

# Pricing Gamma based Temperature Derivatives

Kondwani Daniel Vwalika\*, Nelson Dzupire

Department of Mathematical Science, University of Malawi, Zomba, Malawi

## ABSTRACT

Farmers are impacted by temperature as high temperatures during the rainy season can lead to a substantial decrease in crop production. To safeguard farmers from this risk, temperature derivatives can be used, but they are frequently mispriced. This study aims to address this issue by developing a Stochastic Differential Equation (SDE) for temperature, with the assumption that it conforms to a gamma distribution. A synthesis technique that effectively manages the auto correlation within the data is employed to deduce the SDE. The resulting pricing formula is based on the anticipated value derived from the SDE. Notably, the formulated equation's outcome is not linked to the expected temperature itself, but rather hinges on the gamma distribution parameters and the trigger temperature. This approach yields accurate forecasts for both price predictions and temperature projections. The model is found to predict temperature with  $R^2=91\%$ ,  $MSE=0.14$  and  $MAPE=1.3\%$ . When used to price call option, the prices decrease with increase in trigger value, which is more realistic. Thus, the model is more flexible.

**Keywords:** Gamma derivatives; Temperature; Synthesis method; Trigger temperature

## INTRODUCTION

Temperature plays a vital role in agricultural growth and production, impacting crop development, seed germination, flowering and fruit ripening. Extreme temperatures, whether too high or too low, can adversely affect crop yields and quality, causing heat stress, reduced photosynthesis, water loss, slowed growth, frost damage and diminished pest control effectiveness. To mitigate temperature-related risks, various approaches are employed, including irrigation, breeding for heat and drought tolerance and the use of shade nets and greenhouses. An emerging and popular method for risk management is the use of weather derivatives financial contracts linked to specific weather events or indices. These derivatives offer advantages over traditional insurance, being more cost-effective, flexible and transparent, based on objective weather data. The weather derivatives market has seen significant growth since its inception in 1996, primarily driven by the energy sector in developed countries. In developing countries, weather derivatives are gradually gaining traction, especially in agricultural and renewable energy sectors. However, challenges remain, such as the lack of standardized weather data and regulatory barriers. Call options are more commonly used in developing countries

for weather hedging due to their simplicity and accessibility. The study aims to price weather derivatives using data from Chitedze research station, providing coverage for farmers against unfavorable temperature conditions and ensuring payout security.

The majority of authors who employ the Ornstein-Uhlenbeck (OU) model assume that the noise follows Gaussian Wiener, Levy or fractional Brownian motion, as noted by Dzupire et al. Using the first two processes implies independent temperature changes, but authors note correlations in changes. Temperature reverts to seasonal trend, confirming the argument. Mean-reverting parameter captures this behavior. Certain researchers utilize fractional Brownian motion as an alternative to assuming independence. However, it remains uncertain whether the issue is completely resolved.

It is reasonable to assume independence in the random noise process while still retaining the characteristic of reverting to the regular trend. Nonetheless, it is justifiable to consider temperature as Markovian, implying that the mean-reversion parameter should encompass autocorrelation. The authors highlight that this parameter reflects the rate at which temperature returns to its seasonal average. The question not

**Correspondence to:** Kondwani Daniel Vwalika, Department of Mathematical Science, University of Malawi, Zomba, Malawi; E-mail: kondwanivwalika@gmail.com

**Received:** 06-Mar-2024, Manuscript No. ME-24-30034; **Editor assigned:** 08-Mar-2024, PreQC No. ME-24-30034 (PQ); **Reviewed:** 22-Mar-2024, QC No. ME-24-30034; **Revised:** 04-Mar-2025, Manuscript No. ME-24-30034 (R); **Published:** 11-Mar-2025, DOI: 10.35248/1314-3344.25.15.244

**Citation:** Vwalika KD, Dzupire N (2025) Pricing Gamma based Temperature Derivatives. Math Eter. 15:244.

**Copyright:** © 2025 Vwalika KD, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

answered by many authors is how to ensure that this parameter also captures information about autocorrelation. Additionally, using the Ornstein-Uhlenbeck (OU) process implies the underlying probability distribution of temperature follows a Gaussian family, like the normal or inverse one.

The Kolmogorov forward equation (Fokker-Planck equation) can be satisfied to acquire the particular probability distribution resulting from the fitting of the OU differential equation. Satisfying these additional partial differential equations addresses nonlinearity and non-stationarity problems, emphasizing the Gaussian family probability distribution. Nonetheless, historical temperature datasets frequently exhibit deviations from a normal distribution. This discrepancy has prompted authors to utilize the normal inverse Gaussian distribution in order to depict data that displays skewness and possesses heavy-tailed characteristics [1].

The gamma distribution, a widely used probability distribution for modeling weather variables, such as rainfall, evaporation and temperature, has two parameters (shape and scale) that can be estimated from historical temperature data. Modeling temperature using the gamma distribution helps estimate the likelihood of extreme temperature events and corresponding financial losses, which is relevant for pricing temperature derivatives like call options. While the gamma distribution has been used to model temperature, its application to weather derivatives remains less explored.

Several studies have used the gamma distribution as a flexible and accurate model for temperature. For instance, Andrews and Shivamoggi [2], found the gamma distribution provided a good fit to temperature data and effectively estimated temperature, while Moriarty et al. [3], found it suitable for estimating expected temperature in the future. Haddad [4], shows the gamma distribution is appropriate for daily temperature readings, accommodating both maximum and minimum temperatures and non-negative values. However, the literature does not address how to ensure an associated stochastic differential equation satisfies the Fokker Planck equation and Markovian auto-correlation. Stochastic Differential Equations (SDEs) are widely used mathematical models in various fields such as physics, finance and engineering, derived through methods dependent on stochastic calculus type or moments.

Three techniques for deriving diffusion time evolution are commonly utilised. The first is Ito calculus, based on the Ito formula, which defines SDEs using underlying physical or economic processes and is widely used in finance and economics. The second is the Stratonovich calculus, similar to Ito calculus but with a different interpretation of drift and diffusion terms, often applied in physics and engineering. The third is the Fokker-Planck equation, a partial differential equation describing the evolution of the probability density function of a stochastic process, useful for studying long-term behavior. However, converting partial differential equations into SDEs can result in information loss. Another approach is the method of moments, which involves finding moments like mean and covariance from SDEs, derived based on physical intuition, data analysis and domain knowledge. However, this method

relies on assumptions of stationary moments and independence, which may not be valid in some cases.

