African Journal of Education Science and Technology (AJEST) Vol. 8 No.3 (October, 2025)
University of Eldoret, Kenya, Mount Kenya University, Kenya, Chukwuemeka Odumegwu Ojukwu University, Nigeria, Kyambogo University, Uganda and University of Makeni, Sierra Leone

Pricing Rainfall Derivatives Based on a Hidden Markov Model to Mitigate Poor Yield

*Khumbo Thindwa and Nelson Dzupire
Mathematical Science Department, University of Malawi

*Corresponding Email: thindwakhumbo@gmail.comlk

Abstract

The study derives a hidden Markov model (HMM) in order to describe the occurrence process of daily rainfall. The amount of monthly rainfall on wet days is obtained using auto-regressive moving average (ARIMA) models because the model is linear, since the future values are constrained to be linear functions of past data. Monthly rainfall data from 1995 to 2015 were taken from the Kamuzu International Airport (KIA) Meteorology station in order to assess the model performance. The selected model corresponds to the ARIMA (1,0,1) (2,0,1)12, and was validated by another historical rainfall data under the same conditions. The results obtained prove that the model could be used to model and forecast the future rainfall variability in Malawi. The steady state matrix from hidden Markov model (HMM) shows likelihood that rainfall pattern at Kamuzu International Airport (KIA) would follow the pattern of the given probabilities (0.396 0.604) which implies that the probability that it would rain on a particular day in future at KIA is 0.369 and the probability that it would be dry on a particular day is 0.604. As such, the study suggests the rainfall weather derivative as a solution to mitigate poor yield due to erratic rainfall. Therefore, the study suggests buying basket options. The utility indifference approach is used to derive the buyer's and seller's prices. Reasonably, the indifference approach starts with an interesting idea that the amount of money at which a possible buyer of a claim is indifferent in terms of expected utility between buying and selling creates an upper or lower limit for the contract price.

Keywords: *Hidden Markov model, Incomplete market, utility indifference, Hedging, Crop insurance, Weather derivatives, ARIMA*

Introduction

Malawi's agricultural sector is widespread and largely rain-fed. The agricultural land is measured at 5,580,000 hectares, whilst the irrigated area is 29,000 hectares. This, therefore, testifies to how Malawi is dependent on rainfall. Ngongondo et al. (2011) cite late onset of rains, erratic rainfall, long-lasting dry spells, and floods as the major weather hazards that affect rain-fed agriculture. In Malawi, the seasonal rainfall depends primarily on two factors, which include: the position of the Intertropical Convergence Zone, (ITCZ), and the Indian Ocean Sea Surface Temperatures (Lazenby & Todd, 2023). These systems differ in their timing and intensity every year. The main rains are experienced mainly from November in the South and progressively spreading to central and northern areas. Meteorologically, the normal growing season is defined as from November to April (Ngongondo et al., 2011). This puts Malawi in a vulnerable position, being a densely populated low-income country in sub-Saharan Africa, which depends on only one rainy season and therefore only one major harvest per year. It is in this vein that crop insurance is one of several promising interventions for overcoming the negative impacts of climate risk on households' livelihoods (Mahul & Stutley, 2010). Crop insurance covers the risks of expected loss in yield of various crops. However, poor financial performance, with claims and administrative costs exceeding premiums, for example, unsuitable pricing, uncontrolled moral hazard, and adverse selection, are among the key problems underlying crop insurance programs worldwide, in both developed and developing countries.

Consequently, a weather derivative is a new class of weather insurance that has been a promising field of research to cope with weather risks in agricultural production. It is a financial contract with its income depending on the evolution of an underlying meteorological index. Examples of underlying indices include temperature, humidity, precipitation, wind, and so on. Under crop insurance, payouts are based on a client's loss, whilst with index-based insurance, payouts occur when an index falls below a predetermined threshold (Mensah et al., 202). This means that farmers can receive payouts without verifying the losses. Previously, the requirement for verification of loss limited the feasibility of traditional indemnity insurance. The objective nature of the pay-out also means the index-based insurance is more resilient to moral hazard, with the pay-out no longer dependent on the crop, and farmers remain incentivized to ensure its success in difficult conditions (Mensah et al., 202). Skees (2000) also observed that the weather derivatives have received considerable attention in the literature as potential risk management instruments for agricultural production. As such, farmers may buy a weather derivative to hedge against poor yields caused by too much or too little rainfall, sudden temperature changes, or destructive winds (Carmona, 2008).

However, it is difficult to price weather derivatives, due to the incompleteness of weather derivative markets, because weather indices are non-tradable (Li et al., 2020). Additionally, a lack of reliable weather data in sub-Saharan regions may limit the ability of an insurer to set an appropriate threshold for an index-based policy (Li et al., 2020). It is possible for a farmer to experience crop failure despite the threshold for payout not being exceeded (Mensah et al., 2027). This study deal with rainfall as the underlying asset, and there are several challenges in modelling it. Firstly, challenges are associated with the data quality, availability, and the spatial-temporal resolution of rainfall data. Secondly, model accuracy involves the evaluation of the accuracy and reliability of different rainfall models in diverse geographical regions and climatic conditions. However, recent advancements such as the integration of machine learning techniques with traditional models, ensemble forecasting, and big data analytics have been explored (Hussein et al., 2020). Many researchers have also used the hidden Markov model (HMM) to describe the occurrence of rainfall. However, their HMMs could not directly model the seasonal pattern of rainfall. As such, this study derives the seasonal hidden Markov model (SHMM) that would account for seasonal rainfall whilst taking into account the factors such as ITCZ, tropical cyclones, and Congo Air mass, which affect the seasonal pattern of rainfall in Malawi as unobserved states, while the dry state, heavy rainy state, and moderate rainy state as observable states. Later, the study prices the rainfall weather derivatives using the utility indifferent pricing approach. It is built upon the investors' preference towards risks that cannot be eliminated because of market incompetence, commonly applied in pricing the traditional financial derivative market.

Materials And Methods

Mathematical model (Hidden Markov model)

Let K and d be positive integers. The hidden Markov model is described as M(K,d) such that $d = 0,1,2,3,.....n$ days and K for hidden states. Let $(X_t)_{t\geq 1}$ be a first order homogeneous Markov chain with state space (hidden states) $X = \{1,2,....,,K\}$ which include; ITCZ, Congo air mass, tropical cyclones and increase in average daily temperature. Let π be the initial $(Y_t)_{t\geq 1}$ are Y-valued random variables that are independent conditionally on $(X_t)_{t\geq 1}$ and such that for all $j \geq 1$, the conditional distribution of Y_j given $(X_t)_{t\geq 1}$ only depends on X_j, such that it is zero if day t is "dry" and 1 if day t is rainy. so the observation probability $P[Y_t|X_t = k]$ and transition probability (moving from one state to another state) $\pi = P(X_t|X_{t-1})$. We can get the joint probability (probability of all the states from time $t = 1$ to $T-1$ and probability of the observations given the states from time $t' = 1$ to T);

$$p(X,Y) = p(X_1)\Pi_{t=1}^{T-1}P(X_{t+1}|X_t)\Pi_{t'=1}^{T}p(Y_{t'}|X_{t'}) \quad (1)$$

We would now want to infer the hidden state, that is probability in hidden state given all observations, $p(X|Y)$ which would be written as $p(X_t|Y)$ which are defined as $\alpha_t(X_t) = p(X_t, Y_1,......., Y_t)$ (joint probabilities between the current state x_t up until this time).

