MATEMATIKA, MJIAM, Volume 41, Number 3, 283–308

Enhancing Drought Forecasting Accuracy: A Wavelet-ARIMA Modelling Approach for Standardized Precipitation Index Data

[1,2]Basiru Yusuf* and [1]Ani Shabri

[1]Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia.
[2]Department of Statistics, Jigawa State Polytechnic-Dutse, PMB 7040, Dutse, Jigawa State, Nigeria.
*Corresponding author: yusuf.basiru@graduate.utm.my

Article history
Received: 26 November 2024
Received in revised form: 9 September 2025
Accepted: 24 September 2025
Published online: 15 November 2025

Abstract Droughts are periods of inadequate precipitation that have severe and diverse impacts on human societies and ecosystems. However, a reliable forecasting technique can improve drought monitoring and mitigate impacts. This research aimed to assess the predictive accuracy of a hybrid forecasting model, Wavelet-ARIMA (W-ARIMA), for drought forecasting, using the Standardized Precipitation Index (SPI) with a traditional ARIMA model as a benchmark. Monthly rainfall data from the Badeggi district in the north-central part of Nigeria, covering the period from January 1968 to December 2018, was utilized for the analysis. Subsequently, SPI values for various time scales (3, 6, 9, and 12) were computed. The Wavelet Transform was then employed to decompose the data series into L components, encompassing details and approximations from A to D ($A_n, D_1, D_2, ..., D_{L-1}$) of the SPIs, respectively. Furthermore, an ARIMA model was fitted to each of these details and approximations, and their sum constituted the forecasted W-ARIMA values. Evaluation based on the performance metric, the RMSE of W-ARIMA (0.4889, 0.4326, 0.3566, and 0.2177) while that of ARIMA (0.7771, 0.6667, 0.4648, and 0.3212), revealed that the hybrid model (W-ARIMA) consistently outperformed the ARIMA model among the metrics for all the SPI 3, 6, 9, and 12, respectively.

Keywords SPI, ARIMA, SARIMA, Drought forecasting, Wavelet transform, Wavelet-ARIMA.

Mathematics Subject Classification 62M10, 62P12, 37M10, 42C40, 86A40.

1 Introduction

Drought is a natural disaster that negatively impacts various aspects of human life and the ecosystem, including agriculture, water resources, socio-economic conditions, and policymaking [1, 2]. Recurrences of drought are unavoidable, occur randomly, and are only noticeable after extended periods of below-average precipitation. Although its characteristics vary from place

to place, it affects almost everyone in all climatic zones. As such, it is challenging to determine with precision when a drought will start and end [3, 4].

When drought is prolonged due to inadequate precipitation, unfavorable conditions such as climate change, rising temperatures, deforestation, and ocean temperature fluctuations are expected consequences [5]. These issues have attracted researchers from various fields, including hydrology, environmental science, meteorology, agriculture, and ecology [6]. As an integral part of climate variability, drought is one of the foremost natural disasters, impacting extensive regions and posing significant threats to the environment and human life [3]. Its unpredictable nature, coupled with variable characteristics across different locations, makes it challenging to predict the onset and duration of droughts. Given its far-reaching impacts on human societies and ecosystems, effective water resource planning and management are imperative. Addressing this challenge necessitates the continued enhancement of forecasting techniques to predict hydrological and meteorological variables, such as precipitation, as a proactive measure to mitigate the effects of drought [7].

Drought forecasting is crucial for implementing preemptive measures to mitigate the risks and impacts associated with droughts. The effectiveness of drought preparedness and mitigation strategies hinges on timely and accurate information about drought onset and progression. Such information is typically derived from continuous drought monitoring, which relies on the use of drought indices [8]. Drought indices are metrics designed to capture various drought characteristics, including magnitude, duration, severity, and spatial extent. The Standardized Precipitation Index (SPI) is the most widely used tool for measuring meteorological drought [9–11]. It was originally proposed by [12]. The SPI is favored for its simplicity and versatility [13], requiring only precipitation data to effectively monitor and predict drought conditions in a given region.

Research on drought forecasting has highlighted the effectiveness of hybrid models that combine wavelet analysis and autoregressive integrated moving averages (ARIMA) for improved predictions. Studies by various authors have shown that hybrid models like Wavelet-MLR, Wavelet-GPR, EWT-ARIMA, and Wavelet-ARIMA outperform standalone models in forecasting drought conditions [2, 14–20]. These hybrid models have been found to improve the accuracy of forecasting multi-scalar Standardized Precipitation Evapotranspiration Index (SPEI) values over different lead times, with the choice of mother wavelet significantly impacting the model's performance. The combination of wavelet analysis with ARIMA in these studies has demonstrated superior results in predicting drought events, providing valuable insights for decision-makers and planners in managing hydrological droughts and implementing early warning systems effectively.

Wavelet-ARIMA models have gained significant attention due to their effectiveness in handling non-stationary time series data. Several studies highlight the significance of integrating wavelets with other methods. [21] Wavelet Methods for Time Series Analysis provides a comprehensive overview of wavelet methods applied to time series analysis, laying the groundwork for integrating wavelets with ARIMA models. [22] combined wavelet transforms with neuro-fuzzy models, and the results demonstrated significant improvements in precipitation forecasting accuracy. [23] employed a hybrid ARIMA-neural network model; the study underscores the potential of combining traditional models with advanced techniques, inspiring subsequent wavelet-ARIMA research. [24] used wavelet analysis to identify climate oscillations, paving the way for applying wavelet transforms in climate-related time series forecasting. Recently, several

studies have employed hybrid models to improve drought forecasting accuracy, including but not limited to [2, 25–28].

In trying to mitigate the consequences of drought, advanced actions and plans need to be put in place. To this end, several forecasting methods have been developed to predict drought occurrences. Among the statistical methods is the conventional autoregressive integrated moving average (ARIMA), which has been widely applied by many researchers to forecast drought using the SPI series [29–31]. ARIMA models are linear and assume stationary data sets; therefore, they are inadequate for handling non-linear and non-stationary series in hydrological processes [29].

The primary aim of wavelet transformation is to analyze time series data in both the time and frequency domains by breaking down the original time series into different frequency bands. Wavelets serve as essential tools in time series forecasting, utilizing wavelet functions for this purpose. Mathematically, wavelet transforms are functions that enable the analysis of time series containing non-stationarities. This technique facilitates the use of long-time intervals to capture low-frequency information and short-time intervals to capture high-frequency information, thereby revealing data characteristics such as trends. Additionally, wavelet analysis offers the flexibility to select the mother wavelet based on the specific attributes of the investigated time series. Furthermore, the wavelet transform provides the advantage of examining diverse independent behaviors at distinct time scales [32].

The aim of this research (paper) is to investigate the performance of predictive accuracy of a proposed hybrid wavelet-autoregressive integrated moving average (W-ARIMA) model for drought forecasting using SPI data series with the traditional ARIMA model as the benchmark. Time scales consisting of 3, 6, 9, and 12 periods were utilized in the SPI for the drought analysis. These intervals were selected because they are widely recognized in the literature for capturing different types of drought impacts: shorter scales such as 3- and 6-month SPI reflect short-term precipitation variability relevant to short term meteorological droughts, while longer scales such as 9- and 12-month SPI are more effective in classifying long-term drought conditions. The adoption of these time scales is consistent with recommendations by [13, 33] and the World Meteorological Organization, thereby ensuring comparability with previous studies and providing a comprehensive assessment of drought dynamics in the study region.

2 Study Area and Data

With a total area of 923,768 km^2 (356,669 sq mi), Nigeria ranks as the 32nd largest country globally and is situated on the Gulf of Guinea in Western Africa. The country's diverse landscape includes the Sahara Desert to the north, dense forests with numerous rivers and streams in the south, and fertile land in the center. Nigeria is home to an estimated 220,000,000 people and is bordered by four countries: Cameroon to the east, Chad and Niger to the north, the Benin Republic to the west, and the Atlantic Ocean to the south. Niger State experiences an average annual temperature of 34°C, contributing to its classification as an extremely warm state. While there are not many traditionally tropical or hot months, the year remains consistently warm or hot, with occasionally uncomfortably high humidity from June to September. Figure (1) shows the distribution of the study area by state and location.

The dataset used in this study comprises rainfall data sourced from the Badeggi irrigation station in Niger State, located in the north-central part of Nigeria, spanning 600 entries

from January 1968 to December 2018. The data was divided into a training set consisting of 98% (1968 to 2017) and a validation set comprising 2% (January - December 2018) for both ARIMA and Wavelet-ARIMA modelling. This data was then transformed into a Standardized Precipitation Index (SPI) series, serving as the foundation for model construction and drought forecasting. For model development and validation, MATLAB software was used for the decomposition of the various SPIs, and R software was utilized for ARIMA model development and validation.