The Kalman filter [5], is a widely used recursive algorithm in stochastic control theory and engineering that estimates the system state based on measurements and a model. However, its numeric drift and diffusion coefficients make it challenging to incorporate temperature data's seasonality. To address this, the Bayesian estimation method utilizes Bayes' theorem to update probabilities of system states, commonly employed in decision making and machine learning.

Nevertheless, recurrent neural networks require large sample sizes for accurate seasonality capture. Existing methods fail to ensure that differential equations capture nonlinearity, nonstationarity and auto-correlation in data. To address this, Primak et al., introduced a synthesis method for deriving Ito's stochastic differential equations, satisfying Fokker-Planck equation and auto-correlation intervals, commonly used in telecommunication and non-seasonal data. This study intends to use the method to derive stochastic differential equations for seasonal temperature data.

Weather derivative pricing commonly involves Esscher transforms, stochastic optimization and discounting expected payoff [6]. The reflection method, constructing Brownian motion from temperature and reflecting it to estimate prices, is difficult to incorporate temperature readings in Brownian motion noise.

Esscher transform uses probability density to price derivatives but may lose time-dependent parameters. Geman and Yor [7], proposed a Brownian motion-based equilibrium approach for pricing temperature derivatives, extending the model with stochastic volatility for heating degree day options using Monte Carlo simulation. Scher and Messori examined temperature forecast uncertainty's impact on weather derivative pricing, proposing a model incorporating both forecast uncertainty and temperature volatility, although lacking the most accurate pricing. Hess [8], introduced a forward-looking weather model for pricing rainfall derivatives, showing improved accuracy over models ignoring rainfall-temperature correlation.

Ahčan [9], proposed a new model using the variance gamma process for pricing temperature derivatives, capable of pricing both heating and cooling degree day options but with weak estimations. Another early model is the Seasonal Autoregressive Integrated Moving Average (SARIMA) introduced by Liu et al. [10], effective for pricing simple temperature derivatives like Heating Degree Day (HDD) and Cooling Degree Day (CDD) swaps, capturing seasonality and long-term trends. Another approach is using the Ornstein-Uhlenbeck (OU) process, introduced by Zhu and Bauer, which is a mean-reverting process suitable for pricing complex temperature derivatives, including HDD and CDD swap options. Additionally, other proposed models for temperature derivatives pricing include the Regime-Switching model (RS), allowing changes in statistical properties over time, the Jump-Diffusion model (JD) capturing sudden jumps in temperature data and the Stochastic Volatility model (SV) handling temperature data volatility.

Pricing temperature derivatives remains challenging in finance, with model choice depending on temperature data characteristics and derivative complexity. Ongoing research aims to enhance accuracy in temperature derivative pricing models. Alaton et al. [11], applied martingale methods to value weather call options; however, this approach contradicts the Black-Scholes assumption of market completeness due to the non-tradability of temperature. This study uses the method of discounting expected payoff, incorporating temperature model information and probability density [12].

In summary, the existing literature has extensively explored the development of a more precise and efficient stochastic model for temperature forecasting, aiming to enhance the pricing of weather derivatives. However, the potential of the derived stochastic model for forecasting temperature values remains unaddressed in the literature. Furthermore, there is a lack of research on how the derived stochastic model can be utilized to price temperature derivatives effectively. Additionally, the literature has not provided a comprehensive understanding of how the mean-reverting Ornstein-Uhlenbeck (OU) process, employed by Scher and Messori, Hess and Zapranis and Alexandridis, tackles non-linearity and non-stationarity in temperature data through the satisfaction of the Fokker-Planck equation. Specifically, the limitations of the OU process in dealing with autocorrelation intervals in temperature data remain unexplored. Therefore, it is crucial to investigate how the proposed stochastic model can be constructed to incorporate autocorrelation interval parameters in both diffusion and drift components, enabling it to effectively capture autocorrelation, nonlinearity and non-stationarity in temperature data. This endeavor will provide insights into how the inclusion of autocorrelation interval parameters in the proposed model enhances the accuracy of temperature predictions.

## MATERIALS AND METHODS

To solve this problem, this paper derives a stochastic model from temperature that follows a gamma distribution. By utilizing the synthesis method for deriving diffusion and drift coefficients, we incorporate a parameter representing the autocorrelation interval in the model. The satisfaction of the Fokker-Planck equation enables the parameters to capture nonlinearity and non-stationarity in the data, resulting in more accurate daily temperature predictions. With this model, we can predict temperature values that are subsequently used to estimate prices of gamma weather derivatives, whose underlying asset is temperature.

### Preliminary concepts

Gamma distribution was fitted to daily temperature data. Below is the provided description of the gamma distribution as outlined in the work by Haddad [4].

**Definition 1:** A gamma distribution is defined for a random variable  $Y$  when its probability density function is expressed in the following manner.

$$f_Y(y) = \frac{(\beta\theta^{-1})^\alpha}{e^{\frac{y}{\theta}}\beta\Gamma(\alpha)}, \quad y > 0, \theta > 0, \alpha > 0 \quad (1)$$

When value of  $\Gamma(\alpha)$  is determined as follows:

$$\Gamma(\alpha) = \int_0^\infty \frac{s^{\alpha-1}}{e^s} ds \quad (2)$$

and

$$ig(\alpha, w) = \int_w^\infty \frac{s^{\alpha-1}}{e^s} ds \quad (3)$$

The term used to denote this concept is referred to as the upper incomplete gamma.

**Lemma 1:**  $\lim_{\alpha \rightarrow \infty} \Gamma(\alpha) = \infty$ ,  $\lim_{\alpha \rightarrow \infty} ig(\alpha, w) = \infty$  and  $\lim_{w \rightarrow 0} ig(\alpha, w) = \Gamma(\alpha)$

The significance of this lemma is that when applying gamma distributions to model temperature, it's important for the  $\alpha$  values to be moderate in size. Furthermore, if the  $w$  values are small, employing the gamma function is also a viable option.