We may also compute future probability as $\beta_t(X_t) = p(Y_{t+1},......, Y_T|X_t)$, which means the probabilities of future observation given the current state. Combining $\beta_t(X_t)$ and $\alpha_t(X_t)$, we obtained;

$$\alpha_t(X_t)\beta_t(X_t) = p(X_t, Y_1,......., Y_t)p(Y_{t+1},......, Y_T|X_t). \quad (2)$$

$= p(X_t, Y)$.

Normalising the equation (2) to get conditional probabilities;

$$\alpha_t(X_t)\beta_t(X_t) = p(X_t, Y) \propto p(X_t|Y) \quad (3)$$

This means the probability of current state and all the observations is directly proportional to the probability of the current state given all the observations.

Now our goal is on how to compute parameters α_t and β_t, this would be achieved by using forward inference.

Since $\alpha_t(X_t) = p(X_t, Y_1,......., Y_t)$ initializing we have

$p(X_1, Y_1) = p(X_1)p(Y_1|X_1) = \alpha_1(X_1)$.
we have

$$p(X_2, Y_1, Y_2) = \Sigma_{x1} p(X_1, Y_1)p(X_2|X_1)p(Y_2|X_2)$$

$$= \alpha_2(X_2) \quad (4)$$

$$= \Sigma_{x1}\alpha_1(X_1)p(X_2|X_1)p(Y_2|X_2).$$

In general, the joint probabilities for the next time step state and all the evidences including the next time step would be written as the following expression;

$p(X_{t+1}, Y_1, \ldots\ldots, Y_{t+1}) = \alpha_{t+1}(X_{t+1})$
(5)

$=\Sigma_{x\,t}\, \alpha_t\, (X_t)p(X_{t+1}|Xt)p(Y_{t+1}|X_{t+1})$.

Now we want to compute the parameter β_t using the back inference $\beta_t(X_t) = p(Y_{t+1}, \ldots\ldots, Y_T|X_t)$ this refers to the probability of the future evidence. Hence, we can re-write it as

$$p(\{\}|X_T) = 1 = \beta_T(X_T)$$

We also have that

$$p(X_t|Y) = \frac{\alpha_t(X_t)\beta_t(X_t)}{\Sigma_{x_t'}\alpha_t(X_t')\beta_t\,(X_t')}$$

So, the probability of future evidence could be found as

$B_{t-1}(X_{t-1}) = p(Y_t, \ldots\ldots, Y_T|X_{t-1})$

$$= \Sigma_{xt} p(X_t|X_{t-1})p(Y_t, \ldots\ldots, Y_T|X_{t-1}) \quad (6)$$

$= \Sigma_{xt} p(X_t|X_{t-1})p(Y_t|X_t)p(Y_{t+1}, \ldots\ldots Y_T|X_t)$.
This yield

$$B_{t-1}(X_{t-1}) = \Sigma_{xt}\, p(X_t|X_{t-1})p(Y_t|X_t)\beta_t(X_t\,)$$

We can also compute the probability of current and next state given all the observations as;

$$p(X_t, X_{t+1}|Y) = \frac{p(X_t, X_{t+1}, Y_1, \ldots., Y_t, Y_{t+1}, Y_{t+2}, \ldots., Y_T)}{p(Y)} = \frac{p(X_t, Y_1, \ldots, Y_t)p(X_{t+1}|X_t)p(Y_{t+2}, \ldots., Y_T|X_{t+1})p(Y_{t+1}|X_{t+1})}{\Sigma_{X_T} p(X_t, Y)}$$

(7)
Therefore

$$p(X_t, X_{t+1}|Y) = \frac{\alpha_t(X_t)p(X_{t+1}|X_t)\beta_{t+1}(X_{t+1})p(Y_{t+1}|X_{t+1})}{\Sigma_{X_T}\alpha_T(X_T)}$$

(8)
This summarises the forward-backward inferences. In order to avoid underflow, we re-normalize (we avoid as much as we can for parameters to go to zero) at each time step.

$$\widehat{\alpha_t}(X_t) = \frac{\alpha_t(X_t)}{\Sigma_{X_t'}\alpha_t(X_t')}$$

And

$$\hat{\beta}_t(X_t) = \frac{\beta_t(X_t)}{\Sigma_{X'_t}\beta_t(X'_t)}$$

So our observation model is given by

$$p(Y_t|X_t)$$

. and Conditional transition matrix is given by

$$P(X_t|X_{t-1})$$

Since X is hidden, we treat it as latent variable, therefore, we carry out the expectation maximization using Baum-Welch Algorithm. This algorithm estimates the parameters of HMM. It is the expectation-maximization algorithm that re-estimates the hidden parameters by calculating forward, backward and posterior probabilities of sequences in each iteration. We can compute $p(X_t|Y)$ and $p(X_t, X_{t+1}|Y)$ using forward- backward and we can also maximise weighted (expected) log-likelihood as

$$p(X_1) \longleftarrow \frac{1}{T}\Sigma_{t=1}^{T} p(X_t|Y)$$

or
$p(X_1|Y)$
the transition probabilities can be computed as

$$p(X'_{t+1}|X'_t) \longleftarrow \frac{\Sigma_{t=1}^{T-1} p(X_{t+1}, X_t|Y)}{\Sigma_{t=1}^{T-1} p(X_t|Y)}$$

The probability of Y given X can also be computed using multinomial models as

$$p(y|x) \longleftarrow \frac{\Sigma_{t=1}^{T} p(X_t = x|Y)/(y_t=y)}{\Sigma_{t'=1}^{T} p(x_t = x|Y)}$$

The second part of the model is to determine the intensity of rainfall. The study focuses on the statistical modelling. For time series analysis, the dependency precludes the use of traditional statistical tests. The important assumption for statistical testing which is the independence of the error in the data, could usually not be met. ARIMA models have proven especially useful within time series analysis because they provide a basic methodology to model the effects of dependency from the data series (Velicer & Fava, 2003) and allow valid statistical testing. The ARIMA models would account for various patterns, for example, linear or nonlinear trend, constant or varying volatility and seasonal or non-seasonal fluctuations. In short, ARIMA model deals with data that is complex, non-stationary, or has multiple patterns and features.

