuals graph.	39
Figure 4:7, Empirical volatility	40
Figure 4:8. Shows the backward forecasted temperature of Kasungu district.....	41
Figure 4:9. Shows the forward forecasted Temperature of the Kasungu district.	41

LIST OF TABLES

Table 4.1: Descriptive statistics of temperature for Kasungu.....	32
Table 4.2: Estimated parameter values for the seasonal mean	34
Table 3: Specification of GDD options for the year 2021	44

LIST OF ACRONYMS

AR (1):	Auto-regression of the first order
ARIMA:	Auto-regressive Integrated Moving Average
ACF:	Auto-Correlation function
CAT:	Cumulative Average Temperature
CDD:	Cooling Degree Day
CDF:	Cumulative Density Function
DAT:	Daily Average Temperature
FBM:	Fractional Brownian Motion
GARCH:	Generalised Auto-regressive Condition Heteroskedastic
GDD:	Growing Degree Day
HDD:	Heating Degree Day
MPE:	Mean Performance Error
MPR:	Market Price of Risk
O-U:	Ornstein Uhlenbeck
PDF:	Probability Density Function
RMSE:	Root Mean Square Error

CHAPTER ONE

INTRODUCTION

This chapter will discuss the study's background in section 1.1. Section 1.2 looks at the problem statement, followed by the study's objective in section 1.3. The study's impact has been discussed in section 1.4. This is followed by its impact, limitations, and structure in sections 1.5, 1.6, and 1.7, respectively.

1.1 Background

Crops are susceptible to rising atmospheric CO₂ concentrations and climate change, including temperature and precipitation variations (Rosenzweig, 2014; Wheeler, 2013). of all the changes, rising temperatures have the most significant potential to affect agricultural yields negatively (Ottman et al., 2012; Porter, 1999). Climate models can better predict changes in regional temperatures than changes in precipitation. According to meteorological data, the mean annual temperature in regions where soybean, wheat, rice, and maize are grown has risen by about 1 °C over the past century. It is predicted to rise even more if greenhouse gas emissions rise (Zhao, 2017). To feed a growing global population, it is imperative first to evaluate the danger to global food security and then estimate the impact of temperature increase on crop yields worldwide, considering any geographical variations (Nelson, 2010).

Agriculture is the backbone of the Malawian economy (Ministry of Agriculture, 2016). It is an important sector that generates income for most of the population, helps Malawi earn money through exports, and supplies most manufacturing industries. Many smallholder farmers in Malawi primarily get their income from the agriculture and agribusiness sectors, which comprise an essential portion of the country's critical economic activity. Most farmers grow maize, soya beans, tobacco, rice, and ground nuts. Maize is the main crop and staple food in Malawi.

Smallholder farmers grow maize for food, and they sell the excess. Low yields of maize are a yardstick of hunger in the country. According to reports, it contributes to more than 25% of Africa's GDP and roughly 70% of the labour force (Gyamrah, 2019). According to (the Ministry of Agriculture, 2016), It is reported that the agriculture sector supports about 85% of the population in terms of employment. It accounts for over one-quarter of Malawi's gross domestic product and 90% of the foreign exchange. As a result, the development of the African economies, of which Malawi is a member state, has agriculture as one of its most significant and prominent fields.

Malawi's agricultural production heavily depends on weather elements like rainfall, temperature, wind, etc. Any variations in these elements significantly affect harvest. For instance, the production of maize is greatly influenced by the weather. The stages of plant growth affect how each crop species responds to temperature variations. A specified range of maximum and minimum temperatures for each species determines the limits of visible growth. Extreme temperature makes it difficult for maize to develop. Extremely high temperatures during the reproductive stage will impact the viability of the pollen, the process of fertilization and the development of grains or fruits (Hatfield et al., 2011, 2015). The yield potential of initial grains or fruits will be reduced by repeated exposure to extremely hot or cold conditions during pollination. However, the most detrimental effects of acute exposure to vital events may occur during the reproductive stages. Because this measure is the one that producers and consumers are most concerned about, the effects of climate change are most noticeable in maize crop productivity. Maize grows well under temperatures of 21⁰c to 27⁰c. According to (Hatfield et al., 2015), the base temperature below which crop growth ceases is 10⁰c and above 38⁰ C.

This underlines the significance of contingency planning, which includes setting up backup funds, strengthening currency reserves to protect against external shocks, and validating the usage of weather insurance (Dzupire et al., 2019). Like insurance against poor crop yields, which can be used at the family level, weather derivatives offered to small farmers may do the same.

According to (Islam & Chakraborti, 2015), derivatives are "financial instruments linked to a specific monetary instrument, indicator, or commodity and through which specific risks can be traded in the capital markets in their own right." The value of a financial derivative is derived from the price of an underlying item, such as an asset or index. Unlike debt securities, no principal is advanced to be repaid and no investment income accrues." The underlying asset may take on a variety of shapes, including goods such as grain, coffee, and orange juice; precious metals such as gold and silver; currency exchange rates; and bonds of many kinds, such as medium- to long-term transferable debt securities issued by governments, businesses, etc. Finally, securities of corporations are marketed on reputable transactions of stocks and the Stock Index, including shares and share warrants.

Financial derivatives are essential for controlling the financial risk that corporations face. They have been highly successful, widely used, and valuable advances in the capital markets. Financial derivative markets have recently been discovered to be operating actively in both developed and developing nations. The principal application is hedging, often known as a function of price insurance, risk shifting, or risk transference. They give traders a means to manage their risks or shield themselves from unfavourable changes in the value of the underlying assets they work in. For instance, a farmer may take the market's risk and sell a futures contract to offset the risk.

According to Prabakaran (2018), a class of financial derivatives known as "weather derivatives" are those whose payment is based on weather factors like temperature, precipitation, wind, and snowfall. Organizations or individual farmers can use this financial tool as risk management to lower the risks associated with unforeseen weather situations. The derivative's seller accepts disaster risk in exchange for a premium. The seller will earn if no damages happen before the contract's expiration. The derivative buyer is entitled to the agreed sum in the event of unforeseen or unfavourable weather. Using temperature-based derivatives, one may now buy and trade a natural event. To do so, two things are necessary: a price enabling a transaction and a unit of measurement for the natural occurrence that everyone can agree on. According to A. K. Alexandridis and A. D. Zaprani (2012), contracts created for the weather derivative market based on temperature make up 98–99% of the trading asset.

The fundamental assets in temperature contracts are temperature indices like heating degree days (HDD) and cooling degree days (CDD).

In principle, insurance and derivatives may seem identical, yet they differ significantly in the buyer's motivation. A derivative is an investment vehicle in which the investor may stand to gain or lose money depending on whether they exercise their call option to redeem their investment before the derivative contract expires. Depending on the strike price, the investor could either make or lose money if the asset's value changes throughout the contract's period. A derivative so derives its value from the underlying asset or securities that are similar to it. In contrast, insurance is a means of risk transfer in which the insured assigns a portion of the risk to a third party and experts who will only be compensated in the event of a loss (Nicholson, 2018).

Insurance contracts and derivatives differ primarily in that the holder of an insurance contract must establish that he has experienced a monetary loss to be entitled to compensation. The insurance provider will withhold his payment if he cannot demonstrate this. No matter how the weather impacts the derivative's holder, payouts for weather-related contracts are only based on the weather's actual outcome. A weather derivative can be purchased and used to one's advantage without having any weather-sensitive production, for instance. These contracts can be purchased only for speculative purposes, just like other derivatives (Alaton; Djehiche; & Stillberger, 2010).

It is difficult for insurance contracts to function when there are uncertainties in average weather because they are often created to protect the holder from significant weather occurrences like earthquakes and typhoons. On the other hand, weather derivatives can be designed to have pay-outs in any weather scenario. Derivative contracts have another significant advantage over insurance contracts. Two actors might be available, one of whom will profit from a freezing winter while the other will profit from a warm one. These two parties can negotiate in a futures market and sign a contract to share risk-hedging responsibilities. In the insurance industry, this is not feasible (Alaton, Djehiche; & Stillberger, 2010).

In Malawi, where farming is an investment, derivatives make sense. Investors typically anticipate some return, though this is not always true. For instance, unlike crop insurance, no proof of abnormally high or low temperatures is required in order to get compensation. No matter how the temperature impacts the derivative holder, the payouts of temperature derivatives are only based on the temperature's actual result. Regardless, the farmer will always have access to derivatives to protect him against losses. If there are no losses, the farmer will make money. Derivatives are, therefore, more farmer-friendly than crop insurance.

1.2 Problem statement

Crop production faces several challenges, mainly due to exposure to unfavourable weather conditions, including temperature changes. Agriculture yield is significantly influenced by weather and climate (Geng et al., 2019). However, any restrictions are unacceptable at a time when agricultural yields are crucial to the security of the world's food supply. The introduction of farm management strategies, including controlled environment farming, is the consequence of this search for solutions. A form of controlled environment farming called greenhouse farming enables optimum management and lessens the risks associated with inclement weather. However, most farmers cannot use greenhouses due to their expensive cost. Agriculture has not adopted financial weather derivatives to the same extent as the energy sector. However, studies have looked at contracts based on historical data based on rainfall or heat index (Sun et al., 2013). The relationship between temperature and maize output is not always clear-cut because various products, stages of maize development, and soil textures, to name a few, respond to temperature differently. In order to focus on pricing financial derivatives for hedging maize crop output, Kooten (2015) analysed various methods of pricing weather derivatives options based on Growing Degree Day (GDD). Patricia P. (2021) developed the Growing Degree Day (GDD) European put option for rice in Laguna by simulating the daily average temperature. Weather derivatives and weather insurance may be used as tools for agricultural risk management, according to (Turvey, 2001), who looked into the price of weather derivatives in Ontario.

Zong, and Ender (2016) created a climatic zone-based growth degree-day contract, a new weather derivatives contract. Their goal was to reduce weather risk in mainland China's agriculture industry by providing new forms of temperature indexes. Although various weather risk management technologies have recently been brought into the Weather Derivatives (WD) market for smallholder farmers in the majority of nations across the world, purchases have been fewer than predicted. These studies ignored that such agriculture may also be done by irrigation, where temperatures are not regulated, as opposed to agriculture during the rainy season, where raindrops control the air temperature (Means, 2018). According to Quindala and Cuaresma (2021), every 1°C increase during the dry season might result in a 10% decrease in yield. As a result, this study focuses on hedging crop production against excessive temperatures during irrigation farming, which is done without a greenhouse since it is prone to extreme temperatures, which damage crop yields.

1.3 Research objective

1.3.1 Main objective

The main objective is to hedge crop yields using temperature derivatives

1.3.2 Specific objective

Specifically, the objectives of the study are:

1. Develop a daily average temperature stochastic model
2. Derive statistical properties of the model based on historical data
3. Pricing temperature derivatives to hedge crop yield

1.4 Impact of the study

This study will help companies and individuals overcome crop yield loss due to temperatures. They can explore the possibility of hedging the risks due to weather by buying temperature derivatives that guarantee a payoff once there is forecasted erratic temperature. The seller of the temperature derivative agrees to bear the risk of disasters in return for a premium. If no damage occurs before the expiration of the contract, the

seller will make a profit. In the event of unexpected or adverse weather, the buyer of the derivative claims the agreed amount, thereby being protected from the loss. This means that providing the derivatives means increasing cash holding and food, which aligns with Malawi's 2063 vision MIPI, thereby improving the poverty situation.

1.5 Limitation of the study

There still needs to be an organised market for derivatives in Malawi. However, the results of this study can open new products that can be offered by agencies that already offer agricultural insurance products.

1.6 Structure of the study

The study is structured as follows. We begin with the next chapter, which reviews related literature to see what other scholars and researchers have said about modelling temperatures. The third chapter discusses the development of daily average temperatures, followed by the stochastic method for simulating daily average temperatures and a description of a weather index distribution method to price weather derivatives. We end with the fourth and fifth chapters by discussing and analysing the results and making some concluding remarks.

CHAPTER TWO

LITERATURE REVIEW

This chapter will review the historical background of weather derivatives in section 2.1. It also looks at how temperature and pricing derivative models have been reviewed in sections 2.2 and 2.3, respectively.

2.1 Historical background of weather derivatives

Weather is a significant factor in many commercial operations, and unforeseen weather catastrophes can result in substantial financial losses. For instance, mild summers lessen demand for power and hence utility businesses' profit margins; less snowfall in the winter raises ski resort running expenses; and droughts result in poorer agricultural output. Weather risk refers to the unfavourable financial effects of climatic variation. As per (Allianz, 2013), weather and climate directly or indirectly impact nearly 30% of US GDP (\$5.7 trillion out of \$15.7 trillion), and extreme weather risk affects 70% of US enterprises. The same survey also shows how vulnerable many firms are to even minute variations in the weather. According to estimates, regular weather volatility impacted \$534 billion, or 3.4%, of the US GDP in 2012.

Weather derivatives, catastrophe bonds, and insurance can all be used to reduce weather risk. Traditionally, insurance has handled severe weather risk on an indemnity basis. The disadvantage of this approach is that filing claims is typically expensive and time-consuming. Furthermore, as (Hess, et al., 2002) point out, there is a shortage of acceptable forms of collateral, government interventions, and information asymmetry in private insurance, which results in high unit transaction costs, restricted distribution of institutions, and restricted access to people experiencing poverty.

Catastrophe bonds were first introduced in the mid-1990s. Through these bonds, weather risk can be shifted to capital market investors, who can access a considerably greater pool of resources than the insurance sector. Bond payments are correlated with an industry-wide loss or a weather index in catastrophe bonds, which mitigates the knowledge asymmetry issue. In comparison to insurance contracts, they also offer more prompt payments.

Weather derivatives convey weather risk to capital market investors in a manner akin to catastrophe bonds. Their profits are reliant on easily measurable climatic occurrences. Weather derivatives are intended to hedge the adverse financial effects of regular weather variance rather than those of exceptional weather-related occurrences, in contrast to catastrophe bonds. As a weather option included in a power contract, the first weather derivative transaction was carried out by Aquila in 1997 (Considine, 2000). Since then, the deregulation of the US energy sector has led to a sharp expansion of the weather derivatives market.

Therefore, improved weather risk management is necessary when deregulation increases competition among energy suppliers. The range of meteorological variables covered by weather derivatives has increased to include temperature, precipitation, snowfall, wind speed, humidity, and other factors to meet the demands of diverse market players. They have durations ranging from a week to several years, and their transaction sizes cover tiny risks up to several hundred million dollars or more for more extensive exposures. The London International Financial Futures and Options Exchange (LIFFE), the Chicago Mercantile Exchange (CME), and the Intercontinental Exchange (ICE) now offer standardized weather derivatives, even though the majority of weather derivative transactions still take place over the counter.

We focus on the most traded weather derivatives in the market, temperature derivatives, and this study focuses on the GDD Index. Two main steps are usually involved in pricing temperature derivatives. Firstly, a model for temperature dynamics must be built. Next, the second phase involves utilizing the presumptive model to obtain analytical or numerical estimations of weather derivative pricing to hedge crop yield.

2.2 Temperature process models

According to (Dzupire et al., 2018; Alaton, 2002), the temperature process's four characteristics include seasonality, long-term trends, unpredictability, and mean reversion. The temperature process varies in value throughout the year due to seasonality. A well-defined model should include all of these temperature characteristics.

Dornier and Queruel, (2000) modelled temperature fluctuations as a regression between daily deseasonalized temperatures. The suggested model divides the daily average temperature evolution into two sections, namely seasonal trend and random walk. Seasonal change and global warming are included in the formulation of seasonality as a sine function. However, this model was lacking and did not consider the seasonality and instability of temperature.

By describing volatility as a piece-wise constant function that represents the monthly variance in volatility, (Alaton, 2002) updated the model by Dornier and Queruel (2000). Their choice of volatility was confirmed by the observation that the quadratic variation of the volatility was practically constant over each month in the data set. Despite the lack of a statistical test for normality, the Wiener process was chosen as the driving noise in the model since the observed temperature differences were almost normal. Possible time dependencies in the residuals seen in the regression models are not mentioned in either study.

According to (Brody, 2002), temperature dynamics are represented using a stochastic process known as fractional Brownian motion, where the temperature change is regressed on the deseasonalized temperature from the day before. Based on a data series of daily average temperatures from Central England, it was possible to see evidence of fractional behaviour in temperature swings after subtracting the seasonal mean. The authors did not conduct the same fractional analysis for the model's residuals, as noted

by (Benth, & Šaltyt-Benth, 2005). Hence, it is unclear if the residuals' time dependence will exhibit characteristics of fractional noise. According to the argument (Dornier & Queruel, 2000), one should include seasonal variation changes in the model to obtain a reliable mean reversion model.

A mean-reverting model powered by a Lévy process was developed by (Benth, 2003; Dornier & Queruel, 2000) based on the Ornstein-Uhlenbeck (O-U) model. This came about because actual data from Norway rejected the normality test. In addition, the model's variance is a function with an empirical foundation estimated from observed variances. A flexible distribution that can capture the semi-heavy tails and skewness shown in the data is recommended: a generalized hyperbolic distribution. The Lévy process complicates the concept, which assumes constant mean reversion speed. Additionally, the application of AR (1) misses the gradual decrease in temperature autocorrelations, which may materially undervalue weather derivatives.

In a related study (Benth & Šaltyt-Benth, 2005), the authors developed an O-U mean reversion model with Brownian motion as the driving noise and the seasonal mean and volatility as truncated Fourier series. Both series' order was decided upon randomly, and no statistical analysis was done to determine the importance of each parameter. A closed-form solution for pricing weather derivatives was discovered to be possible using the model, which could also capture temperature dynamics. In Benth and Šaltyt-Benth (2011), a different model from the O-U is put forth as a continuous time autoregressive model for the temperature dynamics, with the volatility function being the product of the seasonal function and a stochastic process using the Barndoff-Nielsen and Shephard model for stochastic volatility.

Alexandridis and Zapranis (2006) extended the O-U mean reverting model developed by (Benth & Šaltyt-Benth, 2005) with seasonality in the level and volatility, validated by more than 100 years of temperature data collected in Paris. Unlike (Benth, & Šaltyt-Benth, 2005) here, wavelet analysis is used to identify the seasonality component in the temperature process and the volatility of the temperature residuals. It was observed that the distribution statistics of the residuals of AR (1) showed the presence of negative

skewness and positive kurtosis (> 3), indicating a significant deviation from the normal distribution. In addition, the effect of replacing the AR (1) process with ARMA, ARFIMA, and ARFIMA-FIGARCH was also explored. However, all these processes failed to capture the slow time decay of the autocorrelations of temperature.

In order to study the time dependence of the mean reversion parameter $k(t)$ on time, Alexandridis (2008) created an O-U stochastic temperature model driven by the Wiener process. He then employed neural networks to analyse the model's results. The model is discretized as an AR (1) model and is an extension of the (Benth & Šaltyt-Benth, 2005) generalization of (Dornier & Queruel, 2000) work.

A series of daily values of $k(t)$ are obtained by non-parametrically estimating the temperature process using neural networks and computing the derivative of the network output concerning the input. This eliminates the restriction of a constant mean reversion speed in some models. The results indicate that the speed of mean reversion varies quite a bit daily; therefore, expressing the speed as a function of time increases the model's accuracy and dramatically lowers the cost of weather derivatives.

For the Zhengzhou region, Wang et al. (2015) created a workable model of the daily average temperature that is used in weather derivative pricing. Then, using 62 years of daily historical data, they applied the mean-reverting Ornstein-Uhlenbeck process to characterize the temperature evolution after researching the history of the weather derivatives market. Chen et al., . (2018) used SARIMA (Seasonal Autoregressive Integrated Moving Average) techniques to analyse the monthly mean temperature in Nanjing, China, from 1951 to 2017. The training set includes data from 1951 to 2014, and the testing set includes data from 2015 to 2017.

Additionally, Nury and Koch (2013) established ARIMA (Auto Regressive Integrated Moving Average) models and applied them to provide short-term forecasts of the monthly maximum and minimum temperatures in the northeast Bangladeshi districts of Sylhet and Moulvibazar. Like Benth and Šaltyt-Benth (2005) and Alaton, (2002), this

O-U mean reversion model uses the truncated Fourier series to express volatility as a cyclic function.

2.3. Pricing models of weather derivatives

The literature empirically demonstrates the presence of climate modelling and its impact on the electricity market, as Weron (2014) demonstrates. This study's findings imply that temperature may partially explain the predictable behaviour of power pricing, which has been consistently proven by the literature on electricity price forecasting to be governed by geographical and temporal variables.

Uncertainty was reduced using an earlier strategy (Atalon et al., 2002; Richards, 2004) and later adopted by (Svec & Stevenson, 2007; Taştan & Hayfavi, 2017). The Sydney accumulated heating degree day (HDD) and cooling degree day (CDD) index levels were forecasted by Svec and Stevenson (2007) using time series and stochastic techniques to model pricing. Similarly, (Alaton, 2002) developed a pricing model for weather derivatives with pay-outs based on temperature, for which they conducted Monte Carlo simulations on a mean-reverting process. Using a mean-reverting process driven by a Lévy process to depict jumps and other temperature aspects that may play a significant role in the deviations for specific places (Taştan & Hayfavi, 2017) computed the temperature index. In conclusion, a generic pricing strategy for weather derivatives was proposed by (Richards, 2004) for cooling degree day weather options in Fresno, California. The temperature follows a mean-reverting Brownian motion process with discrete jumps and autoregressive conditional heteroscedastic errors, which they discovered using specification tests. They developed an equilibrium price model based on this procedure for the cooling degree day weather possibilities.

Weron (2014) offers more justification for his projection of the power price. Weron (2014) sought to clarify the complexity of the currently available solutions, their advantages and disadvantages, and the potential dangers that forecasting tools present.

The Black-Scholes equation was then used by (Prabakaran, 2018) to construct an option-pricing model for energy derivative markets.