**Lemma 2:** If  $Y$  follows gamma probability distribution with parameters  $\alpha$  and  $\theta$ , then the mean of  $Y$  is given by,

$$E(Y) = \alpha\theta$$

its second moment is given by

$$E(Y^2) = \alpha\theta^2(\alpha + 1)$$

and its variance is

$$\text{Var}(Y) = \alpha\theta^2$$

**Definition 2:** The method of moment estimators of the parameters  $\alpha$  and  $\theta$  are given by

$$\hat{\alpha} = \frac{E(Y)}{\hat{\theta}} \quad (4)$$

and

$$\hat{\theta} = \frac{\text{Var}(Y)}{E(Y)} \quad (5)$$

This study focuses on univariate stochastic differential equations of Ito's type to model temperature given that it follows gamma distribution.

**Definition 3:** Ito's stochastic differential equation is given by

$$dY_t = \mu(Y_t, t)dt + \sigma(Y_t, t)dW_t \quad (6)$$

$\mu(Y_t, t)$  is the drift coefficient,  $\sigma(Y_t, t)$  is the diffusion coefficient

**Definition 4:**  $W_t$  is the standard Brownian motion and is defined as the stochastic process  $\{W(t), t \geq 0\}$  such that  $P(W(0)=0)=1$ ,  $E(W(t))=0$ ,  $\text{Var}(W(t))=t$ ,  $W(t) \sim N(0, t)$  and  $\{W(t)-W(s), 0 \leq s < t\}$  is independent process.

**Lemma 3:** It can be proved that variance of  $\{W(t)-W(s), 0 \leq s < t\}$  is  $t-s$ .

**Theorem 1:** If the probability density function of an Ito's stochastic differential equation is given by  $f_Y(y)$  and that  $\tau$  is auto correlation function of the process  $\{Y_t\}$ , then synthesis

method of finding the drift and the diffusion coefficient is given by

$$\sigma^2(Y_t, t) = \frac{-2}{\tau f_Y(y)} \int_0^y (s - \mathbb{E}(Y)) f_Y(s) ds \quad (7)$$

$$\mu(Y_t, t) = \frac{\sigma^2(Y_t, t)}{2} \frac{\partial}{\partial y} [\log(\sigma^2(Y_t, t) f_Y(y))] \quad (8)$$

**Definition 5:** As the data exhibited a seasonal pattern, a truncated Fourier series was employed to estimate certain parameters. Here is the formal definition of the truncated Fourier series:

$$g(t) = \sum_{n=0}^M a_n \cos \frac{2\pi t}{L} + b_n \sin \frac{2\pi t}{L} \quad (9)$$

where  $L$  represents the duration of the seasonal pattern and  $2M + 1$  denotes the count of parameters involved.

The majority of stochastic differential equations do not have analytical solutions. In this study, the Euler-Maruyama method was employed for discretization to obtain an approximation for a stochastic differential equation.

**Definition 6:** The Euler-Muryama method given by  $Y_{n+1} = Y_n + a(Y_n, \tau_n) \Delta t + b(Y_n, \tau_n) \Delta W_n$ , where  $\Delta W_n = W_{t_{n+1}} - W_{t_n}$  and recursively defined  $Y_n$  for  $0 \leq n \leq N-1$

In the context of this analysis, where approximations are necessary due to the inherent complexity of most stochastic differential equations, three key metrics to assess the accuracy and precision of our models were employed. The Mean Square Error (MSE) was used to quantify the extent of both overestimation and underestimation between predicted values and raw data. Meanwhile, the Mean Absolute Percentage Error (MAPE) was used to gauge the reliability of the models in terms of prediction accuracy. Lastly, the coefficient of determination ( $R^2$ ), was used to evaluate the models' effectiveness in describing variations within the dataset. These measures collectively provided a comprehensive evaluation of the approximations' performance, which helped to address the challenges presented by stochastic differential equations.

**Definition 7:** To estimate auto-correlation parameters, both Runge-Kutta and optimization methods were employed.

$$MSE = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n} \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12)$$

**Definition 8:** The form of the fourth-order Runge-Kutta method is as follows:

$$\begin{aligned} k_1 &= hf(t_j, y_j) \\ k_2 &= hf(t_j + 0.5h, y_j + 0.5k_1) \\ k_3 &= hf(t_j + 0.5h, y_j + 0.5k_2) \\ k_4 &= hf(t_j + h, y_j + k_3) \\ y_{j+1} &= y_j + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

The pricing approach employed the method of expected payoff, as defined in the following description.

**Definition 9:** The price of the call option, determined by discounting the expected payoff using the available temperature data, is expressed as follows:

$$c(t) = e^{-r(t_n - t)} \mathbb{E}^Q[\max\{Z_t - K, 0\} | \mathcal{F}_t] \quad (13)$$

## Temperature model derivation

The general expression for an Ito stochastic differential equation is provided as follows:

$$dY_t = \mu(Y_t, t)dt + \sigma(Y_t, t)dW_t \quad (14)$$

$\mu(Y_t, t)$  denotes the drift coefficient,  $\sigma(Y_t, t)$  is the diffusion coefficient and  $W_t$  is the standard Brownian motion also called the stochastic process  $\{W(t), t \geq 0\}$  in which  $P(W(0)=0)=1$ ,  $E(W(t))=0$ ,  $\text{Var}(W(t))=t$ ,  $W(t) \sim N(0, t)$  and  $\{W(t)-W(s), 0 < s < t\}$  is independent process. It is the noise only that is normally distributed and not the temperature. This assumption is realistic.

Utilising the synthesis method outlined in Bykhovsky [13] and Kanellopoulos et al. [14], for deriving diffusion equations, we stem the stochastic differential equation governing daily temperature. During this derivation, we make two fundamental assumptions. Firstly, we presume that temperature values exhibit autocorrelation. Secondly, we assume that the normalized covariance function is expressed as

$$\frac{C_\tau}{\sigma^2} = R(\tau) = \exp[-\gamma\tau]$$

and we define the auto-correlation interval as,

$$\tau_{corr} = \int_0^\infty |R(\tau)| d\tau = \int_0^\infty |\exp[-\gamma\tau]| d\tau = \frac{1}{\gamma} \quad (15)$$

$$\begin{aligned} \Rightarrow \sigma(Y_t, t) &= \sqrt{2\gamma_1(t)\theta Y_t} \\ \mu(Y_t, t) &= \frac{\sigma^2(Y_t, t)}{2} \frac{\partial}{\partial y} [\log(\sigma^2(Y_t, t) f_Y(y))] \\ &= \frac{\theta y}{\tau_{corr}} \frac{\partial}{\partial y} \left[ \log \left( \frac{2\theta}{\tau \theta^\alpha} \right) - \frac{y}{\theta} + (\alpha - 1) \log(y) \right] \\ &= \frac{\theta(\alpha - 1) - y}{\tau} \\ \Rightarrow \mu(Y_t, t) &= \gamma_1(\theta(\alpha - 1) - Y_t) \end{aligned} \quad (16)$$