Our ARIMA model given as

$$Y_t = c + \epsilon_t + \theta_1 y_{t-1} + \beta_1 y_{t-1}.$$

Where y_t is the observable value, c is a constant, ϵ_t is an error function (white noise), β is the constant for auto-regressive model (order 1) and θ_1 is a constant for moving average model (order 1).
So adding another term for comparison for 12 months, we get

$$y_t = c + \epsilon_t + \theta_1 y_{t-1} + \beta y_{t-1} + \beta_{12} y_{t-12}.$$

where β_{12} is constant for autoregressive model for 12 months.

Data source

The study used historical monthly rainfall data for 20 years covering the period from 1995 to 2015 collected from Chitedze Research Station in Lilongwe District and from the Department of Climate Change and Meteorological Services (DCCMS) in the Ministry of Natural Resources, Energy and Environment in Malawi. The main reasons for the choice of Lilongwe District are because of its dominance in cereal production with a humid sub-tropical climate (Msowoya et al., 2016) and highly population. Lilongwe District's predominant soil type is sandy clay loam. Over the past decades, average maize yield per hectare in Lilongwe have nearly equalled the national maize yields, making it suitable region for hedging against the effects of climate change on maize production (Msowoya et al., 2016).

Pricing the rainfall weather derivatives

In this study, we adopt the general method for pricing a derivative contract for the rainfall amount which is given as:

$$F(t;\tau_1,\tau_2) = E^Q[I(\tau_1,\tau_2)|f_t] = E^Q\left|\sum_{t=\tau_1}^{\tau_2} R_t \,|f_t\right| \quad (12)$$

where $F(t;\tau_1,\tau_2)$represents a future contract priced at time point t for a contract period from time point τ_1 until time point τ_2 which is the period of three months after rains start. For this, *it* does not have to be equal to τ_1, because future are priced for a future date $E^Q\left[\Sigma_{t=\tau_1}^{\tau_2} R_t|f_t\right]$ $\tau_1 till\ \tau_2$ is the expected amount of rainfall given the prevailing amount of rainfall today. It should be noted that Q which is the risk-neutral measure does not have anything to do with the physical probability of occurrence of scenarios. Q is a probability measure on the set of scenarios, which is a bet on the occurrence of this event. The rainfall estimates $I(\tau_1, \tau_2)$ are considered the expected price under the physical measure P , but are within the 'risky' world. **Ω** is the sample space, a set of all possible outcomes, *f* is a set of events, where each event is a set containing zero or more outcomes and P the assignment of probabilities to the functions. The study prices a contingent claim written on rainfall which is an underlying asset that is not traded but has great influence on agriculture products. We assumed that the option is static that could not be changed once the buyer and the seller agree on the prices. The option is held for the whole growing season duration regardless of the current weather data being experienced. We generalize the geometric Brownian motion price model by considering the coefficients μ and σ to depend on another process Y. The market model consists of a risky asset Y that is traded and presence of rainfall indexes on which a European option is written. The tradable asset has a price dynamic as follows:

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t \quad (13)$$

$$Y_0 = y \in \mathrm{R} \quad (14)$$

where $Y = \{Y_t\}$ is a continuous Markov process in R^n independent of the Wiener process W defined on a filtered probability space $[\Omega, \Im, (\Im)]_{0\le t \le T}$, P] where $(\Im)$ is σ- algebra ($\Im_t$ is the information available to us at time t) , and $\mu(Y_t)$and $\sigma(Y_t)$ are measurable functions. Such models are common in financial applications, that is regime switching models and stochastic volatility models. The relevant results for our purposes are stated in **proposition 1** and followed by the corollary.

Proposition 1 Suppose that there exists a constant K such that the following conditions are satisfied for all x,y and t

$$\|\mu(t,x_1) - \mu(t,x_2)\| \le K\|x_1 - x_2\| \quad (15)$$

$$\|\sigma(t,x_1) - \sigma(t,x_2)\| \le K\|x_1 - x_2\| \quad (16)$$

$$\|\mu(t,x) + \sigma(t,x)\| \le K(1 + \|x\|) \quad (17)$$

then there exists a unique solution to Equations (13 and 14) that $\Im^w$ is adapted and a Markovian process. The contingent claim for indifference price $W_T\left(\Sigma_{t=\tau_1}^{\tau_2} R_t\right)$ is constructed from three stochastic optimal control problems as follows: The first problem is where the agent maximizes the expected utility wealth without the contingent claim $W_T\left(\Sigma_{t=\tau_1}^{\tau_2} R_t\right)$, commonly known as the classical **Merton model** of investment with the value function defined as

$$V(x,t) = sup\ E[-e^{\gamma X_T}|X_t = x] \quad (18)$$

where γ is the risk aversion of the agent representing his or her attitude towards the risk that cannot be eliminated and $X_t = x$ is the initial wealth. Taking into account the contingent claim $W_T(\Sigma_{t=\tau_1}^{\tau_2} R_t)$, we have the following value functions for the seller and the buyer, respectively,

$$V^s(x,t) = supE\left[-e^{-\gamma[X_T - W_T]}\middle|X_t = x, \Sigma_{t=\tau_1}^{\tau_2} R_t\right] \quad (19)$$

$$V^b(x,t) = supE\left[-e^{-\gamma[X_T + W_T]}\middle|X_t = x, \Sigma_{t=\tau_1}^{\tau_2} R_t\right] \quad (20)$$

Musiela & Zariphopoulou (2004) define the indifference price of the seller of the weather derivative as a function $p^s(x,t)$ such that the investor is indifferent towards the following two cases: optimizing the expected utility without employing the derivative, and optimizing it taking into account on the one hand the liability $W_T(\Sigma_{t=\tau_1}^{\tau_2} R_t, T)$ at expiration T and on thk6te other hand the compensation $p^s(x, t)$ at inscription t. Mathematically,

$$V(x,t) = V^s(x + P^s(x,t),t) \quad (21)$$

and similarly, the buyer's indifference price $p^b(x, t)$ satisfies

$$V(x,t) = V^b(x - P^b(x, t),t) \quad (22)$$

Definition A process π in the form $\pi_t = a(t, X(t))$ for measurable function.

$$a: [0,T] \times R^n \rightarrow R^n$$

is called *Markovian.*

However, optimization of the value function in equations 15 - 18 to Markovian admissible controls leads to deriving the Hamilton-Jacobi-Bellman (HJB) equation. There is a vast array of literature on the specification of V and on the study of value functions in complete market settings. For value functions of problems in incomplete markets, like the optimization, martingale methods seem to produce limited results whilst the analysis via the HJB equation appears more appealing. Therefore, HJB equation for the classical Mertion problem value function V is given as;

$$V_t + max_\pi\left(\mu\pi V_x + \frac{1}{2}\sigma^2\pi^2 V_{xx}\right) = 0$$

$$V(x,T) = -e^{-\gamma x}$$

Proposition 2: Without weather hedging, the optimal value function and control are

$$V(x,t) = e^{\left\{-\gamma x + \frac{1\mu^2}{2\sigma^2}(T-t)\right\}}$$

$$\pi^* = \frac{\mu}{\sigma^2\gamma}$$

Proposition 3: The seller's utility indifference price is

$$p^s = \frac{1}{\gamma_s} In\left(E_P\left[e^{\gamma_s W_T(R)}\middle|X_t = x, \Sigma_{\tau=\tau_1}^{\tau_2} R_t\right]\right)$$

The utility indifference prices are clearly independent of initial wealth x and is nonlinear.