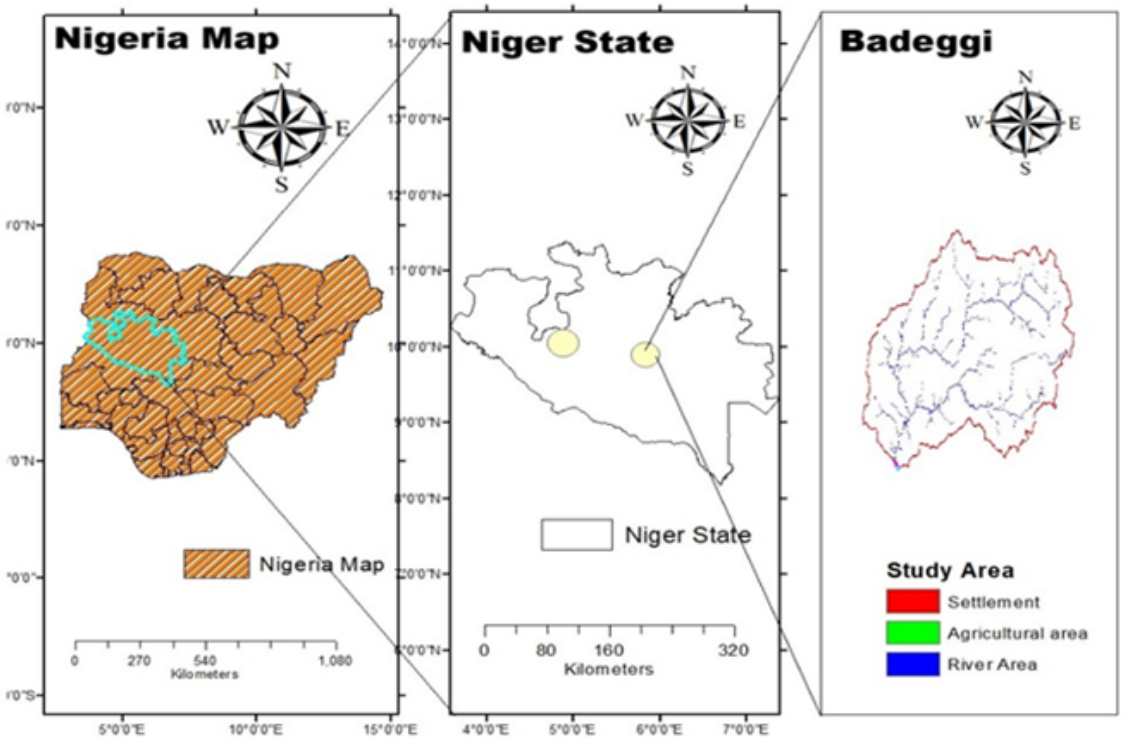

Figure 1: Spatial location map showing Nigerian Map by States, Niger State the Study State and Badeggi Study Area. The map was generated using ArcGIS version 10.8 (Esri Inc.,`https://www.esri.com`).

3 Methods

3.1 Introduction

To address the inadequacies of ARIMA models, specifically their limitations in handling non-linear and non-stationary data series typical of drought data, a model integrating wavelet and ARIMA, called Wavelet-ARIMA, is proposed to enhance forecasting accuracy. This section will present the data source along with its size. The methods used to model and forecast the data are ARIMA and SARIMA, and the performance evaluation metrics - RMSE, MAE, and MAPE - will also be presented.

3.2 Standardized Precipitation Index (SPI)

The Standardized Precipitation Index (SPI) was originally developed by [33] with the primary purpose of defining and monitoring droughts in the United States. It is a simple, basic drought index used to identify drought in many regions. Due to its characteristics, it offers several advantages. First, it is easy to evaluate since its values are derived solely from precipitation [34]. Second, it can be used to represent drought conditions in a region over different time scales [34–36]. Third, the standardization of SPI makes it suitable for comparing drought conditions across different regions, times, and climates; it is geographically independent and can describe both dry and wet periods. The details of its computational procedure can be found in [34].

The value of SPI is calculated by fitting a density of Gamma Probability distribution given in equation 1:

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{\alpha}{\beta}} \text{ for } x,\alpha,\beta > 0, \tag{1}$$

where, α is the shaping parameter, β is the scaling parameter and,x is the amount of precipitation which are all positive integers $(\alpha > 0; \beta > 0, x > 0); \Gamma(\alpha)$ is the value of alpha obtained over the integral of the gamma function as presented in equation (2).

$$\Gamma(\alpha) = \lim_{n\to\infty} \prod_{v=0}^{n-1} \frac{n! n^{y-1}}{y+v} \equiv \int_0^{\infty} y^{\alpha-1} e^{-y} dy. \tag{2}$$

SPI has no dimension; its values range from with the negative ending indicating extreme drought and the positive ending representing an extremely wet period. The classification of SPI based on drought intensity is given in Table 1. The effect of different time scales on the SPI lies in their ability to represent distinct drought characteristics. Shorter time scales, such as the 3- and 6-month SPI, are sensitive to short-term precipitation variability and are therefore effective in detecting meteorological and agricultural droughts, which directly influence soil moisture and crop conditions. In contrast, longer time scales, such as the 9- and 12-month SPI, smooth short-term fluctuations and highlight hydrological droughts, reflecting cumulative precipitation deficits that affect groundwater recharge, reservoir storage, and streamflow.

Table 1: Classification of Drought based on SPI values

SPI Values	Class of Drought	SPI Values	Class of Drought
0 to -0.99	Mildly dry	0 to -0.99	Near Normal
-1.0 to -1.49	Moderately dry	1.0 to 1.49	Moderately wet
-1.50 to -1.99	Severely dry	1.50 to 1.99	Severely wet
$\leqslant$ -2.0	Extremely dry	$\leqslant$ 2.0	Extremely wet

3.3 Autoregressive integrated moving average (ARIMA) model

As reported by [29–31, 34, 37] the autoregressive integrated moving average (ARIMA) model is the most widely used linear model stochastic model for drought forecasting. [38] described the general form of the non-seasonal ARIMA model for any time series X_t, represented by:

$$X_t = \frac{\theta(B)\varepsilon_t}{\phi(B)\nabla^d} \tag{3}$$

where:

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - ...\phi_p B^p), \tag{4}$$

and

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - ...\theta_q B^q), \tag{5}$$

where X_t is the value of the observed time series, B is the backshift operator, p is the order of the autoregressive,q is the order of the moving average, $\phi(B)$ and $\theta(B)$ are polynomials of order p and q respectively,d is the number times of differencing, ∇^d differencing operator and ε_t is the time-independent uncorrelated random variables assumed to be white noise.

3.4 Seasonal Autoregressive Integrated Moving Average SARIMA

While ARIMA models are best suited for modelling a wide range of linear and stationary data, many seasonal and non-stationary data patterns exist in hydrological processes. To handle these seasonal and periodic properties, [39] improved the ARIMA model by including a seasonal component to capture the seasonal characteristics that may exist in the data. This modified model, called the Seasonal Autoregressive Integrated Moving Average (SARIMA), is written as ARIMA $(p, d, q)(P, D, Q)s$. For a given time series X_t, the SARIMA model can be written as follows:

$$X_t = \frac{\theta_q(B)\Theta_Q(B^s)\varepsilon_t}{\phi_p(B)\Phi_p(B^s)\nabla_d\nabla_s^D} \tag{6}$$

where:

$$\Phi_P(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - ...\Phi_p B^{ps}), \tag{7}$$

$$\Theta_Q(B^s) = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - ... - \Theta_q B^{qs}), \tag{8}$$

From the above equations [6–8], Where X_t is the value of the observed time series, s is the seasonal length,(p, d, q) is the Non-seasonal part of the model,(P, D, Q) is the seasonal component of the model, ∇_s^D is the seasonal differencing part with degree D,$\phi(B)$ and $\theta(B)$ are polynomials of order p and q respectively, ∇^d differencing operator, and are the time-independent uncorrelated random variables assumed to be white noise. The seasonal part of the model is made up of terms that are very similar to the non-seasonal components of the model, however, they involve backshifts of the seasonal period.

3.5 ARIMA Model Development

The important steps in the implementation of ARIMA/SARIMA modelling include model identification and the selection of suitable models for time-based data. Three essential steps make up the expansion of the ARIMA/SARIMA structure: model identification, parameter estimation, and diagnostics checking to confirm the model's suitability. Automatic functions, the Hyndman-Khandakar algorithm for automated ARIMA modelling in R software, is used to complete these phases. The sections that follow provide more details on these phases [40–43].

3.5.1 Model Identification

The main aim of model identification is to meticulously choose a subset from the SARIMA/ ARIMA model family that accurately captures the inherent characteristics of the time series data. According to [44], the identification phase entails the initial step of defining a model structure that closely corresponds to the observed data collected. The specific configuration of the model is ascertained through a two-step procedure:
(1) The examination of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the data transformation to identify the temporal correlation structure [29].
(2) The application of appropriate differencing to the series, if necessary, to attain stationarity, which is subsequently employed to delineate the overall form of the model for fitting purposes.

Studies have indicated that estimates with longer delays lack statistical reliability [45]. The PACF helps to diminish the impact of short-lag autocorrelation, thus decreasing correlation estimates at extended delays. The PACF quantifies the correlation not explained by lower-order delays between a variable and its lag. By employing goodness of fit, the precision of the ultimate model will be evaluated using established metrics such as the Akaike information criterion (AIC) and the Schwarz Bayesian criterion (BIC) which are given in equation (19) and (20) respectively [46].

$$AIC = -2log(L) + 2k, \tag{9}$$

$$BIC = -2\log(L) + k\ln(L). \tag{10}$$

The illustration of the number of parameters within the model can be found at the point where the probability function of the SARIMA/ARIMA model is situated, in conjunction with the quantity (n) of data within the model. Opting for the model that exhibits the lowest AIC and BIC values is recommended as the most optimal choice.

3.5.2 Parameter estimation

For estimating unknown parameters various techniques are employed which comprise least squares, moments, and maximum likelihood. In this specific investigation, the approach delineated in [47], by employing the method of maximum likelihood estimation that will estimate the unknown parameters in the model. One advantageous characteristic of the maximum likelihood technique is its flexibility; in contrast to other methodologies such as the Cochrane-Orcutt approach, it can be utilized in scenarios where the autocorrelation pattern of the error is more intricate than basic first-order autoregressive structures [48].

