Check for updates

Forecasting monthly rainfall using hybrid time-series models and Monte Carlo simulation amidst security challenges: a case study of five districts from northern Nigeria

Salim Jibrin Danbatta[1] · Ahmad Muhammad[2] · Asaf Varol[3,4] · Daha Tijjani Abdurrahaman[5]

Abstract

Nigeria's agricultural sector relies heavily on rainfall, but insecurity in various regions poses significant challenges. This study aims to address this issue by identifying secure, rain-rich areas in northern Nigeria to support sustainable agriculture. Two models, one integrating classical statistical methods (polynomial and Fourier series fittings) and another using a hybrid approach (artificial neural networks, polynomial, and Fourier series fittings), were employed to analyze historical rainfall data from 1981 to 2021 in the selected districts (Kano, Zaria, Bida, Nguru, and Yelwa) known for their rainfall levels and security stability. The study demonstrates that the machine learning-classical hybrid model outperforms existing models, including the classical-classical hybrid and benchmark models like Iwok's (2016) model, Fourier series, and SARIMA models. Multi-step ahead forecasting with this hybrid model reveals potential changes in rainfall patterns. Notably, Kano, Zaria, Bida, and Yelwa are expected to experience increased rainfall from 2022 to 2026, while Nguru may initially witness decreased rainfall, with improvement in the final year (2026). In conclusion, this study introduces an effective approach for rainfall modeling and forecasting, facilitating the identification of secure agricultural regions in northern Nigeria. These findings carry implications for crop production and agricultural development, contributing to climate resilience efforts and assisting stakeholders in strategic decision-making for regional agricultural investments.

Keywords Hybrid modeling · Agricultural sustainability · Rainfall forecasting · Climate resilience

1 Introduction

The National Population Commission (NPC) of Nigeria recently announced plans to conduct a national census in 2023 (NPC, 2021). This marks a significant milestone for the country, as it will be the first time that Nigeria will conduct a census using digital technology (NPC, 2021). The census is expected to provide more accurate and reliable data that can be used for planning and decision-making at the national, state, and local levels

---

Extended author information available on the last page of the article

🖄 Springer

(National Bureau of Statistics, 2021). However, security concerns may prevent data collection in some areas, resulting in population projections instead.

In parallel to responses to global crises like the COVID-19 pandemic, which have prompted investigations into sustainable supply chains and recovery networks (2023a; Abbasi & Erdebilli, 2023; Abbasi et al., 2022b, 2023b; Abbasi et al., 2022a), the NPC cautioned that insecurity could hinder access to certain areas during the 2023 census. This early information availability has led to researchers using population projection models, such as Iyanda (2019) who projected Nigeria's population from 2006 to 2031 using the Malthus model.

Oluwole et al. (2022) found Nigeria's urban population growing exponentially, while Olanrewaju et al. (2020) identified rapid overall population growth, suggesting the need for government intervention. Other studies, including Nwogu and Okoro (2017), Clementina et al. (2018), and Effiong and Ekpe (2022), concentrate on Nigeria's projected population growth. Onyeoma and Omotsefeodejimi (2021) highlighted the impact of internal migration on instability, and Effiong et al. (2022) discussed the adverse effects of rural-to-urban migration on food security.

According to projections and the country's population, Nigeria must produce sufficient food for its population. However, as noted by Dickson (2020), there are issues with population displacement and instability in the nation's primary agricultural regions, particularly in the northern part of the nation. According to Ayinde et al. (2020), some important agricultural districts in the northern region of the country are deemed unreachable because of insecurity issues. Security challenges in northern Nigeria have become a major concern in recent years. The region has been plagued by various forms of insecurity, including insurgency (Peace, 2022), banditry (Egwu, 2021), and communal clashes (Solomon, 2020).

In Nigeria, especially in the northern region, the majority of crop production is predominantly rain-fed (Xie et al., 2017), implying that it relies on natural precipitation rather than irrigation. Perhaps this is one of the factors contributing to Nigerian farmers' willingness to pay for rainfall-based forecasts (Awolala et al., 2023). The rainy season in Nigeria typically lasts from April to September, with a peak in rainfall occurring in June and July (Ayinde et al., 2020). During this period, crops are planted and rely on rainfall for growth and development (FAO, 2020). Modeling and predicting rainfall for Nigeria is one technique that can help the nation's agricultural stakeholders and policymakers make better agricultural decisions because the country's crop production is primarily rain-fed. This is crucial if modeling and forecasting methods are used to analyze rainfall data from districts with stable security that have an equal amount of rainfall precipitation as unstable agriculturally productive districts in the northern region.

This study focuses on identifying alternative, secure districts that receive similar amounts of rainfall to agriculturally prosperous but insecure districts, as opposed to the above-mentioned studies and many other studies that focus on population estimates and dynamics in the northern region's inaccessible districts due to security concerns. Rainfall data in Nigeria are seasonal, implying that they usually adhere to a foreseeable cycle of wet and dry seasons (FAO, 2020). Given the predictable pattern of wet and dry seasons in Nigeria's rainfall data, the majority of studies that utilize these data employ models that are specifically tailored to handle seasonal data.

As an example, Akinbobola (2018) employed the SARIMA model to examine Nigeria's rainfall data gathered from 14 stations located in the country's forest and savanna eco-climatic regions. Additionally, the SARIMA model has been applied to rainfall data from different regions and districts, such as Osun state (Adams et al., 2019), Abuja (Adams & Bamanga, 2020), Umuahia (Nwokike et al., 2020), and Warri (Eni & Adeyeye, 2015).

Some studies, such as Iwok (2016), Onwukwe and Ikpang (2015), and Umeh Constance Nnenna et al. (2023), implemented models that, in part, might explain some trends in the data.

A research gap exists as there are no models that we could find in the literature that simultaneously consider the trend, seasonality, and residual components in Nigerian rainfall data, encompassing elements like seasonality, linearity, and residuals in its analysis. To identify alternative, secure regions in northern Nigeria that receive ample rainfall to support increased or rehabilitated agricultural practices, we introduce two innovative modeling approaches designed to capture both trend and seasonality patterns in Nigeria's rainfall data. These techniques seamlessly integrate classical and machine learning methods. The classical modeling technique combines polynomial and Fourier series fittings, while the machine learning-based approach incorporates artificial neural networks (ANN), polynomials, and Fourier series fittings. This method can be seen as a fusion of the modeling approach proposed by Shabri et al. (2020) involving Fourier-ANN. The key distinction from the approach of Shabri et al. (2020) lies in our utilization of the ANN model to accommodate the residuals derived from the polynomial-Fourier series model.

We utilized the models to analyze Nigeria's monthly rainfall data spanning from 1981 to 2021. Our study involved modeling and predicting monthly rainfall in five specific districts in northern Nigeria: Kano, Zaria, Bida, Yelwa, and Nguru, as indicated in Fig. 1.

We chose these districts because of their secure conditions and varying rainfall levels, as discussed in our previous conversations and also highlighted in Ayinde et al. (2020). The results from SARIMA, the most widely used model to analyze data in several studies, are compared to those from our proposed models to assess their effectiveness in capturing and forecasting rainfall data. The results of the proposed models were benchmarked using the Fourier series model and Iwok's (2016) model. To account for the uncertainty in the rainfall estimates, a Monte Carlo simulation scheme was applied to the forecasts of the proposed models. The goal is to simulate a variety of potential values that can produce various
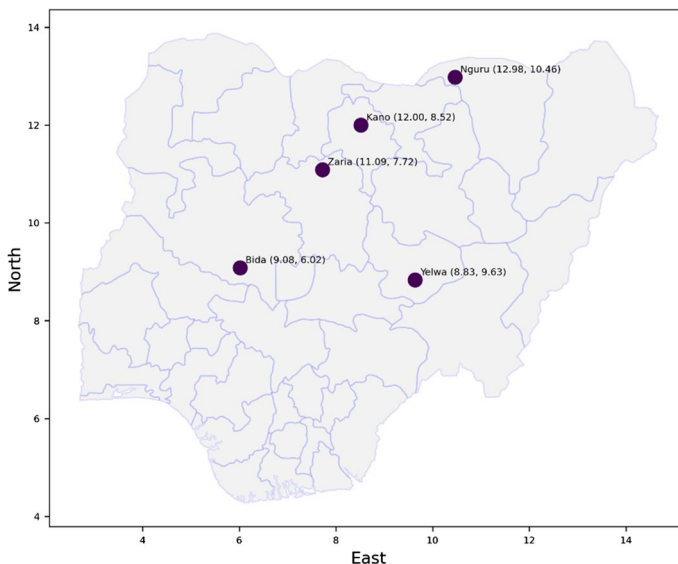

Fig. 1 The selected districts (Kano, Zaria, Bida, Nguru, Yelwa) of the northern Nigeria region

results for a specific prediction period (Danbatta & Varol, 2022). The fundamental concept behind the Monte Carlo simulation is that the accuracy of the findings increases with the number of simulations conducted (Külahcı et al., 2020; Muhammad et al., 2020).