Many writers have utilized the Black-Scholes model (Black & Scholes, 1973), which can be employed independently or as part of a portfolio, to estimate option pricing based on features. The Black-Scholes model's applicability to weather derivatives has been examined in the literature (Botoş & Ciumaş, 2012). The Black-Scholes model, however, did not perform well with a weather index, according to the results, which were attributed to mathematical and economic discrepancies. According to these authors (Botoş & Ciumaş, 2012), the models could be applied to portfolios that include weather derivatives. In addition, (Brody, 2002) proposed a fractional version of the Ornstein-Uhlenbeck process. By taking into account the memory effect and partial differential equations, they were able to derive the cost formulas for the most widely used temperature indices, such as heating Degree Day (HDD), cooling Degree Day (CDD), and cumulative average temperature (CAT). The same fractional Ornstein-Uhlenbeck process was then used in (Benth, 2003) to propose an arbitrage-free model for derivatives based on temperature, together with quasi-conditional expectation and Wick Itô Skorohod (WIS) integrals of the fractional Brownian motion (fBm). Oksendal (2014) created the WIS integrals of the fBm. In addition to the pricing for contracts as in (Benth, 2003), the dynamics of the option values are derived. The autocorrelation feature of the return time series is attempted to be modelled using various approaches. Fractional Brownian motion (fBm) is one of these attempts. The primary source cited was Oksendal (2014) (Prabakaran et al., 2020). Fraction Brownian Motion (FBM) was utilized to price temperature derivatives on energy using option pricing.

Although financial weather derivatives in agriculture have not been as widespread as in the energy industry, studies have examined using historical data to create heat or rainfall index-based weather derivatives (Sun, & Lou, 2013). This study proposes weather derivatives in agriculture hedging crop yields using temperature derivatives developed by (Patricia, 2021) that modelled the average temperature and designed a growing degree day (GDD) European put option for rice in Laguna. Patricia's (2021) study concentrated on modelling temperature during the rainy season, which is similar to what (Kooten, 2015 and Sun, 2017) did. However (Kooten, 2015) also concentrated

on pricing using financial derivatives to protect the output of corn crops by contrasting several approaches to calculate the price of weather derivatives using Growing Degree Day (GDD)

It has been observed that these studies (Sun, 2017; Kooten, 2015; Patricia, 2021; Sun, & Lou, 2013) did not take into consideration that such agriculture can also be done through irrigation where temperatures are not controlled, as compared to agriculture done in the rainy season where the rain drops to control the air temperature (Means, 2018). It has also been observed that their pricing is just a vanilla option, which is also expensive. Hence, this study focuses on hedging crop yield against extreme temperatures during irrigation farming, which is being done without a greenhouse as it is prone to an extreme temperature that affects crop yields. Furthermore, it involves down-and-out barrier option pricing, which is cheaper than the vanilla option.

To sum up this chapter, these models reviewed will help develop temperature stochastic models and pricing models in the agricultural sector in the next chapter. This will involve utilizing the presumptive model to obtain analytical or numerical estimations of weather derivative pricing to hedge crop yield.

CHAPTER THREE

METHODOLOGY

This chapter will discuss different methods and principles used in the study. These will include the sampling procedure, data collection tool and techniques and method data of analysis. Such that in section 3.1 we look at daily average temperature data. This is followed by Temperature indices in section 3.2. In section 3.3, we have a daily average temperature stochastic model. While pricing weather derivatives in an incomplete market and temperature barrier option pricing will be discussed in sections 3.4 and 3.5, respectively

3.1 Daily average temperature data

The data in this study is from the Kasungu district. It has been chosen because it is the top-list district that grows maize under irrigation during the rainy and dry seasons, which feeds a good percentage of Malawians (Phiri, 2016). Low maize yields in this district will lead to hunger and affect the country's economy.

From January 1 1990, through December 31 2020, historical data on daily minimum and maximum temperatures were gathered from the Department of Climate Change and Meteorological Service's headquarters. This sample size of 31 31-year years will ensure we accurately capture the long-term trend to forecast future temperatures. The missing data was treated by finding the average temperature adjacent to the missing data, carrying forward the last observation, and carrying backwards the following observation. All the leap years have been dropped.

3.2 Temperature indices

The temperature information collected from a given station in a given area is the foundation for temperature indices. If T_i^{max} and T_i^{min} are the highest and lowest temperatures recorded at a meteorological station on day i , respectively, the Daily Average Temperature (DAT) on that day is calculated as

$$T_i = \frac{(T_i^{max} + T_i^{min})}{2} \quad (1)$$

A base temperature, the required temperature for a crop to germinate, and the daily average temperature (DAT) are two variables that determine how many degrees a day there are. The heat that must build up each day for a crop to grow, sprout new leaves, reach the reproductive stage, and eventually mature is known as the growing degree day (Patricia P. et al., 2021). We consider maize crops to be produced under irrigation from August to December. Maize requires at least 10°C. GDD is defined as:

$$GDD_i = \max\{T_i - 10, 0\} \quad (2)$$

and the GDD for the whole 153-day period as,

$$GDD = \sum_{i=1}^{153} GDD_i \quad (3)$$

The 153-day period spans the period for maize crops from germination up to harvest.

3.3 Daily average temperature Stochastic model

Modelling daily average temperature is a hazardous task because multiple variables govern weather. We will obtain information about temperature behaviour, which is possible to assume as regular because temperature changes follow a cyclical pattern, although with some variability. Hence, it is essential to carefully validate the specified model before putting it into practical use for pricing weather derivatives. The modelling approach specifies a stochastic process of temperature evolution by selecting it from a parameterized family of processes.

3.4 A Gaussian Ornstein-Uhlenbeck Model for Temperature

Temperature has a mean-reverting evolution, meaning it can only briefly vary from its mean value. The mean-reverting Ornstein-Uhlenbeck process predicts temperature

behaviour by considering seasonal variation and long-term trends. The daily average temperature variation will return to the mean in the long run.

We consider the stochastic process model by putting together seasonality, long-term trends, unpredictability, mean reversion, and temperature.

$$dT_t = dT_t^m + k(T_t^m - T_t)dt + \sigma_t dW_t \quad (4)$$

Proposition 3.1 (Alaton, 2010; Brody, 2002; Peter Alaton, 2002): If the Daily Average Temperature (DAT) follows an Ornstein-Uhlenbeck process with a mean reversion that is mean-reverting and whose speed varies over time as well as a seasonal mean and variance:

$$dT_t = dT_t^m + k(T_t^m - T_t)dt + \sigma_t dW_t$$

The Ito formula yields the following implicit solution:

$$T_t = T_t^m + e^{\int_0^t k(u)du} (T_s^m - T_s) + e^{\int_0^t k(u)du} + \int_s^t \sigma_s e^{-\int_s^t k(u)du} dW_t$$

which is the same as the following:

$$T_t = [T_s^m - T_s]e^{-k(t-s)} + T_t^m + \int_s^t e^{-k(t-r)} \sigma_r dW_r \quad (5)$$

Proof: Let us rewrite $dT_t = dT_t^m + k(T_t^m - T_t) + \sigma_t dW_t$ as

$$d\hat{T}_t = k\hat{T}_t dt + \sigma_t dB_t$$

Where $\hat{T}_t = (T_t - T_t^m)$. The following transformation is practical for solving the stochastic equation above:

$$G(\hat{T}, t) = e^{-\int_0^t k(u)du} \hat{T}_t$$

Making use of the Ito lemma

$$\frac{\partial G}{\partial \hat{T}} = e^{-\int_0^t k(u)du}, \quad \frac{\partial^2 G}{\partial \hat{T}^2} = 0, \quad \frac{\partial G}{\partial t} = -k e^{-\int_0^t k(u)du} \hat{T}_t$$

and $a = k(t)\hat{T}_t$

$$b = \sigma_t$$

We have that

$$dG_t = \left(k e^{-\int_0^t k(u) du} \hat{T}_t - k e^{-\int_0^t k(u) du} T_t \right) dt + \sigma_t e^{-\int_0^t k(u) du} dB_t$$

Which reduces to

$$dG_t = \sigma_t e^{-\int_0^t k(u) du} dB_t$$

Suppose $s < t$, then integrating the above equation, we have that

$$G_t - G_s = \int_s^t \sigma_t e^{-\int_0^t k(u) du} dB_t$$

And replacing G , we have that $e^{-\int_0^t k(u) du} \hat{T}_t - \hat{T}_s = \int_s^t \sigma_t e^{-\int_0^t k(u) du} dB_t$

By rearrangement, we achieve that

$$\hat{T}_t = e^{-\int_0^t k(u) du} \hat{T}_t + e^{-\int_0^t k(u) du} \int_s^t \sigma_t e^{-\int_0^t k(u) du} dB_t$$

Since $\hat{T}_t = (T_t - T_t^m)$,

$$(T_t - T_t^m) = e^{-\int_0^t k(u) du} (T_s - T_s^m) + e^{-\int_0^t k(u) du} \int_s^t \sigma_t e^{-\int_0^t k(u) du} dW_s$$

Finally, by rearranging, we prove the proposition

$$T_t = T_t^m + e^{-\int_0^t k(u) du} (T_s - T_s^m) + e^{-\int_0^t k(u) du} + \int_s^t \sigma_s e^{-\int_s^t k(u) du} dW_s$$

Since $s = t - 1$ and $u = t - s$ hence

$$T_t = [T_{t-1} - T_{t-1}^m] e^{-k(t-s)} + T_t^m + \int_s^t e^{-k(t-r)} \sigma_r dW_r \quad \blacksquare$$

In equation (5), W_r is the Brownian motion, σ_t is the deterministic function of time t , the random variable $\int_s^t e^{-k(t-r)} \sigma_r dW_r$ is normally distributed with mean zero and a variance $\int_s^t e^{-2k(t-r)} \sigma_r^2 dW_r$.

The Brownian motion's characteristic of independent increments forms the foundation of the proof. The filtration $\mathcal{F}_s$ leads us to the conclusion that T_t is normally distributed, with mean and variance given by:

$$E^p[T_t|\mathcal{F}_s] = [T_s - T_t^m]e^{-k(t-s)} + T_t^m \quad (6)$$

$$V_t^2 = \text{var}[T_t|\mathcal{F}_s] = \int_s^t e^{-2k(t-r)} \sigma_r^2 dr. \quad (7)$$

We now explicitly describe the equation's (3) expression when $s = t - 1$. I get

$$T_t = [T_{t-1} - T_{t-1}^m]e^{-k(t-s)} + T_t^m + \int_{t-1}^t e^{-k(t-r)} \sigma_r dW_r. \quad (8)$$

The variable $\int_{t-1}^t e^{-k(t-r)} \sigma_r dW_r$ is a random variable having a Gaussian distribution, mean zero, and variance:

$$V_t^2 = \text{var}[T_t|\mathcal{F}_{t-1}] = \int_{t-1}^t e^{-2k(t-r)} \sigma_r^2 dr. \quad (9)$$

Therefore, we can write (8) as

$$T_t = [T_{t-1} - T_{t-1}^m]e^{-k} + T_t^m + \sigma_t \varepsilon_t. \quad (10)$$

This means that (6) has a simple discrete time representation with an autoregressive structure of order 1 (AR (1)). This result has significant implications for the estimate of the unknown. This indicates that (6) has a straightforward discrete time representation with an order 1 autoregressive structure (AR (1)). This outcome significantly impacts the estimation of the unknown parameter in the equation above. Since the variable in the case of T_t is observed at discrete points in time rather than constantly, estimating in continuous time is generally quite challenging. The likelihood function can only be analytically expressed for a small subset of processes. The Ornstein-Uhlenbeck process is one of them and is illustrated in (10). These models can be calculated with precision using techniques like maximum likelihood. Although they use a two-step estimation method, Atalon (2002) never assume that a generalized Ornstein-Uhlenbeck process permits flawless discretization. The model of average temperature T_t in day t is given as follows as a result of applying the parameterized maximum likelihood method for equation (10)

$$T_t = \alpha[T_{t-1} - T_t^m] + T_t^m + \sigma_t \varepsilon_t, \quad (11)$$

with $\alpha = e^{-k}$ and where, $\varepsilon_t \sim N(0, 1)$, the deterministic portion of the temperature is T_t^m , the volatility is σ_t , and α is the mean reversion speed. When coming up with this deterministic part, we consider the combination of seasonality, trend, and expression of the sine function shift. Hence, it is given by the expression:

$$T_t^m = A + Bt + C \sin(\omega t + \rho) \quad (12)$$

$$\omega = \frac{2\pi}{365}$$

where T_t^m shows the predicted temperature for a day with t in 2021, where the number of days in a year is given by t (1–365), Variable A represents the average daily air temperature for the period from 1.1.1990 to 31.12.2020, variable B describes the impact of an annual global warming trend, and variable C establishes the seasonality of temperatures throughout the year, or how much the winter and summer temperatures deviate from the annual temperature mean. Since the highest and lowest temperatures neither happen at the beginning nor the middle of the year, this phenomenon is known as phase shift, denoted by ρ .

Using trigonometric formula, we have

$$\begin{aligned} T_t^m &= A + Bt + C \sin(\omega t + \rho) \\ &= A + B + C_1 \sin(\omega t) + C_2 \sin(\omega t) \end{aligned}$$

Where

$$c = \sqrt{c_1^2 + c_2^2}$$

$$\rho = \arctan\left(\frac{c_2}{c_1}\right) - \pi$$

Similarly, based on practical findings, we define the cyclic nature of the function δ^2 (Benth & Benth, 2005; Zapranis & Alexandridis, 2014).

$$\delta^2 = c + \sum_{i=1}^I c_i \sin\left[\frac{2\pi i}{365}\right] + \sum_{j=1}^J c_j \cos\left[\frac{2\pi j}{365}\right] \dots\dots\dots (13)$$

Based on past temperature records, variables A, B, and C, along with the shift ρ , will be determined. In equation (12)(Alexandridis A. K. and Zapranis A. D, 2014; Benth, F.E., & Šaltyt-Benth, 2005)

$$\delta^2 = c + \sum_{i=1}^I c_i \sin\left[\frac{2\pi i}{365}\right] + \sum_{j=1}^J c_j \cos\left[\frac{2\pi j}{365}\right] \dots\dots\dots (13)$$

Variables A, B and C and the shift ρ will be calculated based on historical temperature data.

In equation (12), Alaton et al. (2002) concentrated on the seasonality, trend and the sine function's shift expression. Alaton et al. (2002) modelled the long-term temperature model, which may be less applicable in the short term. They made the assumption known as homoscedasticity that variance would be constant over time. However, according to (Niyitegeka & Tewari, 2013), empirical evidence has disputed this supposition. When calm and volatile periods are observed in time series, volatility clustering is known to occur, making the variance at least appear predictable. Since temperature data is a time series, it may display significant auto-correlation. Hence, we propose to extend this (Alaton et al., 2002) by relaxing the assumption concerning ε_t to (12) and allowing ε_t to be autocorrelated. So it will be

$$T_t^m = A + B + C_1 \sin(\omega t) + C_2 \sin(\omega t) + \varepsilon_t, \varepsilon_t | F_{t-1} \sim N(0, \sigma_t^2) \quad (14)$$

Assume ε_t follows the GARCH model (Generalized Autoregressive Conditional Heteroskedastic). According to (Niyitegeka & Tewari, 2013). GARCH forecasts future variability and addresses the issue regarding heteroskedasticity, or the time series' non-linear variability. The conditional variance can depend on prior lags in the GARCH model, which uses the maximum likelihood method. The following is how the conditional variance equation is written.

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \dots\dots\dots (15)$$

This replaces equation (13) where α_0 has been substituted by c . σ_t^2 is the volatility at time t . α_1 has also been substituted by $c_i \sin(\omega t)$, and β has been substituted by

$c_i \cos(wt) \cdot \varepsilon_{t-1}^2$ is the previous period's squared error term. σ_{t-1}^2 , is the previous period's volatility.

Proposition 3.2. (Brody, 2002). If $T_t^m = E[T_t]$, the process

$$dT_t = [\alpha(T_t^m - T_t) + \frac{dT_t^m}{dt}] dt + \sigma_t dW_t \text{ reverts to } T_t^m.$$

Proof. : Let $Z_t = e^{\int_0^t \alpha ds} (T_t^m - T_t)$

Ito's lemma

$$\begin{aligned} dZ_t &= e^{\int_0^t \alpha ds} T_t^m dt + \alpha e^{\alpha t} (T_t^m - T_t) dt - e^{\int_0^t \alpha ds} dT_t \\ &= e^{\int_0^t \alpha ds} [(T_t^m + \alpha(T_t^m - T_t)) dt - (\alpha(T_t^m - T_t) + T_t^m) dt - \sigma_t dW_t] \end{aligned}$$

$$\therefore Z_t = Z_0 - \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

$$\therefore e^{\int_0^t \alpha ds} (T_t^m - T_t) = T_0^m - T_0 - \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

Now $T_0^m - T_0 = c$ gives

$$T_t^m - T_t = -e^{\int_0^t \alpha ds} \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

$$\therefore T_t = T_t^m + e^{\int_0^t \alpha ds} \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

$$\Rightarrow E[T_t] = T_t^m \quad \blacksquare$$

3.5 Pricing weather derivatives in an incomplete market

The weather derivatives market is an example of an incomplete market because the fundamental variable temperature cannot be traded. In order to establish distinct pricing for such contracts, a risk's market value is added. The cost of risk is assumed to be constant to maintain simplicity. Furthermore, the tick price is determined at \$1 per degree day, and the risk-free interest rate, r , is assumed to remain constant. Typically, a risk-neutral valuation approach is used to price financial derivatives. According to the

financial theory, a contingent claim's cost F , which is based on stochastic variable I , can be determined as follows (Dzupire et al., 2019; Xu et al., 2007):

$$F = E_Q(D, W_T(I)) \quad (16)$$

I can be a traded asset, like a stock, or untraded, like a weather index. At expiry time, T , W represents the payoff of the derivative, and D represents a discount factor e^{-rT} with a free rate. The subscript Q indicates that the expectation of the derivative payoff is to be calculated by replacing a real-world probability with risk-neutral probability measurements, and E denotes an expectation conditional on the information now available (Dzupire et al., 2019)

Equation (18) can be written as

$$F = e^{-rT} E_P\left(\frac{dQ}{dP} W_T(I)\right) \quad (17)$$

$\frac{dQ}{dP}$ shows the Radon-Nikodym derivative of Q concerning P .

With the change of neutral measure, the stochastic process of I becomes a martingale. If the stock moves with a geometric Brownian motion, reducing the drift to a risk-free rate can lead to a change in measure.

However, the market is insufficient if indexes cannot be traded since a self-financing portfolio cannot reproduce the derivative. As a result, it is impossible to use no-arbitrage pricing methods for weather derivatives because we cannot create a portfolio free of risk that combines weather index and derivative (Xu et al., 2007). Additionally, the no-arbitrage criterion does not produce a distinctive pace because there are numerous martingale measures. Therefore, only contingent claims secured by bonds may be achieved (Dzupire et al., 2019). Formally, we have a range

$$\left[\inf_Q e^{-rT} E_P\left(\frac{dQ}{dP}, W_T(I)\right), \sup_Q e^{-rT} E_P\left(\frac{dQ}{dP}, W_T(I)\right) \right] \quad (18)$$

The interval in equation (14) is exceedingly significant and, therefore, useless, where Q indicates the set of all equivalent martingale measures (Xu, Wei, Martin Odening, 2007). By engaging in dynamic trading, the investor seeks to maximize the anticipated utility of the final wealth and minimize risks associated with the uncertain reward in an incomplete market. The objective is to find a method that maximizes the expected utility of terminal wealth under the physical measure while minimizing risks as measured by a risk measure.

3.6 Temperature barrier option pricing

A barrier option operates similarly to a standard option until the underlying asset's price, X , crosses a predetermined barrier, B (Primajati, 2020). The feature of the selection is either a knock-in or a knock-out. If an option is knocked in, it has no value unless the asset price passes the threshold. When an option is knocked out, it loses its value once the asset's price crosses the threshold. The barrier must be crossed in the direction indicated by the arrows going up and down. In addition to X and B , the strike price K , interest rate r , time to maturity T , dividend rate q , and volatility σ are the input arguments utilized to determine the value of barrier options.

Barrier options are one of the most frequently traded derivatives on the financial markets, claim (Wang & Wang, 2011); they stand out from standard solutions thanks to their unique qualities. The fact that barrier options are typically less expensive than standard options is one reason for the payoff; the asset price must pass a particular threshold first. The third factor is that barrier options better-fit risk hedging requirements than conventional options.

Farmers may buy a call option if the price is anticipated to be higher. Sun (2017) and Kooten (2015) provide the payoffs for the call contracts from the purchasers' perspective.

$$P(x)_{call} = \begin{cases} 0, & x < k \\ D(x - k), & x \geq k \end{cases} \quad (19)$$

$P(x)$ denotes the option payoff, D is the tick size (the amount of money for each weather index unit), and k is the strike (trigger) value. Equation (19) also is the payment of the barrier down and out call option for the barrier temperature option, where the k is now the barrier level.

For barrier it is

$$P(x)_{call} = \begin{cases} 0, & x < k \text{ and } x < B \\ D(x - k), & x \geq k \text{ and } x > B, 0 \leq t \leq T \end{cases} \quad (20)$$

B is the barrier level, and x is the weather index (GDD).