3.5.3 Diagnostic Checking

After the appropriate model has been chosen and the unknown parameters have been estimated, it is critical to evaluate the model's suitability and, if needed, recommend modifications. This phase's main goal is to confirm that the model's residuals show characteristics of independence, have a normal distribution, and are homoscedastic. These standards are essential for verifying that the current time series model is adequate. It is recommended to use residual plots and various examinations to achieve this. A white noise process serves as a sign that the model is adequate; if the residuals follow a white noise process, which is defined by a distribution around zero and a lack of correlation, then the model is suitable.

Analyzing the residuals of ACF and PACF is one way to assess the series' independence. There is no significant link between the residuals if the ACF and PACF residuals fall inside the confidence interval. To achieve this goal, an alternate method that can be used is the Ljung-Box-Pierce (LBQ) test. One statistical test designed to determine residual independence is the LBQ test. The test statistic is formulated in equation (11).

$$Q = n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k}. \tag{11}$$

In this case, n is the number of samples of data and the number of autocorrelation lags, and r_k is the sample autocorrelation at *lag* k. The LBQ statistic's null hypothesis is that the residuals are independent. Q is distributed approximately as a chi-square with $m-p-q$ degrees of freedom. If the derived value of Q is excessively large, it suggests that the model is unsatisfactory.

3.6 Wavelet analysis

Recently, researchers have turned their attention to the use of wavelets due to their ability to show information within a given signal in both time and scale domains [49]. Wavelet is a mathematical tool that was developed for digital signal processing and image compression. Wavelet is defined as a mathematical function used in decomposition of original signal usually in the time domain to various scales of interest for processing and or analysis [50].

The wavelet transform is the procedure that decomposes the original data into various components given a time scale. It is categorized into discrete wavelet transform (DWT) or continuous wavelet transform (CWT). The latter is not commonly used due to its computational compatibility and time requirement [51]. Hence, the DWT which is given in equation (12) is frequently used in time series forecasting due to its simplicity and fewer time requirements.

$$\psi_{m,n}(t) = \frac{1}{\sqrt{s_o^m}}\psi\left(\frac{t - n\tau_o s_o^m}{s_o^m}\right), \tag{12}$$

where, $\psi(t)$ is the mother wavelet, and m and n are integer values that control the scale and time respectively. The usual choices for the parameters are $S_0 = 2$ and $\tau_0 = 1$. However, Mallat's wavelet transform states that the inverse of the discrete wavelet transform (DWT), given in equation (12), is used to decompose the original discrete time series X_t into linearly independent components at various scales of approximation and detail signals [52]. The inverse is expressed in equation (13):

$$x(t) = T\sum_{m=1}^{M}\sum_{t=0}^{2M-m-1} W_{m,n}2^{-\frac{m}{2}}\psi(2^{-m}t-n), \tag{13}$$

where: $W_{m,n} = 2^{-\frac{m}{2}}\sum_{t=0}^{N-1}\psi(2^{-m}t-n)x(t)$ represents the wavelet coefficient for the discrete wavelet transform at scale $s = 2^m$ and $\tau = 2^m n$.

To achieve the necessary level of decomposition [3, 6, 9, 12, ...] of an original data set into wavelets, [53] proposed the following formula, presented in equation (14):

$$L = \text{int}[\log(N)]. \tag{14}$$

In this formula,L represents the level of decomposition, and N denotes the total number of elements in the SPI data. The original SPI drought data series is decomposed into L components, comprising both details and approximation from A to $D_{L-1}(A_n, D_1, D_2, ...D_{L-1})$ These components correspond to different frequency elements of the original data, each playing a unique role and impacting the original SPI drought data differently. The sum of these components $(A_n, D_1, D_2, ...D_{L-1})$ is equal to the original SPI, as presented in equation (15). Rather than using the components individually as input for the model, we combine the appropriate components. This combined approach is more effective and enhances the forecast performance of hybrid models.

$$S_n = A_n + D_1 + D_2 + ... + D_{L-1}. \tag{15}$$

3.7 Wavelet-ARIMA

The combination of the Wavelet Transform (WT) with the Auto-Regressive Integrated Moving Average (ARIMA) model has gained prominence for its effectiveness in predicting phenomena such as wind speed, temperature, and drought events. The hybrid method involves integrating wavelet decomposition and empirical mode decomposition to stabilize original time series data into various components comprising an approximation (A_n) and details $(D_{n-1}, ..., D_1)$ representing low and high frequencies patterns of the data. This is followed by the application of ARIMA models to each of the decomposed subseries for model selection and predictions; and subsequently summing up the fitted values from each of the fitted ARIMA made up of the forecast wavelet-ARIMA.

Notice that these forecasted SPI values are obtained based on the inverse transform using wavelets, leveraging the strength of both the wavelet analysis and the ARIMA Methods in improving the predictive accuracy for modelling complex phenomena like drought. In essence, wavelet-ARIMA involves three (3) fundamental stages: decomposition of data, fitting ARIMA Model, and construction of signals to obtain Wavelet-ARIMA forecasts. The wavelet-ARIMA framework is presented in Figure 2.

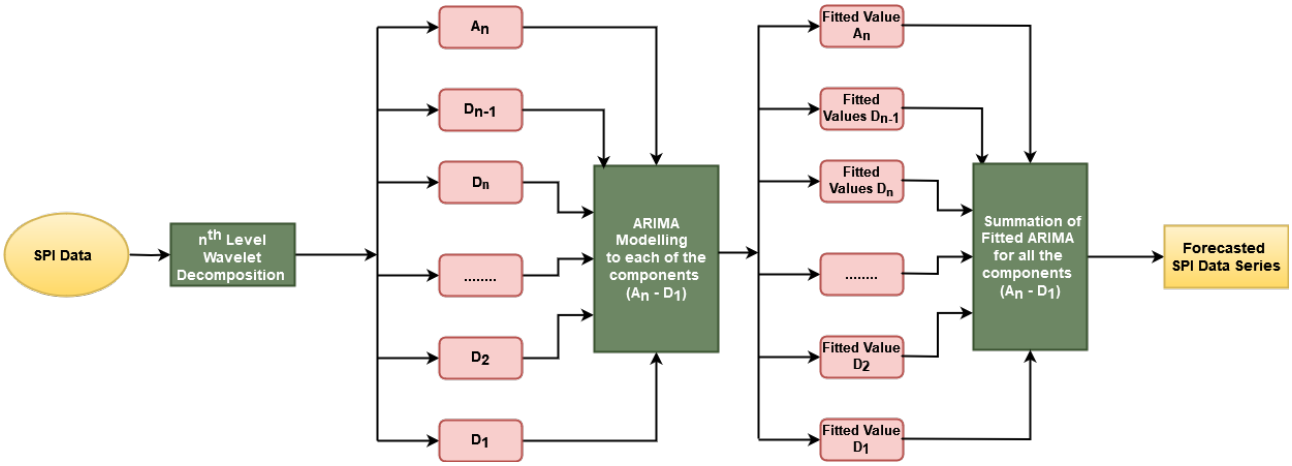

Figure 2: Wavelet – ARIMA Framework

1) The first stage is to apply wavelet transform to decompose the SPI values using MATLAB software. The process enables the data to be transformed to appropriate approximate and detailed component levels capturing the low and high frequency patterns of the data. Selecting a suitable decomposition method is essential in wavelet analysis. Hence, scholars suggested various methods which include but are not limited to Symlet, Meyer, Daubechies, and Morlet. For all these methods, the type of mother wavelet is dependent on the characteristic of the data [54], [55]. For this study, therefore, the Daubechies function of order 2 and decomposition level 3 (which correspond to integer log (600)) were used, as describe in equation (16).

$$f(t) = D_1 + D_2 + D_3 + A_3. \tag{16}$$

The choice of the Daubechies wavelet is justified because of its property of maximal number of vanishing moments and has superior ability to model complex data, its orthogonality and hence beneficial in denoising signals typical of drought data. More so, its flexibility in handling various applications, makes it an ideal candidate for enhancing the accuracy of drought forecasting models.

2) The optimal ARIMA models are fitted into each deconstructed layer for each SPI series in the second phase. All of the iterative processes outlined in Section (3.5) are used to create a unique model for every decomposed component in order to arrive at the appropriately selected ARIMA model.

3) Thirdly, equation (17) will be used to rebuild the signal utilizing these expanded and decomposed signals on various scales. This is achieved by summation of all subseries projections of each deconstructed layer for each SPI, hence it yields the anticipated value of W-ARIMA.

$$\hat{y} = \hat{D}_1 + \hat{D}_2 + \hat{D}_3 + \hat{A}_3. \tag{17}$$

3.8 Performance Evaluation Criteria

To ascertain the performance of the forecasting accuracy of the estimated hybrid models, we employ the use of evaluation metrics; which include the RMSE defined as the standard deviation of the model's forecasted values and the mean absolute error MAE that computes how closely

the forecasted values with the observed values; while the mean absolute percentage error MAPE measure how accurate the forecasted values are. These measures are defined in the equations [18–20] respectively. Details of these measures of performance evaluation metrics are details discussed in the literature: [56–60]. By these parameters, superior model performance is achieved by choosing any metric with a lower value.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y_i})^2}, \tag{18}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y_i}|, \tag{19}$$

$$MAPE = \frac{1}{n}\left|\frac{Y_i - \hat{Y_i}}{Y_i}\right| \times 100. \tag{20}$$

For equations (18) to (20), the number of sample data is symbolized as n, where Y signifies the recorded data and indicates the predicted values.