In pursuit of our study's objective, we conducted five-year multi-step ahead forecasting for all chosen districts by utilizing the obtained results of the machine learning-classical model. The multi-step ahead forecasting results are presented in three cases: the worst-case scenario (simulation path 30), best-case scenario (simulation path 65), and most-probable scenario (Monte Carlo mean). Rainfall forecasts for Kano, Zaria, Bida, Yelwa, and Nguru in Nigeria suggest an increase in rainfall for the next five years, starting in 2022, under both the best-case and most-probable scenarios. However, Nguru is expected to experience a decrease in rainfall for the first four years, with only the best-case scenario predicting an increase in the last year. These forecasts are valuable for the Nigerian government, agricultural communities, and policymakers, as they can help them make informed decisions about water management, crop planting, and other important matters.

2 Methodology

This study analyzed monthly rainfall data for the northern Nigerian districts of Kano, Zaria, Bida, Yelwa, and Nguru from 1981 to 2021. The data were sourced from the Central Bank of Nigeria's (CBN) statistical database and collected using rain gauges. This time-series dataset, which can be accessed at http://statistics.cbn.gov.ng/cbn-onlinestats/DataBrowser.aspx on the CBN statistics section's website, has three attributes: month, rainfall (in millimeters), and year. The dataset includes a total of 372 recorded observations, each representing the rainfall measurement for a specific month. According to the weather and climate website (Climate, 2023), Kano district normally receives 49.8 mm of precipitation annually, and there are 62.99 wet days (17.26% of the time). Zaria experiences 128.23 wet days (35.13% of the time) each year and averages about 89.92 mm of precipitation. While Yelwa normally receives about 153.16 mm of precipitation and experiences 178.81 wet days (48.99% of the time) annually, Bida receives 121.81 mm of precipitation and experiences 141.56 rainy days (38.78% of the time) yearly. Nguru usually experiences 67.45 wet days per year, or about 48.46 mm of precipitation, for a percentage of 18.48%.

Rainfall data for Nigeria have been shown to be seasonal based on studies such as Chibuike et al. (2014), Sawa and Ibrahim (2011), and Agogbuo et al. (2017). This matches the selected northern Nigerian district data plots shown in Fig. 2. However, just because the time-series data do not show a clear trend does not mean that there is no trend component. Time series typically exhibit a dominant pattern, such as the seasonal component in this case.

2.1 Polynomial fitting

Polynomial fitting is a time-series data modeling and forecasting technique that uses polynomial functions to approximate the underlying trend of a time series. By fitting a polynomial function to the data, we can create a model that can be used to predict future values of the time series. The polynomial function is mathematically represented as the sum of the powers of the independent variable (time), each with a different coefficient. The idea behind polynomial fitting is to find the combination of coefficients that best approximates the time-series data. This is done by finding the polynomial function that best fits the
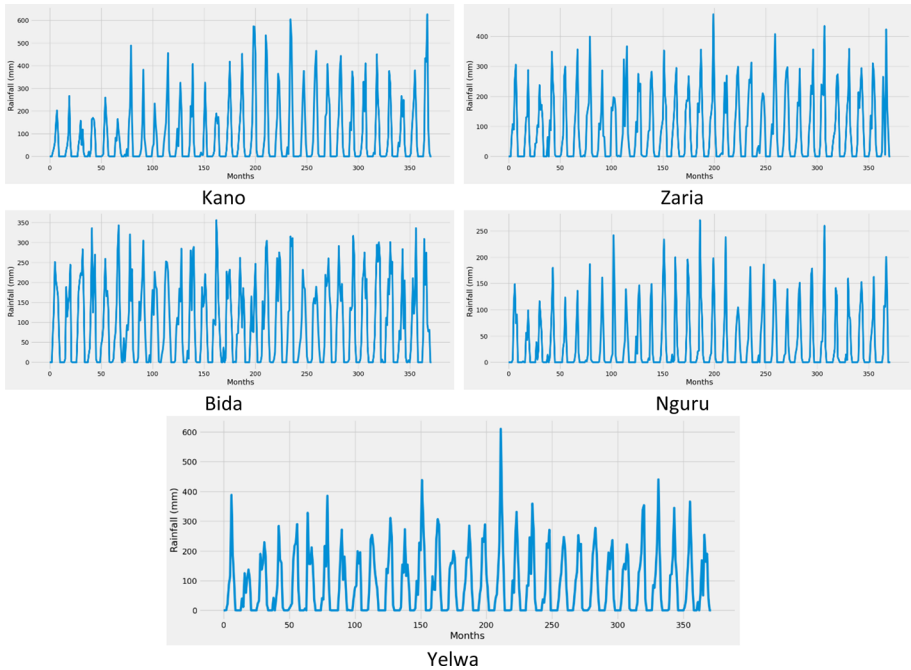
Fig. 2 Data plot for the selected districts of the northern Nigeria region, on the x-axis are months, and on the y-axis are the rain indices (in mm)

observed time-series data, as measured by the smallest possible difference between the two (Chatfield, 2003).

If y_n is a set of observations and x_n is count lag, where n represents the number of observations, then x and y relationship is expressed as $y^{r(x)}(r < n)$;

$$y^{r(x_i)} = a_0 + a_i x_i + a_2 x_i^2 + \cdots + a_{r-1} x_i^{r-1} + a_r x_i^r \tag{1}$$

where $i = 1, \ldots, n$, and the regression coefficients $a_0, a_1, \ldots, a_r$ are calculated using least squares error fitting (Gao et al., 2020; Liu & Wang, 2014). The residual $e_{pi} = |y_i - y_{(x_i)}^r|$ is the difference between the fitted value $y_{(x_i)}^r$ and the observed value y_i. The sum of squared errors for all predicted values is calculated using the following equation: $e_p = \sum_{i-1}^{n} e_{pi}$ (Danbatta & Varol, 2021), when n is the number of observations.

2.2 Fourier series fitting

Fourier series fitting is a technique used in time-series data modeling and forecasting (Wang et al., 2022). This is based on the idea that any periodic time-series data can be represented as a sum of sine and cosine functions (Hippenstiel, 2017). The technique has several advantages, including its ability to capture and model the underlying periodicity of time-series data and its versatility to handle nonlinear and nonstationary time-series data. It can be implemented using various techniques, and once fitted, it can be used to generate

forecasts for future values of the time series (Wang et al., 2022). The Fourier series fitting is given by (Talbot et al. (2020) and Wilson et al. (2018) as follows:

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega x) + \sum_{n=1}^{\infty} b_n \sin(n\omega x) \tag{2}$$

where the number of terms of the series is denoted by n, and ω represents the data frequency, while a_0, a_n, and b_n are the coefficients of the parameters of the Fourier series determined using the least squares error method. The Fourier fit error is represented as $e_{fi} = |y_i - y^r_{(x_i)}|$, and the sum of Fourier errors is represented as $e_f = \sum_{i=1}^{n} e_{fi}$.

2.3 The artificial neural networks (ANN) models

Artificial neural networks (ANNs) belong to a category of machine learning models frequently applied in time-series modeling and forecasting. These networks draw inspiration from the structure and function of the human brain, consisting of artificial neurons in the form of interconnected nodes that can be trained to accomplish various tasks (Haykin & Elektroingenieur, 2012). ANNs are powerful tools for time-series modeling and forecasting. They can model relationships that are nonlinear in data, which is particularly useful for time-series data that exhibit complex patterns (Bengio, 2009).