The GDD option depends on the sum of the GDD across the growth season, where each temperature process follows the Gaussian process, represented by $T_t \sim N(\mu_t, \nu_t)$. We

can proceed once we get the conditional mean and variance of the GDD. The conditional mean and variance of GDD_n for time $t < t_1$ can be calculated as follows:

$$GDD_n \sim (N(\mu_u, \sigma_n^2))$$

Where,

$$\mu_u = E^{\mathbb{Q}}[GDD_n | \mathcal{F}_s] = \sum_{i=1}^n E^{\mathbb{Q}}[T_{ti} | \mathcal{F}_t] - k, \quad (21)$$

$$\sigma_n^2 \equiv \text{var}^{\mathbb{Q}}[GDD_n | \mathcal{F}_t] = \sum_{i=1}^n \text{var}^{\mathbb{Q}}[T_{ti} | \mathcal{F}_t] + 2 \sum \sum_{i < j} \text{cov}^{\mathbb{Q}}[T_{ti}, T_{tj} | \mathcal{F}_t] \quad (22)$$

The anticipated return is the following, assuming a normal distribution for the weather indicator used in a financial instrument:

$$E_p = \int_a^b f(x) p(x) dx, \quad (23)$$

Where $p(x)$ is the payment associated with the financial instrument for the potential outcome, x is the weather index. This x , at some point, will reach the barrier level. The weather index's probability density function (PDF) is denoted by $f(x)$. When the weather index is transformed into a regular normal distribution, let $z = \frac{x-\mu}{\sigma}$, and the expected pay-out function is as follows:

$$E_p = \int_a^b \phi(z) p(z) dz = \frac{1}{\sigma} \int_a^b \phi(z) p(x) dx \quad (24)$$

From equation (24), $\phi(z)$ signifies the PDF of the typical normal distribution, and σ is the standard deviation of the weather index.

Inserting the payoff function for the call contract in the corresponding uncapped call options with closed-form functions are as follows when the expected pay-out function is entered:

$$E_{p,CALL} = \frac{1}{\sigma} \int_k^\infty D(x - k) \phi\left(\frac{x-\mu}{\sigma}\right) dx = D\sigma\phi\left(\frac{k-\mu}{\sigma}\right) + D(\mu - k) \left[1 - \Phi\left(\frac{k-\mu}{\sigma}\right)\right], \quad (25)$$

Multiplying the above by the difference of x and B ($x - B$) gives us the call payoff of the barrier call option to be

$$= (x - B) \left[D\sigma\phi\left(\frac{k-\mu}{\sigma}\right) + D(\mu - k) \left[1 - \Phi\left(\frac{k-\mu}{\sigma}\right)\right] \right] \quad (26)$$

Proposition 3.3: Therefore, the cost of the GDD call option at time $t < t_1$ is as follows (Hung-Hsi Hua, 2008; Patricia, 2021; Peter, 2002; Sun, 2017):

$$\begin{aligned}
c(t) &= e^{-r(t_n-t)} E^Q[\max(H_n - k, 0) | \mathcal{F}_t] \\
&= e^{-r(t_n-t)} \int_{k_1}^{\infty} (x - k) f_{H_n}(x) dx \\
&= e^{-r(t_n-t)} \left\{ (\mu_n - k) \Phi\left(-\frac{k-\mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}} \right\}
\end{aligned} \tag{27}$$

Here t_n is time to maturity. The normal distribution's probability density function is f_{H_n} , and Φ represents the cumulative distribution function for the standard normal distribution.

Proposition 3.4 The price of the down and out barrier call option is

$$= e^{-r(t_n-t)} \left\{ (\mu_n - k) \Phi\left(-\frac{k-\mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}} \right\} (x - B) \tag{28}$$

Proof of Proposition 3.3

We know that $H_n \equiv x \sim (N(\mu_n, \sigma_n^2))$. Let $u = \left(\frac{x-\mu_n}{\sigma_n}\right)$,

then $du = \frac{dx}{\sigma_n}$, $dx = \sigma_n du$, and $x = \mu_n + \sigma_n u$.

$$\int_k^{\infty} (x - k) f_{H_n}(x) dx = \int_k^{\infty} x f_{H_n}(x) dx - k \int_k^{\infty} f_{H_n}(x) dx$$

It is simple to calculate the right-hand side's second term in the manner shown below:

$$\begin{aligned}
k \int_k^{\infty} f_{H_n}(x) dx &= k \left(1 - \int_k^{\infty} f_{H_n}(x) dx \right) \\
&= k \left(1 - \Phi\left(\frac{k - \mu_n}{\sigma_n}\right) \right) \\
&= k \Phi\left(-\frac{k - \mu_n}{\sigma_n}\right)
\end{aligned}$$

The first term $\int_k^{\infty} x f_{H_n}(x) dx$ can be calculated as the follows:

$$\int_k^{\infty} x f_{H_n}(x) dx = \int_k^{\infty} \frac{x}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-\mu_n)^2}{2\sigma_n^2}} dx$$

$$\begin{aligned}
&= \int_{\frac{k-\mu_n}{\sigma_n}}^{\infty} \frac{\mu_n + \sigma_n u}{\sqrt{2\pi}\sigma_n} e^{-\frac{u^2}{2}} du \\
&= \int_{\frac{k-\mu_n}{\sigma_n}}^{\infty} \frac{1}{\sqrt{2\pi}} (\mu_n + \sigma_n u) e^{-\frac{u^2}{2}} du \\
&= \int_{\frac{k-\mu_n}{\sigma_n}}^{\infty} \mu_n \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du + \int_{\frac{k-\mu_n}{\sigma_n}}^{\infty} \sigma_n u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\
&= \mu_n \Phi\left(-\frac{k-\mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} \int_{\frac{k-\mu_n}{\sigma_n}}^{\infty} u e^{-\frac{u^2}{2}} du \\
&= \mu_n \Phi\left(-\frac{k-\mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}}
\end{aligned}$$

The following can be used to generate the second term of the final equation:

$$\int_{\frac{k-\mu_n}{\sigma_n}}^{\infty} u e^{-\frac{u^2}{2}} du = -e^{-\frac{u^2}{2}} \Big|_{\frac{k-\mu_n}{\sigma_n}}^{\infty}$$

Therefore, $\int_k^{\infty} (x - k) f_{H_n}(x) dx = (\mu_n - k) \Phi\left(-\frac{k-\mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}}$ ■

If equation (27), which determines the call price $c(t)$ is used, then

$$\begin{cases} \frac{\partial c_t}{\partial \mu_n} > 0 \\ \frac{\partial c_t}{\partial \sigma_n} > 0 \end{cases} \quad (29)$$

The claim above states that a rise in the mean μ_n will increase the price of temperature calls. Furthermore, the call prices will rise in proportion to the standard deviation σ_n

To summarise this chapter, we have seen that the temperature and pricing model has been modelled. Hence, in the next chapter, we will see whether these models predict the future temperature and price temperature derivative as a way of hedging maize crop yields using the given temperature data.

CHAPTER FOUR

RESULTS AND DISCUSSION

This section analyses and discusses the study's objectives and results. The interpretation of the data was also tied to the literature reviewed. The data has been analysed mainly by R. Hence in section 4.1 we look at descriptive statistics. In section 4.2, we look at trend and seasonality followed by estimation of mean reverting speed in section 4.3. Volatility of the temperature process has been calculated in section 4.4 and lastly, we look at hedging maize crop yield using modelled pricing model.

4.1 Descriptive statistics

This study examines the daily average temperature data recorded in degrees Celsius in Malawi's Kasungu area. The data spans the years 1990 to 2020, and 11315 data series are included. The temperature readings on February 29 of every leap year have been dropped to preserve consistency over time. Kasungu district is located in Malawi's central area.

A significant sample, which risks estimating parameters being changed by dynamics that no longer represent future temperature behaviours, including urban influences, is considered a worse sample to research temperature dynamics than 31 years. On the other hand, if the period is short, crucial dynamics may not be shown, which could lead to a flawed model (Alexandridis & Zapanis, 2006).

From our data set of 31 years, we plot the average temperature graph as shown in Figure 4.1. This graph shows the average temperature of Kasungu district from 1990-2020. It illustrates the seasonality in average daily temperature movements, indicating its similarity to a sine function in particular. The daily average temperature moves repeatedly and regularly through periods of high temperature (summer) and low temperature (winter). From our data set of 31 years, we plot the average temperature graph as shown in Figure 4.1. This graph shows the average temperature of Kasungu district from 1990-2020. It highlights the seasonality of daily average temperature changes, mainly how it resembles a sine function. The daily average temperature fluctuates often and predictably between hot summer and cold winter months.

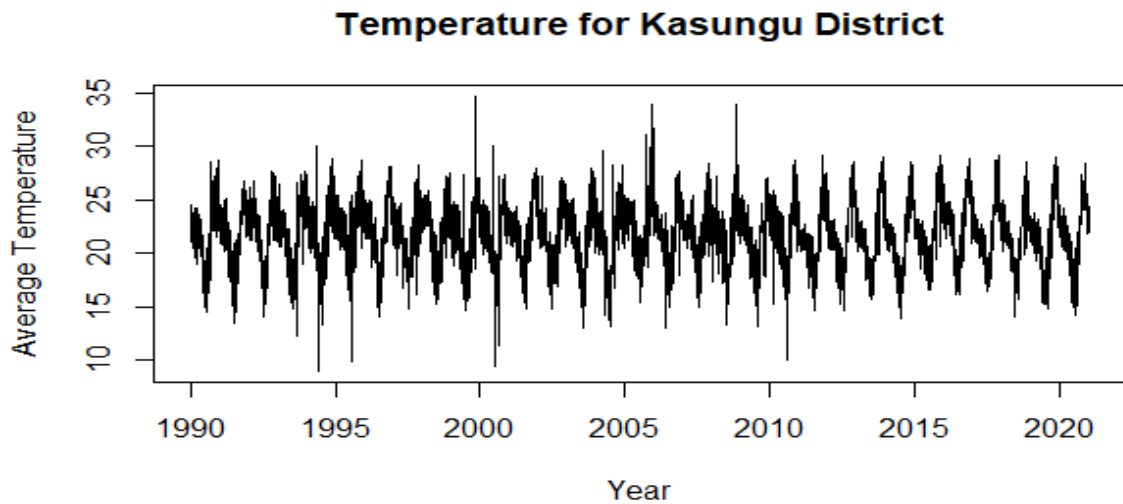


Figure 4.1 shows a stochastic model of average daily temperature in Kasungu from 1990 to 2020.

The use of the Ornstein-Uhlenbeck procedure to mean-revert simulate temperature behaviour is justified in light of seasonal fluctuation and long-term trends in temperature. The daily average temperature variation will gradually return to the mean over time. The Anderson-Darling test was used to determine whether the varying temperatures for the Kasungu district were average. The hypothesis was rejected with an $A = 0.83279$, $P. value = 0.03181$. Although it is not normally distributed, we shall treat the daily variation in temperature as a Brownian motion. This is the case because when the histogram is used to check the temperature difference's normality, it finds that it is roughly average. This agrees with what (Wang et al., 2015) state that when data is

enormous, then it has to be approximately normal and Brownian motion should be considered. Figure 4.2 illustrates that the temperature differential is roughly average.

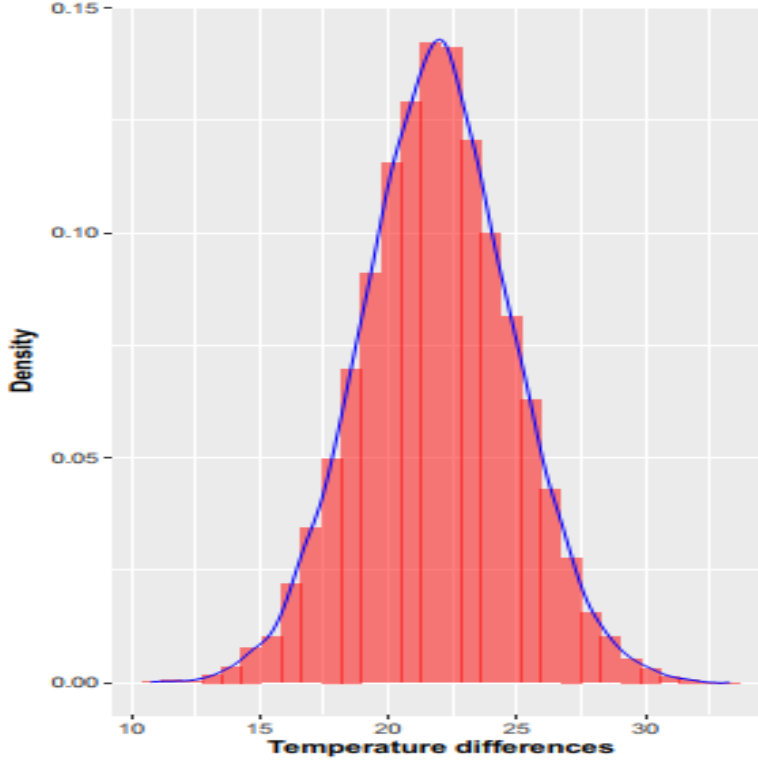


Figure 4.2: Histogram of temperature differences of Kasungu district, Malawi.

The stochastic differential equation characterizes the motions of the process if we indicate the average temperature at the date t by T_t in equation (4) of which is

$$dT_t = dT_t^m + k(T_t^m - T_t)dt + \sigma_t dW_t$$

where W_t is the Brownian Motion, σ_t is the volatility square root time of the fluctuations, T_t^m is the long-term mean, and k is the mean reversion rate. To better understand the temperature dynamics, we computed the descriptive statistics of the data, as provided in Table 1.

Table 4.1: Descriptive statistics of temperature for Kasungu

Maximum	34.7
Minimum	9
Mean	21.88196
Median	22.1
Mode	23.25
Variance	8.504897
Standard Deviation	2.916316
Skewness	-0.1304493
Kurtosis	2.961842
Coefficient of variation	13 %

According to a descriptive evaluation of the information, the modal temperature is 23.25°C, while the average daily temperature is 21.8833°C. The coefficient of variation (CV) is the difference between the standard deviation and the mean. The mean dispersion becomes more pronounced with an increase in the coefficient of variation. Usually, it is expressed as a percentage. The estimated average of 21.88330C (in our example) is shown to be representative of the data by the coefficient of variation. The CV of 13% demonstrates the data's minimal variability. As a result, a dataset's estimated average of the values is more reliable.

The observed distribution exhibits a small amount of leftward asymmetry, with a skewness (asymmetry coefficient) smaller than 0 (-0.1304493). The data's median temperature is 22.1°C, whereas the mean is 21.88195°C, making the data's mode, 23.25°C, more excellent. The mode must be more significant than the median and the arithmetic mean for left-skewed asymmetry, which is proven (Kovac, 2020).

The size of the two tails' combined lengths is measured by kurtosis. It calculates how likely the tails are. The amount is typically contrasted with the normal distribution's

kurtosis, which is 3. Our data are typically distributed because the observed kurtosis is 2.961842, which is roughly 3

4.2 Trend and seasonality

The temperature process has a seasonal trend, as seen in Figure 4:1. Consequently, we create a seasonal mean that also considers patterns as

$$T_t^m = A + Bt + C\sin(\omega t + \rho), \quad \omega = \frac{2\pi}{365}$$

where t is a calendar year's range of days (1–365). T_t^m is the predicted temperature for a given day in 2021. $A + Bt$ reflects the tendency brought on by urban effects and global warming since extreme temperatures do not necessarily occur at the beginning and middle of the year. C , which stands for amplitude, determines when we feel the highest or lowest temperature.

Using trigonometric formulae, we have

$$\begin{aligned} T_t^m &= A + Bt + C\sin(\omega t + \rho) \\ &= A + B + C_1 \sin(\omega t) + C_2 \sin(\omega t) \end{aligned}$$

Where

$$c = \sqrt{c_1^2 + c_2^2}$$

$$\rho = \arctan\left(\frac{c_2}{c_1}\right) - \pi$$

To determine the parameters $A = \{A, B, c_1, c_2\}$ that solve the optimization problem $\min_A ||T_t^m - X(t)||$, we therefore fit T_t^m using least squares methods. $X(t)$ is the data vector. We present estimated values for parameters in Table 4:2, all of which suggest they are significant. When the estimated values are inserted into T_t^m we get

$$T_t^m = 22.04 - 0.0000286t + 2.89929\sin\left(\frac{2\pi}{365}t + 1.60807\right) \quad (30)$$

where A is the coefficient, set to 22.04 °C and is computed as the average daily temperatures between 1.1.1990 and 31.12.2020. Coefficient B, which has a value of -0.0000284t°C, represents the expected rise in the annual average temperature in 2021 compared to the average of the average temperatures for the base period (1990-2020).

Given that it assumes the highest average daily temperature in 2021 will be around 24.93929 °C (22.04 °C + 2.89929 °C) and the lowest will be approximately 19.14071 °C (22.04 °C - 2.89929 °C), the value of Factor C is 2.89929 °C, which is realistic and predicted.

As a result, the average daily temperature, previously computed at 22.04 °C, can be anticipated at the end of March (2.85 months after the year's start), and all average daily values beginning on January 1 should be lower than 22.04 °C with an upward tendency. The sine function shift is 1.60807, or roughly three months of a year, in terms of the sine function, or 2.85 months. In March, the yearly average temperature for the months that are not susceptible to significant temperature swings changes.

The predicted values for 2021 were derived using the O-U procedure and presented in Figure 3:9 by substituting the 365 calendar days for t.

Table 4:2. Estimated parameter values for the seasonal mean

Parameter	Estimated	Std Error	t value	Pr(> t)
A	2.204e+01	3.600e-02	612.248	< 2e-16 ***
B	-2.864e-05	5.512e-06	-5.196	2.07e-07 ***
C1	-8.671e-01	2.546e-02	-34.058	< 2e-16 ***
C2	2.986e+00	2.545e-02	117.331	< 2e-16 ***

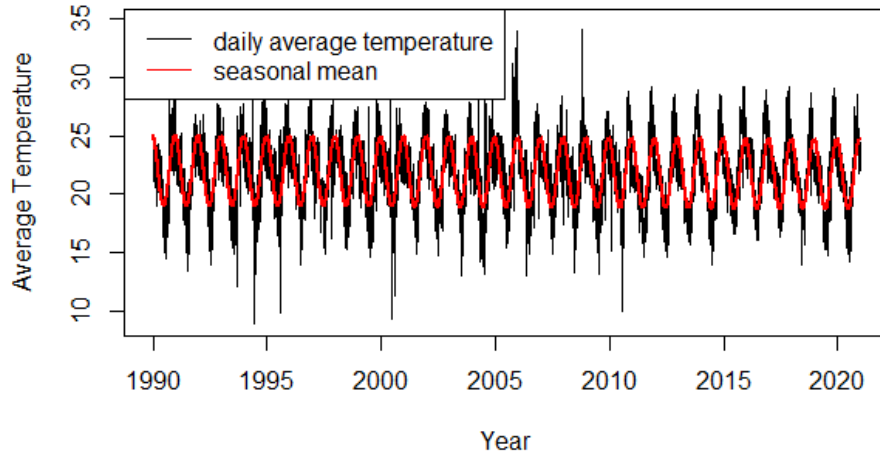


Figure 4:3 shows the daily average temperature and predicted season means.

Figure 4:3 depicts the yearly mean and the daily average temperature. The mean reasonably fits the data.

4.3 Estimation of mean reverting speed

According to Patricia et al. (2021) & Wang et al. (2015), derived the most straightforward Ornstein-Uhlenbeck process, often known as a mean-reverting process, is as follows.

$$dT_t = [\alpha(T_t^m - T_t) + \frac{dT_t^m}{dt}] dt + \sigma_t dW_t \quad (31)$$

Here, T_t^m is the normal level of T_t , to which T_t tends to revert, and α is the reversion speed. Remember that the difference between T_t and T_t^m determines the anticipated change in T_t . If T_t is higher (less) than T_t^m is more likely to decrease (increase) over the following brief period. Therefore, despite satisfying the Markov characteristic, this process lacks independent increments.

If the value of T_t is currently T_0 and T_t is calculated using equation (18), then its anticipated value at any point in the future is given by

$$E[T_t] = T_t^m + (T_0 - T_t^m)e^{-\alpha t} \quad (32)$$

Also, the variance of $(T_t - T_t^m)$ is

$$V[T_t - T_t^m] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) \quad (33)$$

We return to the easy O-U (mean-reverting) approach of equation (32) for the derivation of equations 32 and 33. Set T_t^m to 0 to make the calculation more straightforward, which results in the equation: $dT_t = -\alpha T_t dt + \sigma_t dW_t$ (34)

We said that equations (32) and (33) provide the text's mean and variance of T_t . To support this assertion, we can apply the Kolmogorov forward equation.

Write $M(\theta, t) = E[e^{-\theta T}]$ as the moment-generating function for T_t

$$= \int_{-\infty}^{\infty} \phi(T_0, t_0; T, t) e^{-\theta T} dt \quad (35)$$

Then,

$$\frac{\partial M}{\partial t} = \int_{-\infty}^{\infty} \frac{\partial \phi}{\partial t} e^{-\theta T} dT \quad (36)$$

The Kolmogorov forward for this process is

$$\frac{\partial \phi}{\partial t} = \frac{1}{2} \sigma^2 \frac{\partial^2 \phi}{\partial T^2} - \alpha T \frac{\partial \phi}{\partial T} + \alpha \phi \quad (37)$$

The following equation for $M(\theta, t)$ is obtained by substituting this for $\frac{\partial \phi}{\partial t}$ in equation (17) and integrating by parts:

$$\frac{1}{2} \sigma^2 \theta^2 - \alpha \theta \frac{\partial M}{\partial \theta} = \frac{\partial M}{\partial t} \quad (38)$$

Boundary conditions must be met in order to solve this partial differential:

$$M(0, t) = 1, -M_\theta(0, 0) = T_0 \text{ and}$$

$$V[T_0] = M_{\theta\theta}(0, 0) - T_0^2 = 0$$

Hence, the equation has the following equation

$$M(0, t) = e^{\frac{\sigma^2 \theta^2}{4\alpha}} \left[1 - T_0 \theta e^{-\alpha t} + \left(\frac{1}{2} T_0^2 - \frac{\theta^2}{4\alpha} \right) \theta^2 e^{-2\alpha t} \right] \quad (39)$$

Using the fact the $E[T_t] = -M_\theta(0, t)$ and $E[T_t^2] = M_{\theta\theta}(0, t)$ verifies equation (32) and (33) ■

The expected value of T_t converges to $\frac{\sigma^2}{2\alpha}$, as seen from these formulae. Additionally, as $\alpha \rightarrow \infty$, $V[T_t] \rightarrow 0$, this denotes that T_t can never stray from T_t^m , not even momentarily. In the end, T_t becomes a straightforward Brownian motion and $V[T_t] \rightarrow \sigma_t^2$. The first-order autoregressive process in discrete time is represented by equation (32) in continuous time. Equation (32) is the limiting case for the following AR (1) process as $\Delta t \rightarrow 0$.

$$T_t - T_{t-1} = T_t^m(1 - e^{-2\alpha}) + (e^{-2\alpha} - 1)T_{t-1} + \epsilon_t \quad (40)$$

where ϵ_t has the mean of a normal distribution zero and a standard deviation of σ_t and

$$\sigma_t^2 = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) \quad (41)$$

Therefore, by doing the regression and utilizing discrete temporal data (the only data ever accessible), one might estimate the parameters of equation (34):

$$T_t - T_{t-1} = a + bT_{t-1} + \epsilon_t \quad (42)$$

Afterwards, calculating $T_t^m = -\frac{a}{b}$,

$$\alpha = -\log(1 + b) \quad (43)$$

And

$$\sigma = \sigma_t \sqrt{\frac{\log(1+b)}{(1+b)^2 - 1}} \quad (44)$$

where σ_t is the regression's standard error.