Basket options for KIA

We consider defining *rainfall defice*, denoted as RD, as the number of millimetres by which the average rainfall X_t is below the base rainfall X_{ref} based on the literature (Kimaro & Sibande, 2008) we take it as $600mm$, that is,

$$f(X_t) = (600 - X_t)^+ \tag{26}$$

We define rainfall excess, denoted as RE, as the number of millimetres by which the average rainfall X_t is above the base rainfall X_{ref}, that is,

$$f(X_t) = (X_t - 600)^+ \tag{27}$$

An investor can protect himself against low level of rains by taking a position on RD contracts and payoff will be calculated according to the (26). In case of higher levels of rain, an investor can protect himself by taking a position on RE contracts and the payment has a payoff defined according to the equation (27). Consequently, according to the forecast rainfall data, the investors must take position on both RD and RE contracts. The rainfall risk can be managed by buying RD or RE 278 (American, Asian or European) options, taking short or long positions. In this regard our rainfall options are defined as;

$$W(x, y_T, T) = tick \times (K - y_T)^+, \quad or \quad W(x, y_T, T) = tick \times (y_T - K)^+ \tag{28}$$

where y_T is the value of RD or RE index at maturity, K is the strike level and where tick is a pre-agreed upon constant factor that determines the amount of payment per unit of weather index (that is per millimetre of the RE or RD). The grain producer is (partially) insured against revenue losses due to little precipitation in the growing season. However, we need to determine the premium based on the rainfall index which would determine the pay-off in case of erratic rains. Alaton et al., (2002) suggest the price of the premium that we adopt in the study is calculated from the expected payoff as;

$$c = (1 + R)e^{-r(T-S)W_T} \tag{29}$$

where c is the premium that the hedgers (buyers) need to pay for contract, r is a risk- free periodic interest rate, S is the date that the contract is issued (purchased) and T is the date the contract is claimed or the expiration date.

Results And Discussion

4.1 Rainfall Occurrence by HMM

The state matrix approaches the steady state matrix (0.396 0.604) regardless of initial state matrix values. From the steady state matrix (0.396 0.604), there is likelihood that rainfall pattern in KIA would follow the pattern of the given probabilities.

$$P = \begin{bmatrix} 0.713 & 0.287 \\ 0.188 & 0.812 \end{bmatrix}$$

$$P^4 = \begin{bmatrix} 0.4416908 & 0.5583092 \\ 0.3657217 & 0.6342783 \end{bmatrix}$$

$$P^6 = \begin{bmatrix} 0.408441 & 0.591559 \\ 0.387502 & 0.612498 \end{bmatrix}$$

$$P^{12} = \begin{bmatrix} 0.396 & 0.604 \\ 0.396 & 0.604 \end{bmatrix}$$

Rainfall intensity

The **Figure** 1 shows that there are harmonically fluctuations on the amount of precipitation for some years. Evidently, there has been a decline in the amount of rainfall for years 2012 to 2015. One of the contributing factors is the location of Malawi. Since Malawi is located between two regions of ENSO episodes, it is often not easy to predict the influences of ENSO on the climate of Malawi. The influence of ENSO events has a significant influence on the climate of Malawi. The lake also influences rainfall in areas along the lakeshore

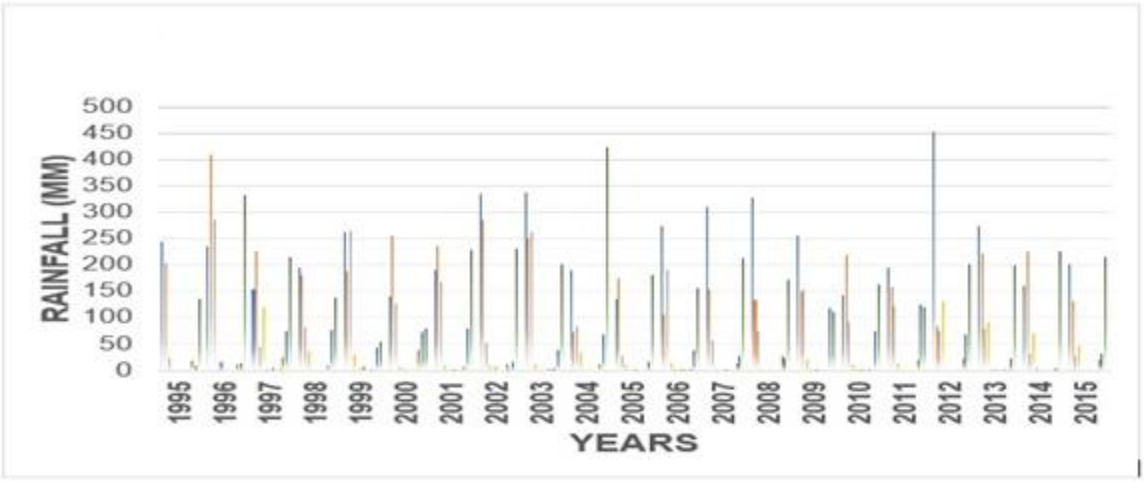


Figure 1: Monthly rainfall time series plot of KIA

The **Figure 2** shows the standardized values of rainfall for the rainy season of historical data.

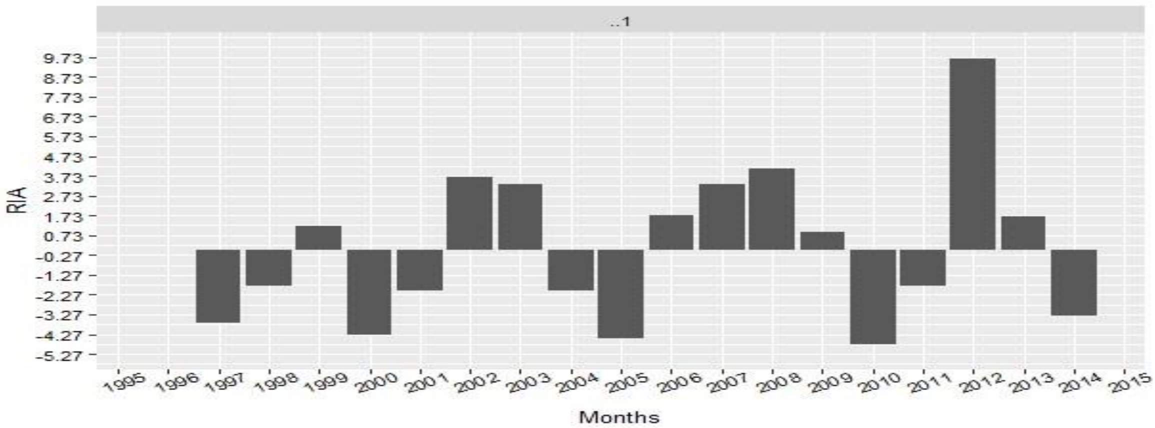


Figure 2: The Rainfall Anomaly Index for historical data from Kamuzu International Airport (KIA) in Lilongwe from 1995 – 2015

In figure 2, the RAI ranges from ≥ 9.73(extreme wet) to ≤ −4.27 (extreme dry) which are used to determine extreme conditions of rainfall at KIA in Lilongwe. However, these extreme conditions are not favourable for rain-fed tropical agricultural production because extreme wetness is associated with flooding that destroys plants while extreme dry conditions are associated with severe water deficits that cannot support maize production.