Following the derivation of the stochastic equation presented earlier, we have determined the expressions for drift and the



diffusion coefficient as 15 and 16 respectively. Now, by substituting these in 14 we obtain the result as below:

$$dY_t = \gamma_1(t)(\theta(\alpha-1)-Y_t)dt + p_2\gamma_1(t)\theta Y_t dW_{(1)}(t) \quad (17)$$

This differs slightly from the Ornstein-Uhlenbeck differential equation in that the diffusion coefficient in 30 is contingent upon the random variable  $Y_t$ .

### Analysis of the temperature model parameters

**Lemma 4:** The expected value of  $(Y)$  is given by the following differential equation

$$dz = \gamma(\theta(\alpha-1)-z)dt, \text{ where } z = E(Y_t) \quad (18)$$

Proof

We apply Ito isometry in equation 30

$$dE(Y_t) = \gamma(\theta(\alpha-1)-E(Y_t))dt, \quad dz = \gamma(\theta(\alpha-1)-z)dt$$

**Lemma 5:** Suppose that the model in equation 30 satisfies the following FPE.

$$\frac{\partial f_{Y_t}(y)}{\partial t} = -\frac{\partial}{\partial y}[\mu(Y, t)f_{Y_t}(y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(Y, t)f_{Y_t}(y)] \quad (19)$$

Then  $\gamma_1$  is given by the following equation

$$\gamma_1 = \frac{\left(\log\left(\frac{y}{\theta}\right)\frac{\partial\alpha}{\partial t} + \frac{y-\alpha\theta}{\theta^2}\frac{\partial\theta}{\partial t} - \frac{1}{\Gamma(\alpha)}\frac{\partial\Gamma(\alpha)}{\partial t}\right)\theta y}{(\alpha^2\theta^2 - 2\alpha\theta y - \theta^2 - \theta y + y^2)} \quad (20)$$

Proof

We assume that  $\theta(t)$  and  $\alpha(t)$  are  $\{F_t\}$ -measurable

$$\begin{aligned} \text{Let } f_Y(y) &= \frac{y^{\alpha-1}e^{-\frac{y}{\theta}}}{\Gamma(\alpha)\theta^\alpha}, \text{ then by logarithmic differentiation, we have} \\ \frac{\partial}{\partial t} \log(f_Y(y)) &= \frac{\partial}{\partial t} \left( (\alpha-1)\log(y) - \frac{y}{\theta} - \log(\Gamma(\alpha)) - \alpha\log(\theta) \right) \\ \frac{1}{f_Y(y)} \frac{\partial}{\partial t} (f_Y(y)) &= \left( \log(y) \frac{\partial\alpha}{\partial t} + \frac{y}{\theta^2} \frac{\partial\theta}{\partial t} - \frac{1}{\Gamma(\alpha)} \frac{\partial\Gamma(\alpha)}{\partial t} - \frac{\alpha}{\theta} \frac{\partial\theta}{\partial t} - \log(\theta) \frac{\partial\alpha}{\partial t} \right) \\ \frac{\partial}{\partial t} (f_Y(y)) &= \left( \log(y) \frac{\partial\alpha}{\partial t} + \frac{y}{\theta^2} \frac{\partial\theta}{\partial t} - \frac{1}{\Gamma(\alpha)} \frac{\partial\Gamma(\alpha)}{\partial t} - \frac{\alpha}{\theta} \frac{\partial\theta}{\partial t} - \log(\theta) \frac{\partial\alpha}{\partial t} \right) f_Y(y) \end{aligned}$$

Applying laws of logarithms and simplifying fraction, we get

$$\frac{\partial}{\partial t} (f_Y(y)) = \left( \log\left(\frac{y}{\theta}\right)\frac{\partial\alpha}{\partial t} + \frac{y-\alpha\theta}{\theta^2}\frac{\partial\theta}{\partial t} - \frac{1}{\Gamma(\alpha)}\frac{\partial\Gamma(\alpha)}{\partial t} \right) f_Y(y) \quad (21)$$

The second operator in the FPE can be expressed as follows

$$-\frac{\partial \mu(Y_t, t) f_{Y_t}(y)}{\partial y} = \frac{-(\gamma_1 y^{\alpha-1} e^{-\frac{y}{\theta}} (\alpha^2 \theta^2 - 2\alpha \theta y + \theta^2 + \theta y + y^2))}{\theta^{\alpha+1} \Gamma(\alpha)}$$

Applying the third operator of the FPE to the stochastic differential equation gives

$$\frac{\partial^2 \sigma^2(Y_t, t) f_{Y_t}(y)}{\partial y^2} = \frac{-(2\gamma_1 y^{\alpha-2} e^{-\frac{y}{\theta}} (-\alpha^2 \theta^2 + \alpha \theta^2 + 2\alpha \theta y - y^2))}{\theta^{\alpha+1} \Gamma(\alpha)}$$

Therefore, the right hand side of the FPE can be expressed as follows

$$\begin{aligned} -\frac{\partial}{\partial y}[\mu(Y, t)f_{Y_t}(y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(Y, t)f_{Y_t}(y)] &= \frac{-(\gamma_1 y^{\alpha-2} e^{-\frac{y}{\theta}} (-\alpha^2 \theta^2 + \alpha \theta^2 + 2\alpha \theta y + \theta^2 + \theta y - y^2))}{\theta^{\alpha+1} \Gamma(\alpha)} \\ \text{Since } f_Y(y) &= \frac{y^{\alpha-1} e^{-\frac{y}{\theta}}}{\Gamma(\alpha)\theta^\alpha}, \text{ we have the following} \\ -\frac{\partial}{\partial y}[\mu(Y, t)f_{Y_t}(y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(Y, t)f_{Y_t}(y)] &= \frac{-(\gamma_1 (-\alpha^2 \theta^2 + 2\alpha \theta y + \theta^2 + \theta y - y^2))}{\theta y} f_Y(y) \quad (22) \end{aligned}$$