4 Results and Discussions

4.1 The preliminary Data Analysis

Preliminarily, a plot of different values of all the SPI was constructed and presented in Figure 3. This is to observe any underlying pattern contained in the various time series. Examining the plots, we can examine that there are very good number of values that shows greater portion are moderately dry i.e., the drought occurrences is noticed and can lead to measures, there are however for all the SPIs an extremely dry values occurring consistently during the 1993 (100 month) and for about 175 months in SPI 3 and 6. This is justified by the drought events experienced in the area in the early 1990s. onward. The drought index value exhibited is not surprising due to susceptible climate variation of Badeggi area, and that has significantly affected the agricultural and other livelihood activities.

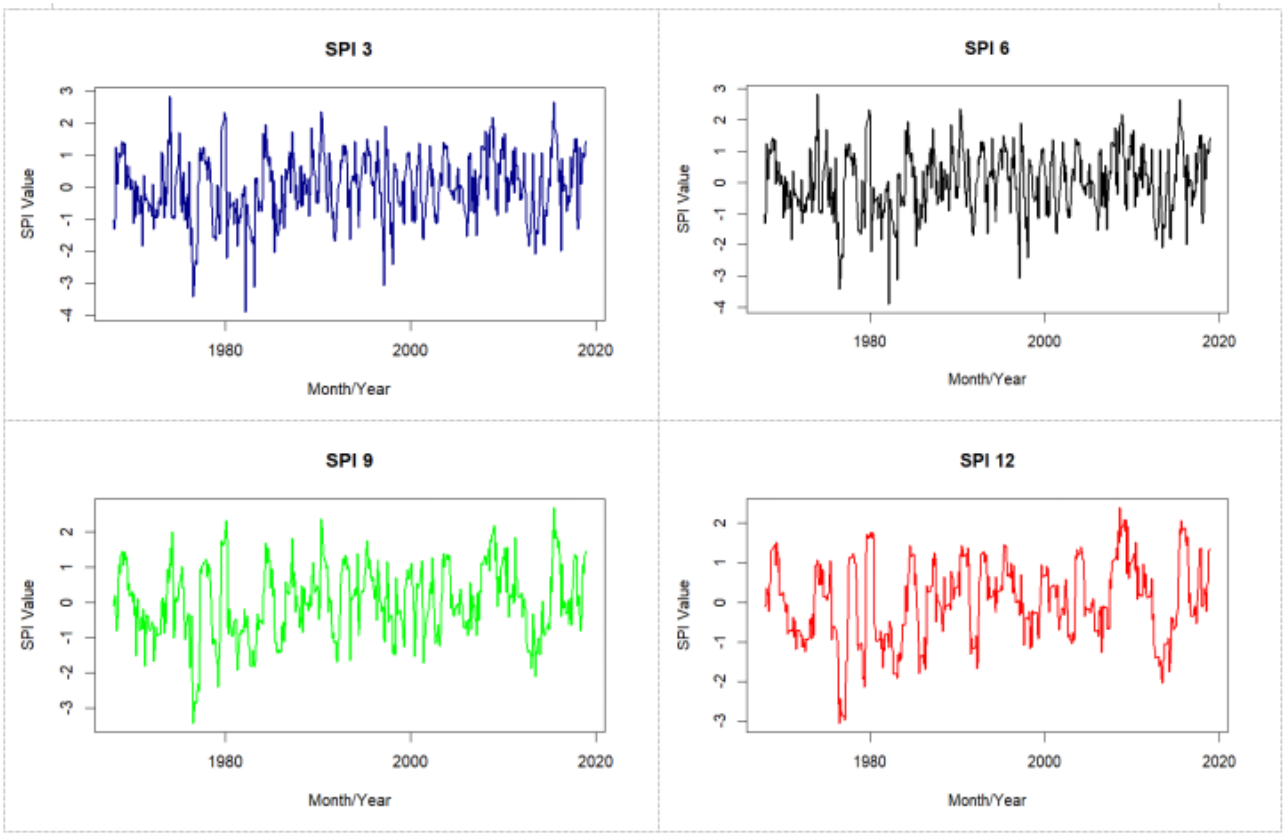

Figure 3: Plots of 600 Data Sets for Various SPIs (3,6,9 and 12) Used for the Study

4.2 The ARIMA Model

The three model development stages, the model identification, parameter estimation, and diagnostic checking of modelling ARIMA are executed herein the following subsections.

Model Identification

The identification stage is the initial phase of ARIMA modelling, during which the model's order is established. To do this, the autocorrelation function (ACF) and the Partial autocorrelation function (PACF) of the original SPI data series are observed and studied. Figure 4 displays the ACF and PACF charts for each SPI. The model parameters p, d and q values are frequently determined based on the plots' significant lags.

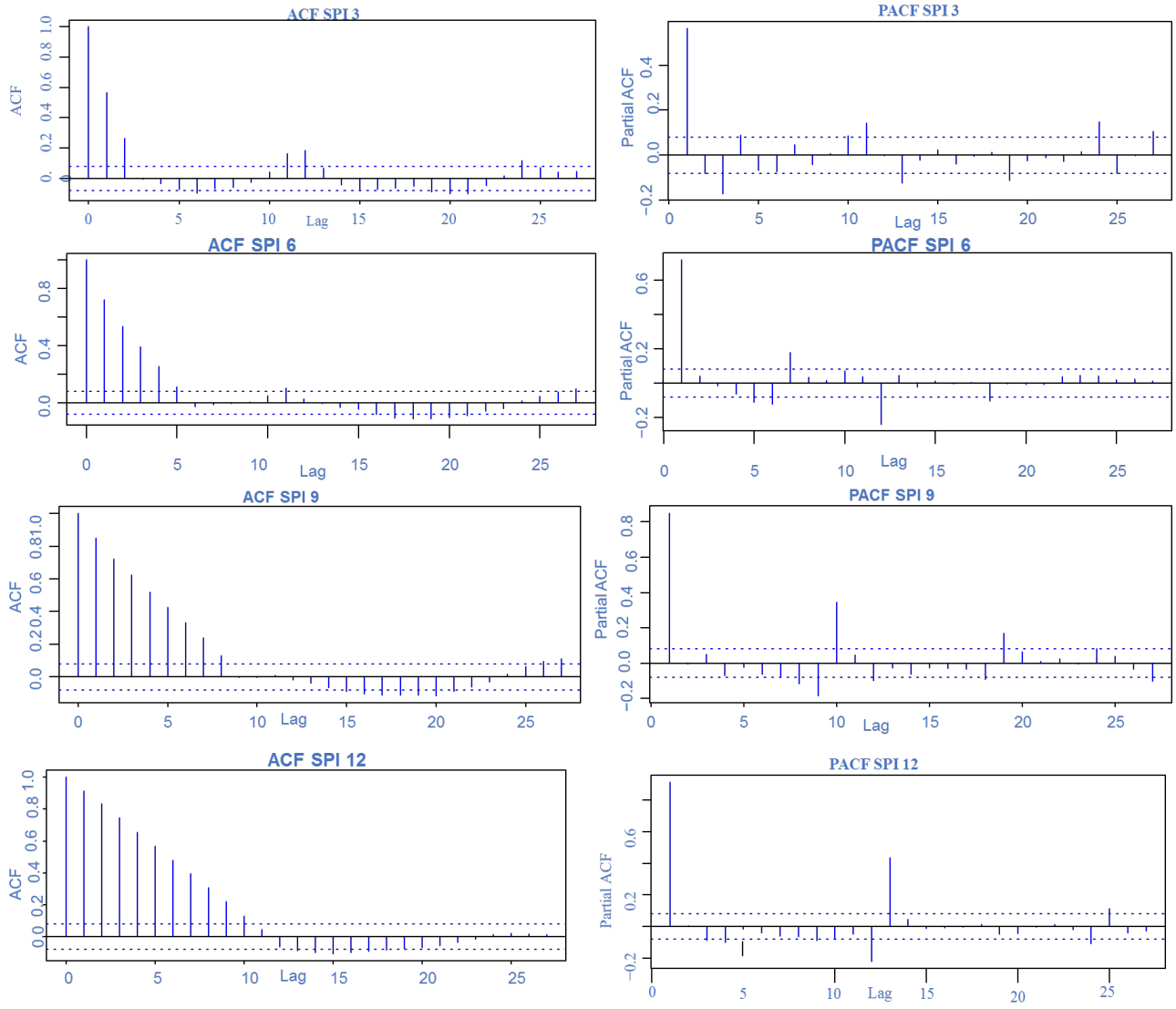

Figure 4: Plots of ACF and PACF Utilized for Models' Selection in SPI 3, 6, 9, 12

In ARIMA modelling, ensuring that the data is stationary is crucial [61]. Before estimating parameters, it's common to assess both stationarity and seasonality in the time series. One can check for stationarity by looking at the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. If the time series is nonstationary, techniques like differencing can be used to achieve stationarity. To identify the correct order for the ARIMA model, ACF and PACF plots are also utilized. To find the best-fitting model, researchers rely on goodness-of-fit measures such as the Akaike Information Criterion (AIC) and the Schwarz-Bayesian Criterion (SBC). The model with the lowest AIC and SBC values is typically considered the best option [46, 49, 62].