Time Series plots for historical rainfall data from Kamuzu International Airport (KIA)

The figure 3 shows the time series plot for an observed data from 1995 to 2015.

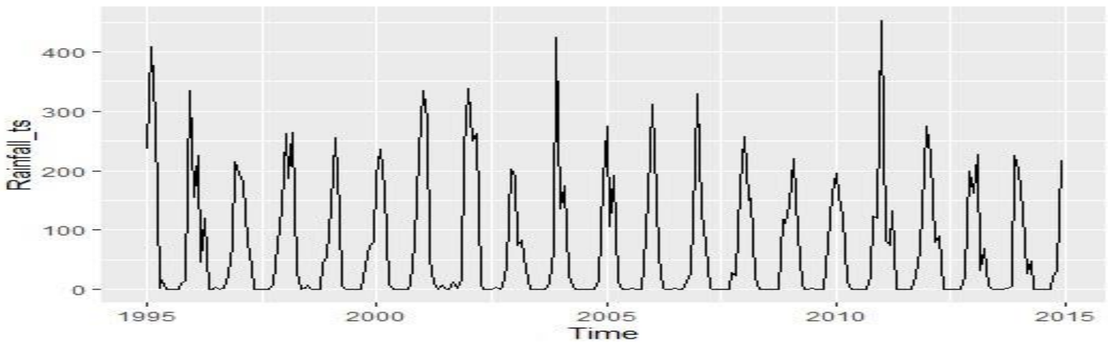


Figure 3: Time series plot for the observed yearly rainfall data from 1995 to 2015

The time series plot (figure 3) reveals that there is no cyclical component either as data does not display rises or falls around trend levels. As a matter of fact, the stationarity in variance and mean is a requirement for a time series data before an ARIMA model is fitted on the data. The method to test for stationarity is by computing the autocorrelation function (ACF). The autocorrelation function is generally expressed as lag k, to the variance. At lag k, p_k denotes the ACF and is defined as follows;

$$p_k = \frac{y_k}{y_0}$$

where y_k is the covariance at lag k and y_0 is the variance.

Therefore**, figure 4** shows most of the lags in the autocorrelation function (ACF) and partial autocorrelation function (PACF) are significant and therefore the time series data is non-stationary (this shows that the data has trend and seasonality).

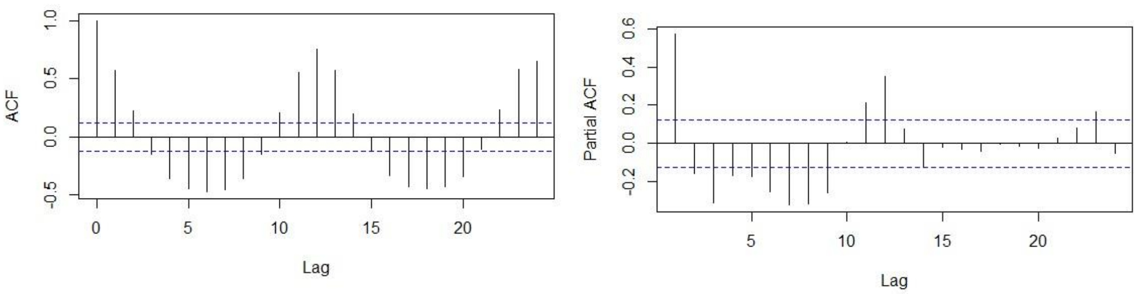
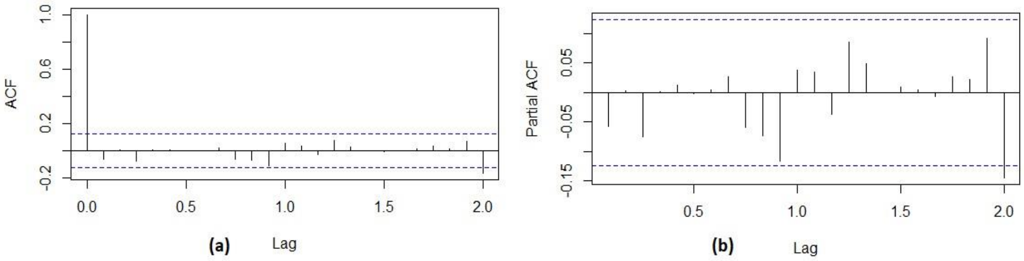


Figure 4: autocorrelation function (ACF) and partial autocorrelation function (PACF) for the rainfall data

After removing the trend and seasonality rainfall data, the plots for the ACF and PACF correlograms in figure 5 (a) and 5(b) resembles white nose which shows that the time series is stationary. A significant spike at the first lag in ACF, combined with the exponential decline as lags increase in the PACF, suggest a MA (1). To conform stationary de-seasonalized rainfall data, formal tests for example, Augment Dickey-Fuller (ADF), Phillips-Perron (PP) and Kwiatkowski-Philips- Schmidt-Shin (KPSS) can be employed.



The Best fitted autoregressive integrated moving average (ARIMA) Model

The best fitted ARIMA model for the deseasonalised data according to R-studio, is ARIMA (1,0,1)(2,0,1)$_{12}$. Table 1 shows the estimate results of ARIMA model chosen. The model has three auto-regressive components, one moving average component and an integrated of zero order. By default, step-wise selection is used in R for the model search. The models are selected on Akaike information criterion (AIC) by default (See table 1) but approximation could also be used for the models tried in the model selection step.

Table 1: The parameter estimates of Seasonal ARIMA model (1,0,1)(2,0,1)$_{12}$

	ar1	aa1	Sar1	sar2	sma1	intercept	sigma^2	Log likelihood	aic
	-0.380	0.313	-0.567	-0.145	0.678	69.064	2123	-1323.9	2 2661.75
s.e	1.689	1.7441	0.1760	0.0832	0.1696	2.7201			

Table 2: Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.171862	46.07933	24.92925	8.089056	0.7097126	-0.005586848

The ARIMA Model (1,0,1)(2,0,1)$_{12}$ model is given as

$$y_t - y_t{-}1 - y_t{-}12 + y_t{-}13 = c + \epsilon_t + \beta\epsilon_t{-}1 + \Theta\epsilon_t{-}12 + \beta\Theta\epsilon_t{-}13$$

$$y_t = 69.0643 - 0.3804y_{t-1} + 0.3127\epsilon_{t-1} - 0.5665s_{t-1} + -0.1449s_{t-2} + 0.6775z_{t-1} + \varepsilon_t$$

where the ε_t is white noise with standard deviation of $\sqrt{2123} = 46.08$.