Putting the results in Equation (21) and (22) in the FPE and then solving for  $\gamma_1$  gives

$$\gamma_1 = \frac{\left(\log\left(\frac{y}{\theta}\right)\frac{\partial\alpha}{\partial t} + \frac{y-\alpha\theta}{\theta^2}\frac{\partial\theta}{\partial t} - \frac{1}{\Gamma(\alpha)}\frac{\partial\Gamma(\alpha)}{\partial t}\right)\theta y}{(\alpha^2\theta^2 - 2\alpha\theta y - \theta^2 - \theta y + y^2)} \quad (23)$$

### Numerical estimation of trigger value

A trigger value refers to a specific threshold or level that, when crossed or reached, initiates a specific action or event in a financial instrument or contract. It serves as a condition or criterion that determines when certain provisions or features associated with the instrument come into effect. It varies depending on the context and the specific financial instrument or contract involved. It could be a predefined price level, a specified time period or value of a particular variable, a particular event occurrence or a combination of factors. For example, in options contracts, a trigger value may represent the strike price at which the option is exercised or the barrier level at which a knock-out or knock-in feature is activated.

In order to apply the Runge-Kutta method in equation 18 we let

$f(t_i, z_i) = \gamma_1(\theta(\alpha-1)-z_i)$  then,

$$\begin{aligned} k_1 &= -\gamma_1 h_i (z_i - \theta_i (\alpha_i - 1)) \\ k_2 &= \frac{(\gamma_1 h_i (\gamma_1 h_i (z_i - 2)(\theta_i + z_i - \alpha_i \theta_i)))}{2} \\ k_3 &= \frac{-(\gamma_1 h_i (\gamma_1^2 h_i^2 - 2\gamma_1 h_i (z_i + 4)(\theta_i + z_i - \alpha_i \theta_i)))}{4} \\ k_4 &= \frac{(\gamma_1 h_i (\theta_i + z_i - \alpha_i \theta_i) (\gamma_1^3 h_i^3 - 2\gamma_1^2 h_i^2 + 4\gamma_1 h_i - 4))}{4} \end{aligned}$$

$$\begin{aligned} z_{i+1} &= z_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ &= z_i + \frac{\gamma_1^2 h_i^2 \theta_i}{2} - \frac{\gamma_1^3 h_i^3 \theta_i}{6} + \frac{\gamma_1^4 h_i^4 \theta_i}{24} + \frac{\gamma_1^2 h_i^2 z_i}{2} - \frac{\gamma_1^3 h_i^3 z_i}{6} + \frac{\gamma_1^4 h_i^4 z_i}{24} \\ &\quad - \frac{\gamma_1^2 h_i^2 z_i}{6} + \frac{\gamma_1^4 h_i^4 z_i}{24} - \gamma_1 h_i \theta_i - \gamma_1 h_i z_i + \alpha_i \gamma_1 h_i \theta_i - \frac{\alpha_i \gamma_1^2 h_i^2 \theta_i}{2} + \frac{\alpha_i \gamma_1^3 h_i^3 \theta_i}{6} - \frac{\alpha_i \gamma_1^4 h_i^4 \theta_i}{24} \end{aligned}$$

$i = 1, 2, 3, \dots, N-1$ , and the time step given by  $h_i = \Delta t_{N-i}$

**Forecasting temperature:** The parameters  $\alpha$  and  $\theta$  governing scale and shape respectively in the Stochastic differential equation 30, estimated through a combination of Fourier series as outlined in definition 9 and optimization. Subsequently, the Euler-Murayama method as defined in definition 6 and supported by Mao [15], was employed to estimate temperature values. MSE given by equation 10 was used to account for discrepancies between raw and predicted values Willmott and Matsuura, MAPE given by equation 11 was used to assess model

reliability, in accordance with Khair et al., and equation 12 was used to gauge the model's suitability for the given dataset, aligning with the methods presented by Allison et al. Ultimately, pricing was determined by discounting the expected payoff, as specified in equation 13 based on available temperature data at any given time  $t$ , ensuring that our model incorporates real-time information.

**Pricing temperature derivatives:** Taking into account the fact that  $Y_t$  follows a gamma distribution, thus, the call option's price at  $t \leq t_1$  is given by:

$$c(t) = e^{-r(t_n-t)} E^Q[\max\{Y_t - K, 0\} | \mathcal{F}_t] \quad (24)$$

$$= e^{-r(t_n-t)} \int_K^\infty (y - K) f_{Y_t}(y) dy, \quad (25)$$

where  $K = Z_t$

Here,  $Z_t$  represents the mathematical expectation, which serves as the solution to Ordinary Differential Equation 23. We designate  $K$  as the agreed-upon critical or trigger value, which we take to be the predicted average of  $Z_t$  and  $F_t$  signifies temperature information at time  $t$  also known as the filtration. Upon evaluating the integral in Equation 3 the outcome is as follows:

$$c(t_n) = -e^{-r(t_n-t)} (K\Gamma(\alpha, K/\theta) - \theta\Gamma(\alpha + 1, K/\theta)) / \Gamma(\alpha) \quad (26)$$

where  $\Gamma(\alpha)$  is given by

$$\Gamma(\alpha) = \int_0^\infty s^{\alpha-1} e^{-s} ds \quad (27)$$

$$\Gamma(\alpha, w) = \int_w^\infty s^{\alpha-1} e^{-s} ds \quad (28)$$

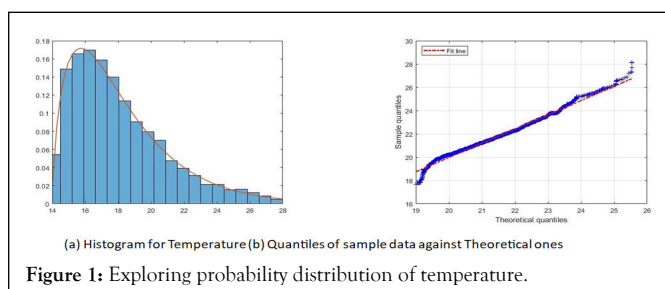
This is referred to as the upper incomplete gamma function, with  $r$  denoting the risk-free interest rate, while  $t$  and  $t_n$  represent the initial and final times of the agreed contract, respectively.