Parameter Estimation

The second stage of ARIMA Modelling involves the estimating of parameters of the various identified models. After identifying the most suitable ARIMA model for different temporal scopes of the Standardized Precipitation Index (SPI), the next step involved estimating the model parameters. The maximum likelihood method of estimation is employed for estimating the unknown parameters in the model due to its flexibility [47]. The results of these estimations are presented in Table 2, which includes the estimated parameters, standard error, t-statistic, and p-value for each SPI. It is worth noting that a significant proportion of the parameters show a lower standard error than the parameter value itself. Furthermore, except SPI 9, all p-values demonstrate statistical significance, suggesting that the estimated parameters are valid for inclusion in the models.

Table 2: Summary of Selected ARIMA Model Parameters for All the SPIs

Fund	Model Parameters	Variables contained in the Model			
		Estimated Parameter	Standard Error	t-Statistic	p-Value
SPI 3	ψ_1	0.6263	0.03377	18.5846	0.0000
	θ_1	-0.2591	0.0433	-5.9838	0.0000
	ε	0.144	0.0635	2.2677	0.0236
SPI 6	ψ_1	1.1549	1.3018	0.8872	0.3752
	ψ_2	-0.2593	1.0624	-0.2441	0.8072
	θ_1	-0.4394	1.2966	-0.3389	0.7348
	θ_2	-0.0184	0.1459	-0.1261	0.8997
	θ_3	0.3109	0.1687	1.8429	0.0656
	θ_4	-0.6247	0.1542	4.0512	0.0001
SPI 9	ψ_1	0.9136	0.0173	52.80925	0.0000
	ψ_2	-0.8453	0.1336	-6.3271	0.0000
	θ_1	0.3208	0.1291	2.484895	0.0130
	θ_2	-0.5231	0.0619	-8.45073	0.0000
SPI 12	ψ_1	0.9768	0.0090	108.5333	0.0000
	θ_1	-0.8215	0.0243	-0.8458	0.3977

Diagnostic Checking

Following identification and evaluation, the models are chosen based on the forecasting accuracy results for the different SPIs that were calculated. Table 3 presents the summaries of the models that were chosen based on the performance evaluation metrics that showed the lowest performance and AICs. All four (4) SPIs, however, exhibit seasonal data series, according to the ARIMA Model's results; SPI 6 differs from the others with AR(2) and MA(2), respectively.

Table 3: Summary of the Selected Best ARIMA Models and Metrics for the Various SPIs

		TRAINING		TESTING			
DATA	**MODEL**	**RMSE**	**MAE**	**RMSE**	**MAE**	**AIC**	**BIC**
SPI-3	SARIMA(1,0,0)(0,0,1)$_{[3]}$	0.7771	0.6030	0.8568	0.7504	1380.72	1398.23
SPI-6	SARIMA(2,0,2)(1,0,1)$_{[6]}$	0.6667	0.479	0.5027	0.4116	1207.4	1238.04
SPI-9	SARIMA(1,0,0)(1,0,2)$_{[9]}$	0.4648	0.3367	0.4806	0.3906	782.62	804.5
SPI-12	SARIMA(1,0,0)(0,0,1)$_{[12]}$	0.3212	0.2191	0.2277	0.1735	353.59	366.72

4.3 Assessing Model Adequacy

The final stage in ARIMA modelling involves assessing the adequacy of the model. After selecting an appropriate model and estimating its parameters, it is critical to evaluate how well the model fits the data and to propose adjustments if necessary. This evaluation aims to ensure that the residuals of the model exhibit key characteristics such as independence, normal distribution, and homoscedasticity.

To conduct this assessment, various diagnostic tests are employed, including the analysis of residual plots, and presented in Figure 5. An indication of model adequacy is when the residuals behave like white noise, characterized by a mean centered around zero and an absence of correlation among them. Standard practices for evaluating the goodness-of-fit for ARIMA models involve examining the Partial Autocorrelation Function (PACF) and Autocorrelation Function (ACF) of the residuals. In this context, the observed residuals appear to be normally distributed, which supports the conclusion that the model is adequate for the data at hand.

4.4 Proposed W-ARIMA Model

To enhance forecasting accuracy, a combined model named the Wavelet-ARIMA (W-ARIMA) model, integrating a wavelet and ARIMA, is introduced to address the limitations of standalone ARIMA models when dealing with nonstationary data. In this approach, the DWT is applied to decompose the SPI series, generating suitable approximate and detailed components. Inverse wavelet transform is subsequently used to reconstruct the decomposed series. Optimal ARIMA/SARIMA models are fitted to each of the decomposed elements, and the W-ARIMA predictions are obtained by summing the forecasted values from all the decomposed components.

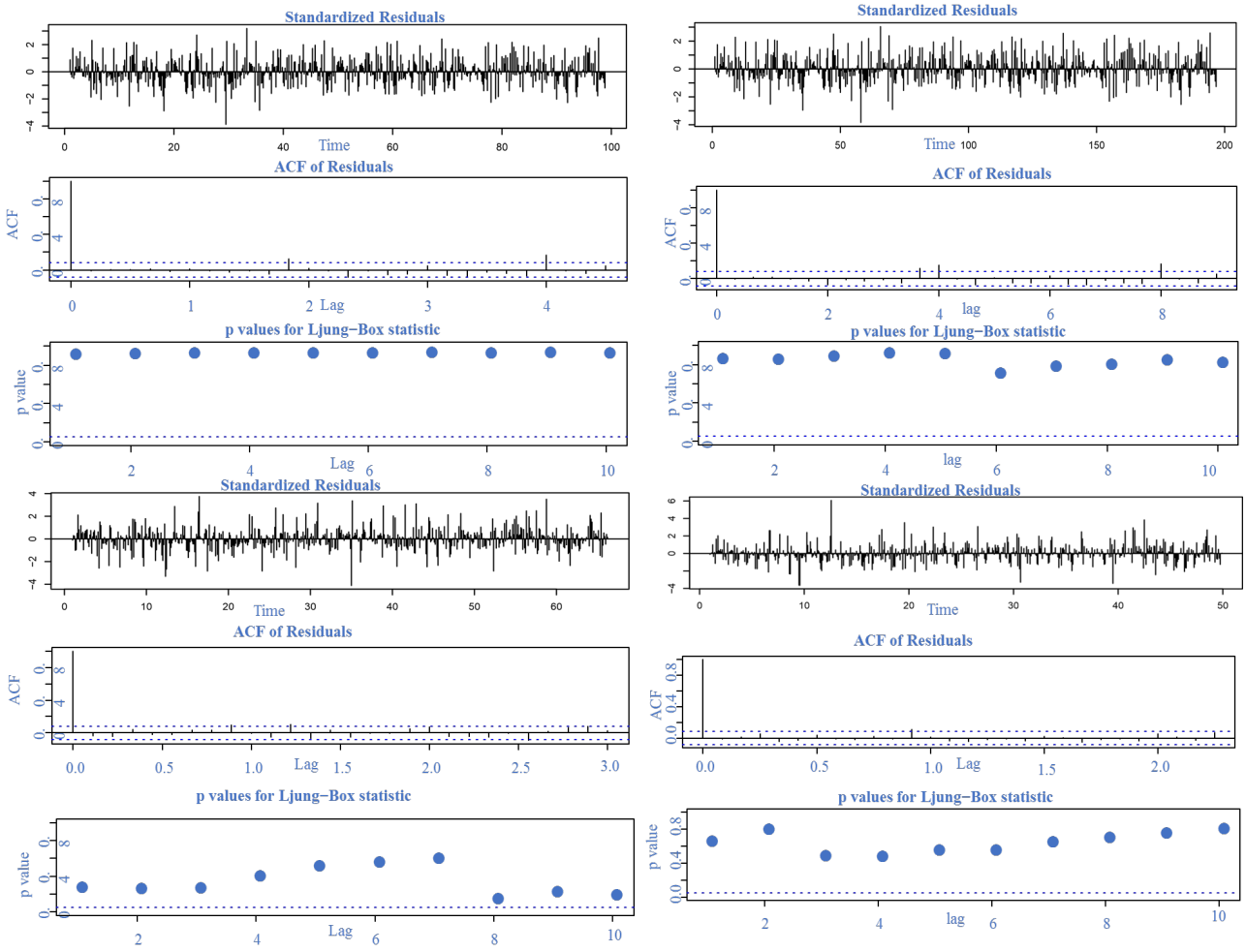

Figure 5: Plots of Standardized Residual, ACF, and P-values of the Residual for the SPIs

Each SPI series undergoes a six-level decomposition utilizing the DWT with the application of the Daubechies function of order 2 (db2). This decomposition results in the extraction of components including an approximate representation (A_3) and various detail components (D_3, D_2, and D_1). Figure 6 illustrates the decomposition of SPI 6 into approximate and detailed components.

