Available online at www.garmian.edu.krd

Passer Journal of Basic and Applied Sciences

Passer 8 (Issue 2) (2026) 794-805

Original Article

A Hybrid Wavelet–SARIMA–NARNN Model for Time Series Forecasting

Huda Kardagh Yalda[*1], Saman Hussein Mahmood[1]

[1]*Department of Statistics and Information/ College of Administration and Economics / Salahaddin University-Erbil, Erbil, Iraq.*

Received 18 January 2026; revised 5 May 2026;
accepted 7 May 2026; available online 15 July 2026

DOI: 10.24271/PSR.2026.571932.2527

ABSTRACT

This research aims to develop a hybrid forecasting framework to improve the accuracy of time series models by integrating Wavelet Transform with Seasonal Autoregressive Integrated Moving Average (SARIMA) and Nonlinear Autoregressive Neural Networks (NARNN). The proposed method relies on using wavelet shrinkage techniques to remove noise from time series data before applying the forecasting models. Several families of wavelets, such as Daubechies, Coiflets, and Symlets, were tested along with different threshold estimation methods like Universal, Minimax, and SURE to select the best model configuration. The methodology was applied to real data representing the maximum monthly electricity demand in the Kurdistan region of Iraq for the period (2014–2024). Simulation experiments were also conducted using data generated from the SARIMA model, with 250 repetitions for each case. The results showed that the proposed hybrid models (Wavelet-SARIMA and Wavelet-NARNN) clearly outperformed the traditional models in terms of accuracy metrics RMSE, MAE, and MAPE, confirming the effectiveness of wavelet integration in improving forecasting accuracy in time series analysis. The researchers conducted all analyses through MATLAB software.

Keywords: Time series, NAR neural network, SARIMA, Wavelet Transform, Peak Electricity Demand.

1 Introduction

The ongoing rise in energy consumption creates difficulties for predicting peak electricity demand because of its unpredictable nature which makes forecasting peak demand extremely complicated. The system plays an important role to maintain grid stability because it protects against outages while reducing operational expenses [1]. Time series electricity demand data is influenced by seasonal, climatic, and behavioral factors which require advanced predictive models and sophisticated analytical techniques to create effective planning and operational efficiency solutions [2].

Time series analysis, the wavelet transforms and artificial neural networks (ANN) are all examples of useful techniques that can be used to develop effective time series forecasting models [3]. Although traditional time series models such as ARIMA (Autoregressive Integrated Moving Average) developed by Box and Jenkins have been popular methods to forecast time series data, they are limited when dealing with nonlinear and/or non-stationary time series data [4,5]. To address these limitations, ANNs

* Corresponding author
E-mail address Huda.yalda@su.edu.krd (Instructor).
Peer-reviewed under the responsibility of the University of Garmian.

The system needs to deliver a flexible solution that shows the ability to learn advanced nonlinear patterns according to complex requirements, traditional forecasting models and neural forecasting models both experience performance improvement through the application of wavelet transforms techniques [6]. The combination of wavelet analysis and neural networks has shown promising results in forecasting peak electricity demand, particularly in high-variability datasets [7].

Several studies and proposals have been conducted, including a study by [8] that this approach provided higher accuracy compared to traditional models, highlighting its effectiveness in addressing load forecasting challenges in power systems [9]. developed an AWNN model which uses Adaptive Wavelet Neural Networks to predict short-term electricity prices. The "Mexican hat" wavelet served as the activation function. The researchers achieved better accuracy results than all three models which included Wavelet-ARIMA and MLP and RBF and FNN when they evaluated electricity markets in Spain and PJM [10] Researchers need to conduct their temperature change study in northeastern Bangladesh by applying Mann-Kendall analysis. The researchers created a temperature forecasting model which uses wavelet technology together with ARIMA and artificial neural networks. The results showed that Wavelet-ARIMA model provides more accurate predictions than Wavelet-ANN model, which helped researchers better understand climate change effects on the area. Researchers need to conduct their temperature change study in northeastern Bangladesh by applying Mann-Kendall analysis [11]. The researchers created a hybrid model which combines Discrete Wavelet Transform (DWT) with nonlinear Recurrent Neural Networks (RNNs) to analyze financial time series data. The model uses wavelet-decomposed data as its input to the neural network which enhances its ability to recognize patterns. Its performance was benchmarked against ARIMA, traditional neural networks, and autoregressive models. The model showed better performance than its competitors when applied to euro/dollar exchange rates and Brent crude oil prices and the Nasdaq index, which all used R software for their implementations [12]. The researchers created a Legendre-type Multi-Wavelet Neural Network (MWNN) model which predicts nonlinear wind speed changes. The system operates using Legendre-based multi-wavelets which function as activation functions to decrease computational requirements and improve learning performance. The system gains accuracy through a hybrid reconstruction technique. The model shows strong performance through its ability to predict multiple future time points in wind energy optimization when tested with actual data.