To find this mean reversion in our case, we run the ACF of the decomposed data. The ACF of our data is our b . The same can also be found by running a regression. $b = -0.2137$. Substituting this into (43) gives us the mean reversion of 0.2404. Fitting in mean reversion and σ_t in (44) gives us 0.8943.

Kasungu's average temperature data was decomposed under the additive model to find the mean reversion. **Figure 4:4** shows the decomposed plot of the statistics on Kasungu's average temperature. **Figure 4:5** shows the Auto Correlation Function (ACF) plot.

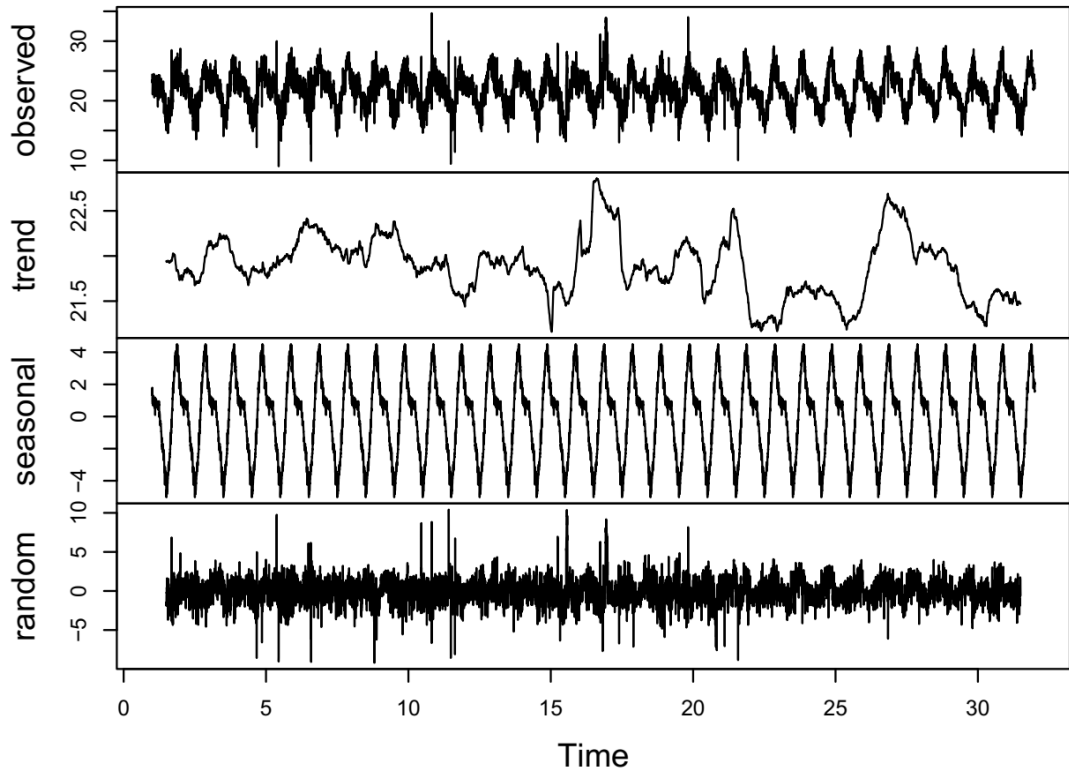


Figure 4:4: Decomposed plot of the Average Temperature of Kasungu

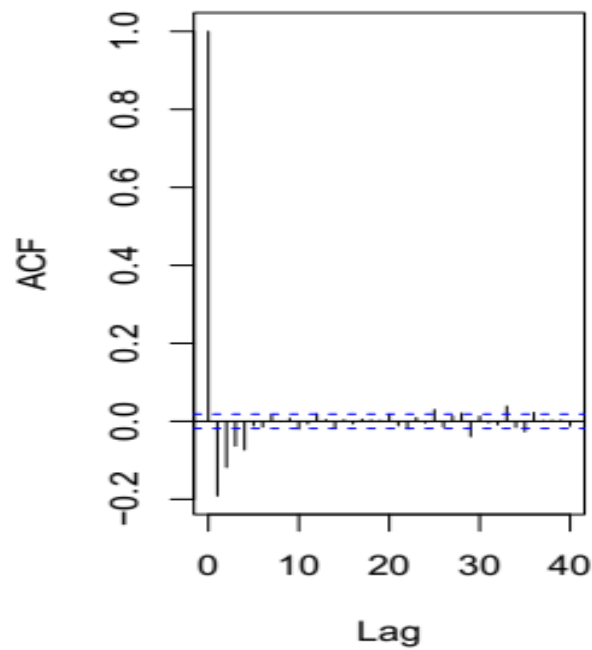


Figure 4:5: shows the Auto Correlation Function (ACF) plot

There are significant autocorrelation values for ACF for several lags, which may call for adopting more advanced autoregressive models. The ACF also reveals the presence of seasonality in both the residuals and time dependence in the variance of residuals.

To extract the σ_t^2 from the residuals. We must first analyse and simulate the random noise process—but testing for stationarity Augmented Dickey-Fuller (ADF) test, the *Dickey – Fuller* = -7.3612 , *P. value* = 0.01. Since the P-value is less than 0.05, we cannot rule out the possibility that the data are stationary and that mean reversion is constant.

Calculating the volatility of the temperature process

The residual graph in Figure 4:6 below shows that the residuals are normally distributed, as many points do not vary from the regular line.

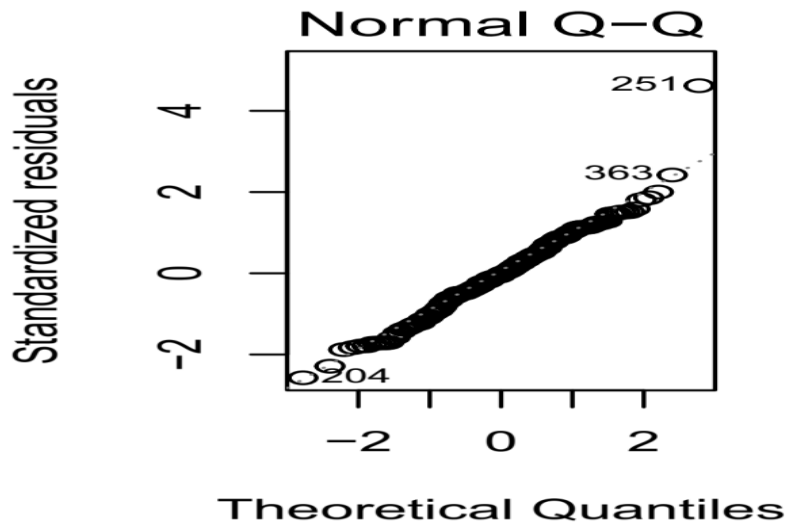


Figure 4:6, a residuals graph.

According to the methods outlined in (Benth, & Šaltyt-Benth, 2011; Dzupire et al., 2019; Wang et al., 2015). the volatility is extracted. After that, the residuals are split into 365 categories, each representing a day of the year for the last 31 years (1990-2020). Then, we calculate the mean of the squared residual in each set of expected daily residuals:

$$\sigma_t^2 = E[(\sigma_t \varepsilon_t)^2] \quad (45)$$

As seen in the figure below, these numbers are used as observable estimates of the daily variance based on years of observations for a specific day.

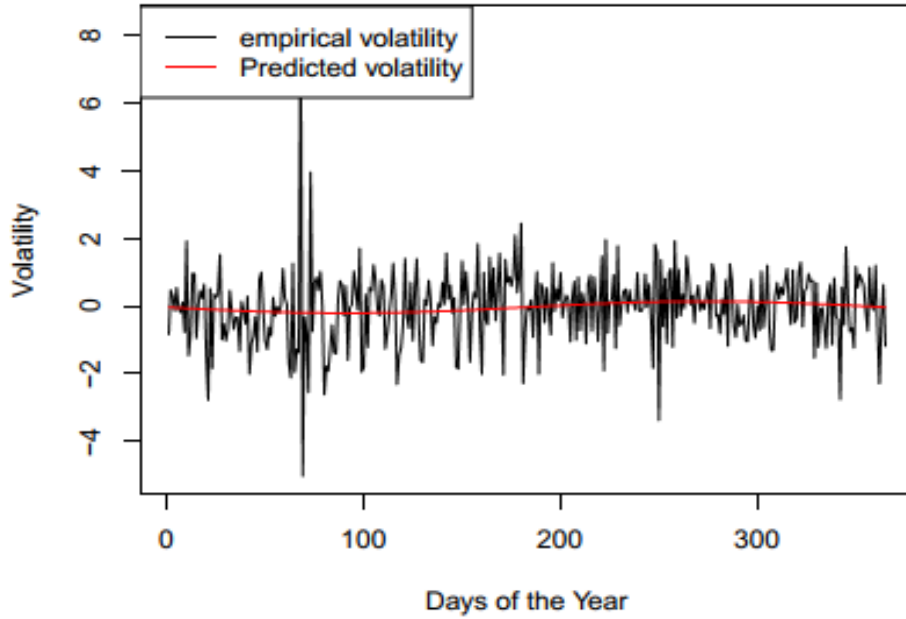


Figure 4:7, Empirical volatility

As can be observed, the wet season has bigger temperature fluctuations than the dry season. In order to derive the volatility model, we fit the data using the volatility model (44)

$$\sigma_t^2 = -0.02603 - 0.17421 \sin \left[\frac{2\pi t}{365} \right] \quad (45)$$

Figure (3:7) compares the estimated model and the empirical volatility and demonstrates how closely the estimated model matches the empirical model.

Comparing the actual average temperature to the outcomes of the stochastic models under ARIMA to determine how well the model performed. The calculated Root Mean Square Error (RMSE), which evaluates how well the temperature model performs, has a value of approximately 1.158505, and the Mean Performance Error (MPE), which compares the actual temperature to temperature models, was roughly -0.3110809. This means that the model is over-forecasted by around 0.311%. Hence, figure 4:8 below is the graph of the forecast temperature for backwards. This has been done on the decomposed temperature without seasonality and trend (random). Figure 4.9 shows the forward forecasted temperature from 2021 to 2024 of Kasungu district using our temperature model. These forecasted temperature figures tell that no matter the

situation is, this model can support the option provider in setting the derivative price for a specific time frame regardless of anything.

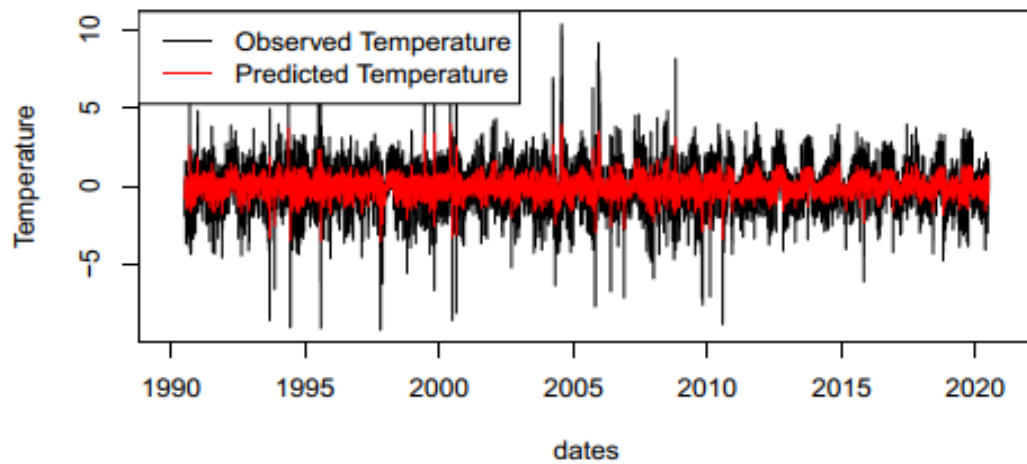


Figure 4:8. Shows the backward forecasted temperature of Kasungu district.

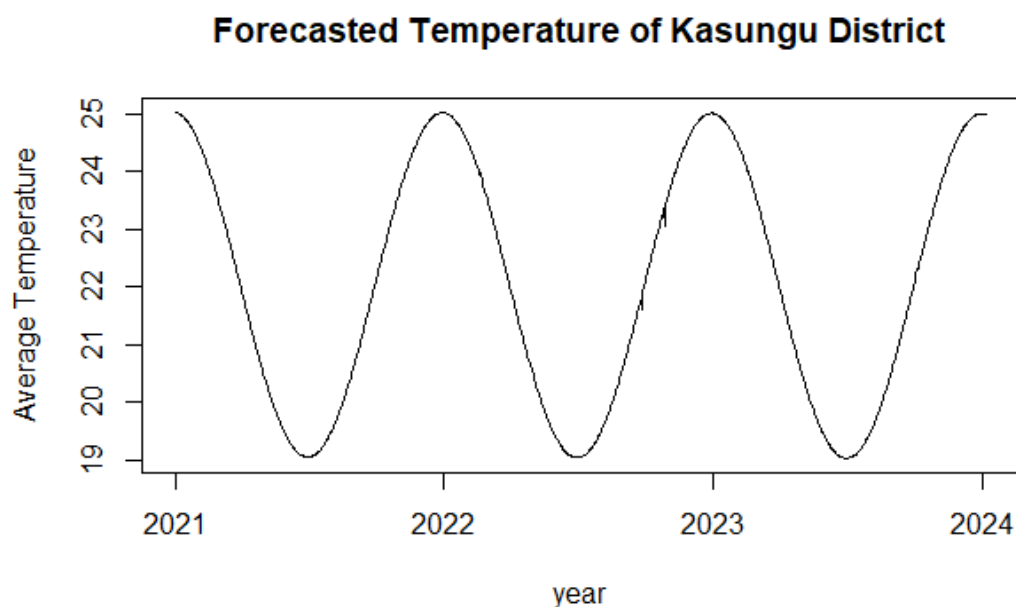


Figure 4:9. Shows the forward forecasted Temperature of the Kasungu district.

The study results prove that the average temperature has increased since 1990. This is evidenced in Figure 4:5, which tells us a story of an increase in average temperature. The results show that there will be a slow but steady increase in temperature. This

scenario does not offer a positive outlook for agriculture production since they are vulnerable to increased temperature. In relation to (Quindala, 2021), this temperature rise will lead to low maize yields as it says 1⁰C will lead to a 10% decrease in crop yields. As the case with this study where it looks into agriculture done during irrigation which is more vulnerable to high temperature since according to (Means, 2018) we will not experience raindrops to control air temperature. Hence, with this rising temperature that cannot be controlled, the idea of hedging crop yield against extreme temperatures during irrigation will help farmers have something at the end of the farming season.

According to our model, if the parameters remain unchanged in the next 30 years or so, the results of this study provide an early view of possible temperature patterns. As stipulated earlier, temperatures of 38⁰c and above and 10⁰c and below can trigger maize sterility. These forecasts can particularly threaten agriculture. This tells us that agriculture done during irrigation is much more vulnerable. This can help the government and non-governmental organisations devise contingency measures. One of these measures is to consider taking temperature derivatives to hedge maize crop yields into agriculture policy into this 2063 agenda.

This is why farmers should use weather derivatives to protect their yields from rising temperatures. This is especially true for Malawi as one of the developing nations, where a sizable section of the populace still depends on agriculture and where government insurance and other support forms still need improvement. Indeed, this study reveals what other research has revealed that, for instance, farmers in central and northwestern China are drawn to weather-indexed insurance (Sun & Kooten, 2015).

4.5 Hedging maize crop yield

Hedging is an advanced risk management strategy involving buying or selling an investment to help reduce the chance of loss of an existing position. When purchasing call options, farmers use temperature as the underlying asset. A call option gives its owner the right to pay a premium and purchase the underlying asset from the provider for the specified period at the specified price (Berislav & Matej, 2020). If the buyer

expects unfavourable weather that could increase the price of derivatives due to monetary gain generated by the set price the supplier has to offer the customer, they may decide to buy call options.

The monetary value of each GDD and the weather derivative is calculated based on the index's value. The derivative buyer wagers that his company will experience an unfavourable change in air temperature throughout the duration of the contract. The provider will pay the buyer a set fee if his prediction comes true.

Weather derivatives must be introduced as a contingency mechanism to assist farmers in hedging possible losses in maize yield due to climate change, particularly temperature fluctuation. The paper offers temperature weather derivatives with temperature as their underlying index. Temperature is one of the most critical elements affecting the variability of maize output.

Now, taking from **Proposition 3.3**: *Therefore, the cost of the GDD call option at time $t < t_1$ is as follows:*

$$\begin{aligned} c(t) &= e^{-r(t_n-t)} E^Q[\max(H_n - k, 0) | \mathcal{F}_t] \\ &= e^{-r(t_n-t)} \int_k^\infty (x - k) f_{H_n}(x) dx \\ &= e^{-r(t_n-t)} \left\{ (\mu_n - k) \Phi\left(-\frac{k - \mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k - \mu_n)^2}{2\sigma_n^2}} \right\} \end{aligned}$$

$C(t)$ is the price or premium that the hedger (the buyer of a call option) needs to pay for the contract, r is a risk-free periodic market interest rate, t is the date the contract is issued (purchased), and t_{hen} is the date the contract is claimed or the expiration date. E^Q is the expected payoff based on the predicted historical mean temperature value. The seller of the option would expect a reward for taking a risk loading, which is often between 20% and 30% of the payoff (Kooten, 2015). In the current application, we set the risk loading at 20% of the expected payoff of the contract.

From the forecasted temperature of 2021, the GDD has been calculated from August 1 to December 31 since this is the time when maize crops are mainly irrigated. It has been observed that the temperatures are too low in early August. This gives the farmer insight into how to hedge their crop yield against low temperatures. The temperatures are also high in November and December; this also gives a farmer insight into how to hedge his or her crop yield against high temperatures.

To price the financial weather derivatives, we assume a tick size $D = \$1$ and risk-free interest rate $r = 0.08$, $\Delta t = \frac{3}{4} \text{ year}$ (August 1 to December 31, 2021), making 153 days. August to October is hot-dry season and Nov/Dec to March/April is hot wet dry season in Malawi. But this study assume the stated date by considering the effects of climate change where most part of Malawi are still dry up to mid-December. Our risk loading $b = 20\%$, $\mu = 12.68^\circ\text{C}$. The forecasted GDDs and standard deviation for 2021 are 1937.43°C and 1.830°C , respectively, and are used to calculate the actual premiums for the contracts.

Table 3: Specification of GDD options for the year 2021

Items	call options
Weather index	GDD(x)
Strike Level	12.68-2*1.83
Barrier level	B = 18 ⁰ c
Tick Size (D)	\$1
Premium	\$3.50
PayoffPayoff	Max(GDD-K, 0)
Issue Date	August 1 2021
Maturity date	December 31 2022

The strike value is $\mu - 2\sigma$ for above 2 standard deviations of the forecasted GDDs; this $2 = \frac{x-\mu}{\sigma}$, then $\phi(2) = 0.9772$ and $\Phi(2) = 0.9772$. The premium is calculated from equation 33. The premium is \$3.50 because we assume the GDD is above the barrier level with one difference. This is where the barrier option is equal to the vanilla option. In this case, the barrier option does not knock out. Hence, it gives the holder the right to buy the underlying asset since it does not exceed a predefined level over the option's lifetime.

The premium has been calculated to be \$3.50 because we assume that the GDD is above the barrier level with one difference. This is where the barrier option is equal to the vanilla option. In this case, the barrier option does not knock out. Hence, it gives the holder the right to buy the underlying asset since it does not exceed a predefined level over the option's lifetime. Now, if the GDD is within the barrier level, the farmer will not have to pay the premium of \$3.50 because the option is invalid. On the same note, the farmer will have to exercise his right by paying the premium of this calculated premium when the GDD exceeds the barrier level. In return, the farmer gets paid off, which is the difference between GDD and barrier. This will be taken daily within the contract.

In summary, this chapter shows that the corrected data fit well into the modelled temperature model. Using the model, we can predict future temperature indices. Taking the future temperature index (GDD) into the pricing model can also hedge maize crop yields.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

This chapter examines the conclusion of this study in section 5.1, the recommendation in section 5.2, and the further studies in section 5.3.

5.1 Conclusion

In Malawi, most people, especially small-holder farmers, rely on the agricultural sector as their primary source of income. A trustworthy and effective insurance product (weather derivative) is required for small-holder farmers and stakeholders because this industry is susceptible to climatic shocks. Due to high basis risks in the product's pricing and design, most farmers would prefer to purchase this product. Because weather indices are not traded assets, the market for weather derivatives is a classic incomplete market, making it impossible to price weather derivatives using traditional no-arbitrage techniques like the Black-Scholes formula.

Given that maize is a staple crop for farmers in Malawi, it was decided to utilize maize yield as a proxy for crop yields. Each weather variable was given a significance score based on the feature's importance before being included in the ensemble learning model. The study demonstrates that the average temperature was the most critical weather variable for developing weather derivatives. This reduced the risk associated with the product design foundation.