A variety of ANN architectures are commonly used for time-series modeling and forecasting, including feedforward neural networks (FNNs), recurrent neural networks (RNNs), and long short-term memory networks (LSTMs). FNNs are the simplest ANN architecture, and they are often used for supervised learning and pattern recognition tasks (Haykin & Elektroingenieur, 2012). Typical ANN networks have three layers: an input layer, hidden layers (optional), and an output layer (Shabri et al., 2020). This study used a feedforward single-layer perceptron ANN model, which has three layers, as shown in Fig. 3.

The feedforward ANN is mathematically represented as:

Fig. 3 Three-layered feedforward ANN model (Palmer et al., 2006)
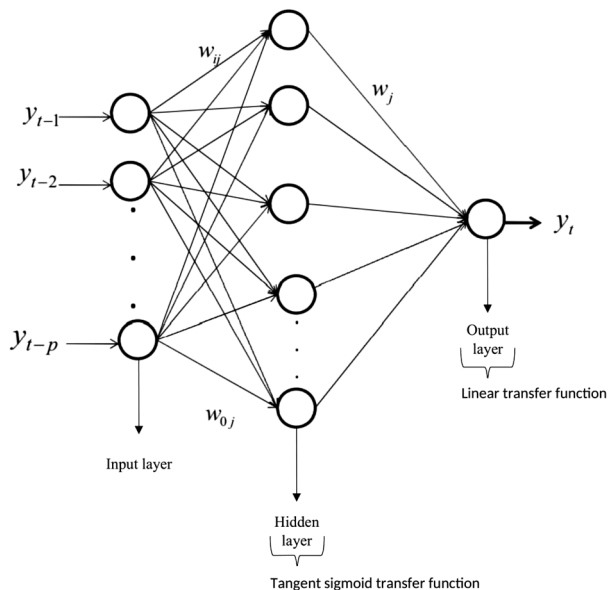
$$y_t = g(w_0 + \sum_{j=1}^{q} w_j^\circ f(\sum_{i=1}^{p} w_{ij} y_{t-i} + w_{oj})) \tag{3}$$

where the number of input observations at a time is represented as p, and the model observation inputs are denoted as $y_{t-1}, y_{t-2}, \ldots y_{t-p}$, while y_t is the output for the given set of inputs.

2.4 The proposed classical-classical model (polynomial-Fourier series)

The classical-classical polynomial-Fourier series model fits the trend of the time series using the first term of a polynomial ($a_1 X$) and captures seasonality using the Fourier series terms, specifically the 11th expansion. To balance the trend and seasonal components, a second polynomial term ($a_2 X^2$) is incorporated as a weighting function. This hybrid model, which is the first proposed in the classical-classical category, is defined as follows:

$$f(x) = a_0 + a_1 X + a_2 X^2 + \sum_{n=1}^{11} a_n cos(n\omega X) + \sum_{n=1}^{11} b_n sin(n\omega X) \tag{4}$$

where the polynomial coefficients are a_0, a_1, and a_2, and the Fourier series coefficients are a_n and b_n, while n and ω represent the number of the Fourier series terms and the data frequency, respectively.

2.5 The proposed machine learning-classical model (ANN-polynomial-Fourier series)

Classical-classical polynomial-Fourier series hybrid model in Eq. 4 is applied to fit the data. The residual Re of the model in Eq. 4 is then modeled using the ANN model described in Eq. 3. The modeled residue, represented as Re (ANN), is incorporated into Eq. 4 to formulate the proposed machine learning-classical model, denoted as the ANN-polynomial-Fourier series model, as shown in Eq. 5.

$$f(x) = a_0 + a_1 X + a_2 X^2 + \sum_{n=1}^{11} a_n \cos(n\omega X) + \sum_{n=1}^{11} b_n \sin(n\omega X) + Re(ANN) + e \tag{5}$$

where Re (ANN) is the ANN modeled residue, and e is the model error.

2.6 Monte Carlo simulation scheme

To anticipate, improve, and identify potential unusual rainfall events in the forecast, we conducted 100 Monte Carlo simulations for each projected model curve, within $a \pm 2\sigma$ (95%) range. We used Monte Carlo simulations to account for forecast uncertainties and simulate all possible future rainfall scenarios. In the design of this study's Monte Carlo simulation scheme, we randomly picked the training residual and incorporated it into the model prediction or forecast trajectory, generating a distinct outcome for every simulation run. Any Monte Carlo-generated paths that surpass the prediction boundaries were disregarded in the analysis. This method allowed us to explore a wide range of potential rainfall outcomes, considering the inherent uncertainties in weather predictions.

The mechanism is straightforward: The Monte Carlo simulation begins by assembling the model prediction and the residual into separate datasets. Next, the mean and standard deviation of the residuals are calculated. The data, model prediction, and index are then fed into a function that calculates and returns the residual upper and lower bounds. Before simulating the model residual, we used a distribution filter to identify the distribution that best fits the residual in cases where the residual is not normally distributed. After determining the data distribution, we run a Monte Carlo simulation, which returns all Monte Carlo simulation paths and the mean of the simulation paths.

3 Findings and limitation

3.1 Models and Monte Carlo forecast results

To assess the effectiveness of the models introduced in this study, we utilized two key performance metrics: the root mean square error (RMSE) and the R-square (R^2) estimate. These measures were employed to gauge the accuracy and reliability of the proposed models, providing a comprehensive evaluation of their performance. According to Pierce (1979), when working with time-series data, R^2 is a reliable estimator despite the potential for overfitting. We also compared the results of the proposed model to those of the model of Iwok (2016), the Fourier series, and the SARIMA models as a benchmark. For brevity and to make it simpler and more concise, we will use the following shorthand notation in the rest of the study: **M1** for the proposed machine learning-classical model (ANN-polynomial-Fourier series), **M2** for the proposed classical-classical model (polynomial-Fourier series), **M3** for the model proposed by Iwok (2016), **M4** for the Fourier series model, and **M5** for the SARIMA model.

To provide a comprehensive and clear understanding of the study's findings, the results of the proposed models and the benchmark models have been systematically tabulated and presented in Table 1. This allows for easy comparison and analysis of the results and enables the reader to gain a thorough understanding of the performance of the various models examined in the study (see Table 1).

The models were trained and evaluated using the historical rainfall data collected from the five districts of the northern Nigeria region between 1981 and 2021. The benchmark models (M3, M4, and M5) were also applied to the same dataset for comparison. The M1 model, as proposed in this study, demonstrated superior performance across all five districts in the northern region, surpassing both the M2 method and

Table 1 Models forecast results

Models	Kano		Zaria		Nguru		Bida		Yelwa	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
M1	3.22	0.93	2.54	0.94	2.51	0.94	3.81	0.92	2.08	0.96
M2	4.01	0.89	3.33	0.91	3.88	0.90	4.65	0.88	2.99	0.94
M3	7.93	0.70	7.45	0.72	9.76	0.65	10.34	0.61	5.05	0.84
M4	16.87	0.49	14.09	0.53	15.07	0.50	17.38	0.45	7.33	0.71
M5	10.31	0.60	9.81	0.64	9.11	0.66	11.71	0.58	5.51	0.79

benchmark models. Specifically, in Kano, the M1 model accurately captured 93% of the variation in rainfall data, with a low RMSE of 3.22 mm. Similarly, in Zaria and Nguru, the model achieved remarkable R^2 estimate values of 94%, coupled with precise RMSE values of 2.54 and 2.51 mm, respectively. Additionally, the M1 model accounted for up to 92% of the variation in Bida's rainfall data and achieved an impressive R^2 estimate value of 96% in Yelwa. The forecasting errors were notably low, with 3.81 and 2.08 mm for Bida and Zaria, respectively (refer to Table 1). These results underline the effectiveness of the M1 model in accurately predicting rainfall patterns in the studied regions.

To account for uncertainty in the rainfall estimates, 100 Monte Carlo simulations are applied to the proposed M1 and M2 models' forecasts. Specifically, the mean of the 100 Monte Carlo simulations was used to report the most-probable amount of expected rainfall. This value is considered the best estimate of the expected rainfall, as it takes into account the uncertainty associated with the M1 and M2 models' forecasts. Additionally, the best-case and worst-case scenarios of the expected amount of rainfall were also reported. Simulation path 65 produced the best output among the 100 Monte Carlo simulation paths generated; hence, it is used for reporting the best-case scenario of the expected amount of rainfall. On the other hand, simulation path 30 showed the least performance among the 100 simulation paths generated, and thus, it is used for reporting the worst-case scenario of the expected amount of rainfall.