## RESULTS AND DISCUSSIONS

This section provides results on modeled temperature, how gamma distribution fits into data collected. It also indicates parameter estimates both alpha and theta as in scale and shape respectively. A sample contract has been demonstrated with an African Risk Company (ARC) and Malawi government. Finally,

**Table 1:** Parameter estimates.

Mean	Variance	$\theta^*$	$\alpha^*$
21.67	31.89	1.4718	14.72



the paper has also demonstrated the relationship between the call option and trigger value.

## Estimating a gamma distribution for temperature data

Temperature ( $Y$ ) in the area conforms to a gamma distribution. The conventional method typically involves assuming that temperature adheres to the probability distribution described by the Ornstein-Uhlenbeck stochastic process which posits that the residuals, when the trend reverts to the mean, follow a Gaussian probability distribution. However, the research done by Gyamerah et al., suggests that there is a non-negligible probability that temperature residuals may not adhere to a normal distribution.

In previous research, a normal probability distribution was found to be a good fit primarily for maximum temperatures. Haddad demonstrate that the gamma distribution is more appropriate when considering daily temperature data, which incorporates both the maximum and minimum temperatures recorded on a given day. The probability density function of the gamma distribution is described by the function (1). The mean of  $Y$ , second moment and variance are given by results in Lemma 2. Similarly, the method of moment estimators of the parameters are given in definition 2.

In this scenario, we estimate  $E(Y)$  by the sample mean and  $\text{Var}(Y)$  by sample variance. To illustrate, temperature data obtained from the Chitedze research station in Malawi, spanning the years from 1981 to 2021 (as available during the writing of this paper) was used. The resulting statistics were as follows.

This study incorporated temperature data collected concurrently with rainfall data. Specifically, all temperature values on days with zero rainfall were excluded from the dataset. Subsequently, the collected data was fitted to a gamma distribution, as illustrated in Figures 1a and 1b where it is evident from Figure 1b that temperature values closely align with the best-fit line, indicating a strong adherence to the gamma distribution [16]. It's worth noting that the histogram slightly resembles a normal curve due to the presence of a large scale parameter, as detailed in Table 1.

To evaluate the goodness of fit of the model, a *Chi-square* test was employed, as detailed by Balakrishnan et al. utilizing the parameter estimates computed, the chi-square test is defined as follows:

$$\chi^2 = \sum_{i=1}^N \frac{(O_i - E_i)^2}{E_i} \quad (29)$$

The outcomes indicated that temperature conforms to a gamma distribution as a random variable, with a p-value of 0.9261. This took into consideration the fact that temperature was categorized into seasons like "winter," "spring," "summer" and "autumn." This categorization allows to treat temperature as a categorical variable for the *chi-square* test.

**The temperature model:** The derived temperature model is given by

$$dY_t = \gamma_1(t) (\theta(\alpha-1) - Y_t) dt + p_2 \gamma_1(t) \theta Y_t dW_{(1)}(t) \quad (30)$$

The properties of this equation are derived in section 3.1. Such properties give conditions that must be satisfied by the autocorrelation interval coefficients  $\gamma_1$  for the model to satisfy the Fokker-Planck equation. This approach is also found in the work of Dzupire et al. According to Tabandeh et al., Fokker-Planck equation is responsible for describing how uncertainly spreads in dynamic systems that are influenced by random events. The solution to this equation is a Probability Density Function (PDF) that changes over time and is often of large size and scope. In addition, there are certain characteristics of the joint and any marginal solution PDF that need to remain constant over time. Furthermore, meeting the requirements of the Fokker-Planck equation guarantees that the coefficients  $\gamma_1$  address issues of non-linear and non-stationary behavior within the data set that may have arisen as a result of climate change.

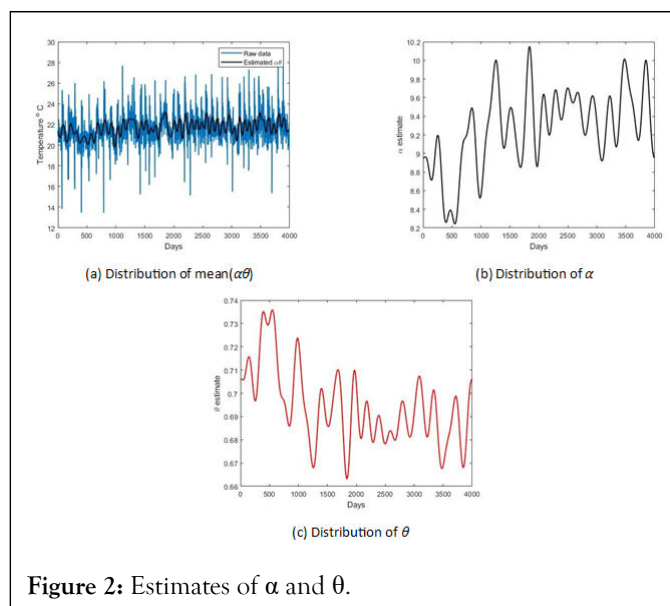
**Parameter estimation:** The estimate for the mean  $\alpha(t)\theta(t)$  was derived using a truncated Fourier series in equation 9. The parameters  $a_n$  and  $b_n$  were determined through the application of the least square method, as described by Yazdi et al. To estimate variance  $\alpha(t)\theta(t)^2$ , the mean square error was computed from

$$\frac{1}{N - \eta - 1} \sum_{\forall t} (Y(t) - \alpha(t)\theta(t))^2 \quad (31)$$

**Table 2:** Estimates of  $\gamma$ .

Shape parameter( $\alpha$ )	Scale parameter ( $\theta$ )	$\gamma_1$ estimate	
9.54	0.734	0.048	0.045
9.49	0.735	0.043	0.039
9.46	0.736	0.04	0.042
9.44	0.736	0.039	0.041
10.15	0.721	0.503	0.507
10.04	0.724	0.145	0.138
9.92	0.726	0.112	0.11

where  $\eta$  is the number of parameters. This value was subsequently multiplied by  $1-h_{ij}$ , where  $h_{ij}$  are representing influential points in the Fourier series. The obtained results were further modeled using the Fourier series, following a similar approach to that employed by Benth et al. [17], in estimating variance. As a result, we derived estimates for  $\theta^* = \alpha(t)\theta(t)^2 / (\alpha(t)\theta(t))$  and  $\alpha^* = \alpha(t)\theta(t) / (\theta^*)$ . The summarized outcomes are presented in Figure 2. Notably, the values obtained in Figures 2a, 2b and 2c correspond to the values listed in Table 1, demonstrating the consistency of this method.