5.2 Recommendation

With the agenda of Malawi 2063, this study is good enough to be incorporated into agriculture policy as it presents a way to protect farmers from financial consequences due to extreme temperatures. This will bring an increase in the income of individual farmers, thereby mitigating poverty. We assume there is no correlation between the tradable asset and weather indexes, considering that we are interested in how a farmer can hedge temperature-related weather risks. The pricing model developed can be used in the agriculture industry where a farmer is interested in hedging weather risks due to temperature. It can also price weather derivatives in other weather-related industries affected by temperature. The results of this study can help insurance providers and the government to design products that can help the farmers. Projecting future temperatures and growing degree days is uncertain, so farmers wish to hedge against weather risk. However, markets must provide farmers with attractive, practical hedges representing producers' risks.

With efficient and reliable pricing models, basis risks would decrease. As a result, the farmer would be more willing to pay for the contracts and trading activities in the market for weather derivatives.

5.3 Further studies

Further research should be conducted on barrier option pricing of weather derivatives under the basket. This pricing should consider both the lower and upper barrier of the weather index, which may lead to lower yields.

REFERENCES

- Alexandridis, A.K., & Zapranis, A.D. (2006). Wavelet analysis and weather derivatives pricing. *5th Hellonic Finance and Accounting Association (HFAA), Thessaloniki*, 15–16.
- Alexandridis, A.K. & Zapranis, A.D. (2012). *Weather Derivatives: Modeling and Pricing Weather-related Risk*. Springer Science & Business Media.
- Alexandridis, A.K., & Zapranis, A.D. (2014). Wavelet Neural Networks: With Applications in Financial Engineering, Chaos, and Classification. *John Wiley & Sons*, 61–80.
- Alaton, P. Djehiche, B., & Stillberger, D. (2010). On modelling and pricing weather derivatives. *Applied Mathematical Finance.*, Vol. 9., 1–20. *Appl. Math. Finance*, 9.
- Alexandridis A. K., & Z. A. D. (2008). Modelling the temperature time dependent speed of mean reversion in the context of weather derivative pricing. *Appl. Math. Finance*, 15(4), 355–386.
- Allianz Commercial (2013). The Weather Business - How companies can protect against increasing weather volatility. *Retrieved from* <Http://Www.Agcs.Allianz.Com/Assets/PDFs/Reports/2484%20Allianz%20Weather%20Risk%20LR.Pdf>.
- Alaton, P. (2002). On Modelling and pricing Weather Derivatives. *Appl. Math. Finance*, 9(1), 1–20.
- Benth, F.E., & Šaltyt-Benth, J. (2005). Stochastic modelling of temperature variations with a view towards weather derivatives. *Appl. Math. Finance*, 12(1), 53–85.
- Benth, F.E., & Šaltyt-Benth, J. (2011). Weather Derivatives and Stochastic Modelling of Temperature. *Int. J. Stoch. Anal.*
- Benth, F. E. (2003). On Arbitrage-Free Pricing of Weather Derivatives Based on Fractional Brownian Motion. *Appl. Math. Finance*, 10(4), 303–324.
- Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *J. Political Econ*, 81 (3), 637–654.
- Botoş, H.M. & Ciumaş, C. (2012). The Use of the Black-Scholes Model in The Field

- of Weather Derivatives. *Procedia Econ. Finance*, 3, 611–616.
- Brody, D., Syroka, J., & Zervos, M. (2002). Dynamical Pricing of Weather Derivatives. *Quantitative Finance*, 2, 189–198. <https://doi.org/10.1088/1469-7688/2/3/302>
- Chen, P., Niu, A., Liu, D., Jiang, W., & Ma, B. (2018). Time Series Forecasting of Temperatures using SARIMA: An Example from Nanjing. *IOP Conference Series: Materials Science and Engineering*, 394, 052024. <https://doi.org/10.1088/1757-899X/394/5/052024>
- Considine, G. (2000). *Introduction to weather derivatives*. Retrieved from https://www.cmegroup.com/Trading/Weather/Files/WEA_intro_to_weather_der.Pdf.
- Dornier, F. & Queruel, M. (2000). Caution to the Wind. In: *Weather Risk Special Report. Energy & Power Risk Management.*, 30–32.
- Dzupire N., Ngale, P., & Odongo, L., O. (2018). Levy process based on Ornstein-Uhlenbeck temperature model with time varying, speed of mean reversion. *Advances and Applications in Statistics*. <https://doi.org/DOI:10.17654/as053030199>
- Dzupire N., Ngale, P., & Odongo, L., O. (2019). Pricing Basket Weather Derivatives on Rainfall and Temperature Process. *International Journal of Financial Studies.*, 7, 35. <https://doi.org/Doi:10.3390/ijfs7030035>.
- Geng, X., Wang, F., Ren, W., & Hao, Z. (2019). Climate Change Impacts on Winter Wheat Yield in Northern China. *Advances in Meteorology*, 2019, 2767018. <https://doi.org/10.1155/2019/2767018>
- Hatfield, J. L., & Prueger, J. H. (2011). .Agroecology: Implications for Plant Response to Climate Change. In: Yadav, S. S., et al, (Eds.),. *Wiley-Blackwell, West Sussex, UK*, 27–43.
- Hatfield, J. L., & Prueger, J. H. (2015). *Temperature extremes; Effect on plant growth and development*. Elsevier.
- Hess, U., & Richter, K. S. A. (2002). Weather risk management for agriculture and agribusiness in developing countries. *R.S. Dischel (Ed.), Climate Risk and the Weather Market, Risk Books, London*.

- Hung-Hsi Huang, Yung-Ming Shiu, P.-S. L. (2008). HDD and CDD Option pricing with Market Price of Weather Risk for Taiwan. *Wiley Periodical, Inc. Jrl Fut Mark*, 28, 790–814.
- Kemuel Quindala, Diane Carmeliza Cuaresma, and J. M. (2021). Modelling the Historical Temperature in the Province of Laguna Using the Ornstein-uhlenbeck process. *New York Business Global*, 14 (1), 327–339.
- van Kooten, B. S. (2015). Financial weather derivatives of corn production in northern China; A comparison of pricing methods. *Journal of Empirical Finance*, 32, 201–209.
<https://doi.org/https://dx.doi.org/10.1016/j.jemption.2015.03.014>
- Kovac, B. Z. M. (2020). Ornstein-Uhlenbeck process and GARCH model for temperature forecasting in weather derivatives valuation. *CREBSS*, 6, 27-42.
<https://doi.org/DIO:10.1515/crebss>.
- Islam, M. & Chakraborti J. (2015). *Futures and Forward Contracts as a route of Hedging the risk. Risk Governance & Control*. Financial Market & Institutions.
- Means, T. (2018). Understanding the Various temperatures of raindrops. How air conditions affect droplets' warmth or coolness. Retrieved from <Http://Www.Thought.Com/What-Dertemines-Rain-Temperature-3443616>.
- Ministry of Agriculture (2016). *National Agriculture Policy*.
- Nelson, G., Rosegrant, M., Palazzo, A., Gray, I., Ingersoll, C., Robertson, R., Tokgoz, S., Zhu, T., Sulser, T., Ringler, C., Msangi, S., & You, L. (2010). *Food Security, Farming, and Climate Change to 2050: Scenarios, Results, Policy Options*.
<https://doi.org/10.2499/9780896291867>
- Nicholson, L. (2018). Derivative vs Insurance. Retrieved from <Https://Medium.Com/@larrynicholson/Derivative-vs-Insurance-32399640c8ce>.
- Nury, A., & Koch, M. A. (2013). Time Series Analysis and Forecasting of Temperatures in the Sylhet Division of Bangladesh. *Environmental Science*. 43(3–4), 471–488
- Oksendal, B. (2014). *Stochastic Differential Equations*. Springer.

- Niyitegeka, O. & Tewari, D.D (2013). Volatility clustering at the Johannesburg Stock Exchange: Investigation and Analysis. *Mediterranean Journal of Social Sciences*. <https://doi.org/Doi:10.5901/mjss.2013.v4n14p621>
- Ottman MJ, Kimball BA, & White JW, W. G. (2012). Wheat growth response to increased temperature from varied planting dates and supplemental infrared heating. *Agron J*, 104, 7–16.
- Patricia P., & C. D. (2021). Modelling the Daily Average Temperature and Designing a Growing Degree Day (GDD)European put option for rice in Laguna. *Journal of Nature Studies*, 19(2), 52–61.
- Peter Alaton, B. D. and D. S. (2002). On modelling and pricing weather derivatives. *Applied Mathematical Finance.*, Vol. 9., 1–20.
- Phiri, U. M. (2016). Politics of Conservation Agriculture: Kasungu Rural District, Malawi (Masters thesis). *International Institute of Social Studies*.
- Porter JR, G. M. (1999). Temperatures and the growth and development of wheat. *Eur J Agron*, 10, 23–36.
- Prabakaran, S., Carcia I., & Mora, J. (2020). A temperature stochastic model for option pricing and its impacts on the electricity market. *Elsevier, Economic A*, 58–77.
- Prabakaran, S. (2018). Modelling and Pricing of Energy Derivative Market. *Int. J. Eng. Technol.*, 7 (4.10)., 1005–1017.
- Primajati G., R. M. ;& A. H. (2020). Barrier Options Models for Agricultural Commodity Price Protection. *Sinta*, 4, 71–78.
- Richards, T. J. (2004). Pricing weather derivatives. *Am. J. Agric. Econ.*, 86 (4), 1005–1017.
- Rosenzweig C. (2014). We are assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc Natl Acad Sci USA*, 111, 3268–3273.
- Samuel A.G.. (2019). Hedging Crop Yields Against Weather Uncertainties Weather Derivatives Perspective. *Mathematical and Computational Applications*.
- Sun, W., & Lou, S. (2013). Design of Agriculture insurance policy for tea tree

- freezing damage in Zhejiang Province, China. *Theoretical and Applied Climatology*, 111, 713–728.
- Sun, B. (2017). Financial weather Derivatives for Corn Production in Northeastern China: Modelling the underlying Weather Index. *Working Papers 257083*, University of Victoria, Resource Economics and Policy.
- Sun, B. (2017). Financial weather derivatives of corn production in northern China; Modelling the underlying weather index. *Resource Economics & Policy Analysis Research Group, Department of Economics, University of Victoria*.
- Svec, J. & Stevenson, M. (2007). Modelling and Forecasting Temperature Based Weather Derivatives. *Glob. Finance J.*, 18 (2), 185–204.
- Taştan, B & Hayfavi, A. (2017). Modelling Temperature and Pricing Weather Derivatives Based On Temperature. *Adv. Meteorol.*, 2017, 1–14.
- Turvey, C. D. (2001). Weather Derivatives for Specific Events Risks in Agriculture. *Review of Agricultural Economics*, 23(2), 333–351.
- Wang, B. & Wang, L. (2011). *Pricing Barrier Options Using Monte Carlo Methods*.
- Wang, Z. (2015). Modelling and Forecasting Average Temperature for Weather Derivative Pricing. *Adv. Meteorol.*
- Weron, R. (2014). Electricity Price Forecasting: A Review of the State-Of-The-Art with A Look into The Future. *Int. J. Forecast*, (30), 1030–1081.
- Wheeler T, von B. J. (2013). Climate change impacts on global food security. *Science*, 341, 508–513.
- Xu, Wei, M.O & MuBhoff, O (2007). Inference pricing of weather insurance. Paper presented at 101st Seminar of the European Association of Agricultural Economists (EAAE). *Management of Climate Risks in Agriculture"*, Berlin, Germany, July 5-6.
- Zhao C. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Alizona State University*.
- Zong, L. & Ender, M. (2016). Spatially-aggregated temperature derivatives: Agricultural Risk Management in China. *International Journal of Financial Studies.*, 4(3), 99–117.

MANUSCRIPT
HEDGING CROP YIELDS USING TEMPERATURE DERIVATIVES

Prince Blackson Dennis Chirwa* & Nelson Christopher Dzupire

Mathematical Science Department, University of Malawi.

*prince_chirwa@yahoo.com

ndzupire@unima.ac.mw

ABSTRACT

Agriculture production yield varies with weather changes. This causes farmers to incur losses. For instance, extreme temperature leads to low maize yield. This study describes incomplete temperature weather derivatives in agriculture markets and applies risk management hedging techniques. It focuses on hedging crop yield against extreme temperatures during irrigation farming, which is done without a greenhouse. This study primarily aims to hedge maize crop yields using temperature derivatives. This is achieved by (i) developing a daily average temperature stochastic model. (ii) Deriving statistical properties of the model based on the historical data of 31 years of our sample space (1990 – 2020 Kasungu District Temperature data). (iii) Pricing temperature derivatives to hedge maize crop yield. To achieve this, a stochastically Ornstein-Uhlenbeck process with the time-varying speed of reversion, seasonal mean, and local volatility that depends on the local average temperature was proposed. Based on the average temperature model, down and output, option pricing models for average temperature and growing degree day are presented. The study's findings suggest that the temperature will rise gradually but steadily. This scenario does not offer a positive outlook for agriculture production since a temperature rise can damage it. The premium for weather derivative options has been calculated as \$3.50 per GDD index contract. Farmers and agricultural stakeholders can hedge their crops against extreme temperature-related weather risks with these models. In line with Malawi's 2063 Millennium Development Goals (MDGs), this study acts as an eye opener for the government to put a policy on whether derivatives should be practised in our country hence, increase cash holding by improving the situation of the farmer and country.

Key words: Incomplete market, Weather derivatives, Growing Degree Day, temperature barrier option pricing.

1. INTRODUCTION

Agriculture is the backbone of the Malawian economy (Ministry of Agriculture, 2016). It is an important sector that generates income for the majority of the population, helps Malawi earn money through exports, and supplies the majority of manufacturing industries. A large percentage of smallholder farmers in Malawi primarily get their income from the agriculture and agribusiness sectors, which make for an important portion of the country's key economic activity. Most of these farmers grow maize, soya beans, tobacco, rice, and ground nuts. Maize is the main crop and it is the staple food in Malawi. Smallholder farmers grow maize for food and they sell the excess. Low yields of maize are a yardstick of hunger in the country. According to reports, it contributes to more than 25% of Africa's GDP and roughly 70% of the labour force. (Samuel Asante Gyamerah, 2019). According to Ministry of Agriculture, (2016), It is reported that agriculture sector support about 85% of the population in terms of employment. It accounts for more than one quarter of the Malawi gross domestic product and account for 90% of the foreign exchange. As a result, the development of the economies in Africa, of which Malawi is a member state, has agriculture as one of its most significant and largest fields.

Agriculture production in Malawi is heavily dependent by weather elements like rainfall, temperature, wind etc. any variations in these elements greatly affects harvest. For instance, the production of maize is greatly influenced by the weather. The stages of plant growth affect how each crop species responds to temperature variations. The limits of visible growth are determined by a specified range of maximum and minimum temperatures for each species. Extreme temperature makes it difficult for maize to develop. Extremely high temperatures during the reproductive stage will have an impact on the viability of the pollen, the process of fertilization and development of grains or fruits. (Hatfield, J. L., & Prueger, 2011, 2015). The yield potential of initial grains or fruits will be reduced by repeated exposure to extremely hot or cold conditions during pollination. However, it's possible that the most detrimental effects of acute exposure to strong events will occur during the reproductive stages. Because this measure is the one that both producers and consumers are most concerned about, the

effects of climate change are most noticeable in maize crop productivity. Maize grows well under temperatures of 21⁰c to 27⁰c. According to Hatfield, J. L., & Prueger, (2015) the base temperature below which crop growth cease is 10⁰c and above 38⁰ C.

This underlines the significance of contingency planning, which includes setting up backup funds, strengthening currency reserves to protect against external shocks, and validating the usage of weather insurance (Dzupire et al., 2019). Similar to how insurance against poor crop yields can be used at the family level, weather derivatives offered to small farmers may do the same.

According to Islam and Chakraborti, (2015), derivatives are described as “financial instruments that are linked to a specific financial instrument or indicator or commodity and through which specific risks can be traded in financial markets in their own right. The value of a financial derivative derived from the price of an underlying item, such as an asset or index. Unlike debt securities, no principal is advanced to be repaid and no investment income accrues.” The underlying asset may take on a variety of shapes, including: goods such as grain, coffee, and orange juice; precious metals such as gold and silver; currency exchange rates; and bonds of many kinds, such as medium- to long-term transferable debt securities issued by governments, businesses, etc. Finally, securities of corporations marketed on reputable transactions of stocks and Stock Index, including shares and share warrants.

Financial derivatives are essential for controlling the financial risk that corporations face. All throughout the world, they have been extremely successful, widely used, and valuable advances in the capital markets. Financial derivative markets have recently been discovered to be operating actively in both developed and developing nations. The principal application is hedging, often known as a function of price insurance, risk shifting, or risk transference. They give traders a means through which to manage their risks or shield themselves from unfavourable changes in the value of the underlying assets they work in. For instance, a farmer may take the market's risk and sell a futures contract to offset the risk.

Due to the crops' frequent exposure to unfavourable weather, crop production is a difficult endeavour which temperature is amongst of it. Crop production rate is significantly influenced by weather and climate conditions. However, restrictions have

no place when agricultural production plays a significant role in ensuring global food security. As a result, farming in a controlled environment is now practiced as a result of the quest for alternatives (Ines, 2017). One of the fundamental types of farming in a controlled environment is greenhouse farming. Effective management is made possible by greenhouse farming, which also lowers hazards brought on by adverse weather conditions. (Ines, 2017). It manages heating to suit the crops. Since greenhouse are costly, cannot be managed by most farmers.

Although the use of financial weather derivatives in agriculture has not been as widespread as it is in the energy industry, studies have examined the use of historical data to create heat or rainfall index-based weather derivatives (Sun, W., & Lou, 2013). For agriculture production the relationship is not always as straight forward since differences in products, crop growth phases and soil textures just to mention a few have different responses to the temperatures. Baojing & Kooten,(2015) concentrated on pricing using financial derivatives to protect the output of corn crops by contrasting several approaches to calculate the price of weather derivatives using Growing Degree Day (GDD). Patricia P. et al., (2021) modelled the daily mean temperature and created the Growing Degree Day (GDD) European put option for rice in Laguna.

These studies did not put into consideration that such agriculture can also be done through irrigation where temperatures are not controlled, as compared to agriculture done in the rainy season where the rain drops control the air temperature (Means, 2018). Hence this study focus hedging crop yield against extreme temperatures during irrigation farming which is being done without green house as it is prone to extreme temperature that affects their crop yields.

Four sections make up the structure of this study. The temperature-based weather derivatives are covered in section 2. Hedging maize crop production is covered in Section 4 after Section 3's talks of simulation and parameter estimation. The research was finally concluded in Section 5.

2. TEMPERATURE-BASED WEATHER DERIVATIVES

2.1 Daily average temperature data

The data in this study is from Kasungu district. It has been chosen because is the top-list district that grows maize both during rainy season and dry season under irrigation which feeds a good per cent of Malawians (Phiri, 2016). This means that low maize yield in this district will lead to hunger and hence affect the economy of the country.

From 1 January 1990 through 31 December 2020, historical data on daily minimum and maximum temperatures were gathered from the Department of Climate Change and Meteorological Service's headquarters. This sample size of 31 years' period will ensure that we are capturing the overall long term trend to have a real picture to forecast future temperature. The missing data has been treated by either finding the average temperature adjacent to the missing data or last observation carried forward and next observation carried backwards. All the leap years have been dropped.

2.2 Temperature indices

The temperature information collected from a given station in a given area is the foundation for temperature indices. If T_i^{max} and T_i^{min} are the highest and lowest temperatures recorded at a meteorological station on day i , respectively, the Daily Average Temperature (DAT) on that day is calculated as

$$T_i = \frac{(T_i^{max} + T_i^{min})}{2} \quad (1)$$

a base temperature, as well as the daily average temperature (DAT) are two variables that determine how many degrees a day there are. The amount of heat that must build up each day for a crop to grow, sprout new leaves, reach the reproductive stage, and eventually mature is known as the growing degree day (Patricia P. et al, 2021). We consider maize crop to be produced under irrigation for the period of August to December. Maize requires at least 10°C . GDD is defined as:

$$GDD_i = \max\{T_i - 10, 0\} \quad (2)$$

a full 153-day period's GDD is defined as,

$$GDD = \sum_{i=1}^{153} GDD_i \quad (3)$$

2.3 Daily average temperature stochastic dynamics

Putting together seasonality, long-term trends, unpredictability, and mean reversion, temperature has been modelled by a stochastic process

$$dT_t = dT_t^m + k(T_t^m - T_t)dt + \sigma_t dW_t \quad (4)$$

Proposition 2.2.1: If the Daily Average Temperature (DAT) follows an Ornstein-Uhlenbeck process with a mean reversion that is mean-reverting and whose speed varies over time as well as a seasonal mean and variance:

$$dT_t = dT_t^m + k(T_t^m - T_t)dt + \sigma_t dW_t$$

The Ito formula yields the following implicit solution:

$$T_t = T_t^m + e^{\int_0^t k(u)du} (T_s^m - T_s) + e^{\int_0^t k(u)du} + \int_s^t \sigma_s e^{-\int_s^t k(u)du} dW_t$$

which is the same is:

$$T_t = [T_s^m - T_s]e^{-k(t-s)} + T_t^m + \int_s^t e^{-k(t-r)} \sigma_r dW_r \quad (5)$$

In equation (5) W_r is the Brownian motion, σ_t is the deterministic function of time t , the random variable $\int_s^t e^{-k(t-r)} \sigma_r dW_r$ is normally distributed with mean zero and a variance $\int_s^t e^{-2k(t-r)} \sigma_r^2 dW_r$.