The numerical values in Table 2 when compared to the models' results reported in Table 1 showed that for both models, M1 and M2, the Monte Carlo simulations were able to capture more uncertainty in the model's predictions. This was evident by the decrease in forecast error (RMSE) and increase in R^2 estimate values in the case of the simulation mean and simulation path 65 (see Table 2). On the other hand, simulation path 30 could not reach the models' performances with increasing forecast errors and decreasing R^2 estimate values in both models (see Table 2). It can be inferred that simulation path 30 has a lower level of accuracy when compared to the other simulation paths, and therefore, it is not an optimal representation of the predictions. Nevertheless, if it were to occur, it can serve as a valuable source of information for understanding the worst-case scenario in terms of the expected amount of rainfall.

The forecast results provide a nuanced perspective on expected rainfall across districts, emphasizing the trade-off between accuracy and uncertainty. Monte Carlo simulations enhance our understanding of uncertainty, with simulation path 30 serving as a reference for the potential worst-case scenario. These findings provide crucial information for understanding and planning for various rainfall scenarios in the studied districts.

Upon evaluating the performance of both the proposed and benchmarked models, as detailed in Tables 1 and 2, it has been concluded that the Monte Carlo simulation forecast results for the M1 model stand out as the most appropriate for in-depth analysis and multi-step ahead forecasting. This choice is driven by its superior performance compared to all benchmarked models. It is crucial to emphasize, however, that this does not imply the lack of statistical significance in the performance of the M2 model.

The selection of the M1 Monte Carlo forecast results is based on the recommendations outlined in the study conducted by Makridakis et al. (2018). The research systematically assessed and contrasted the efficiency of both regression and machine learning models across various forecasting periods, encompassing a substantial subset of 1045 monthly time-series forecasting challenges. Significantly, the study underscored the significance of employing outcomes from traditional methods as a benchmark while assessing any machine learning-based time-series forecasting model. This comparative

Table 2 The 100 Monte Carlo simulations forecast results for methods **M1** and **M2**

Models	Simulation path	Scenario	Kano district		Zaria district		Nguru district		Bida district		Yelwa district	
			RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
M1	Mean	Most-probable	2.62	0.95	2.04	0.95	1.91	0.97	2.81	0.94	1.98	0.97
	Path 65	Best-case	1.09	0.96	0.90	0.97	0.97	0.98	1.14	0.95	0.88	0.99
	Path 30	Worst-case	17.29	0.65	13.71	0.76	13.04	0.78	18.64	0.61	12.45	0.80
M2	Mean	Most-probable	4.01	0.89	3.33	0.91	3.88	0.90	4.65	0.88	2.99	0.94
	Path 65	Best-case	1.57	0.93	1.29	0.95	1.71	0.94	2.46	0.92	1.18	0.97
	Path 30	Worst-case	20.76	0.60	18.14	0.66	19.38	0.62	13.53	0.71	10.61	0.75

approach enables a clear determination of whether the increased intricacy of machine learning models indeed contributes to enhancing forecasting accuracy.

In the context of our study, the hybrid classical and machine learning-based model M1 yielded better performance than the hybrid classical and classical-based M2 model. This further emphasizes the importance of utilizing classical methods as a baseline when evaluating the performance of any time-series forecasting model.

Figure 4 showcases the outcomes derived from the mean of the Monte Carlo simulation scheme and the 100 simulation paths generated by employing the M1 model for the entire dataset encompassing the five districts of northern Nigeria. The visual representation in the figure delineates the forecast results of the 100 Monte Carlo simulation paths, where each slender line signifies an individual simulation trajectory. In contrast, the bold line signifies the mean of these simulation paths. This method was employed to enhance the model's performance across all five districts, as evident from the detailed findings presented in Table 2.

In addition to evaluating the proposed M1 model, an in-depth examination of the model's forecast residue distribution was carried out for all five selected districts in the northern Nigeria region. This analysis is visually represented in the form of a histogram, depicted in Fig. 5. The histogram demonstrates the favorable fit of the M1 model with the collected rainfall data from the chosen districts.

The model's forecast residue distribution, illustrated in Fig. 5, offers valuable insights into the error distribution. It reveals that the majority of errors cluster around the mean, with only a small number of data points deviating slightly from this central tendency. This observation is indicative of the M1 model's accuracy in representing the data. Moreover,
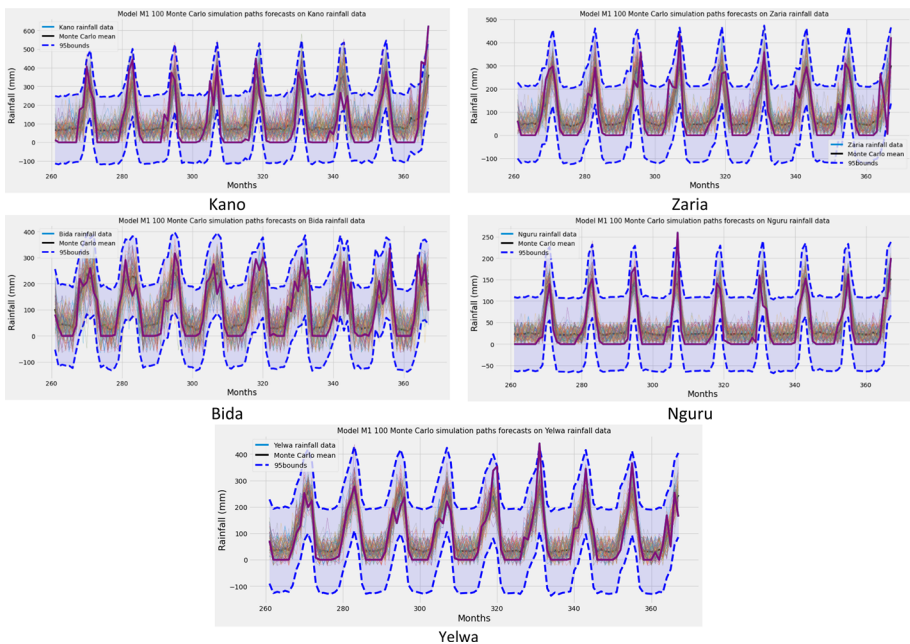

Fig. 4 M1 model's mean and 100 Monte Carlo simulation paths for the districts of Kano, Zaria, Bida, Nguru, and Yelwa of northern Nigeria
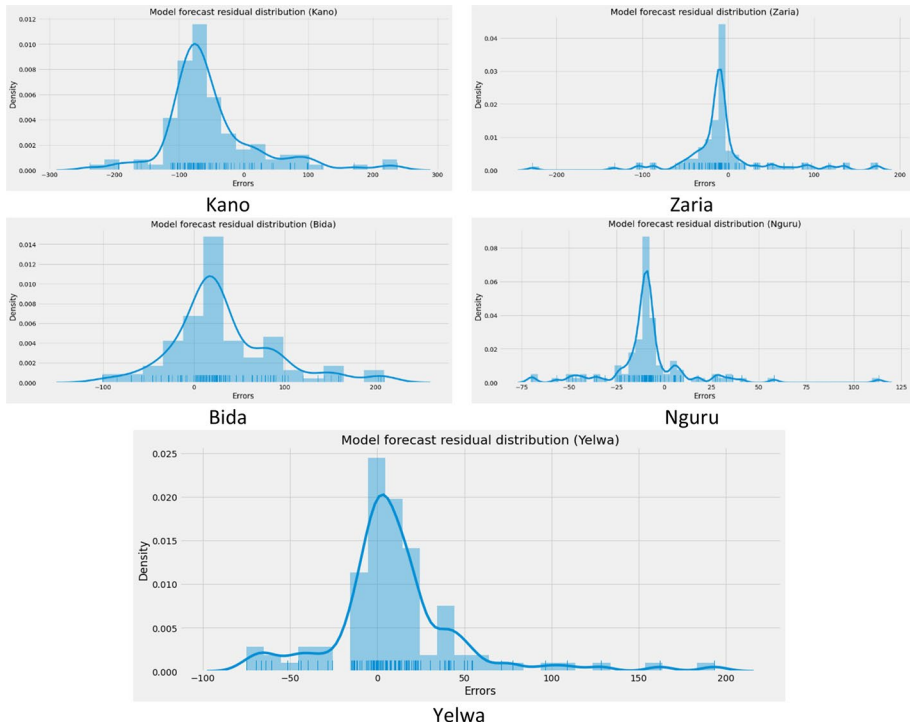
Fig. 5 M1 model's forecast residual distribution for all selected districts of the northern Nigeria region

the presence of random errors and the absence of systematic patterns signify the model's reliability in providing an authentic representation of the rainfall data for these districts.