Researcher [13], suggest a new method for short-term electric load forecasting that utilizes a recurrent neural network (RNN) combined with wavelet transform and enhanced by the Sakho Kho adaptive algorithm. it has outperformed six modern methods, achieving the lowest MAE and RMSE. The researchers [14] proposed a novel method for estimating the state of charge (SOC) of supercapacitors used in electric vehicles, based on wavelet transformation followed by feeding low-frequency components into a back-propagation neural network (BPNN) for one-step prediction. The experiments demonstrated the reliability of hybrid energy storage systems.

The main contributions of this research can be summarized as follows:

1. Proposing a hybrid forecasting framework that integrates wavelet transform, the SARIMA model, and nonlinear autoregressive neural networks (NARNN) to improve forecasting accuracy in time series analysis.

2. Conducting a systematic investigation of different wavelet families (Daubechies, Coiflets, and Symlets) along with several threshold estimation techniques (Universal, Minimax, and SURE) to determine the most effective approach for noise reduction.

3. Evaluating the performance of the proposed framework using real electricity demand data as well as simulated datasets with different sample sizes and noise levels.

4. Demonstrating that wavelet-based hybrid models achieve higher forecasting accuracy compared with traditional models such as SARIMA and standalone neural networks.

2 Methods and Materials

The following sections provide a theoretical overview of some basic concepts related to time series analysis, wavelet transforms, and artificial neural networks (ANNs).

2.1 Time series

A time series is a set of apparent values arranged over time, such as annual, quarterly, monthly, or daily. It can be analysed to understand changes in the phenomenon's value over time. There are two types of changes: general changes over the long term and seasonal trends. Periodicity is determined by developments, economic and political factors, and occasional changes, other random or systematic fluctuations [15], Components of Time Series can be decomposed into four main components: Long-term movements in the mean are referred to as a trend (T_t). Seasonal effects (I_t), or calendar-related cyclical changes, Cycles (C_t): other cyclical fluctuations (such as business cycles), and the value of the irregular component (Residuals), other random or systematic fluctuations [16]

$$Z_t = T_t + S_t + C_t + I_t \tag{1}$$

The Box-Jenkins methodology is a commonly used approach to modelling time series data. It uses four stages to construct time series models [17]:

- Identification: This involves using autocorrelation functions and partial autocorrelation functions to determine the appropriate ARMA model.
- Estimation: Once the model is identified, its parameters are estimated using methods like maximum likelihood, least squares, or Yule-Walker.
- Diagnostic: The model's validity is checked by analysing residuals, adjusting parameters, and comparing criteria.
- Forecasting: The model is used to forecast the time series' future

One Box-Jenkins model is SARIMA (p, d, q) (P, D, Q), which combines the seasonal and non-seasonal components and adds a differential function for both.