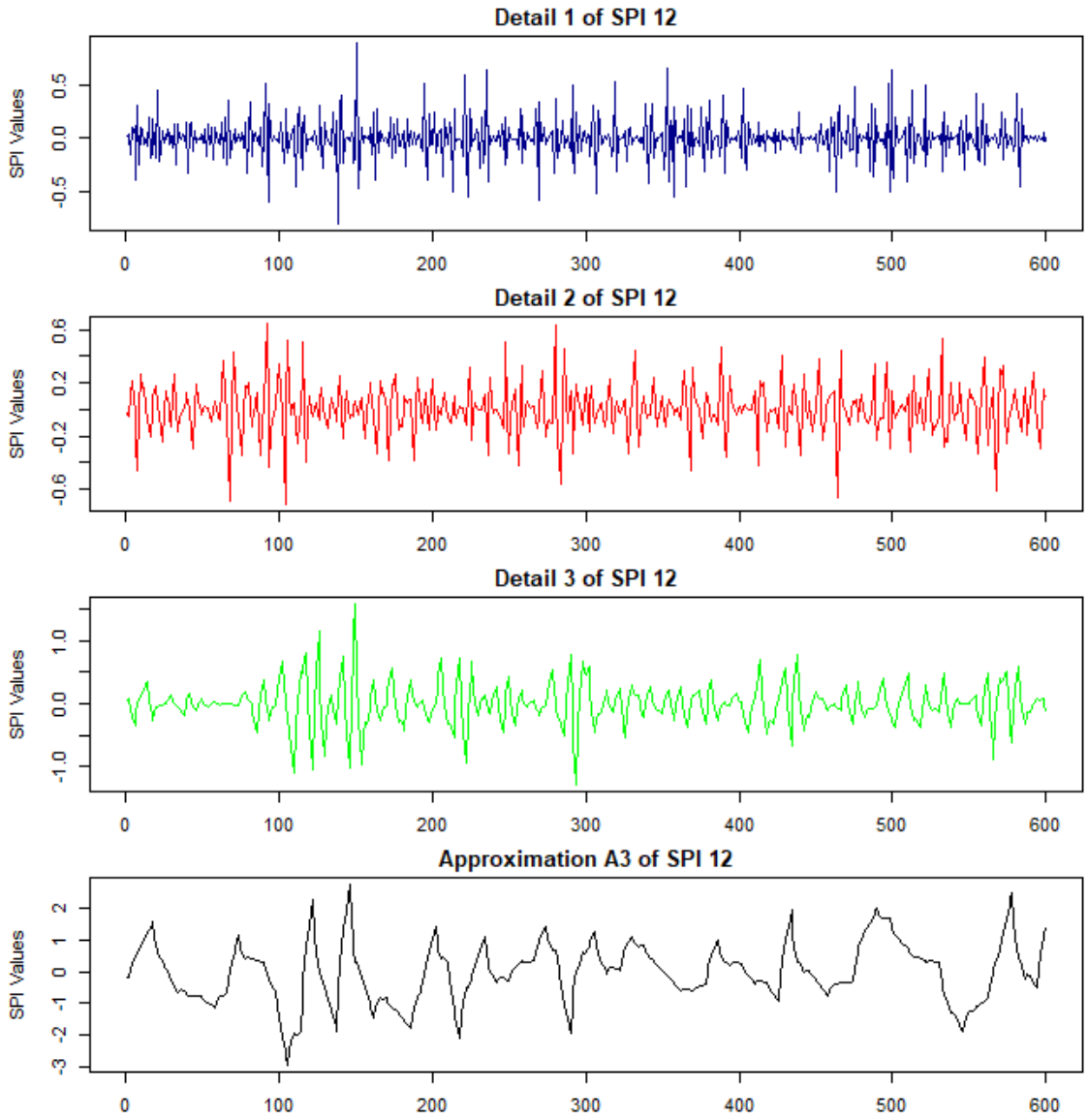

Figure 6: Plots of SPI 12-Time Series DWT Decomposition of Different Levels and Approximation (A_3, D_3, D_2, and D_1)

The proposed W-ARIMA model is easily derived by aggregating the predicted values from each decomposed layer, utilizing the six iterative stages previously elucidated, along with the corresponding ARIMA model for each constituent subseries. Various ARIMA/SARIMA models were applied to each decomposed SPI (A_3, D_3, D_2, and D_1) timescales. The summary of optimal models based on AIC, BIC, and. The summary of optimal models based on AIC, BIC, and p-values is presented in Table 4. The most suitable model for each original SPI is identified by selecting the models with the lowest AIC and BIC values.

Table 4: Summary of the Selected W-ARIMA Models

SPI	**Model**	**Testing Set**			
		X^2	**p-value**	**AIC**	**BIC**
SPI 3	ARIMA(2,0,2) with zero mean	0.0403	0.8408	-3478.3	-3456.45
SPI 6	ARIMA(2,0,2) with zero mean	0.0356	0.8503	-3012.55	-2990.66
SPI 9	ARIMA(2,0,2) with zero mean	0.0842	0.7717	-2622.43	-2600.54
SPI 12	ARIMA(2,0,2) with non-zero mean	0.1329	0.7154	-2377.37	-2351.11

The forecasting ability of the ARIMA model was improved upon by the application of Wavelet analysis with respect to all the SPI as shown in Table 5 containing the values of performance metrics of root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). From the values, it can also be observed and conclude that the application of wavelet has improved the forecasting performance of the model.

Table 5: Summary of the Statistical Performance Metrics (SPM) of W-ARIMA Models

SPI	**Model**	SPM		
		RMSE	MAE	MAPE
SPI 3	ARIMA(2,0,2) with zero mean	0.4889	0.3715	53.4192
SPI 6	ARIMA(2,0,2) with zero mean	0.4326	0.3236	46.4089
SPI 9	ARIMA(2,0,2) with zero mean	0.3566	0.2607	35.5696
SPI 12	ARIMA(2,0,2) with non-zero mean	0.2177	0.1551	19.9942

All of the SPIs had the characteristics of a white note, according to the LQT test result shown in Table 4, with very small x-squared values and high p-values. These indicate a suitable model and residual independence. To evaluate the normality assumption of the residual, Figure 7 provides a scatter plot, probability plot, histogram, and residuals plot against fitted values of SPI 12 as an example. The residuals follow a white noise process, by studies of since the histogram shows a normal distribution and the values of the residuals against the fitted plot are spread around zero mean value with no indication of correlation [63, 64]. The majority of the residuals in the QQ probability plot, on the other hand, align with the diagonal line, indicating that the normality assumption is satisfied [63, 65]. The residual and fitted values are plotted in a scatter plot to confirm the residual's Homoscedasticity. The result indicates that the values have a constant variance of around and no clear pattern. And as a result, it can be said that there is a constant variance that would guarantee the preservation of the predictive accuracy.

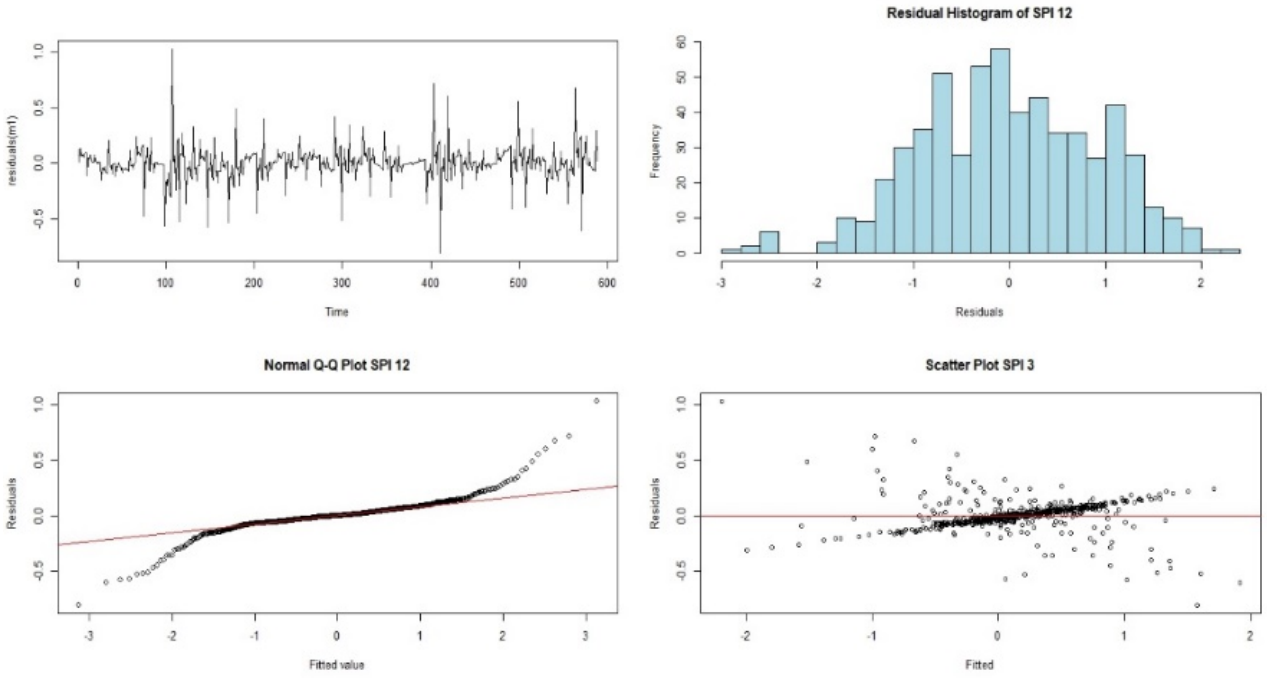

Figure 7: Residual Plot, Histogram, Residual vs Fitted, and Normal Probability Plot of the Residuals for the Selected Model SPI 12

In order to measure the predictive accuracy of the two models (ARIMA and W-ARIMA), a comparison of their performance evaluation metrics is employed. The values RMSE and MAE is presented in Table 6. Based on these values, its evidently clear that the hybrid model (Wavelet-ARIMA) outperformed the traditional ARIMA model in both training and testing sets across all the SPIs.