3.2 Multi-step ahead forecasting

The forecast results, despite some prediction errors, provide a solid foundation for informed decision-making across the five regions in northern Nigeria. Agriculture, a vital sector in this region, can optimize crop planting and harvesting times, ultimately enhancing food security. Water resource managers can allocate water resources more efficiently, considering forecasted rainfall patterns. Disaster preparedness agencies can use the forecasts to anticipate and mitigate flood or drought risks. The M1 model's rainfall forecast results for Kano, Zaria, Bida, Nguru, and Yelwa regions offer a valuable tool for regional planning and resource management. These forecasts contribute to increased resilience in the face of varying weather conditions, benefitting both local communities and the broader northern Nigeria region.

Since the M1 model outperformed all other benchmarks in this study, it was selected for further in-depth analysis and multi-step ahead forecasting. To predict rainfall patterns in the selected districts of northern Nigeria for the next five years (2022–2026), a monthly multi-step ahead forecasting approach was deployed on the existing rainfall dataset. It is important to note that we used the multi-step ahead direct algorithm to predict monthly expected rainfall amounts for the next five years, as this algorithm does not assume any

dependency between future and past steps in the series, ensuring a robust and accurate forecasting process.

In addition, Monte Carlo simulations were used to generate best-case (path 65), worst-case (path 30), and most-probable (mean) scenarios for rainfall amounts in the next five years for each of the five districts. Table 3 shows that the best- and most-probable-case scenarios predict a significant increase in rainfall in Kano district from 2022 to 2026. This is positive news for the region, as adequate rainfall is essential for crop growth and overall agricultural productivity. However, it is important to note that the worst-case scenario also predicts a slight decrease in rainfall (see Table 3).

According to the best-case scenario forecast for Kano, a significant surge in yearly rainfall is anticipated within the next five years when compared to the THY. Specifically, the forecast predicts a total annual rainfall of 1876.7 mm for 2025, surpassing the THY total of 1689.50 mm. This represents an impressive increase of about 11.09% in the percentage rate of change (%ROC) for the year 2025. This is a significant and noteworthy increase in the amount of rainfall for the district, and it is crucial that necessary measures are taken to utilize this increase in the best possible way.

On the other hand, the worst-case scenario presents a different story, showing that the year 2026 having a %ROC of −2.29 mm is the year to see the most decreased amount of rainfall within the next five years when compared to the THY. This decrease, although slight, paints a mixed picture for the Kano district over the next five years, with both positive and negative outcomes. Necessary steps must be taken to make the most of the positive outcomes and to minimize the negative impacts of the negative outcomes.

Based on the best-case and most-probable multi-step ahead forecasting values detailed in Table 4, it is evident that the Zaria district can anticipate a substantial rise in rainfall over the next five years. The data show a noteworthy increase, with the highest expected rainfall of 1222.3 mm in 2026, marking a significant 87.42% increase compared to the threshold year (THY). Furthermore, the two forecasting scenarios also indicate an increased rainfall for the next five years in the Zaria region. On the other hand, the worst-case scenario when compared to the THY paints a very slight decrease in the amount of the expected rainfall spanning from 2022 to 2025. However, even in the worst-case scenario, there is a predicted increase in the amount of the expected rainfall for the last year of the multi-step ahead forecasting period, which is 2026, as presented in Table 4.

For the Yelwa district, all forecasting scenarios predict a consistent increase in rainfall over the next five years, compared to the THY (see Table 5). Similar to Zaria district, the best-case scenario forecast for Yelwa predicts that 2026 will have the highest amount of rainfall, at 718.5 mm, which is a significant 29.66% increase from the THY. This can be observed in Table 5, which provides a comprehensive breakdown of the rainfall projections for each year. This steady increase in expected rainfall is a positive sign for the region, as it has the potential to improve agricultural yields and boost economic activity in the area. Additionally, it may also have a positive impact on the local ecosystem, helping to sustain the biodiversity of the region. Furthermore, the forecast results for Yelwa district are encouraging and suggest that the region is on track to experience a significant improvement in rainfall over the next five years.

Likewise, for the Bida district, the multi-step ahead forecasting outcomes outlined in Table 6 depict a continuous and stable rise in the anticipated rainfall levels over the entire five-year duration. This trend is evident across all forecasting scenarios, with the best-case scenario showing the highest increase in rainfall. The best-case scenario of the multi-step ahead forecasting results in Table 6 predicts a consistent increase in total rainfall each year: 1154.04 mm in 2022, 1115.34 mm in 2023, 1139.29 mm in 2024, 1164.51 mm in 2025,

Table 3 Monte Carlo means (most-probable scenario), simulation path 65 (best-case scenario), and simulation path 30 (worst-case scenario); five years of multi-step ahead monthly rainfall forecasting results for Kano

Month	THY	2022			2023			2024			2025			2026		
		Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30
Jan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Feb	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mar	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Apr	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.40	3.76	3.02
May	71.90	73.61	71.99	68.86	70.08	68.78	50.62	79.00	71.64	60.96	82.30	80.31	71.76	70.00	67.17	60.41
Jun	432.60	441.50	439.47	400.11	381.10	378.43	360.52	450.50	445.73	409.65	455.71	450.97	430.63	425.88	420.54	412.92
Jul	416.60	430.11	426.85	420.75	530.43	526.73	511.43	435.08	431.54	400.86	540.22	535.14	531.84	420.54	415.35	406.23
Aug	626.30	659.90	649.72	647.53	710.0	700.47	592.87	665.01	660.82	655.75	670.32	666.45	509.43	580.01	575.56	570.42
Sep	123.50	140.40	139.63	130.28	148.88	140.83	135.45	141.10	145.65	140.49	112.80	110.64	109.64	200.00	198.64	190.43
Oct	18.60	16.32	15.64	13.93	10.11	06.63	02.74	12.77	12.65	11.54	15.44	14.63	13.64	12.86	11.46	07.28
Nov	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Dec	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	1689.50	1761.80	1743.30	1681.50	1854.50	1825.50	1655.30	1783.50	1668.00	1679.30	1876.07	1858.10	1666.90	1713.60	1692.50	1650.70
%ROC		4.28	3.18	-0.47	9.77	8.05	-2.03	5.56	4.65	-0.61	11.09	9.98	-1.34	1.43	0.17	-2.29

Table 4 Monte Carlo means (most-probable scenario), simulation path 65 (best-case scenario), and simulation path 30 (worst-case scenario); five years of multi-step ahead monthly rainfall forecasting results for Zaria

Month	THY	2022			2023			2024			2025			2026		
		Path 65	Mean	Path30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30
Jan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Feb	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mar	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Apr	16.70	18.00	16.91	6.26	19.70	18.51	10.29	14.50	12.78	10.37	11.70	11.50	5.35	16.70	15.98	10.73
May	265.00	255.30	251.26	199.87	290.00	285.22	280.67	240.00	233.54	200.49	265.00	259.82	239.73	280.29	272.83	250.84
Jun	157.70	132.72	127.42	112.33	191.70	188.09	180.43	127.85	126.98	119.08	157.70	155.73	160.36	200.49	195.37	184.85
Jul	5.50	97.51	89.76	54.86	15.50	13.77	3.85	10.56	9.88	7.92	50.00	48.92	60.83	15.50	12.63	09.84
Aug	422.70	452.95	443.03	400.08	489.31	479.23	389.57	500.49	498.46	487.72	480.37	478.84	430.55	422.70	417.84	400.37
Sep	166.30	136.29	122.45	101.28	171.56	166.84	149.46	166.80	160.02	157.26	176.30	172.74	169.74	190.67	184.68	176.74
Oct	84.70	80.08	79.41	60.52	8.90	8.88	7.37	80.70	78.25	76.83	50.41	48.84	47.76	90.93	91.58	87.84
Nov	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.00	4.82	4.18	0.00	0.00	0.00	5.00	5.92	4.85
Dec	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	1118.60	1172.80	1130.20	935.20	1186.60	1160.54	1021.60	1156.40	1134.80	1069.20	1191.50	1176.40	1114.30	1222.30	1196.80	1126.10
%ROC		4.85	1.04	-16.39	6.08	3.75	-8.66	3.37	1.45	-4.42	6.52	5.16	-0.38	87.42	6.99	0.66