The Brownian motion's characteristic of independent increments forms the foundation of the proof. The filtration $\mathcal{F}_s$ leads us to the conclusion that T_t is normally distributed, with mean and variance given by:

$$E^p[T_t | \mathcal{F}_s] = [T_s - T_t^m]e^{-k(t-s)} + T_t^m \quad (6)$$

$$V_t^2 = var[T_t | \mathcal{F}_s] = \int_s^t e^{-2k(t-r)} \sigma_r^2 d_r. \quad (7)$$

We now describe explicitly the expression of the equation, (3), when $s = t - 1$. I get

$$T_t = [T_{t-1} - T_{t-1}^m]e^{-k(t-s)} + T_t^m + \int_{t-1}^t e^{-k(t-r)} \sigma_r dW_r. \quad (8)$$

The variable $\int_{t-1}^t e^{-k(t-r)} \sigma_r dW_r$ is a random variable having a Gaussian distribution, mean zero, and variance:

$$V_t^2 = \text{var}[T_t|\mathcal{F}_{t-1}] = \int_{t-1}^t e^{-2k(t-r)} \sigma_r^2 dr. \quad (9)$$

Therefore, we can write (8) as

$$T_t = [T_{t-1} - T_{t-1}^m]e^{-k} + T_t^m + \sigma_t \varepsilon_t. \quad (10)$$

This means that (6) has a simple discrete time representation with an autoregressive structure of order 1 (AR (1)). This result has very important implications for the estimate of the unknown. This indicates that (6) has a straightforward discrete time representation with an order 1 autoregressive structure (AR (1)). The estimation of the unknown parameter in the equation above is significantly impacted by this outcome. Since the variable in the case of T_t is observed at discrete points in time rather than constantly, estimate in continuous time is generally quite challenging. The likelihood function can only be analytically expressed for a small subset of processes. The Ornstein-Uhlenbeck process is one of them and is illustrated in (10). These models can be calculated with precision using techniques like maximum likelihood. Although they use a two-step estimation method, Atalou, et al., (2002) never assume that a generalized Ornstein-Uhlenbeck process permits flawless discretization. The model of average temperature T_t in day t is given as follows as a result of applying the parameterized maximum likelihood method for equation (10)

$$T_t = \alpha[T_{t-1} - T_{t-1}^m] + T_t^m + \sigma_t \varepsilon_t, \quad (11)$$

with $\alpha = e^{-k}$ and where, $\varepsilon_t \sim N(0, 1)$, the deterministic portion of the temperature is T_t^m , the volatility is σ_t , and α is the mean reversion speed. When coming up with this deterministic part we take into account the combinational of the seasonality, trend and the expression for the sine function shift. Hence it is given by the expression:

$$T_t^m = A + Bt + C \sin(\omega t + \rho) \quad (12)$$

$$\omega = \frac{2\pi}{365}$$

where T_t^m shows the predicted temperature for a day with t in 2021, where the number of days in a year is given by t (1–365). Variable A represents the average daily air temperature for the period from 1.1.1990 to 31.12.2020, variable B describes the impact of an annual global warming trend, and variable C establishes the seasonality of

temperatures throughout the year, or how much the winter and summer temperatures deviate from the annual temperature mean. Since the highest and lowest temperatures neither happen at the beginning nor the middle of the year. This phenomenon is known as phase shift denoted by ρ .

Using trigonometric formula, we have

$$\begin{aligned} T_t^m &= A + Bt + C\sin(\omega t + \rho) \\ &= A + B + C_1 \sin(\omega t) + C_2 \sin(\omega t) \end{aligned}$$

Where

$$\begin{aligned} c &= \sqrt{c_1^2 + c_2^2} \\ \rho &= \arctan\left(\frac{c_2}{c_1}\right) - \pi \end{aligned}$$

Similarly, based on practical findings, we define the cyclic nature of the function δ^2 (Benth and Benth (2005), Zapranis and Alexandridis (2014)).

$$\delta^2 = c + \sum_{i=1}^I c_i \sin\left[\frac{2\pi i}{365}\right] + \sum_{j=1}^J c_j \cos\left[\frac{2\pi j}{365}\right] \dots\dots\dots (13)$$

Based on past temperature records, variables A, B, and C will be determined, along with the shift ρ . In equation (12) (Atalon, et al., 2002) concentrated on the seasonality, trend and the sine function's shift expression. Alaton et al (2002), modelled temperature model which is of a long-term nature and may not be equally applicable in short term. They made the assumption known as homoscedasticity that variance would be constant over time. However, according to (Olivier Niyitegeka and D.D. Tewari, 2013), empirical evidence has disputed this supposition. When calm and volatile periods are observed in time series, volatility clustering is known to occur, making the variance at least appear predictable. Since temperature data is time series data may display significant auto-correlation. Hence, we propose to extend this Alaton et al (2002) by relaxing the assumption concerning ε_t to (12) and allowing ε_t to be autocorrelated. So it will be

$$T_t^m = A + B + C_1 \sin(\omega t) + C_2 \cos(\omega t) + \varepsilon_t, \varepsilon_t | F_{t-1} \sim N(0, \sigma_t^2) \quad (14)$$

Assume ε_t follows the GARCH model (Generalized Autoregressive Conditional Heteroskedastic). According to Olivier Niyitegeka and D.D. Tewari (2013), GARCH forecasts future variability and addresses the issue regarding heteroskedasticity, or the time series' non-linear variability. The conditional variance can depend on prior own lags in the GARCH model, which uses the maximum likelihood method. The following is how the conditional variance equation is written.

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^p \beta \sigma_{t-j}^2 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad \dots\dots\dots (15)$$

This replaces equation (13) where α_0 has been substituted by c . σ_t^2 is the volatility at time t . α_1 has also been substituted by $c_i \sin(\omega t)$ and β has also been substituted by $c_i \cos(\omega t)$. ε_{t-1}^2 is the previous period's squared error term. σ_{t-1}^2 , is the previous period's volatility.

Proposition 3.2. If $T_t^m = E[T_t]$, the process

$$dT_t = [\alpha(T_t^m - T_t) + \frac{dT_t^m}{dt}] dt + \sigma_t dW_t \text{ revert to } T_t^m. \text{ For the proof refer to appendix.}$$

2.4 Pricing weather derivatives in an incomplete market

The weather derivatives market is an example of an incomplete market, due to the fact that the fundamental variable temperature cannot be traded. In order to establish distinct pricing for such contracts, a risk's market value is added. The cost of risk is assumed to be constant for to maintain simplicity. Furthermore, the tick price is determined at \$1 per degree day and a It is assumed that the risk-free interest rate, r , remains constant. Typically, a risk-neutral valuation approach is used to price financial derivatives. According to the financial theory, a contingent claim's cost F , which is based on stochastic variables I , can be determined as follows (Xu, et al., 2007):

$$F = E_Q(D, W_T(I)) \quad (16)$$

I can be an asset that is traded, like a stock, or untraded, like a weather index. At expiry time T , W represents the payoff of the derivative, and D represents a discount factor

e^{-rT} with free rate. The subscript Q indicates that the expectation of the derivative pay-off is to be calculated replacing a real-world probability with risk-neutral probabilities measurements, and E denotes an expectation, conditional on the information now available (Dzupire, et al., 2019)

Equation (18) can be written as

$$F = e^{-rT} E_P\left(\frac{dQ}{dP} W_T(I)\right) \quad (17)$$

$\frac{dQ}{dP}$ shows the Radon-Nikodym derivative of Q with respect to P.

With the change of neutral measure, the stochastic process of I becomes a martingale. If the stock moves with a geometric Brownian motion, reducing the drift to a risk-free rate can lead to a change in measure.

But if indexes cannot be traded, the market is insufficient since a self-financing portfolio cannot reproduce the derivative. As a result, it is impossible to use no-arbitrage pricing methods for weather derivatives because we are unable to create a portfolio free of risk that combines weather index and derivative (Xu et al., 2007). Additionally, the no-arbitrage criterion does not produce a distinctive pace because there are numerous martingale measures, therefore only contingent claims secured by bonds may be achieved (Dzupire et al., 2019). Formally we have the range

$$\left[\inf_Q e^{-rT} E_P\left(\frac{dQ}{dP}, W_T(I)\right), \sup_Q e^{-rT} E_P\left(\frac{dQ}{dP}, W_T(I)\right) \right] \quad (18)$$

The interval in equation (14) is exceedingly big and therefore useless, where Q indicates the set of all equivalent martingale measures (Xu et al., 2007). The investor seeks to maximize the anticipated utility of the final wealth and minimize risks associated with the uncertain reward in an incomplete market by engaging in dynamic trading. Finding a method that maximizes the expected utility of terminal wealth under the physical measure while minimizing risks as measured by a risk measure is the objective.

2.5 Temperature barrier option pricing

A barrier option operates similarly to a standard option up until the time when the price of the underlying asset, X , crosses a predetermined barrier, B (Primajati G., 2020). The

feature of the selection is either a knock-in or knock-out. If an option is knocked in, it has no value unless the asset price passes the threshold. When an option is knocked out, it loses all of its value once the asset's price crosses the threshold. The barrier must be crossed in the direction indicated by the arrows going up and down. In addition to X and B , the strike price K , interest rate r , time to maturity T , dividend rate q , and volatility σ are the input arguments utilized to determine the value of barrier options.

Barrier options are one of the most frequently traded derivatives on the financial markets, claim Wang, B. & Wang, L. (2011). They stand out from standard solutions thanks to unique qualities. The fact that barrier options are typically less expensive than normal options is one reason why an investor prefers them to plain vanilla options. This is so that the option holder can collect the payoff, the asset price must pass a particular threshold first. The third factor is that barrier options might better fit risk hedging requirements than conventional options.

Farmers may buy a call option if it is anticipated to be higher. The payoff serves for the call contracts are provided by Baosung & Kooten, (2015) from the perspective of the purchasers.

$$P(x)_{call} = \begin{cases} 0, & x < k \\ D(x - k), & x \geq k \end{cases} \quad (19)$$

$P(x)$ denotes the option payoff, D the tick size (the amount of money for each weather index unit) and k for the strike (trigger) value. Equation (19) turn also to be the payment of the barrier down and out call option for the barrier temperature option, where the k is now the barrier level.

For barrier it is

$$P(x)_{call} = \begin{cases} 0, & x < k \text{ and } x < B \\ D(x - k), & x \geq k \text{ and } x > B, 0 \leq t \leq T \end{cases} \quad (20)$$

Where B is the barrier level and x is the weather index (GDD).

The GDD option is dependent on the sum of the GDD across the growth season, where each temperature process follows the Gaussian process, which is represented by $T_t \sim N(\mu_t, v_t)$. Once we get both the conditional mean and variance of the GDD, we can proceed. The conditional mean and variance of GDD_n for time $t < t_1$ can be calculated as follows:

$$GDD_n \sim (N(\mu_u, \sigma_n^2))$$

Where,

$$\mu_u = E^{\mathbb{Q}}[GDD_n | \mathcal{F}_s] = \sum_{i=1}^n E^{\mathbb{Q}}[T_{ti} | \mathcal{F}_t] - k, \quad (21)$$

$$\sigma_n^2 \equiv var^{\mathbb{Q}}[GDD_n | \mathcal{F}_t] = \sum_{i=1}^n var^{\mathbb{Q}}[T_t | \mathcal{F}_t] + 2 \sum \sum_{i < j} cov^{\mathbb{Q}}[T_{ti}, T_{tj}] | \mathcal{F}_t] \quad (22)$$

The anticipated return is the following, assuming a normal distribution for the weather indicator used in a financial instrument:

$$E_p = \int_a^b f(x) p(x) dx, \quad (23)$$

Where $p(x)$, is the payment associated with the financial instrument for the potential outcome. x is the weather index. This x at some point will reach the barrier level. The probability density function (PDF) of the weather index is denoted by $f(x)$. When the weather index is transformed into a regular normal distribution, let $z = \frac{x-\mu}{\sigma}$, and the expected pay-out function is as follows:

$$E_p = \int_a^b \phi(z) p(z) dz = \frac{1}{\sigma} \int_a^b \phi(z) p(x) dx \quad (24)$$

From equation (24), $\phi(z)$ signifies the PDF of the typical normal distribution and σ is the standard deviation of the weather index.

Inserting the payoff function for the call contract in the corresponding uncapped call options with closed-form functions are as follows when the expected pay-out function is entered:

$$E_{p,CALL} = \frac{1}{\sigma} \int_k^\infty D(x - k) \phi\left(\frac{x-\mu}{\sigma}\right) dx = D\sigma\phi\left(\frac{k-\mu}{\sigma}\right) + D(\mu - k) \left[1 - \Phi\left(\frac{k-\mu}{\sigma}\right)\right], \quad (25)$$

Multiplying the above by the difference of x and B ($x - B$) gives us the call pay-off of the barrier call option to be

$$= (x - B) \left[D\sigma\phi\left(\frac{k-\mu}{\sigma}\right) + D(\mu - k) \left[1 - \Phi\left(\frac{k-\mu}{\sigma}\right)\right] \right] \quad (26)$$

Proposition 3.3: *Therefore, the cost of the GDD call option at time $t < t_1$ is as follows:*

$$\begin{aligned} c(t) &= e^{-r(t_n-t)} E^{\mathbb{Q}}[\max(H_n - k, 0) | \mathcal{F}_t] \\ &= e^{-r(t_n-t)} \int_{k_1}^\infty (x - k) f_{H_n}(x) dx \end{aligned}$$

$$= e^{-r(t_n-t)} \left\{ (\mu_n - k) \Phi \left(-\frac{k-\mu_n}{\sigma_n} \right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}} \right\} \quad (27)$$

Where t_n is time to maturity. the normal distribution's probability density function is f_{H_n} , and Φ represents the cumulative distribution function for the common normal distribution.

Proposition 3.4 The price of the down and out barrier call option as

$$= e^{-r(t_n-t)} \left\{ (\mu_n - k) \Phi \left(-\frac{k-\mu_n}{\sigma_n} \right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}} \right\} (x - B) \quad (28)$$

3. PARAMETER ESTIMATIONS

3.1 DESCRIPTIVE STATISTICS

This study looks at the daily average temperature data that was recorded in degrees Celsius in Malawi's Kasungu area. The data spans the years 1990 to 2020 and 11315 data series are included. In order to preserve consistency over time, the temperature readings on February 29 of every leap year have been dropped. In Malawi's center area, Kasungu district is located.

A very big sample, which runs the risk of estimating parameters being changed by dynamics that do not represent future behaviors of temperature anymore, including urban influences, is regarded to be a worse sample to research temperature dynamics than a 31-year duration period. On the other hand, if the period is really short, it's possible that crucial dynamics won't be shown, which could lead to a bad model (Alexandridis and Zapranis, 2006).

From our data set of 31 years, we plot the graph of average temperature as shown in figure 4.1. This graph shows the average temperature of Kasungu district from 1990-2020. It highlights the seasonality of daily average temperature changes, particularly highlighting how it resembles a sine function. The daily average temperature fluctuates often and predictably between hot summer and cold winter months.

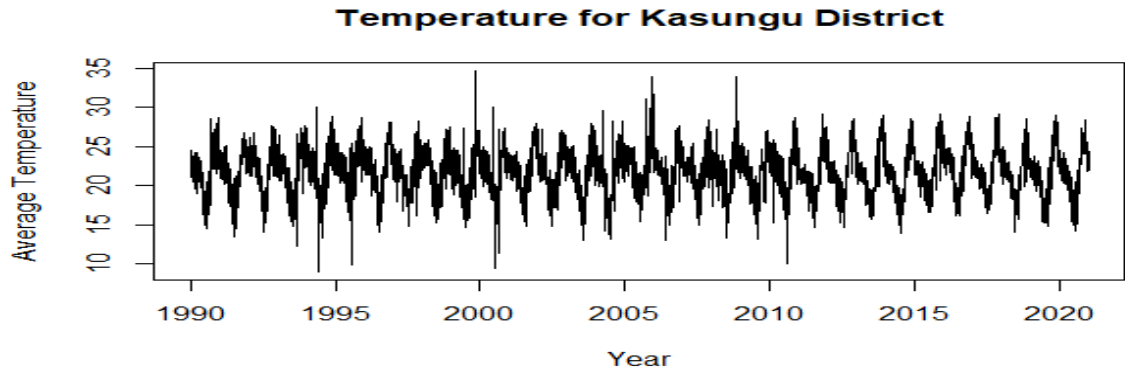


Figure 3.1 shows a stochastic model of average daily temperature in Kasungu from 1990 to 2020.

The use of using the Ornstein-Uhlenbeck procedure to mean-revert simulate temperature behaviour is justified in light of seasonal fluctuation likewise long-term trends in temperature. The daily average temperature variation is meant to gradually return to the mean over time. The Anderson Darling test was used to determine whether the varying temperatures for the Kasungu district are normal. The hypothesis was rejected with an $A = 0.83279, P.value = 0.03181$. Although it is not normally distributed, we shall treat the daily variation in temperature as a Brownian motion. This is the case because when the histogram is used to check the temperature difference's normality, it finds that it is roughly normal. This agrees with what Wang, et al., (2015) who state that when data is huge then it has to be approximate normal and Brownian motion should be considered. Figure 4.2 illustrates that the temperature differential is roughly normal.

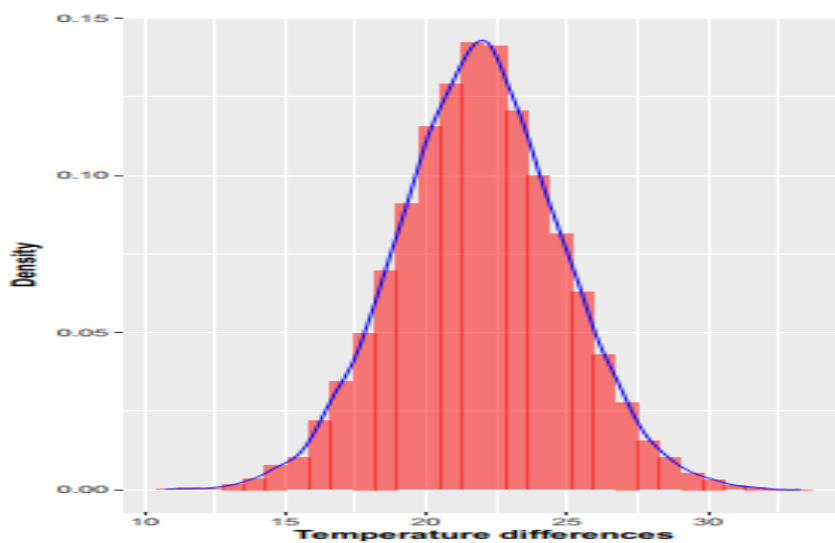


Figure 3.2: histogram of temperature differences of Kasungu district, Malawi.

The motions of the process are characterized by the stochastic differential equation if we indicate by T_t the average temperature at the date t .

$$dT_t = dT_t^m + k(T_t^m - T_t)dt + \sigma_t dW_t$$

where W_t is the Brownian Motion, σ_t is the volatility square root time of the fluctuations, T_t^m is the long-term mean, and k is the mean reversion rate.

To better understand the dynamics of temperature, we computed the descriptive statistics of the data, as provided in Table 1.

Table 3.1: Kasungu's descriptive temperature statistics

Maximum	34.7
Minimum	9
Mean	21.88196
Median	22.1
Mode	23.25
Variance	8.504897
Standard Deviation	2.916316
Skewness	-0.1304493
Kurtosis	2.961842
Coefficient of variation	13 %

According to a descriptive evaluation of the information, the modal temperature is 23.25°C, while the average daily temperature is 21.8833°C. The coefficient of variation (CV) is defined as the difference between the standard deviation and the mean. With an increase in the coefficient of variation, the mean dispersion becomes more pronounced. Usually, it is expressed as a percentage. The estimated average of 21.8833°C (in our example) is shown to be representative of the data by the coefficient of variation. The CV of 13% demonstrates the data's minimal variability. As a result, a dataset's estimated average of the values is more reliable.

The observed distribution exhibits a little amount of leftward asymmetry, with a skewness (asymmetry coefficient) smaller than 0 (-0.1304493). The data's median temperature is 22.1°C, whereas the mean is 21.88195°C, making the data's mode,

23.25°C, greater. The mode must be larger than both the median and the arithmetic mean in order for there to be left-skewed asymmetry, and this is proven (Kovac, 2020).

The size of the two tails' combined lengths is measured by kurtosis. It calculates how likely the tails are. The amount is typically contrasted with the normal distribution's kurtosis, which is 3 in this case. Our data are normally distributed because the observed kurtosis is 2.961842, which is roughly 3.

3.2 TREND AND SEASONALITY

The temperature process has a seasonal trend, as seen in figure 4:1. Consequently, we create a seasonal mean that also considers patterns as

$$T_t^m = A + Bt + C\sin(\omega t + \rho), \quad \omega = \frac{2\pi}{365}$$

where t is the range of days (1–365) in a calendar year. T_t^m is the predicted temperature for a given day in 2021. $A + Bt$, which reflects the tendency brought on by urban effects and global warming, since extreme temperature occurrences do not necessarily occur at the beginning and middle of the year. C , which stands for amplitude, determines when we feel the highest or lowest temperature.

Using trigonometric formulae, we have

$$\begin{aligned} T_t^m &= A + Bt + C\sin(\omega t + \rho) \\ &= A + B + C_1 \sin(\omega t) + C_2 \sin(\omega t) \end{aligned}$$

Where

$$c = \sqrt{c_1^2 + c_2^2}$$

$$\rho = \arctan\left(\frac{c_2}{c_1}\right) - \pi$$

To determine the parameters $A = \{A, B, c_1, c_2\}$ that solve the optimization problem $\min_A ||T_t^m - X(t)||$, we therefore fit T_t^m using least squares methods. $X(t)$ is the data

vector. We present estimated values for parameters in table 4:2, all of which suggested that they are significant. When the estimated values are inserted into T_t^m we get

$$T_t^m = 22.04 - 0.0000286t + 2.89929\sin\left(\frac{2\pi}{365}t + 1.60807\right) \quad (30)$$

where A is the coefficient, which is set to 22.04 °C, and is computed as the average of daily average temperatures between 1.1.1990 and 31.12.2020. Coefficient B, which has a value of -0.0000284t°C, represents the expected rise in the annual average temperature in 2021 as compared to the average of the average temperatures for the base period (1990-2020).