Table 6: Comparison of Evaluation Metrics Between ARIMA and W-ARIMA Models for All the SPIs

PERFORMANCE METRICS	**ARIMA**				**WAVELET-ARIMA**			
	SPI 3	**SPI 6**	**SPI 9**	**SPI 12**	**SPI 3**	**SPI 6**	**SPI 9**	**SPI 12**
RMSE	0.7771	0.6667	0.4648	0.3212	0.4889	0.4326	0.3566	0.2177
MAE	0.603	0.479	0.3367	0.2191	0.3715	0.3236	0.2607	0.1551

Next, to validate the fitted models, a time series plot was employed. Figure 8 presents the visual comparison of the observed and predicted values of the various SPI data series based on the estimated models. From these plots, it shows that the predicted values align closely to the observed values for all the SPIs indicating the validity and adequacy of the models.

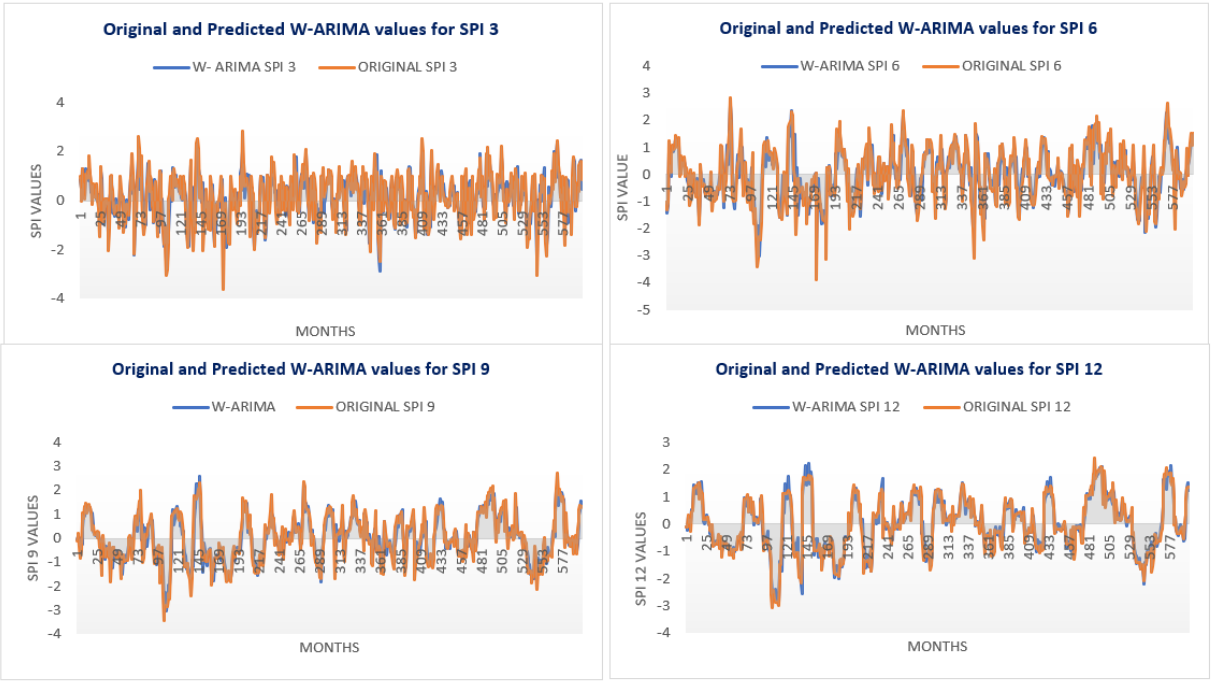

Figure 8: Comparison of Observed and Predicted W-ARIMA Values for all the SPIs

5 Conclusion

In this study, we developed a new hybrid drought forecasting model called Wavelet-ARIMA by combining wavelet transforms with the ARIMA model. Our goal was to improve the prediction of drought occurrences using Standardized Precipitation Index (SPI) data from the Badeggi district in North-Central Nigeria. This data-driven approach addresses the challenges of drought forecasting by leveraging the wavelet's ability to decompose data and extract significant information across different timescales, alongside the accuracy of the ARIMA model. This combination is particularly valuable for modelling Nigeria's diverse climatic conditions and contributes to advancements in drought prediction techniques.

A key aspect of our research was the thorough comparison between the Wavelet-ARIMA model and the traditional ARIMA model, using performance metrics like Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The results showed that the Wavelet-ARIMA model consistently outperformed the traditional ARIMA model across all SPI metrics, highlighting its potential to enhance the accuracy of drought forecasts. This is due to wavelet-ARIMA's ability to capture both high-frequency fluctuations and low-frequency trends in the data. The Wavelet transform acts as a decomposition technique, enabling the separation of time series into different frequency components. This allows the W-ARIMA model to more effectively capture the non-stationary and complex characteristics of drought time series, such as sudden changes in precipitation patterns or long-term drought cycles, which the ARIMA/SARIMA models may struggle with. This explains why W-ARIMA consistently delivered superior results across all SPI metrics.

Initially, the traditional ARIMA model served as a benchmark. We decomposed each SPI series using the Discrete Wavelet Transform (DWT) with the db2 wavelet to capture information at different timescales. The final Wavelet-ARIMA forecast was generated by aggregating the predictions from the optimal ARIMA model for each decomposed component ($A_6, D_6, D_5, D_4, D_3, D_2$, and D_1). Our research introduces an innovative data-driven forecasting technique for drought prediction in north-central Nigeria. The effectiveness of the Wavelet-ARIMA model, demonstrated through its comparison with the traditional ARIMA model, underscores its precision in forecasting droughts.

We believe that the hybrid Wavelet-ARIMA model holds significant potential for improving drought forecasting. Meteorologists, environmental assessors, and administrators can utilize this model for more accurate predictions, which are crucial for effective water resource management and drought preparedness. However, a limitation of this study is that it only applies to SPI data series. We encourage other researchers to expand on this work by applying the model to weekly or daily data and incorporating additional variables, such as streamflow and flood data.

Acknowledgement

The first author acknowledges Jigawa State Polytechnic Dutse (for nomination, and study leave) and the Tertiary Education Trust Fund Nigeria (for sponsorship of the PhD program).

References

[1] Rezaiy, R., & Shabri, A. Enhancing drought prediction precision with EEMD-ARIMA modelling based on standardized precipitation index. *Water Science & Technology*. 2024.89(3): 745-770.

[2] Taylan, E. D. An Approach for Future Droughts in Northwest Trkiye: SPI and LSTM Methods. *Sustainability*. 2024.16(16): 6905. doi:10.3390/su16166905.

[3] Shabri, A. A hybrid wavelet analysis and adaptive neuro-fuzzy inference system for drought forecasting. *Applied Mathematical Sciences*. 2014. 8(139): 69096918. doi:10.12988/ams.2014.48263.

[4] Yksel, G. and Skn, H. EGE BLGES N DNAMK MOD AYRITIRMASI LE KURAKLIK ANALZ. *International Journal of 3D Printing Technologies and Digital Industry*. 2022.6(1): 54-61. doi:10.46519/ij3dptdi.1025073.

[5] Secretariat, W. M. O., Warn, E., Adamo, S. B., & Quality, A. *The journal of the World Meteorological Organization*. 2013.

[6] Mishra, A. K. and Singh, V. P. Review paper A review of drought concepts. *Journal of Hydrology*. 2010. 391(12):202216.doi10.1016/j.jhydrol.2010.07.012.

[7] Wilhite, D. A. *Drought and water crises: science, technology, and management issues*. CRC Press, 2005.

[8] Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., & Stephens, S. The drought monitor. *Bulletin of the American Meteorological Society*. 2002.83(8): 1181-1190.

[9] Wei, X., Huang, S., Liu, D., Li, J., Huang, Q., Leng, G., ... & Peng, J. The response of agricultural drought to meteorological drought modulated by air temperature. *Journal of Hydrology*. 2024.639(7): 131626. doi10.1016/j.jhydrol.2024.131626.

[10] Wang, Z., Chang, J., Wang, Y., Yang, Y., Guo, Y., Yang, G., & He, B. Temporal and spatial propagation characteristics of meteorological drought to hydrological drought and influencing factors. *Atmospheric Research*. 2024.299: 107212. doi:10.1016/j.atmosres.2023.107212.

[11] Arra, A. A., Alashan, S., & iman, E. Trends of meteorological and hydrological droughts and associated parameters using innovative approaches. *Journal of Hydrology*. 2024.640:131661. doi:10.1016/j.jhydrol.2024.131661.

[12] Thom, H.C. A note on the gamma distribution. *Monthly weather review*. 1958.86(4) :117-122

[13] Guttman, N. B. Accepting the standardized precipitation index: a calculation algorithm 1. *JAWRA Journal of the American Water Resources Association*. 1999. 35(2): 311322.

[14] Hinge, G., Piplodiya, J., Sharma, A., Hamouda, M. A., & Mohamed, M. M. Evaluation of hybrid wavelet models for regional drought forecasting. *Remote Sensing*. 2022.14(24): 6381.

[15] Karbasi, M., Karbasi, M., Jamei, M., Malik, A., & Azamathulla, H. M. Development of a new wavelet-based hybrid model to forecast multi-scalar SPEI drought index (case study: Zanjan city, Iran). *Theoretical and Applied Climatology*. 2022.147:499-522.

[16] Katipolu, O. M. Prediction of streamflow drought index for short-term hydrological drought in the semi-arid Yesilirmak Basin using Wavelet transform and artificial intelligence techniques. *Sustainability*. 2023. 15(2):1109.

[17] Xu, D., Ding, Y., Liu, H., Zhang, Q., & Zhang, D. Applicability of a CEEMDARIMA combined model for drought forecasting: A case study in the Ningxia Hui Autonomous Region.*Atmosphere*. 2022.13(7):1109.