To validate the model, we employed back forecast time series against the observed data and figure 6 is plotted.

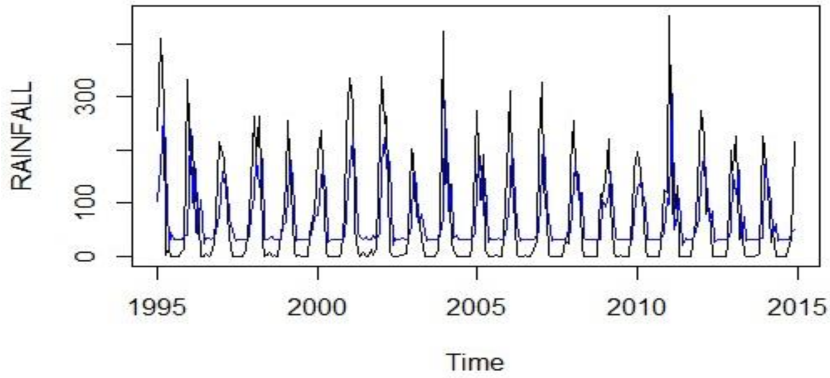


Figure 6: ARIMA fitting plots

Evidence from **Figure 6** shows that the chosen ARIMA model is able to back forecast the data which resembles the observed data which implies that our model is validated.

Rainfall forecasting

Figure 7 shows the forecast result of monthly rainfall data from the fitted ARIMA $(1,0,1)(2,0,1)_{12}$ model from 2015 to 2027 for KIA in Lilongwe.

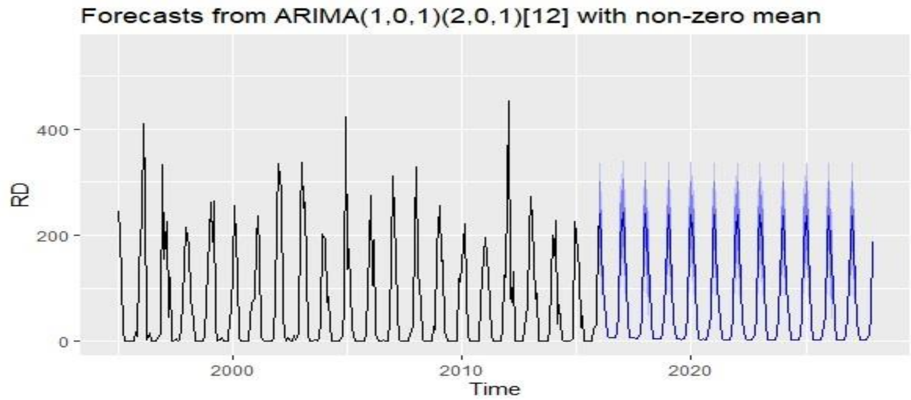


Figure 7: Forecast rainfall results from ARIMA model for Kamuzu International Airport (KIA)

Apparently, with 95% confidence limits from 2015 to December 2027 using ARIMA $(1,0,1)(2,0,1)_{12}$, the forecast plot in figure 9, amount of rainfall in KIA in Lilongwe seems to fluctuates with time. As matter of fact, our simulated data agrees with the actual data. The table 3 shows the simulated rainfall data from 2016 to 2027. It also gives the values of rainfall deficit (RD) and rainfall excess (RE).

Table 3: The results of seasonal amount of rainfall, RE and RD for the forecast rainfall data for 2016 - 2027

YEAR	AMOUNT OF RAINFALL (mm)	RE	RD
2016- 2017	622.662865	22.662865	-
2017-2018	614.91452	14.91452	-

2018-2019	607.6956497	7.6956497	-
2019-2020	611.613282	11.613282	-
2020-2021	610.8775598	10.87755982	-
2021-2022	610.3076209	10.3076209	-
2022-2023	610.3708239	10.3708239	-
2023-2024	610.5575683	10.5575683	-
2024-2025	610.5107952	10.51079522	-
2025- 2026	610.5107952	10.51079522	-
2026-2027	610.5106307	10.5106307	-

Pricing weather derivative results

In spite of the RE or RD calculated above, the contingent claims depend on tick size chosen. The tick size varies based on the asset being traded. For example, in the US stock market, the tick size for stocks priced above 1 USD is 0.01 USD. Taušer & Cajka, (2014) presented the tick size of **£**1 per index point and whilst Xu et al (2006) used 20 USD per monthly index. In addition, Cafiero and Cioffi in 2004 at proceedings of the 86th EAAE seminar, presented the paper under income stabilization in agriculture and the tick size used was £8 per index point. However, tick size requirements are set by stock exchanges and regulatory bodies. Since in Malawi, we do not have set tick size for rainfall index, in this study we adopt a tick size of 20USD.

However, the contingent claim (W) would determine the seller's and buyer's indifference price at a given risk aversion (γ). Utility indifference prices of the buyer and seller were simulated and assumed that the option has a relative risk aversion parameter of $\gamma = 0.1$ because in most cases farmers are risk averse and they would not be willing to take huge risks. The finding in table 4, typically, shows the price buyer would be willing to pay that reflects their risk aversion and expected utility from the derivative.

Table 4: Utility indifference prices for the rainfall derivatives.

Year	Risk aversion	Seller	Buyer
2016- 2017	0.1	453.2	453.2
2017-2018	0.1	298.2	298.2
2018-2019	0.1	153.9	153.9
2019-2020	0.1	232.2	232.2
2020-2021	0.1	217.4	217.4
2021-2022	0.1	206	206
2022-2023	0.1	207.4	207.4
2023-2024	0.1	211	211
2024-2025	0.1	210.2	210.2
2025- 2026	0.1	210.2	210.2
2026-2027	0.1	210.2	210.2

In an incomplete market, it is generally unlikely for a seller's indifference price to be exactly equal to the buyer's indifference price. This is so because indifference prices are influenced by individual risk preferences and the specific risks each party faces. In an incomplete market, not all risks can be perfectly hedged, leading to different valuations based on each party's risk aversion and the specific risks they are exposed to (Musiela & Zariphopouloua, 2004). However, under certain conditions, such as when both parties have similar risk preferences and face similar risks, the indifference prices could be close to each other (Musiela & Zariphopouloua, 2004).