**Figure 2:** Estimates of  $\alpha$  and  $\theta$ .

Nonetheless, employing maximum likelihood estimation entails assuming that there is independent sampling. Given that this study assumes the presence of autocorrelation within the data and follows a gamma distribution, a different approach has been adopted by using Equation (23). Alternatively, Equation 18 can be solved using the Runge-Kutta method to obtain the sequence  $Z_t$ . Following this, the parameters were determined by solving the subsequent minimization problem. The results were not different (Table 2).

$$\text{argmin} L(\gamma) = X(Y_i - Z_i)^2 \quad (32)$$

9.8	0.729	0.089	0.88
9.7	0.731	0.07	0.06
9.61	0.733	0.057	0.053

This provides us with the following estimates for the auto-correlation parameter ( $\gamma$ ). In Table 2 shown above, the auto-correlation parameter ( $\gamma$ ) exhibits a range spanning from 0.039 to 0.503. With these diverse  $\gamma$  values from Table 2, temperature values were estimated utilizing a stochastic differential equation (30), resulting in the estimated values depicted in Figure 3 below.

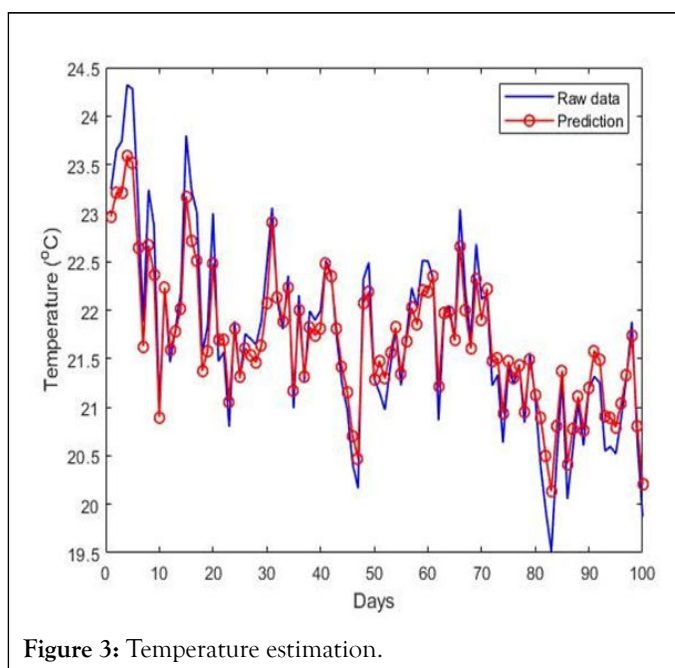


Figure 3: Temperature estimation.

Table 3: Accuracy and reliability.

R-Squared	MSE	MAPE	Reliability
0.9112	0.1429	0.0136	58%

**Pricing derivatives:** The following table shows a sample contract between the African Risk Capacity (ARC) and the Malawi government.

The payout of the contract would be used to support vulnerable communities affected by the temperature variations. In exchange for this coverage, the Malawi government would pay a premium to the ARC, which would be determined based on the level of coverage and the likelihood of the predetermined conditions

Figure 3 depicted above illustrates that the predicted temperature values closely align with the raw data. The predictions and computations encompassed the entire dataset, while the graph specifically displays predictions for the last season, offering a more accessible visual representation. To assess accuracy, we employed the Coefficient of Determination (R-squared), Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) as defined in Equations 10, 11 and 12, respectively. The coefficient of determination gauges the model's suitability for the dataset, with values closer to 1 indicating strong alignment between model variations and raw data variations. MSE aids in identifying whether the predictions involve underestimation or overestimation. MAPE quantifies the reliability of the predictions, measuring the percentage of points for which the distance from the mean falls within the MAPE range, as elucidated by Chicco et al. [18].

With an R-squared value of 91%, it can be inferred that the model can explain approximately 91% of the variations in the data, indicating its suitability for the dataset. The MSE value of 0.1429 suggests that the model's tendency for underestimation or overestimation has been mitigated to within 0.1429 units of the actual data. Furthermore, it's worth noting that approximately 58% of the data points within the model fall within a MAPE of 1.36% as shown in Table 3, signifying a high level of reliability in the predictions. The performance of this model is comparable to that of Schiller et al. and Benth et al., who used fractional integration together with Bsplines to derive temperature model. Since the reliability is much better than their model, then this model is less expensive than their model.

being met [19,20]. The contract would help the Malawi government to manage the financial risk associated with temperature variations and ensure that vulnerable communities are supported in times of need.

Now suppose we have  $\alpha=10.15$ ,  $\theta=0.721$ , then the price of call option per farmer in Table 4 is K9361.

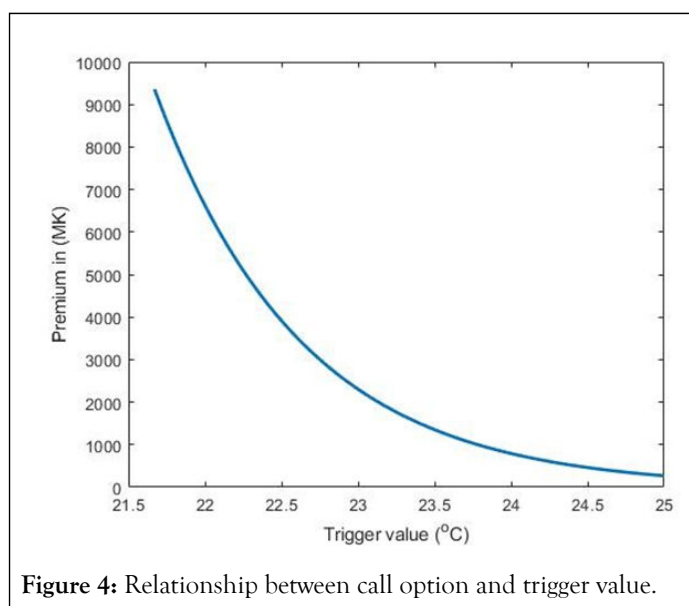
Table 4: Sample contract between (ARC) and government.