Given that it assumes the highest average daily temperature in 2021 will be around 24.93929 °C (22.04 °C + 2.89929 °C) and the lowest will be approximately 19.14071 °C (22.04 °C - 2.89929 °C), the value of Factor C is 2.89929 °C, which is realistic and predicted.

As a result, the average daily temperature, previously computed at 22.04 °C, can be anticipated at the end of March (2.85 months after the year's start), and all average daily values beginning on January 1 should be lower than 22.04 °C with an upward tendency. The sine function shift is 1.60807, or roughly three months of a year, in terms of the sine function, or 2.85 months. In March, the yearly average temperature for the months that are not susceptible to significant temperature swings changes.

The predicted values for 2021 were derived using the O-U procedure and have been presented in Figure 3:9 by substituting the 365 calendar days for t.

Table 3:2 provides the seasonal mean's estimated parameter values.

Parameter	Estimated	Std Error	t value	Pr(> t)
A	2.204e+01	3.600e-02	612.248	< 2e-16 ***
B	-2.864e-05	5.512e-06	-5.196	2.07e-07 ***
C1	-8.671e-01	2.546e-02	-34.058	< 2e-16 ***
C2	2.986e+00	2.545e-02	117.331	< 2e-16 ***

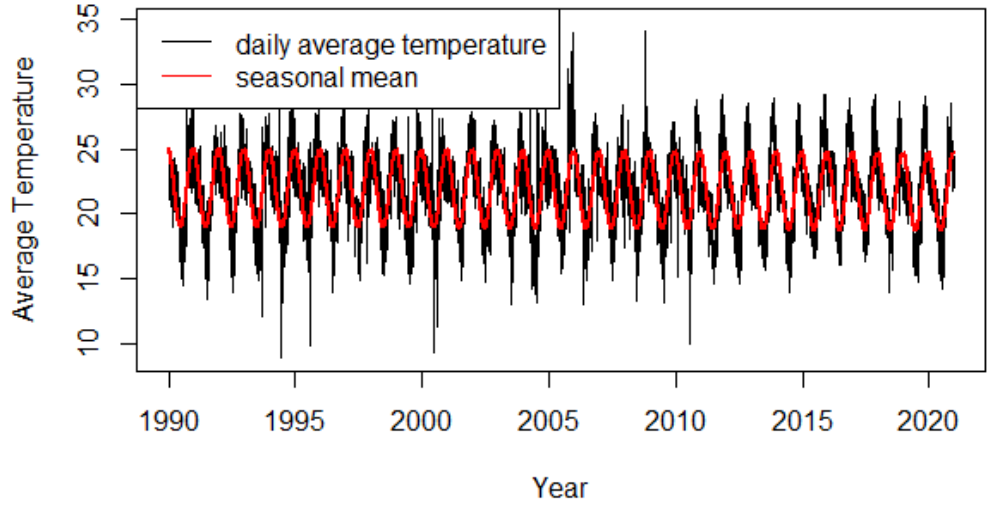


Figure 3:3 shows the daily average temperature and predicted season means.

The yearly mean and the daily average temperature are depicted in Figure 3:3. The mean reasonably fits the data.

3.3 ESTIMATION OF MEAN REVERTING SPEED

According to Patricia P. et al., (2021) and Wang et al.,(2015) derived the simplest Ornstein-Uhlenbeck process, often known as a mean-reverting process, is as follows.

$$dT_t = [\alpha(T_t^m - T_t) + \frac{dT_t^m}{dt}] dt + \sigma_t dW_t \quad (31)$$

Here, T_t^m , is the normal level of T_t , to which T_t tends to revert, and α is the speed of reversion. Keep in mind that the difference between T_t and T_t^m determines the anticipated change in T_t . T_t is more likely to decrease (increase) over the following brief period of time if T_t is higher (less) than T_t^m . Therefore, despite satisfying the Markov characteristic, this process lacks independent increments.

If the value of T_t is currently T_0 and T_t is calculated using equation (18), then its anticipated value at any point in the future is given by

$$E[T_t] = T_t^m + (T_0 - T_t^m)e^{-\alpha t} \quad (32)$$

Also, the variance of $(T_t - T_t^m)$ is

$$V[T_t - T_t^m] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) \quad (33)$$

For derivation of equation 32 and 33, we refer to appendix 4

The expected value of T_t converges to $\frac{\sigma^2}{2\alpha}$, as may be seen from these formulae. Additionally, as $\alpha \rightarrow \infty$, $V[T_t] \rightarrow 0$, this denotes that T_t can never stray from T_t^m , not even momentarily. In the end, T_t becomes a straightforward Brownian motion, and $V[T_t] \rightarrow \sigma_t^2$. The first-order autoregressive process in discrete time is represented by equation (32) in continuous time. Equation (32) is the limiting case for the following AR(1) process as $\Delta t \rightarrow 0$.

$$T_t - T_{t-1} = T_t^m(1 - e^{-2\alpha}) + (e^{-2\alpha} - 1)T_{t-1} + \epsilon_t \quad (40)$$

where ϵ_t has the mean of a normal distribution zero and a standard deviation of σ_t and

$$\sigma_t^2 = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha_t}) \quad (41)$$

Therefore, by doing the regression and utilizing discrete temporal data (the only data ever accessible), one might estimate the parameters of equation (34):

$$T_t - T_{t-1} = a + bT_{t-1} + \epsilon_t \quad (42)$$

Afterwards, calculating $T_t^m = -\frac{a}{b}$,

$$\alpha = -\log(1 + b) \quad (43)$$

And

$$\sigma = \sigma_t \sqrt{\frac{\log(1+b)}{(1+b)^2 - 1}} \quad (44)$$

Where σ_t is the regression's standard error.

To find this mean reversion in our case we run the ACF of the decomposed data. The ACF of our data is our b . The same can also be found by running a regression. $b = -0.2137$. Substituting this into (43) we get the mean reversion of 0.2404. Fitting in mean reversion and σ_t in (44) we get 0.8943.

Kasungu average temperature data was decomposed under the additive model to find the mean reversion. **Figure 3:4**. Shows the decomposed plot of the statistics on the average temperature of kasungu. **Figure 3:5** shows the Auto Correlation Function (ACF) plot.

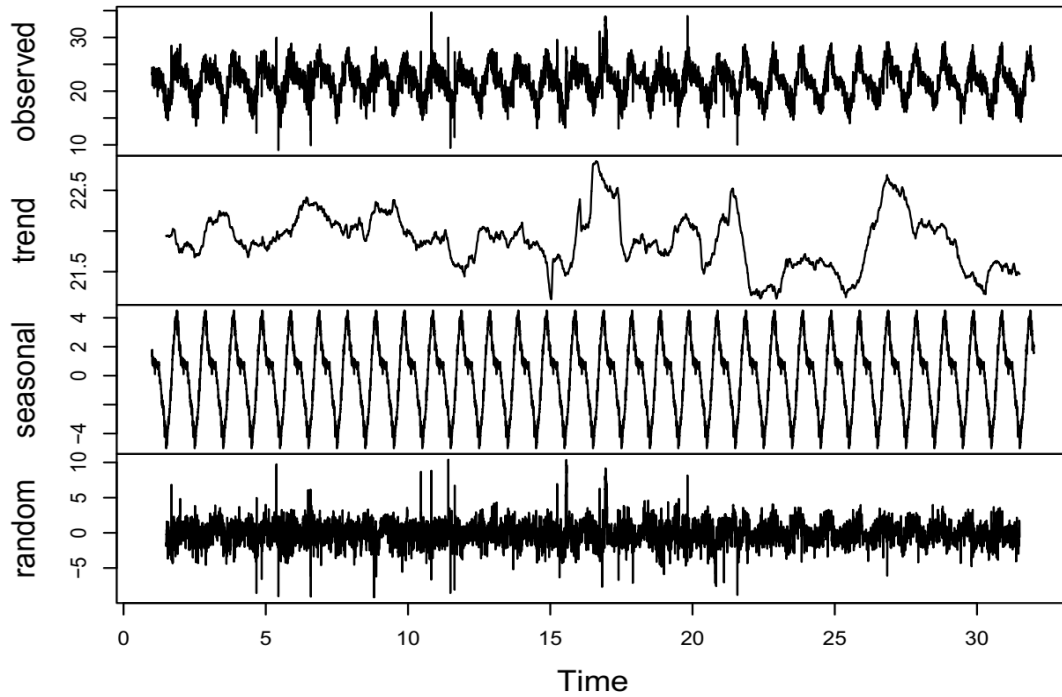


Figure 3:4: Decomposed plot of the Average Temperature of Kasungu

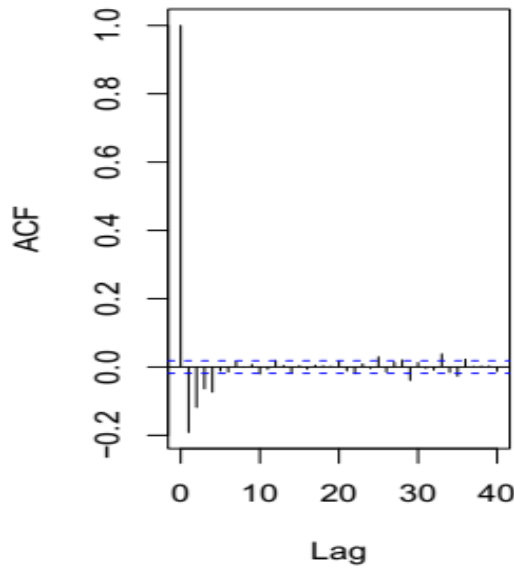


Figure 3:5: shows the Auto Correlation Function (ACF) plot

In ACF, there are significant autocorrelation values for a number of lags, which may call for the adoption of more advanced autoregressive models. The ACF also reveals the presence of seasonality in both the residuals and time dependence in the variance of residuals. To extract the σ_t^2 from the residuals. We must analyse and simulate the random noise process first. But testing for stationarity Augmented Dickey-Fuller (ADF) test, the *Dickey – Fuller* = -7.3612 , *P. value* = 0.01 . Since the P-value is less than 0.05 , we cannot rule out the possibility that the data are stationary and that mean reversion is constant.

3.4 CALCULATING THE VOLATILITY OF THE TEMPERATURE PROCESS

From the residual graph in figure 3:6 below, as many points do not vary from the normal line, it can be said that the residuals are normally distributed.

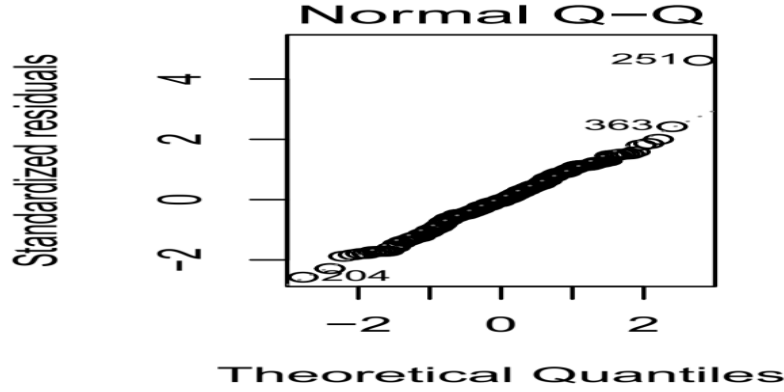


Figure 3:6, a residuals graph.

According to the methods outlined in (Benth, F.E., & Šaltyt-Benth, 2011; Dzupire et al., 2019; Wang, et al, 2015), the volatility is extracted. After that, the residuals are split into 365 categories, each of which represents a day of the year for the last 31 years (1990-2020). Then, we calculate the mean of the squared residual in each set of expected daily residuals:

$$\sigma_t^2 = E[(\sigma_t \varepsilon_t)^2] \quad (45)$$

These numbers are used as observable estimates of the daily variance based on years of observations for the specific day, as seen in the figure below.

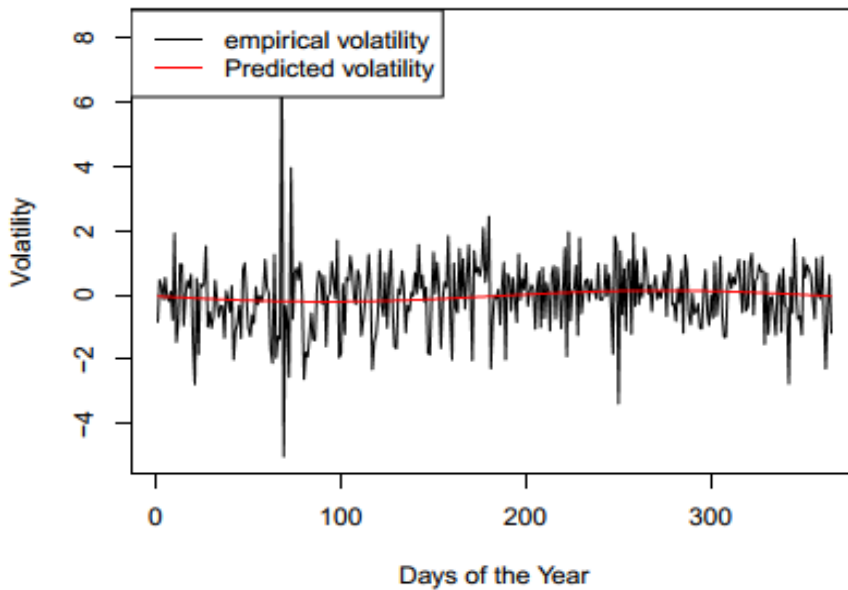


Figure 3:7, Empirical volatility

As can be observed, the wet season has bigger temperature fluctuations than the dry season. In order to derive the volatility model, we fit the data using the volatility model (44)

$$\sigma_t^2 = -0.02603 - 0.17421 \sin \left[\frac{2\pi t}{365} \right] \quad (45)$$

Figure (3:7) illustrates the comparison between the estimated model and the empirical volatility and demonstrates how closely the estimated model matches the empirical model.

Comparing the actual average temperature to the outcomes of the stochastic models under ARIMA to determine how well the model performed. The calculated Root Mean Square Error (RMSE), which evaluates how well the temperature model performs, has a value of approximately 1.158505 and the Mean Performance Error (MPE), which compares the actual temperature to temperature models, was roughly -0.3110809. This means that the model is over-forecasted by around 0.311%. Hence figure 4:8 below is the graph of backward forecast temperature. This has been done on the decomposed temperature that is without seasonality and trend (random). Figure 4.9 shows forward forecasted temperature from 2021 to 2024 of Kasungu district using our temperature model. These forecasted temperature figures tell that no matter the situation is; this model can support the option provider in setting the derivative price for specific time frame regardless of anything.

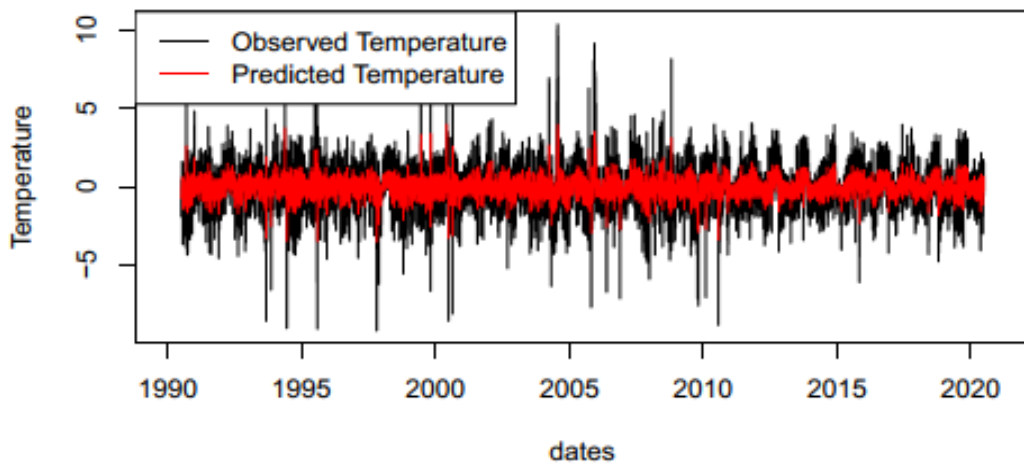


Figure 3:8. shows backward forecasted temperature of kasungu district

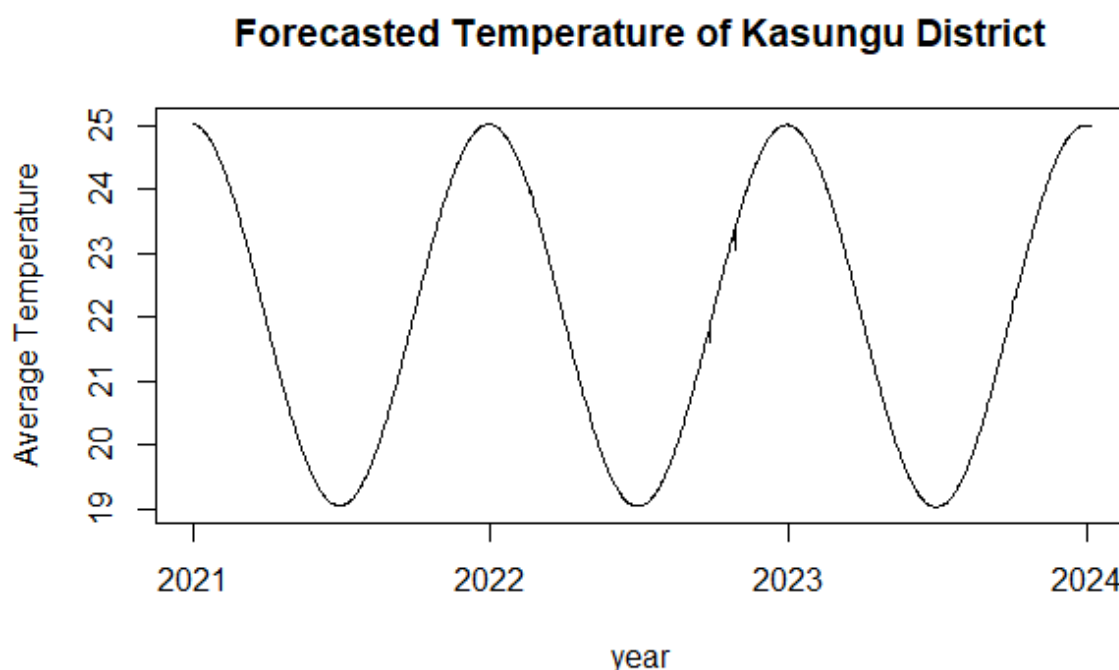


Figure 3:9. Shows forward forecasted Temperature of Kasungu district.

4. HEDGING MAIZE CROP YIELD

Hedging is an advanced risk management strategy that involves buying or selling an investment in order to potentially help reduce the chance of loss of an existing position. When purchasing call options, farmers use temperature as the underlying asset. A call option gives its owner the right to pay a premium and purchase the underlying asset from the provider for the specified time period at the specified price. If the buyer expects unfavourable weather that could increase the price of derivatives due to monetary gain generated by the set price the supplier has to offer the customer, they may decide to buy call options.

Each GDD's monetary value and the weather derivative's monetary value are calculated based on the index's value. The derivative buyer wagers that throughout the duration of the contract, there will be an unfavourable change in air temperature for his company. The provider will pay the buyer a set fee if his prediction comes true.

Weather derivatives must be introduced as a contingency mechanism to assist farmers in hedging possible losses in maize yield due to climate change, particularly the

fluctuation of temperatures. The paper offers temperature weather derivatives, which have temperature as their underlying index. Possibly one of the most important elements affecting the variability of maize output is temperature.

Now taking from **Proposition 3.3**: *Therefore, the cost of the GDD call option at time $t < t_1$ is as follows:*

$$\begin{aligned} c(t) &= e^{-r(t_n-t)} E^Q[\max(H_n - k, 0) | \mathcal{F}_t] \\ &= e^{-r(t_n-t)} \int_k^\infty (x - k) f_{H_n}(x) dx \\ &= e^{-r(t_n-t)} \left\{ (\mu_n - k) \Phi\left(-\frac{k - \mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k - \mu_n)^2}{2\sigma_n^2}} \right\} \end{aligned}$$

$C(t)$ is the price or premium that the hedger (buyer of call option) needs to pay for the contract, r is a risk-free periodic market interest rate, t is the date the contract is issued (purchased), and t_n is the date the contract is claimed or the expiration date. E^Q is the expected payoff based on the predicted historical mean value of the temperature. The seller of the option would expect a reward for taking a risk loading which is often between 20% and 30% of the payoff (Baojing sun and G. Cornelis Van K. 2015). In the current application, we set the risk loading at 20% of the expected payoff of the contract.

From the forecasted temperature of 2021, the GDD has been calculated from August 1 to December 31 since this is the time where irrigation of maize crop is mainly being done. It has been seen observed that the temperatures are too low in the early month of August. This gives an insight to the farmer to hedge their crop yield against low temperature. The temperatures are also high in the months of November and December; this also give an insight to a farmer to hedge his or her crop yield against high temperatures.

To price the financial weather derivatives, we assume a tick size $D = \$1$ and risk-free interest rate $r = 0.08$, $\Delta t = \frac{3}{4} \text{year}$ (1 August to 31 December, 2021) making 153 days. and risk loading $b = 20\%$, $\mu = 12.68^\circ\text{C}$. The forecasted GDDs and standard deviation for 2021 are 1937.43°C , and 1.830°C respectively are used to calculate the actual premiums for the contracts.

Table 3: Specification of GDD options for year 2021

Items	call options
Weather index	GDD(x)
Strike Level	12.68-2*1.83
Barrier level	$B = 18^0 \text{ c}$
Tick Size (D)	\$1
Premium	\$3.50
Payoff	$\text{Max}(\text{GDD}-K, 0)$
Issue Date	1 August 2021
Maturity date	31 December 2022

The strike value is $\mu - 2\sigma$ for above 2 standard deviation the forecasted GDDs, This $2 = \frac{x-\mu}{\sigma}$, then $\phi(2) = 0.9772$ and $\Phi(2) = 0.9772$. The premium is calculated from equation 33. The premium is \$3.50 because we assume that the GDD is above barrier level with a difference of one. This is where the barrier option is equal to vanilla option. In this case, the barrier option does not knock out. Hence gives the holder the right to buy the underlying asset since it does not reach or fall below a predefined level over the option's lifetime.