Table 5 Monte Carlo means (most-probable scenario), simulation path 65 (best-case scenario), and simulation path 30 (worst-case scenario); five years of multi-step ahead monthly rainfall forecasting results for Yelwa

Month	THY	2022			2023			2024			2025			2026		
		Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30
Jan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Feb	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mar	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Apr	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.09	09.09	08.33
May	15.70	16.30	15.43	10.34	18.54	17.99	16.05	14.00	13.51	12.23	17.70	16.29	14.23	15.70	13.56	11.42
Jun	107.30	117.10	114.79	104.26	121.23	119.63	100.75	101.35	99.78	97.53	135.86	134.24	130.65	117.30	112.64	109.96
Jul	105.30	123.30	121.42	110.63	115.75	113.75	97.37	140.19	138.57	137.65	115.30	112.54	110.54	155.90	152.24	159.03
Aug	200.60	225.60	224.42	219.65	230.68	229.29	240.36	260.87	258.32	252.63	220.66	218.47	211.08	257.10	251.73	249.55
Sep	125.20	130.20	128.57	100.01	120.44	118.53	117.52	150.34	148.42	146.92	133.59	130.43	126.40	147.20	139.52	137.84
Oct	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.04	08.32	03.98	15.20	14.26	13.94
Nov	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Dec	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	554.10	612.50	604.6	554.970	611.60	604.20	575.70	666.70	658.60	646.90	633.20	620.30	596.80	718.50	693.00	690.00
%ROC		10.54	9.12	0.16	10.38	9.04	3.91	20.33	18.86	16.76	14.27	11.95	7.72	29.66	25.07	24.54

Table 6 Monte Carlo means (most-probable scenario), simulation path 65 (best-case scenario), and simulation path 30 (worst-case scenario); five years of multi-step ahead monthly rainfall forecasting results for Bida

Month	THY	2022			2023			2024			2025			2026		
		Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30	Path 65	Mean	Path 30
Jan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Feb	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mar	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Apr	64.50	70.21	68.16	60.86	65.84	64.52	63.88	77.45	80.23	79.72	73.21	72.40	70.07	65.88	65.23	64.29
May	308.60	330.41	327.81	322.22	310.01	308.29	307.01	320.26	322.65	319.34	340.04	338.19	333.44	309.57	310.72	309.92
Jun	194.00	209.67	205.65	201.76	196.38	193.54	192.33	200.92	189.43	188.20	220.08	219.64	214.99	195.72	192.73	189.09
Jul	274.30	281.61	276.35	277.75	277.30	276.53	275.39	280.74	277.72	275.99	308.36	308.74	301.12	275.72	275.10	271.29
Aug	95.20	99.20	97.64	98.53	98.71	97.07	96.99	97.78	96.01	95.66	105.04	104.79	102.58	96.74	99.00	96.83
Sep	77.40	77.40	76.63	72.76	80.83	80.31	79.28	79.55	78.46	77.30	67.39	66.94	63.86	79.00	80.12	78.93
Oct	81.40	85.54	83.29	82.65	86.27	85.67	84.84	82.59	81.99	79.08	50.39	45.91	42.41	79.00	77.95	74.29
Nov	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	11.00
Dec	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	1095.40	1154.00	1135.50	1116.50	1115.30	1105.90	1099.70	1139.30	1126.50	1115.30	1164.50	1156.60	1128.50	1101.60	1100.80	1095.60
%ROC		5.35	3.66	1.93	1.82	0.96	0.39	4.01	2.84	0.02	6.31	5.58	3.02	0.56	0.49	0.02

and 1101.6 mm in 2026. These figures signify a significant surge in rainfall compared to the THY total of 1095.4 mm. The forecast results indicate that the next five years will be characterized by an overall increase in the amount of rainfall, with 2025 having the highest expected amount of rainfall.

The multi-step ahead forecasts for the Nguru district, as presented in Table 7, show a consistent decrease in the expected amount of rainfall for the next five years when compared to the THY. This trend is evident across all forecasting scenarios, including the best-case, most-probable, and worst-case scenarios. The best-case scenario, however, does show an increase of 3.57% expected in 2026. The numerical data presented in Table 7 offer additional clarity regarding this pattern. In the most favorable scenario, there is an anticipated decline of -4.19% in 2022, -7.26% in 2023, -1.56% in 2024, and -18.82% in 2025 in the projected rainfall levels.

A representation of the multi-step ahead forecasting results of the chosen districts during the peak rainfall occurring in June and July (Ayinde et al., 2020) spanning the years 2023, 2024, and 2025 is shown in Fig. 6. This visually engaging color map provides viewers with a glimpse into the expected trends, patterns, and variations within these districts in these months over the three years. It not only serves as a sample but also as a valuable analytical resource for stakeholders, equipping them with the ability to discern and interpret the forecasted data showcased in the map, facilitating informed decision-making and strategic planning.

4 Conclusion and limitation

This study is dedicated to identifying secure and rainfall-rich regions in northern Nigeria, crucial for supporting and revitalizing agricultural practices. Accurate rainfall prediction is of paramount importance in various sectors, including agriculture, water resource management, and disaster preparedness. To achieve this, we have proposed a hybrid machine learning-classical rainfall prediction model (M1) and compared its performance against a classical-classical model (M2) and benchmark models (M3, M4, and M5). The study concentrates on five key districts in northern Nigeria: Kano, Zaria, Nguru, Yelwa, and Bida. These districts are pivotal in terms of agriculture, contributing significantly to the region's economy and employing a substantial portion of the population.

One notable feature of this study is the incorporation of Monte Carlo simulations to address the inherent uncertainty in rainfall estimates. We subjected the forecasts generated by the M1 and M2 models to 100 Monte Carlo simulations, using the mean of these simulations to determine the most-probable expected rainfall. Additionally, we have presented both best-case and worst-case scenarios, with simulation path 65 yielding the most favorable outcome and simulation path 30 reflecting the least favorable result among the generated simulations. The results obtained are highly encouraging. In both the best-case and most-probable scenarios, we anticipate a substantial increase in rainfall levels across the Kano, Zaria, Yelwa, and Bida districts over the next five years, spanning from 2022 to 2026. Conversely, the Nguru district is expected to experience a consistent decrease in rainfall throughout this period, with only the best-case scenario showing a potential increase in rainfall in 2026. These findings indicate that, except for the Nguru region, the selected districts could serve as promising alternatives for stakeholders seeking to bolster agricultural practices. This is especially relevant as various critical agricultural regions are grappling with issues related to insurgency and instability.