Discussion Of Findings

The results of the forecast in this study reveal the variability of rainfall at KIA using rainfall data from 1995 to 2015. Although some authors argue that including longer weather observations, spanning 20 to 40 years, is more accurate as it captures more variation (Dischel, 2002), other scholars disagree. Stulec et al. (2016) believed that, given the pronounced climate change, a more recent weather pattern may be a better representation of the expected future weather, implying that a shorter weather observation would provide more credible information. The variations of rainfall were also observed by Kumbuyo et al. (2015), who studied the fluctuation of rainfall time series in Malawi, for the data source from 1981 to 2011. Their results showed that the annual average rainfall varies from 725mm to 2,500 mm. McSweeney et al. (2012) point out that the factors that affect seasonal rainfall in Malawi depend on the position of the Intertropical Convergence Zone (ITCZ), which can vary in its timing and intensity from year to year.
Furthermore, Jury & Mwafulirwa (2002) add that the other main rain-bearing system for Malawi is the northwest monsoon, comprised of tropical Atlantic air that reaches Malawi through the Congo basin. There are times when the country is affected by tropical cyclones from the West Indian Ocean (Jury & Mwafulirwa, 2002). Since the economy of Malawi is largely dependent on agriculture, these climatic factors affect its economy. Therefore, a detailed knowledge of the precipitation regime in Malawi is an important prerequisite for agricultural planning and economic development.

It is therefore important for farmers in Malawi to protect themselves against weather shocks. The study considers the adoption of rainfall weather derivatives, which is a promising field of research to cope with weather risks in agricultural production by developing a pricing model that can be used in the agricultural industry. 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 (Dzupire et al., 2019). The study adopted the utility indifference pricing approach, where the investor's risk preference towards the weather risks that cannot be eliminated is described by the exponential utility function. Reasonably, the approach takes into account price risk, weather risks, and all other risks in the financial capital market (Dzupire et al., 2019). However, the model developed can be refined by either choosing a different utility function, such as power utility, which would account for a different risk preference by the investor, or a quadratic utility function.

Furthermore, the World Food Program (WFP) and Malawi's National Association of Smallholder Farmers (NASFAM) developed an index-based crop insurance contract, which was more efficient and cost-effective than traditional crop insurance. The contract could easily be distributed to individual smallholder farmers to increase their access to finance and to protect farmers and loan providers from weather risk. The contract was designed to provide compensation when rainfall, during a crop growing cycle, was insufficient for farmers to grow and to optimize their yields. However, this study questions the appropriateness of the index-based crop insurance contract that the WFP used since the contracts only worked in the event of low rainfall. As such, this study proposes basket rainfall options. This option works in such a manner that the farmer may receive the payoff in case of low and high rainfall.

Conclusions

The steady state matrix from hidden Markov model (HMM) shows likelihood that rainfall pattern at Kamuzu International Airport (KIA) would follow the pattern of the given probabilities (0.396 0.604) which implies that the probability that it would rain on a particular day in future at KIA is 0.369 and the probability that it would be dry on a particular day is 0.604. A stochastic model for characterizing the properties of rainfall is proposed, where the accumulated rainfall is modelled as an ARIMA process. Therefore, the study used the results from ARIMA model to derive rainfall derivative prices that would mitigate the impending climate shocks, thereby reducing poverty and vulnerability of farmers. As such, the study proposes that the farmer should buy the basket options. This strategy ensures that the farmer benefits from any rainfall beyond or below the normal pattern and provides an effective means of managing weather-related risks. Furthermore, the study also recommends the sensitisation of potential investors on the importance of weather market in Malawi.

References

BALI S (2012). Weather derivatives and pricing approaches hava (iklim) türevleri ve fiyatlama yaklas¸imlari. *Atatürk Üniversitesi Iktisadi ve˙ Idari Bilimler Dergisi; Cilt 26,˙ Sayı 1 (2012): ATATÜRK ÜNIVERS˙ ITES˙ I˙ IKT˙ ISAD˙ I VE˙ IDAR˙ I B˙ IL˙ IMLER DER-˙*
GIS˙ I˙.
Bartkowiak M (2009). Weather derivatives. *Mathematics in Economics*, 5–17.

Benth FE, Benth J (2007). The volatility of temperature and pricing of weather derivatives. *Quantitative Finance*, *7*(5), 553–561.
Benth FE., Karlsen KH (2005). A note on merton's portfolio selection problem for the schwartz mean-reversion model. *Stochastic analysis and applications*, *23*(4), 687–704.
Benth FE, Šaltyte-Benth J (2005). Stochastic modelling of temperature variations˙ with a view towards weather derivatives. *Applied mathematical finance*, *12*(1), 53– 85.
Berhane T, Shibabaw N, Kebede T (2020). Pricing weather derivative using markov chain analogue year daily rainfall model. *SN Applied Sciences*, *2*, 1–17.
Björk T (2009). *Arbitrage theory in continuous time*. Oxford university press.

Brockett PL, Goldens LL, Wen, MM, Yang CC (2009). Pricing weather derivatives using the indifference pricing approach. *North American Actuarial Journal*, *13*(3), 303–315.
Brunel V (2002). Super-replication problem in a jumping financial market. *FINANCE-*

PARIS-, *23*, 29–46.

Cafiero C et al (2006). Income stabilization in agriculture. the role of public policies. In *Proceedings of the 86th eaae seminar «agricultural income stabilization: what role should public policies play* (pp. 21–22).
Cao M, Li A, Wei J Z (2004). Precipitation modeling and contract valuation. *The*

Journal of Alternative Investments, *7*(2), 93–99.

Chen TY, Mei XD, Pan JS, Sun SH (2004). Optimization of hmm by the tabu search algorithm. *J. Inf. Sci. Eng.*, *20*(5), 949–957.
Chu PS, Chen YR., Schroeder TA (2010). Changes in precipitation extremes in the hawaiian islands in a warming climate. *Journal of Climate*, *23*(18), 4881–4900.
Cochand F, Brunner P, Hunkeler D, Rössler O, Holzkämper A (2021). Crosssphere modelling to evaluate impacts of climate and land management changes on groundwater resources. *Science of the Total Environment*, *798*, 148759.
Dischel RS (2002). Climate risk and the weather market: financial risk management with weather hedges. *(No Title)*.
Durbin R, Eddy SR, Krogh A, Mitchison G (1998). *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge university press.
Dzupire NC, Ngare P, Odongo L (2019). Pricing basket weather derivatives on rainfall and temperature processes. *International Journal of Financial Studies*, *7*(3), 35.
González-Hidalgo J, De Luis M, Raventós J, Sánchez J (2001). Spatial distribution of seasonal rainfall trends in a western mediterranean area. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *21*(7), 843–860.

Gül B, Kayaalp N (2023). Estimating idf curves under changing climate conditions for different climate regions. *Journal of Water and Climate Change*, *14*(12), 4527–4543.
Hajani E, Sarma G (2023). Generation of rainfall data series by using the markov chain model in three selected sites in the kurdistan region, iraq. *AI in Civil Engineering*, *2*(1), 5.
Harrison L, Michaelsen J, Funk C, Husak G (2011). Effects of temperature changes on maize production in mozambique. *Climate Research*, *46*(3), 211–222.
Hayaty N, Kurniawan H, Rathomi MR, Chahyadi F, Bettiza M (2023). Rainfall prediction with support vector machines: A case study in tanjungpinang city, indonesia. In *Bio web of conferences* (Vol. 70, p. 01003).