The premium has been calculated to be \$3.50 because we assume that the GDD is above the barrier level with a difference of one. This is where the barrier option is equal to the vanilla option. In this case, the barrier option does not knock out. Hence gives the holder the right to buy the underlying asset since it does not reach or fall below a predefined level over the option's lifetime. Now if the GDD does not exceed the barrier level then the farmer will not have to pay the premium of \$3.50 because the option is invalid. On the same note, the farmer will have to exercise his right by paying the premium of this calculated premium when the GDD exceeds the barrier level. In return for this, the farmer gets paid off which happens to be the difference between GDD and barrier. This will be taken daily within the contract.

5. DISCUSSION AND CONCLUSION

The weather derivative market is a classical incomplete market since the weather indexes are not tradable assets, thus traditional no-arbitrage pricing methods such as the Black–Scholes are not applicable in pricing weather derivatives.

The study results further prove that the average temperature has been increasing since 1990. This is evidenced in Figure 3:5, tells us a story of an increase in average temperature. The results show that there will be a slow but steady increase in temperature. This scenario does not offer a positive outlook for agriculture production since they are vulnerable to an increase in temperature.

According to our model, if the parameters remain unchanged in the next 30 years or so, the results of this study provides an early view of possible temperature patterns. This can help the government and non-governmental organisation to devise contingency measures. Agriculture can be particularly threatened by these forecasts. As already stipulated earlier sterility of maize can be triggered by temperatures of 38⁰c and above and 10⁰c and below. This tells us that agriculture done during irrigation is much more vulnerable.

With the agenda of Malawi 2063, this study is good enough to be incorporated into agriculture policy as it presents a way to protect farmers from financial consequences due to extreme temperatures. This will bring an increase in the income of individual farmers, thereby mitigating poverty. We assume there is no correlation between the tradable asset and weather indexes, considering that we are interested in how a farmer can hedge temperature-related weather risks. The pricing model developed can be used in the agriculture industry where a farmer is interested in hedging weather risks due to temperature. It can also price weather derivatives in other weather-related industries affected by temperature. The results of this study can help insurance providers and the government to design products that can help the farmers. Projecting future temperatures and growing degree days is uncertain, so farmers wish to hedge against weather risk. However, markets must provide farmers with attractive, practical hedges representing producers' risks.

With, efficient and reliable pricing models, basis risks would decrease. As a result, the farmer would be more willing to pay for the contracts and trading activities in the market for weather derivatives.

ACKNOWLEDGEMENT

Temperature data was gathered from the Department of Climate Change and Meteorological Services, Blantyre, Malawi.

REFERENCE

- A. K. Alexandridis and A. D. Zaprakis. (2006). Wavelet analysis and weather derivatives pricing. *5th Hellenic Finance and Accounting Association (HFAA), Thessaloniki*, 15–16.
- Atalon, P. . et al. (2002). On Modelling and pricing Weather Derivatives. *Appl. Math. Finance*, 9(1), 1–20.
- Benth, F.E., & Šaltyt-Benth, J. (2011). Weather Derivatives and Stochastic Modelling of Temperature. *Int. J. Stoch. Anal.*
- Dzupire N., Ngale, P., & Odongo, L., O. (2019). Pricing Basket Weather Derivatives on Rainfall and Temperature Process. *International Journal of Financial Studies.*, 7, 35. <https://doi.org/Doi:10.3390/ijfs7030035>.
- Hatfield, J. L., & Prueger, J. H. (2011). .Agroecology: Implications for Plant Response to Climate Change. In: Yadav, S. S., et al, (Eds.),. *Wiley-Blackwell, West Sussex, UK*, 27–43.
- Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes; Effect on plant growth and development. *Elsevier*. www.elsevier.com/locate/wave
- Ines, H. (2017). *Exceed weather limitations with greenhouse-farming AGRIVI*. <https://www.agrivi.com/blog/greenhouse-farming-exceeds-weather-limitation>
- Kooten, B. S. & Van. (2015). Financial weather derivatives of corn production in the northern China; A comparison of pricing methods. *Journal of Emperical Finance*, 32, 201–209. [https://doi.org/\(https://dx.doing/10.1016/j.jemption.2015.03.014\)](https://doi.org/(https://dx.doing/10.1016/j.jemption.2015.03.014))
- Kovac, B. Z. & M. (2020). Ornstein-Uhlenbeck process and GARCH model for temperature forecasting in weather derivatives valuation. *CREBSS, VOL. 6*, 27–

42. <https://doi.org/10.1515/crebss>
- M. Islam and Chakraborti. (2015). Futures and Forward Contract as a route of Hedging the risk. *Risk Governance & Control: Financial Market & Institutions*.
- Means, T. (2018). *Understanding the various Temperatures of rain drop*.
<https://www.thoughtco.com/what-determines-rain-temperature-3443616>
- Ministry of Agriculture, I. and W. D. (2016). *National Agriculture Policy*.
- Olivier Niyitegeka and D.D. Tewari. (2013). Volatility clustering at the Johannesburg Stock Exchange: Investigation and Analysis. *Mediterranean Journal of Social Sciences*. <https://doi.org/Doi:10.5901/mjss.2013.v4n14p621>
- Patricia P., & C. D. (2021). Modelling the Daily Average Temperature and Designing a growing Degree Day (GDD)European put option for rice in Laguma. *Journal of Nature Studies*, 19(2), 52–61.
- Phiri, U. M. (2016). Politics of Conservation agriculture: Kasungu rural District, Malawi (Masters thesis). *International Institute of Social Studies*.
- Primajati G., R. M. ;& A. H. (2020). Barrier Options Models for Agricultural Commodity Price Protection. *Sinta*, 4, 71–78.
- Samuel Asante Gyamerah, P. N. and D. I. (2019). Hedging Crop Yields Against Weather Uncertainties-A Weather Derivatives Perspective. *Mathematical and Computational Applications*.
- Sun, W., & Lou, S. (2013). Design of Agriculture insurance policy for tea tree freezing damage in Zhejiang Province, China. *Theoretical and Applied Climatology*, 111, 713–728.
- Wang, B. & Wang, L. (2011). *Pricing Barrier Options Using Monte Carlo Methods*.
- Wang, Z. et all. (2015). Modelling and Forecasting Average Temperature for Weather Derivative Pricing. *Adv. Meteorol*.
- Xu, wei, Martin odening, and O. muBhoff. (2007). Inference pricing of weather insurance. Paper presented at 101st Seminar of the European Association of Agricultural Economists (EAAE). *Management of Climate Risks in Agricultere", Berlin, Germany, July 5-6*.

APPENDICES

Appendix 1

Proof of Proposition 3.1: let us rewrite $dT_t = dT_t^m + k(T_t^m - T_t) + \sigma_t dW_t$ as

$$d\hat{T}_t = k\hat{T}_t dt + \sigma_t dB_t$$

Where $\hat{T}_t = (T_t - T_t^m)$. The following transformation is practical for solving the stochastic equation above:

$$G(\hat{T}, t) = e^{-\int_0^t k(u) du} \hat{T}_t$$

Making use of the Ito lemma

$$\frac{\partial G}{\partial \hat{T}} = e^{-\int_0^t k(u) du}, \quad \frac{\partial^2 G}{\partial \hat{T}^2} = 0, \quad \frac{\partial G}{\partial t} = -k e^{-\int_0^t k(u) du} \hat{T}_t$$

$$\text{and } a = k(t)\hat{T}_t$$

$$b = \sigma_t$$

We have that

$$dG_t = \left(k e^{-\int_0^t k(u) du} \hat{T}_t - k e^{-\int_0^t k(u) du} T_t \right) dt + \sigma_t e^{-\int_0^t k(u) du} dB_t$$

Which reduces to

$$dG_t = \sigma_t e^{-\int_0^t k(u) du} dB_t$$

Suppose $s < t$, then integrating the above equation, we have that

$$G_t - G_s = \int_s^t \sigma_t e^{-\int_0^t k(u) du} dB_t$$

And replacing G , we have that $e^{-\int_0^t k(u) du} \hat{T}_t - \hat{T}_s = \int_s^t \sigma_t e^{-\int_0^t k(u) du} dB_t$

By rearrangement, we achieve that

$$\hat{T}_t = e^{-\int_0^t k(u) du} \hat{T}_t + e^{-\int_0^t k(u) du} \int_s^t \sigma_t e^{-\int_0^t k(u) du} dB_t$$

Since $\hat{T}_t = (T_t - T_t^m)$,

$$(T_t - T_t^m) = e^{-\int_0^t k(u)du} (T_s - T_s^m) + e^{-\int_0^t k(u)du} \int_s^t \sigma_s e^{-\int_0^s k(u)du} dW_s$$

Finally, by rearranging, we prove the proposition

$$T_t = T_t^m + e^{-\int_0^t k(u)du} (T_s - T_s^m) + e^{-\int_0^t k(u)du} + \int_s^t \sigma_s e^{-\int_s^t k(u)du} dW_s$$

Since $s = t - 1$ and $u = t - s$ hence

$$T_t = [T_{t-1} - T_{t-1}^m]e^{-k(t-s)} + T_t^m + \int_s^t e^{-k(t-r)} \sigma_r dW_r \quad \blacksquare$$

Appendix 2:

Proof of Proposition 3.2: Let $Z_t = e^{\int_0^t \alpha ds} (T_t^m - T_t)$

Ito's lemma

$$\begin{aligned} dZ_t &= e^{\int_0^t \alpha ds} T_t^m dt + \alpha e^{\alpha t} (T_t^m - T_t) dt - e^{\int_0^t \alpha ds} dT_t \\ &= e^{\int_0^t \alpha ds} [(T_t^m + \alpha(T_t^m - T_t)dt - (\alpha(T_t^m - T_t) + T_t^m)dt - \sigma_t dW_t) \\ &\quad \therefore Z_t = Z_0 - \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s \end{aligned}$$

$$\therefore e^{\int_0^t \alpha ds} (T_t^m - T_t) = T_0^m - T_0 - \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

Now $T_0^m - T_0 = c$ gives

$$T_t^m - T_t = -e^{\int_0^t \alpha ds} \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

$$\therefore T_t = T_t^m + e^{\int_0^t \alpha ds} \int_0^t e^{\int_0^s \alpha ds} \sigma_s dW_s$$

$$\Rightarrow E[T_t] = T_t^m \quad \blacksquare$$

Appendix 3

Proof of proposition 3.3

We know that $H_n \equiv x \sim (N(\mu_n, \sigma_n^2))$. Let $u = \left(\frac{x-\mu_n}{\sigma_n}\right)$,

$$\text{then } du = \frac{dx}{\sigma_n}, dx = \sigma_n du, \text{ and } x = \mu_n + \sigma_n u.$$

$$\int_k^\infty (x - k) f_{H_n}(x) dx = \int_k^\infty x f_{H_n}(x) dx - k \int_k^\infty f_{H_n}(x) dx$$

It is simple to calculate the right-hand side's second term in the manner shown below:

$$\begin{aligned} k \int_k^\infty f_{H_n}(x) dx &= k \left(1 - \int_k^\infty f_{H_n}(x) dx \right) \\ &= k \left(1 - \Phi \left(\frac{k - \mu_n}{\sigma_n} \right) \right) \\ &= k \Phi \left(-\frac{k - \mu_n}{\sigma_n} \right) \end{aligned}$$

The first term $\int_k^\infty x f_{H_n}(x) dx$ can be calculated as the follows:

$$\begin{aligned} \int_k^\infty x f_{H_n}(x) dx &= \int_k^\infty \frac{x}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-\mu_n)^2}{2\sigma_n^2}} dx \\ &= \int_{\frac{k-\mu_n}{\sigma_n}}^\infty \frac{\mu_n + \sigma_n u}{\sqrt{2\pi}\sigma_n} e^{-\frac{u^2}{2}} du \\ &= \int_{\frac{k-\mu_n}{\sigma_n}}^\infty \frac{1}{\sqrt{2\pi}} (\mu_n + \sigma_n u) e^{-\frac{u^2}{2}} du \\ &= \int_{\frac{k-\mu_n}{\sigma_n}}^\infty \mu_n \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du + \int_{\frac{k-\mu_n}{\sigma_n}}^\infty \sigma_n u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\ &= \mu_n \Phi \left(-\frac{k - \mu_n}{\sigma_n} \right) + \frac{\sigma_n}{\sqrt{2\pi}} \int_{\frac{k-\mu_n}{\sigma_n}}^\infty u e^{-\frac{u^2}{2}} du \\ &= \mu_n \Phi \left(-\frac{k - \mu_n}{\sigma_n} \right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k-\mu_n)^2}{2\sigma_n^2}} \end{aligned}$$

The following can be used to generate the second term of the final equation:

$$\int_{\frac{k-\mu_n}{\sigma_n}}^\infty u e^{-\frac{u^2}{2}} du = -e^{-\frac{u^2}{2}} \Big|_{\frac{k-\mu_n}{\sigma_n}}^\infty$$

Therefore, $\int_k^\infty (x - k) f_{H_n}(x) dx = (\mu_n - k) \Phi\left(-\frac{k - \mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{(k - \mu_n)^2}{2\sigma_n^2}}$ ■

Appendix 4

We return to the easy O-U (mean-reverting) approach of equation (32) for the derivation of equations 32 and 33. Set T_t^m to 0 to make the calculation simpler, which results in the equation: $dT_t = -\alpha T_t dt + \sigma_t dW_t$ (34)

We said that equations (32) and (33) provide the mean and variance of T_t in the text. To support this assertion, we can apply the Kolmogorov forward equation.

Write $M(\theta, t) = E[e^{-\theta T}]$ as the moment-generating function for T_t

$$= \int_{-\infty}^{\infty} \phi(T_0, t_0; T, t) e^{-\theta T} dT \quad (35)$$

Then,

$$\frac{\partial M}{\partial t} = \int_{-\infty}^{\infty} \frac{\partial \phi}{\partial t} e^{-\theta T} dT \quad (36)$$

The Kolmogorov forward for this process is

$$\frac{\partial \phi}{\partial t} = \frac{1}{2} \sigma^2 \frac{\partial^2 \phi}{\partial T^2} - \alpha T \frac{\partial \phi}{\partial T} + \alpha \phi \quad (37)$$

The following equation for $M(\theta, t)$ is obtained by substituting this for $\frac{\partial \phi}{\partial t}$ in equation (17) and integrating by parts:

$$\frac{1}{2} \sigma^2 \theta^2 - \alpha \theta \frac{\partial M}{\partial \theta} = \frac{\partial M}{\partial t} \quad (38)$$

Boundary conditions must be met in order to solve this partial differential:

$$M(0, t) = 1, -M_\theta(0, 0) = T_0 \text{ and} \\ V[T_0] = M_{\theta\theta}(0, 0) - T_0^2 = 0$$

Hence the equation has the following equation

$$M(0, t) = e^{\frac{\sigma^2 \theta^2}{4\alpha}} \left[1 - T_0 \theta e^{-\alpha t} + \left(\frac{1}{2} T_0^2 - \frac{\theta^2}{4\alpha} \right) \theta^2 e^{-2\alpha t} \right] \quad (39)$$

Using the fact the $E[T_t] = -M_\theta(0, t)$ and $E[T_t^2] = M_{\theta\theta}(0, t)$ verifies equation (32) and (33) ■

APPENDIX

This section presents the R code used to analyse data in this study.

```
ChitoTemp <- read_csv("C:/Users/Admin/Desktop/R FILES/ChitoTemp.csv")

dates <- seq(as.Date("01/01/1990", format = "%d/%m/%Y"), by = "days", length =
length(ChitoTemp$Temperature))

plot(dates, ChitoTemp$Temperature, type="l", main="Temperature of Kasungu
District", xlab="year", ylab="Average Temperature")

hist(ChitoTemp$Temperature, main = "Temperature Difference Kasungu District",
xlab="Average Temperature Difference", ylab="frequency" )

mean(ChitoTemp$Temperature)

median(ChitoTemp$Temperature)

max(ChitoTemp$Temperature)

min(ChitoTemp$Temperature)

mode(ChitoTemp$Temperature)

sd(ChitoTemp$Temperature)

var(ChitoTemp$Temperature)

summary(ChitoTemp$Temperature)

summary.regression(ChitoTemp$Temperature)

(cor.value<-acf(ChitoTemp$Temperature,plot = FALSE)$acf[2])

fit1=lm(ChitoTemp$Temperature~sin(2*pi*ChitoTemp$`Day
Number`/365)+cos(2*pi*ChitoTemp$`Day Number`/365),ChitoTemp)

summary(fit1)

x<-c(ChitoTemp$`Day Number`)

y<-c(ChitoTemp$Temperature)

regression=lm(y~x)

summary(regression)
```

```

confint(regression)#95% confidence intervals

fit=lm(y~x+sin(2*pi*x/365)+cos(2*pi*x/365), ChitoTemp)

summary(fit)

dates <- seq(as.Date("01/01/1990", format = "%d/%m/%Y"),by = "days", length =
length(ChitoTemp$Temperature))

plot(dates, ChitoTemp$Temperature, type="l", main="Temperature of Kasungu
District", xlab="year",ylab ="Average Temperature")

mu=22.04-0.0000284*ChitoTemp$`Day`                               Number`-
0.08671*sin(2*pi*ChitoTemp$`Day`
Number`/365)+2.986*cos(2*pi*ChitoTemp$`Day` Number`/365)

plot(dates, ChitoTemp$Temperature, type="l", xlab="Year", ylab="Average
Temperature")

lines(dates, mu, type = "l",col="red",lwd=2,lty="solid")

legend("topleft", c("daily average temperature", "seasonal mean"), lty =
c("solid","solid"), col=c("black","red"))

plot(dates, mu, type = "l",col="red",lwd=2,lty="solid")

acf(ChitoTemp$Temperature^2)

plot(ChitoTemp$Temperature, type="l")

attributes(fit)

wt=fit$fitted.values

dat= ChitoTemp-mu

dat = random_Chito

random_Chito

acf(dat, na.action=na.pass)

x<-c(random_Chito)

my_acf=acf(x, na.action = na.pass)

my_acf

```

```

my_acf$acf[2]

acf(dat, na.action=na.pass)

acf(dat^2,na.action=na.pass)

my_acf$acf[2]

write.table(dat, "decompose_chito$random", row.names=FALSE, sep=",")

summary(dat)

hist(dat$Temperature)

plot(dat, type="l")

#dt_1<- diff(random_Chito, lag=1)

dt_1<-lag(random_Chito, 1)

dt_2<-lag(random_Chito,2)

dt_2<-lag(random_Chito,3)

dt_2<-lag(random_Chito,4)

datt=random_Chito[2:length(random_Chito)]

dt_1=dt_1[1:length(dt_1)-1]

fit2=lm(datt~dt_1)

summary(fit2)

sum(((ChitoTemp$Temperature- mean(ChitoTemp$Temperature))^2)

ss= 0.675865*dt_1

dates <- seq(as.Date("01/01/1990", format = "%d/%m/%Y"),by = "days", length =
length(datt))

plot(dates, datt, type="l", ylab="Temperature")

lines(dates, ss, type = "l", col="red")

legend("topleft",          c("Observed          Temperature",          "Predicted
Temperature"),lty=c("solid","solid") , col=c("black","red"))

res=fit2$residuals

res=res/ 1.125

```

```

write.table(res, "decompose_Chito$random", row.names=FALSE, sep=",")

skewness(decompose_Chito$random)

kurtosis(decompose_Chito$random)

plot(decompose_Chito$random, type="l")

plot(decompose_Chito$random^2, type="l")

par(mfrow=c(1,2))

plot(acf(res, plot=FALSE),main="ACF of Residuals",col="red",lwd=4, lty="solid")

plot(acf(res^2, plot=FALSE), main="ACF of Residuals", col="red",lwd=4,
lty="solid")

vol=decompose_Chito$random

summary(vol)

vol=vol[1:365]

x=1:365

#da=lm(vol~sin(2*pi*x/365))

da=lm(vol~sin(2*pi*x/365), subset=(1:length(x)!=306))

summary(da)

par(mfrow = c(2,2))

plot(da)

za=-0.02603 - 0.17421sin(2*pi*x/365)

par(mfrow=c(1,1))

plot(vol, type="l", xlab="Days of the Year",ylab="Volatility")

lines(za, type="l", col="red", lwd=2, lty="solid")

legend("topleft", c("empirical volatility", "Predicted volatility"),lty=c("solid","solid") ,
col=c("black","red"))

x=1:length(res)

res1=res/za^(0.5)

par(mfrow=c(1,2))

```

```

plot(acf(res1, plot=TRUE), main="ACF of Residuals", col="red",lwd=4, lty="solid")
plot(acf(res1^2, plot=FALSE),main="ACF of squared Residuals", col="red",lwd=4,
lty="solid")

dates <- seq(as.Date("01/01/1990", format = "%d/%m/%Y"),by = "days", length =
length(res1))

par(mfrow=c(1,1))

plot(dates, res1, main="Reisiduals Plot", type="l", ylab="Residauls",xlab="Year")

hist(res1)

kurtosis(res1)

skewness(res1)

library(tseries)

jarque.bera.test(res1)

summary(res1)

#Simulation

n.sample <- rnorm(n = 11315, mean = 21.88, sd = 2.916)

#Skewness and Kurtosis

library(moments)

skewness(n.sample)

kurtosis(n.sample)

#Histogram

library(ggplot2)

ChitoTemp <- data.frame(n.sample)

ggplot(ChitoTemp, aes(x = n.sample), binwidth = 2) +

  geom_histogram(aes(y = ..density..), fill = 'red', alpha = 0.5) +

  geom_density(colour = 'blue') + xlab(expression(bold('Simulated Samples')) +

  ylab(expression(bold('Density'))))

```