Table 7 Monte Carlo means (most-probable scenario), simulation path 65 (best-case scenario), and simulation path 30 (worst-case scenario); five years of multi-step ahead monthly rainfall forecasting results for Nguru

| Month	THY	2022			2023			2024			2025			2026		
		Path65	Mean	Path30	Path65	Mean	Path 30	Path65	Mean	Path30	Path 65	Mean	Path 30	Path65	Mean	Path30
Jan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Feb	28.00	20.00	19.45	16.98	25.30	24.19	22.29	26.24	25.76	20.35	34.37	33.12	32.54	29.87	25.46	20.55
Mar	0.00	1.00	0.98	0.32	5.02	4.73	2.54	1.45	1.09	0.92	20.82	18.98	16.65	0.00	1.45	0.00
Apr	56.80	50.91	49.21	45.71	45.80	43.22	41.34	51.43	50.43	45.52	30.95	29.01	22.54	59.21	50.72	48.63
May	168.10	160.78	154.09	150.42	155.95	151.34	149.45	161.36	159.74	170.91	120.43	119.56	117.53	173.47	169.32	160.95
Jun	50.30	49.73	48.64	47.96	41.50	40.98	39.48	53.94	52.87	50.54	20.37	19.11	18.24	53.91	50.52	48.32
Jul	254.30	260.79	257.56	255.63	236.37	232.64	230.93	249.27	247.41	245.63	200.49	189.35	187.64	260.16	261.75	259.92
Aug	163.30	155.14	152.75	150.43	149.27	147.53	146.34	159.47	156.62	154.44	100.28	98.01	97.98	166.93	160.26	158.34
Sep	190.20	180.28	176.77	175.62	186.48	181.43	179.87	191.85	189.56	182.64	230.20	225.54	220.92	198.17	196.87	194.53
Oct	41.20	33.62	32.64	30.58	37.38	43.67	40.24	42.32	43.87	42.45	15.05	14.83	13.43	44.52	43.76	41.09
Nov	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Dec	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	952.20	912.30	892.10	873.60	883.10	869.70	852.50	937.30	927.40	913.40	772.90	747.50	727.47	986.20	947.10	932.30
%ROC		−4.19	−6.31	−8.25	−7.26	−8.6	−10.47	−1.56	−2.61	−4.07	−18.82	−21.49	−23.60	3.57	−0.53	−2.08
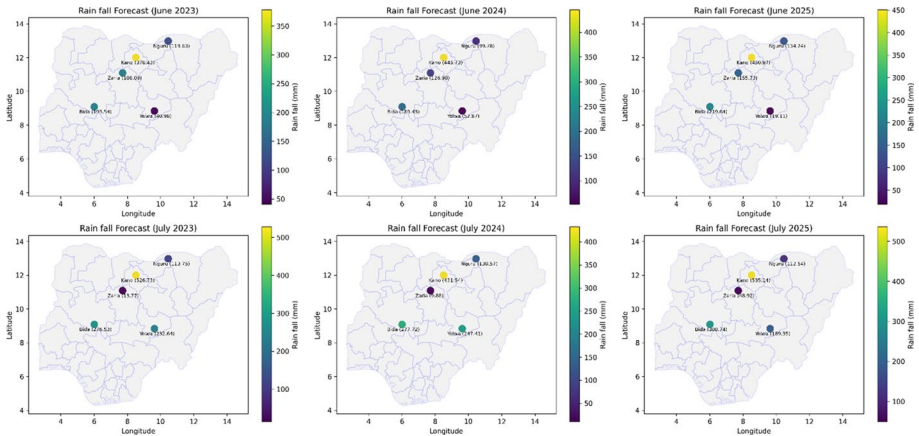
Fig. 6 Multi-step ahead sample forecast results map for June (2023, 2024, and 2025) and July (2023, 2024, and 2025) for all the selected five districts

The study offers a comprehensive approach to rainfall prediction in northern Nigeria, emphasizing the significance of integrating both machine learning and classical methods in rainfall prediction models. The implications of these results extend beyond the scope of this study, as they can guide future planning and decision-making processes in the agricultural sector, water resource management, and disaster preparedness not only within Nigeria but also in other regions facing similar challenges. In light of these findings, we recommend that agricultural organizations, stakeholders, and policymakers consider the insights provided by these models when making decisions related to rainfall forecasting. We believe that the methodologies and results presented here can serve as valuable tools to inform and guide strategic decisions, not only within Nigeria but also in other countries facing similar issues. Additionally, it is essential to acknowledge the limitations of this study, including data constraints and the inherent uncertainties in weather predictions, which should be considered when interpreting and applying the results.

Declarations

References

Abbasi, S., & Erdebilli, B. (2023). Green closed-loop supply chain networks' response to various carbon policies during COVID-19. *Sustainability, 15*(4), 3677. https://doi.org/10.3390/su15043677

Abbasi, S., Daneshmand-Mehr, M., & Ghane Kanafi, A. (2022a). Designing sustainable recovery network of end-of-life product during the COVID-19 PANDEMIC: A real and applied case study. *Discrete Dynamics in Nature and Society*. https://doi.org/10.1155/2022/6967088

Abbasi, S., Khalili, H. A., Daneshmand-Mehr, M., & Hajiaghaei-Keshteli, M. (2022b). Performance measurement of the sustainable supply chain during the COVID-19 pandemic: A real-life case study. *Found Comput Decision Sci, 47*(4), 327–358. https://doi.org/10.2478/fcds-2022-0018

Abbasi, S., Daneshmand-Mehr, M., & Ghane, K. A. (2023a). Green closed-loop supply chain network design during the coronavirus (COVID-19) pandemic: A case study in the iranian automotive industry. *Environ Model Assessment, 28*(1), 69–103. https://doi.org/10.1007/s10666-022-09863-0

Abbasi, S., Sıcakyüz, Ç., & Erdebilli, B. (2023b). Designing the home healthcare supply chain during a health crisis. *Journal of Engineering Research, 23*, 100098. https://doi.org/10.1016/j.jer.2023.100098

Adams, S. O., & Bamanga, M. A. (2020). Modelling and forecasting seasonal behavior of rainfall in Abuja, Nigeria; A sarima approach. *American Journal of Mathematics and Statistics, 10*(1), 10–19. https://doi.org/10.5923/j.ajms.20201001.02

Adams, S. O., Mustapha, B., & Alumbugu, A. I. (2019). Seasonal autoregressive integrated moving average (SARIMA) model for the analysis of frequency of monthly rainfall in Osun State, Nigeria. *Physical Science International Journal, 22*(4), 1–9. https://doi.org/10.9734/psij/2019/v22i430139

Agogbuo, C. N., Nwagbara, M. O., Bekele, E., & Olusegun, A. (2017). Evaluation of selected numerical weather prediction models for a case of widespread rainfall over central and Southern Nigeria. *Journal of Environmental & Analytical Toxicology, 7*(491), 2161–2525. https://doi.org/10.4172/2161-0525.1000491

Akinbobola, A., Okogbue, E. C., & Ayansola, A. K. (2018). Statistical modeling of monthly rainfall in selected stations in forest and savannah eco-climatic regions of Nigeria. *Journal of Climatology & Weather Forecasting*. https://doi.org/10.4172/2332-2594.1000226

Awolala, D., Mutemi, J., Adefisan, E., Antwi-Agyei, P., Taylor, A., Muita, R., Bosire, E., Mutai, B., & Nkiaka, E. (2023). Economic value and latent demand for agricultural drought forecast: Emerging market for weather and climate information in Central-Southern Nigeria. *Climate Risk Management, 39*(January), 100478. https://doi.org/10.1016/j.crm.2023.100478

Ayinde, I. A., Otekunrin, O. A., Akinbode, S. O., & Otekunrin, O. A. (2020). Food security in Nigeria: impetus for growth and development. *Journal of Agricultural Economics and Rural Development, 6*(2), 808–820.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends® in Machine Learning, 2*(1), 1–127. https://doi.org/10.1561/2200000006

Chatfield, C. (2003). The analysis of time series. *Chapman and Hall/CRC*. https://doi.org/10.4324/9780203491683

Chibuike, E. M., Kunda, J. J., Johnson Eze, I., & Adekunle Ayodotun, O. (2014). Mathematical study of monthly and annual rainfall trends in Nasarawa State. *Nigeria. IOSR Journal of Mathematics, 10*(1), 56–62. https://doi.org/10.9790/5728-10135662

Clementina, O., Okechukwu, M., & Chinwe, F. (2018). Development of a new model for population prediction in Anambra State, Nigeria. *International Journal on Recent and Innovation Trends in Computing and Communication, 5*(6), 800–805. https://doi.org/10.17762/ijritcc.v5i6.857

Climate, W. and. (2023). *Nigeria Climate: climate zone and historical climate data*

Danbatta, S. J., & Varol, A. (2021). ANN–polynomial–Fourier series modeling and Monte Carlo forecasting of tourism data. *Journal of Forecasting, 41*(5), 920–932. https://doi.org/10.1002/for.2845

Danbatta, S. J., & Varol, A. (2022). Forecasting foreign visitors arrivals using hybrid model and Monte Carlo simulation. *International Journal of Information Technology & Decision Making, 21*(06), 1859–1878. https://doi.org/10.1142/S0219622022500365

Dickson, C. N. (2020). *North-Eastern Nigeria and Population Displacement Table 1 : Breakdown of the Population affected by the Conflict in North-east Nigeria Source : Culled from Newspapers and Reports , 2020* (Issue c).