Henderson V, Hobson D (2004). Utility indifference pricing-an overview. *Volume on*

Indifference Pricing.

Hussein EA, Ghaziasgar M, Thron C, Vaccari M, Jafta Y (2022). Rainfall prediction using machine learning models: literature survey. *Artificial Intelligence for Data Science in Theory and Practice*, 75–108.
Jewson S, Brix A (2005). *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*. Cambridge University Press.
Jury M, Mwafulirwa N (2002). Climate variability in malawi, part 1: dry summers, statistical associations and predictability. *International Journal of Climatology: A*
Journal of the Royal Meteorological Society, *22*(11), 1289–1302. Karlsson P (2009). Fx basket options-approximation
and smile prices.
Kaul M, Hill RL, Walthall C (2005). Artificial neural networks for corn and soybean yield prediction. *Agricultural Systems*, *85*(1), 1–18.
Kawaye FP, Hutchinson MF (2018). Are increases in maize production in malawi due to favourable climate or the farm input subsidy program (fisp)? *Theory and practice of climate adaptation*, 375–390.
Kumbuyo CP, Yasuda H, Kitamura Y, Shimizu K (2014). Fluctuation of rainfall time series in malawi: An analysis of selected areas. *Geofizika*, *31*(1), 13–28.
Lazenby MJ, Todd MC (2023). Evaluating future changes in the south indian ocean convergence zone projected by cmip5 models and associated uncertainty. *Climate Research*, *91*, 191–209.

Li P, Lu X, Zhu SP (2020). Pricing weather derivatives with the market price of risk extracted from the utility indifference valuation. *Computers & Mathematics with Applications*, *79*(12), 3394–3409.

Mahul O, Stutley CJ (2010). *Government support to agricultural insurance: challenges and options for developing countries*. World Bank Publications.

Martin SW, Barnett BJ, Coble KH. (2001). Developing and pricing precipitation insurance. *Journal of Agricultural and Resource Economics*, 261–274.

Martínez SA, Tarrazon RMA. (2020). Approaching rainfall-based weather derivatives pricing and operational challenges. *Review of Derivatives Research*, *23*(2), 163–190.

Mattfeld C (2014). Implementing spectral methods for hidden markov models with realvalued emissions. *arXiv preprint arXiv:1404.7472*.

Mbano CJ (2009). Impacts of rainfall and forest cover change on runoff in small catchments: a case study of mulunguzi and namadzi catchment areas in southern malawi. *alawi Journal of Science and Technology, 9(1), 11-17*.

Mensah NO, Owusu SE, Adjei C (2023). Revisiting preferences for agricultural insurance policies: Insights from cashew crop insurance development in ghana.
Food Policy, *118*, 102496.

Mohammed Y, Adamu L (2021). A stochastic model for rice yields forecast.

Morishima W, Akasaka I (2010). Seasonal trends of rainfall and surface temperature over southern africa. *African study monographs. Supplementary issue.*, *40*, 67–76.

Msowoya K, Kondwani M, Lund JR (2016). Climate change impacts on maize production in the warm heart of africa. *Water Resources Management 30 (2016): 5299-5312*.

Musiela M, Zariphopoulou T (2004). A valuation algorithm for indifference prices in incomplete markets. *Finance and Stochastics*, *8*, 399–414.

Ngongondo C, Xu CY, Gottschalk L, Alemaw B (2011). Evaluation of spatial and temporal characteristics of rainfall in malawi: a case of data scarce region. *Theoretical and applied climatology*, *106*, 79–93.

Nhangumbe C, Sousa E (2024). Numerical solutions of an option pricing rainfall weather derivatives model. *Computers & Mathematics with Applications*, *153*, 43– 55.

Odening M, Mußhoff O, Xu W (2007). Analysis of rainfall derivatives using daily precipitation models: Opportunities and pitfalls. *Agricultural Finance Review*, *67*(1), 135.

Okereke O, Bernard C (2014). Forecasting gross domestic product in nigeria using box-jenkins methodology. *Journal of Statistical and Econometric methods*, *3*(4), 33– 46.

Ould CA, Yasuda H, Hattori K, Nagasawa R (2008). Analysis of rainfall records (1923–2004) in atar-mauritania. *Geofizika*, *25*(1), 53–64.

Reddy GS, Bhaskar S, Purohit R, Chittora A (2008). Markov chain model probability of dry, wet weeks and statistical analysis of weekly rainfall for agricultural planning at bangalore.

Rondhi M, Pratiwi PA, Handini VT, Sunartomo AF, Budiman SA (2018). Agricultural land conversion, land economic value, and sustainable agriculture: A case study in east java, indonesia. *Land*, *7*(4), 148.

Senthilvelan A, Ganesh A, Banukumar K (2012). Markov chain model for probability of weekly rainfall in orathanadu taluk, thanjavur district, tamil nadu. *International Journal of Geomatics and Geosciences*, *3*(1), 191–203.

Skees JR (2000). A role for capital markets in natural disasters: a piece of the food security puzzle. *Food policy*, *25*(3), 365–378.

Solomon S, Qin D, Manning MR, Chen Z, Marquis M, Averyt K, Miller HL (2007). Climate change 2007: the physical science basis. contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change. summary for policymakers.. Retrieved from https://api.semanticscholar.org/CorpusID:265858649

Stoner O, Economou T (2019). A hierarchical modelling framework for correcting delayed reporting in spatio-temporal disease surveillance data. *arXiv*.

Stoner O, Economou T (2020). An advanced hidden markov model for hourly rainfall time series. *Computational Statistics & Data Analysis*, *152*, 107045.

Stulec I, Petljak K, Bakovic T (2016). Effectiveness of weather derivatives as a hedge against the weather risk in agriculture. *Agricultural Economics/Zemeˇdeˇlská Ekonomika*, *62*(8).

Tang W, Chang S (2016). A semi-lagrangian method for the weather options of mean-reverting brownian motion with jump–diffusion. *Computers & Mathematics with Applications*, *71*(5), 1045–1058.

Taušer J, Cajka R (2014). Weather derivatives and hedging the weather risks.ˇ *Agricultural Economics/Zemedelska Ekonomika*, *60*(7).

Thulu FGD, Katengeza EW, Mkandawire M (2017). Rainfall trends for el niño seasons over malawi from 1970 to 2016 and its impact on crop yield and hydropower generation. *Int. J. Sci. Res. Publ*, *7*(12), 96–104.

Touron A (2017). Modeling rainfalls using a seasonal hidden markov model. *arXiv preprint arXiv:1710.08112*.

Wang H, Asefa T, Sarkar A (2021). A novel non-homogeneous hidden markov model for simulating and predicting monthly rainfall. *Theoretical and Applied Climatology*, *143*(1), 627–638.

Xu W, Odening M, Mußhoff O (2007). Indifference pricing of weather insurance.