Specification	Details
Agreed station	Chitedze research weather station
Calculation of index	Ordinary differential equation and average



Type of derivative	Call option
Initiation	1 <sup>st</sup> Nov 2021
Maturity	1 <sup>st</sup> May 2022
Tick price	K2307.00/degree Celsius
Maximum payout	K50,000.00

Based on the graph in Figure 4 above, it is evident that as the trigger value increases, the premium decreases.



This observation suggests that in financial derivatives, an event is less likely to occur when the average trigger value is higher, resulting in lower premiums.

These results obtained using the model (26) are indicating a similarity with the work of Turvey who derives feasible regions of temperature process space in which a farmer is more likely to be paid by a buyer like ARC. This explains why Johnson suggests that those signing the contract with the ARC on behalf of the government should have sound knowledge about the weather derivatives which are temperature dependent. This model is different from that of Karydas and Xepapadeas in that it is independent of initial temperature and volatility. It is also reflecting well as suggested by Cramer et al., in his study using Machine learning. However, the information about the volatility is contained in the parameters  $\alpha(t)$  and  $\theta(t)$ . Therefore, this model can be more flexible for pricing temperature derivatives.

## CONCLUSION

The objective of this research was to construct a model for estimating call option prices with temperature as the underlying factor. To accomplish this, several steps were undertaken: a Stochastic Differential Equation (SDE) was developed, temperature was modeled and predicted using this SDE, a pricing model was formulated, and ultimately, call option prices were estimated.

A stochastic differential equation was formulated using the synthesis method, with the assumption that temperature exhibited autocorrelation. This equation demonstrated strong predictive accuracy for weather values.

The pricing model does not directly rely on the underlying temperature as a solution to the Stochastic Differential Equation (SDE). Instead, it explicitly hinges on the gamma parameters and the trigger value. As the trigger value increases, the prices generally exhibit a decreasing trend.

The model is much dependent on the gamma parameters and hence it may be used using intermediate and pre-processed data to price temperature derivatives. Hence, we recommend authors as well as traders of weather derivatives to consider using this model for estimating fair prices of derivatives.

In an extension beyond the scope of this paper, it is possible to consider combinations of any two of the following weather conditions: temperature, wind speed and rainfall, as the underlying factors for pricing call options.

## ACKNOWLEDGMENTS

I express my profound gratitude to the Almighty for the gift of life. Additionally, I extend my heartfelt appreciation to Dr. Dzipire for his consistent and invaluable support as my supervisor throughout this journey.

## REFERENCES

1. Gyamerah SA, Ngare P, Ikpe D. Regime-switching temperature dynamics model for weather derivatives. *Int J Stochastic Anal.* 2018;2018(1):8534131.
2. Andrews L, Shivamoggi B. The gamma distribution as a model for temperature dissipation in intermittent turbulence. *Phys Fluids A: Fluid Dyn.* 1990;2(1):105–110.
3. Moriarty M, Murray A, Berx B, Christie A, Munro L, Wallace I. Modelling temperature and fish biomass data to predict annual Scottish farmed salmon, *salmo salar* L, losses: Development of an early warning tool. *Prev Vet Med.* 2020;178:104985.
4. Haddad K. Selection of the best fit probability distributions for temperature data and the use of l-moment ratio diagram method: A case study for NSW in australia. *Theor Appl Climatol.* 2021;143(3):1261–1284.
5. He L, Wang Y, Wei Y, Wang M, Hu X, Shi Q. An adaptive central difference kalman filter approach for state of charge estimation by fractional order model of lithium-ion battery. *Energy.* 2022;244:122627.
6. Cramer S, Kampouridis M, Freitas AA, Alexandridis AK. An extensive evaluation of seven machine learning methods for rainfall

- prediction in weather derivatives. *Expert Syst Appl.* 2017;85:169–181.
7. Geman H, Yor M. Pricing and hedging double-barrier options: A probabilistic approach. *Math Finan.* 1996;6(4):365–378.
8. Hess M. Pricing temperature derivatives under weather forecasts. *Int J Theor Applied Fin.* 2018;21(05):1850031.
9. Ah̃can A. Statistical analysis of model risk concerning temperature residuals and its impact on pricing weather derivatives. *Insurance: Mathemat Econom.* 2012;50(1):131-138.
10. Liu X, Shi L, Qian H, Sun S, Wu P, Zhao X, et al. New problems of food security in Northwest China: A sustainability perspective. *Land Degrad Develop.* 2020;31(8):975-989.
11. Alaton P, Djehiche B, Stillberger D. On modelling and pricing weather derivatives. *Appl Math Fin.* 2002;9(1):1–20.
12. Berhane T, Shibabaw N, Awgichew G, Kebede T. Option pricing of weather derivatives based on a stochastic daily rainfall model with analogue year component. *Heliyon.* 2020;6(1):e03212.
13. Bykhovsky D. Simple generation of gamma, gamma-gamma, and k distributions with exponential autocorrelation function. *J Lightwave Technol.* 2016;34(9):2106-2110.
14. Kanellopoulos SA, Fikioris G, Panagopoulos AD, Kanellopoulos JD. A modified synthesis procedure for first order stochastic differential equations for the simulation of baseband random processes. *Signal Process.* 2007;87(12):3063–3074.
15. Mao X. Convergence rates of the truncated Euler-Maruyama method for stochastic differential equations. *J Comput Appl Math.* 2016;296:362–375.
16. Chidzalo P, Abonongo J, Naryongo R. Enhancing reliability and accuracy in stochastic growth modeling: Method of three selected points approach. *Comput J Math Stat Sci.* 2023;2(2):291-302.
17. Benth FE, Šaltytė Benth JŪ, Koekebakker S. Putting a price on temperature. *Scandinavian J Stat.* 2007;34(4):746–767.
18. Chicco D, Warrens MJ, Jurman G. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *PeerJ Comput Sci.* 2021;7:e623.
19. Guiteras RP, Jack BK. Productivity in piece-rate labor markets: Evidence from rural malawi. *J Dev Econ.* 2018;131:42–61.
20. Johnson L. Rescaling index insurance for climate and development in Africa. *Econ Soc.* 2021;50(2):248–274.