Effiong, U., & Ekpe, J. P. (2022). Fertility and population explosion in Nigeria: Does income actually count. *GPH-International Journal of Business Management, 5*(7), 42–59. https://doi.org/10.5281/zenodo.6954432

Effiong, U. E., Paul, K. A., & Udo, U. H. (2022). Rural population attenuation and food production in Nigeria. *Global Journal of Communication and Humaniies, 1*(1), 64–84. https://doi.org/10.5281/zenodo.7038275

Egwu. (2021). *Nigeria Security Situation* (Issue June). https://doi.org/10.2847/433197

Eni, D., & Adeyeye, F. J. (2015). Seasonal ARIMA modeling and forecasting of rainfall in Warri Town, Nigeria. *Journal of Geoscience and Environment Protection, 03*(06), 91–98. https://doi.org/10.4236/gep.2015.36015

FAO. (2020). *Agriculture in Nigeria*. Suite of Food Security Indicators

Gao, J., Ji, W., Zhang, L., Shao, S., Wang, Y., & Shi, F. (2020). Fast piecewise polynomial fitting of time-series data for streaming computing. *IEEE Access, 8*, 43764–43775. https://doi.org/10.1109/ACCESS.2020.2976494

Haykin, S., & Elektroingenieur, K. (2012). *Course Number and Title: CAP6615 (Section: 1967) – Neural Networks for Computing – SPRING 2012 Meeting Times and Location: MWF, 8. 6615*, 8–11

Hippenstiel, R. D. (2017). Detection theory. *CRC Press*. https://doi.org/10.1201/9781420042047

Iwok, I. A. (2016). Seasonal modelling of fourier series with linear Trend. *International Journal of Statistics and Probability, 5*(6), 65. https://doi.org/10.5539/ijsp.v5n6p65

Külahcı, F., Aközcan, S., & Günay, O. (2020). Monte Carlo simulations and forecasting of Radium-226, Thorium-232, and potassium-40 radioactivity concentrations. *Journal of Radioanalytical and Nuclear Chemistry, 324*(1), 55–70. https://doi.org/10.1007/s10967-020-07059-y

Liu, X., & Wang, Y. (2014). Research of automatically piecewise polynomial curve-fitting method based on least-square principle. *Science Technology and Engineering, 14*(3), 55–58.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: concerns and ways forward. *PLoS ONE, 13*(3), 1–26. https://doi.org/10.1371/journal.pone.0194889

Muhammad, A., Külahcı, F., & Akram, P. (2020). Modeling radon time series on the North Anatolian Fault Zone, Turkiye: Fourier transforms and Monte Carlo simulations. *Natural Hazards, 104*(1), 979–996. https://doi.org/10.1007/s11069-020-04200-8

Nnenna, U. C., James, E., & Edith, E. O. (2023). Modelling an automated rainfall forecasting system using an optimized intelligent agent. *Global Journal of Engineering and Technology Advances, 15*(1), 064–069. https://doi.org/10.30574/gjeta.2023.15.1.0077

NPC. (2021). *2023 Digital Census*.

Nwogu, E. C., & Okoro, C. (2017). Adjustment of Nigeria population censuses using mathematical methods. *Canadian Studies in Population, 44*(3–4), 149. https://doi.org/10.25336/p6vs5k

Nwokike, C. C., Offorha, B. C., Obubu, M., Ugoala, C. B., & Ukomah, H. I. (2020). Comparing SANN and SARIMA for forecasting frequency of monthly rainfall in Umuahia. *Scientific African, 10*, e00621. https://doi.org/10.1016/j.sciaf.2020.e00621

Olanrewaju, S., Olafioye, S., & Oguntade, E. (2020). Modelling nigeria population growth: A trend analysis approach. *International Journal of Innovative Science and Research Technology*, 5(4).

Oluwole, E., Adeoti, T., & Oluwole, A. (2022). Population growth model to project human population in Nigeria. *African Scholars Journal of Science Innovation & Technology Research, 24*(9), 89–98.

Onwukwe, C. E., & Ikpang, I. N. (2015). Modeling rainfall in cross river state, Nigeria, using artificial neural network. *International Journal of Scientific and Engineering Research, 6*(6), 234–239. https://doi.org/10.14299/ijser.2015.06.012

Onyeoma, S., & Omotsefeodejimi, D. (2021). *Population dynamics and displacement in nigeria population dynamics and displacement in Nigeria*. https://doi.org/10.9790/5933-1203012635

Palmer, A., José Montaño, J., & Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism Management, 27*(5), 781–790. https://doi.org/10.1016/j.tourman.2005.05.006

Peace, I. for E. &. (2022). Global Terrorism Index 2022: Measuring the Impact of Terrorism. In *Quantifying Peace and its Benefits*. http://visionofhumanity.org/resources (accessed 24.08.2022 r.).

Pierce, D. A. (1979). R^2 measures for time series. *Journal of the American Statistical Association, 74*(368), 901. https://doi.org/10.2307/2286421

Sawa, B. A., & Ibrahim, A. A. (2011). Forecast models for the yield of millet and sorghum in the semi arid region of Northern Nigeria Using dry spell parameters. *Asian Journal of Agricultural Sciences, 3*(3), 187–191.

Shabri, A., Samsudin, R., & Yusoff, Y. (2020). Combining deep neural network and fourier series for tourist arrivals forecasting. In: *IOP Conference Series: Materials Science and Engineering*, *864*(1). https://doi.org/10.1088/1757-899X/864/1/012094

Solomon, O. C. (2020). Governance, ethnicity and response to conflicts: Deficit to sustainable development in Nigeria. *OIDA International Journal of Sustainable Development, 13*(2), 93–100.

Talbot, P. W., Rabiti, C., Alfonsi, A., Krome, C., Kunz, M. R., Epiney, A., Wang, C., & Mandelli, D. (2020). Correlated synthetic time series generation for energy system simulations using Fourier and ARMA signal processing. *In International Journal of Energy Research, 44*(10), 8144–8155. https://doi.org/10.1002/er.5115

Wang, X., Xie, N., & Yang, L. (2022). A flexible grey Fourier model based on integral matching for forecasting seasonal PM2.5 time series. *Chaos, Solitons and Fractals, 162*, 112417. https://doi.org/10.1016/j.chaos.2022.112417

Wilson, B. T., Knight, J. F., & McRoberts, R. E. (2018). Harmonic regression of Landsat time series for modeling attributes from national forest inventory data. *ISPRS Journal of Photogrammetry and Remote Sensing, 137*, 29–46. https://doi.org/10.1016/j.isprsjprs.2018.01.006

Xie, H., You, L., & Takeshima, H. (2017). Invest in small-scale irrigated agriculture: A national assessment on potential to expand small-scale irrigation in Nigeria. *Agricultural Water Management, 193*, 251–264. https://doi.org/10.1016/j.agwat.2017.08.020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Salim Jibrin Danbatta[1] · **Ahmad Muhammad[2]** · **Asaf Varol[3,4]** · **Daha Tijjani Abdurrahaman[5]**

✉ Salim Jibrin Danbatta
salimdambatta@gmail.com

Ahmad Muhammad
ahmadmuhammad325@gmail.com; a.muhammad@qu.edu.qa

Asaf Varol
asaf-varol@utc.edu; asafvarol@maltepe.edu.tr

Daha Tijjani Abdurrahaman
dahatijjani@gmail.com

[1] Software Engineering Department, Faculty of Engineering and Natural Sciences, Uskudar University, Istanbul, Turkey

[2] College of Art and Sciences, Department of Physics and Material Sciences, Qatar University, Doha, Qatar

[3] Department of Engineering Management and Technology, College of Engineering and Computer Science, The University of Tennessee at Chattanooga, Chattanooga, TN, USA

[4] Computer Engineering Department, College of Engineering and Natural Sciences, Maltepe University, 34857 Maltepe, Istanbul, Turkey

[5] Department of Business Administration, Faculty of Management Sciences, National Open University of Nigeria, Abuja, Nigeria