[18] Alquraish, M., Ali. Abuhasel, K., S. Alqahtani, A., & Khadr, M. SPI-based hybrid hidden MarkovGA, ARIMAGA, and ARIMAGAANN models for meteorological drought forecasting. *Sustainability*. 2021.13(22):12576.

[19] Shaari, M. A., Samsudin, R., & Ilman, A. S. Forecasting drought using modified empirical wavelet transform-ARIMA with fuzzy C-means clustering. *Indonesian J. Electr. Eng. Comput. Sci*. 2018.11(3):1152-1161.

[20] Rezaiy, R., & Shabri, A. Drought forecasting using W-ARIMA model with standardized precipitation index. *Journal of Water and Climate Change*. 2023.14(9): 3345-3367. doi:10.2166/wcc.2023.431.

[21] Percival, D. B. and Walden, A. T. *Wavelet methods for time series analysis*, 4 Edition. Cambridge University Press, 2000.

[22] Partal, T., & Kii, . Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *Journal of Hydrology*. 2007.342(1-2): 199-212.

[23] Zhang, G. P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 2003. 50:159175.

[24] Ghil, M. and Vautard, R. Interdecadal oscillations and the warming trend in global temperature time series. *Nature*. 1991. 350(6316): 324327.

[25] Wu, X., Zhou, J., Yu, H., Liu, D., Xie, K., Chen, Y., & Xing, F. The development of a hybrid wavelet-ARIMA-LSTM model for precipitation amounts and drought analysis. *Atmosphere*. 2021.12(1): 117. doi: 10.3390/ATMOS12010074.

[26] Salim, D., Doudja, S. G., Ahmed, F., Omar, D., Mostafa, D., Oussama, B., & Mahmoud, H. Comparative study of different discrete wavelet based neural network models for long term drought forecasting. *Water Resources Management*. 2023.37(3): 1401-1420.

[27] Rezaiy, R., & Shabri, A. An innovative hybrid W-EEMD-ARIMA model for drought forecasting using the standardized precipitation index. *Natural Hazards*. 2024:130.

[28] Rezaiy, R., & Shabri, A. Improving Drought Prediction Accuracy: A Hybrid EEMD and Support Vector Machine Approach with Standardized Precipitation Index. *Water Resources Management*. 2024:123.

[29] Mishra, A. K. and Desai, V. R. Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment*. 2005. 19(5):326339. doi:10.1007/s00477-005-0238-4.

[30] Mishra, A. K. and Desai, V. R. Drought forecasting using feed-forward recursive neural network. *Ecological Modelling*. 2006. 198(12):127138.

[31] Mishra, A. K., Desai, V. R., & Singh, V. P. Drought forecasting using a hybrid stochastic and neural network model. *Journal of Hydrologic Engineering*. 2007. 12(6): 626-638. doi:10.1061/(asce)1084-0699(2007)12:6(626).

[32] Salisu, A. M., & Shabri, A. A hybrid wavelet-ARIMA model for standardized precipitation index drought forecasting. *Matematika*. 2020. 36(2):141-156. doi:10.11113/matematika.v36.n2.1152.

[33] T. McKee, N. Doesken, J. K.-P. of the 8th, and U. 1993, The relationship of drought frequency and duration to time scales Drought in Africa, *Climate.Colostate.Edu*, no. January, pp. 1722, 1993, [Online]. Available: https://climate.colostate.edu/pdfs/relationshipofdroughtfrequency.pdf

[34] A. Cancelliere, G. Di Mauro, B. Bonaccorso, and G. Rossi, Stochastic Forecasting of Drought Indices, *Methods and Tools for Drought Analysis and Management*, pp. 83100, 2007, doi:10.1007/978-1-4020-5924-7_5.

[35] Tsakiris, G. and Vangelis, H. Towards a drought watch system based on spatial SPI. *Water Resources Management*. 2004. 18:112.

[36] A. K. Mishra and V. R. Desai, Drought forecasting using feed-forward recursive neural network, *Ecological Modelling*. 2006. 198(12):127138. doi:10.1016/j.ecolmodel.2006.04.017.

[37] Han, P., Wang, P. X., & Zhang, S. Y. Drought forecasting based on the remote sensing data using ARIMA models. *Mathematical and Computer Modelling*. 2010.51(11-12): 1398-1403.

[38] Box, G. E. and Jenkins, G. *Time series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1976.

[39] Box, E. P. Jenkins, G. M. and Reinsel, G. C. Time Series Analysis: Forecasting and Control, 3^{rd} ed. New Jersey: Prentice Hall, 1994.

[40] Parmezan, A. R. S., Souza, V. M., & Batista, G. E. Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. *Information sciences*. 2019.484: 302-337. doi:10.1016/j.ins.2019.01.076.

[41] Rezaiy, R., & Shabri, A. Using the ARIMA/SARIMA Model for Afghanistan's Drought Forecasting Based on Standardized Precipitation Index. *Matematika*. 2023. 39(3):239-261.

[42] Adhikari, R., & Agrawal, R. K. An introductory study on time series modelling and forecasting. *arXiv preprint arXiv:1302.6613*. 2013. 1302.6613: 168.

[43] Hyndman, R. J., & Khandakar, Y. Automatic time series forecasting: the forecast package for R. *Journal of statistical software*. 2008. 27(3):1-22. [Online]. Available: http://www.jstatsoft.org/%0Ahttp://www.jstatsoft.org/v27/i03/paper.

[44] Box, G. E. P. and Jenkins, G. M. *Time Series Analysis: Forecasting and Control*. in Holden-Day series in time series analysis and digital processing. Holden-Day, 1976. [Online]. Available: https://books.google.com.my/books?id=1WVHAAAAMAAJ

[45] Durdu, . F. Application of linear stochastic models for drought forecasting in the Byk Menderes river basin, western Turkey. *Stochastic Environmental Research and Risk Assessment*. 2010.24(8):1145-1162. doi:10.1007/s00477-010-0366-3.

[46] Akaike, H. A New look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974. 19(6): 716-723. doi:10.1109/TAC.1974.1100705.

[47] Klein, A. and Moosbrugger, H. Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*. 2000. 65(4): 457474. doi:10.1007/BF02296338.

[48] Kulahci, M. Montgomery, D. C. and Jennings, C. L. *Introduction to Time Series Analysis and Forecasting*. 2nd edition. United Kingdom: Wiley, 2015. [Online]. Available: https://books.google.com.my/books?id=Xeh8CAAAQBAJ

[49] Nourani, V., Alami, M. T., & Vousoughi, F. D. Wavelet-entropy data pre-processing approach for ANN-based groundwater level modelling. *Journal of Hydrology*. 2015. 524:255-269. doi:10.1016/j.jhydrol.2015.02.048.

[50] Zhang, B. L., & Dong, Z. Y. An adaptive neural-wavelet model for short term load forecasting. *Electric Power Systems Research*. 2001.59(2):121-129.

[51] Kisi, O. Wavelet Regression Model as an Alternative to Neural Networks for River Stage Forecasting. *Water Resources Management*. 2011. 25(2):579600. doi:10.1007/s11269-010-9715-8.

[52] Mallat, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*. 1689. 11(7):674693.

[53] Dawson, C. W., Abrahart, R. J., & See, L. M. HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*. 2007.22(7): 1034-1052.

[54] Benaouda, D., Murtagh, F., Starck, J. L., & Renaud, O. Wavelet-based nonlinear multiscale decomposition model for electricity load forecasting. *Neurocomputing*. 2006.70(1-3): 139-154.

[55] Nury, A. H., Hasan, K., & Alam, M. J. B. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University-Science*. 2017.29(1): 47-61.

[56] Woschnagg, E., & Cipan, J. Evaluating forecast accuracy. *University of Vienna, Department of Economics*. 2004: 17

[57] Dawson, C. W., Abrahart, R. J., & See, L. M. HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*. 2007.22(7): 1034-1052. doi:10.1016/j.envsoft.2006.06.008.

[58] Hyndman, R. J. Measuring forecast accuracy.*Business forecasting: Practical problems and solutions*. 2014.177-183. available online at www.otexts.org/fpp/2/5

[59] Shcherbakov, M. V. and Brebels, A. A Survey of Forecast Error Measures. 2013. 24(4):171176. doi:10.5829/idosi.wasj.2013.24.itmies.80032.

[60] Wheelwright, S., Makridakis, S., & Hyndman, R. J. *Forecasting: Methods and Applications*. John Wiley & Sons, 1998.

[61] Zhang, G. P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 2003. 50:159175.

[62] Gharde, K. D., Kothari, M., & Mahale, D. M. Developed seasonal ARIMA model to forecast streamflow for Savitri Basin in Konkan Region of Maharshtra on daily basis.*J. Indian Soc. Coastal Agric*. Res,. 2016.34(1): 110-119.

[63] Mahmud, I., Bari, S. H., & Rahman, M. Monthly rainfall forecast of Bangladesh using autoregressive integrated moving average method. *Environmental Engineering Research*. 2017.22(2): 162-168.

[64] Rahman, M. A., Yunsheng, L., & Sultana, N. Analysis and prediction of rainfall trends over Bangladesh using MannKendall, Spearmans rho tests and ARIMA model. *Meteorology and Atmospheric Physics*. 2017. 129(4): 409-424.

[65] Putro, S. P., Koshio, S., & Oktaferdian, V. Implementation of Arima model to asses seasonal variability macrobenthic assemblages. *Aquatic Procedia*. 2016. 7:277-284.