$$\phi_P(B^s)\varphi_p = \theta_Q(B^s)\vartheta_q(B)e_t \tag{2}$$

2.2 Artificial Neural Network

Among the models of soft computing systems, artificial neural networks (ANNs) operate in a way that mimics the processes of the human brain. Neurons are among the most common essential components in the human body. Neural networks receive data from multiple sources, merge and analyze it in unique ways, and then apply nonlinear operations to the outputs. According to the structure and function of the artificial neural network, it is a system that receives inputs [18], processes the data, and then outputs the results. The network consists of three main layers: the input layer, the hidden layer, and the output layer. And each layer performs a specific function in processing data within the system. with the following function [19,20]

$$Y_j = f(\sum_j W_{ij} X_{ij} + b_j) \tag{3}$$

where, f is the activation function of the output layer neuron

2.2.1 Nonlinear autoregressive (NAR) neural network

The Nonlinear Autoregressive Neural Network (NARNN) is an effective tool for time series prediction based on past values of the same series. It employs a two-layer feedforward neural network with a sigmoid activation function in the hidden layer and a linear transfer function in the output layer [21], The network's outputs y(t) is feedback as inputs through time delays, allowing the model to incorporate historical data. The model behavior is described by the equation:

$$y_t = f(y_{t-1} + \ y_{t-2} + \cdots \ y_{t-d}) + \varepsilon_t \tag{4}$$

where f is a nonlinear function approximated by the neural network, d is the time delay, and ε_t represents the prediction error. The term y represents the combined interference from N−1sources [22].

2.3 Wavelets

A wavelet, also called a "small wave," is a type of mathematical function that breaks the function into different frequency components and studies each one with the right amount of re-design (resolution) at each measurement. Mathematically, we define the wavelet as a real-valued function that is defined on a complete real axis and oscillates up and down regularly around zero. In addition, it represents a distinctive tool and an effective and powerful technique for representing and analyzing data. Because of its small size and ability to distinguish itself from big wave signals like the sine function wave and the cosine wave, it is known as a wavelet [23,24].

2.3.1 The wavelet Function analysis

The Fourier expansion, which represents functions using sines and cosines at different frequencies, is a well-known example of series expansion in terms of orthogonal basis functions. Wavelets, on the other hand, form a set of fundamental functions with a distinct structure.to construct an orthonormal wavelet basis for the space of square-integrable real functions, start with two parent wavelets: the scaling function φ (often called the father wavelet) and the mother wavelet ψ. Additional wavelets are then generated through specific dilation and translation operations φ and ψ, which are governed by these fundamental functions by:

$$\mathrm{\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k) j, k \in z} \tag{5}$$

$$(\mathrm{x}) = 2^{\mathrm{j}/2}\psi(2^{\mathrm{j}}\mathrm{x} - \mathrm{k}) j, k \in \tag{6}$$

Indices j and k are commonly called scale or (dilation) and location (or translation) parameter, and z is the set of all integers. [25], The scaling function is derived from the dilation equation,

$$\varphi(x) = \sum_{k \in z} h_k \varphi(2x - k) \tag{7}$$

and the wavelet function is obtained as

$$\psi(x) = \sum_{k \in z} g_k \ \varphi \ (2x - k) \tag{8}$$

Where h_k is the sequence of the filter coefficients $h_k \in \ell^2(z)$and (i.e., a square summable sequence space), the function has certain properties.

The translation of the scaling and wavelet function on each fixed j form orthonormal basis for a subspace. The spaces constitute a multiresolution analysis which will discuss in the following subsection [25].

Given the above wavelet basis, at a certain (resolution) level, a function

$$f(x) = \sum\nolimits_{k \in z} C_{j,k} \varphi_{j,k}(x) + \sum\nolimits_{j=j_0}^{\infty} \sum\nolimits_{k \in z} d_{j,k} \psi_{j,k}(x) \tag{9}$$

The wavelet coefficients are which represent the approximation coefficients of the function using as basis, at the level, and are the detail coefficients at different level of resolution.

2.3.2 Wavelet Families

The following are some of the wavelet families used in the application for this study:

2.3.2.1 Daubechies Wavelet

The Daubechies wavelet, introduced by Ingrid Daubechies in 1988, belongs to a family of wavelets commonly used in signal processing and time series analysis. Characterized by a specific number of vanishing moments and compact support, these wavelets are well-suited for representing polynomials. Their efficient structure and localized nature have enabled the broad application of discrete wavelet analysis across various domains. The Daubechies wavelets significantly contributed to the practical implementation of discrete wavelet analysis. The system uses D or dbN as its standard names because N represents the wavelet's order which defines the number of vanishing moments. The db2 wavelet represents the second-order Daubechies wavelet which has two vanishing moments. Their formulation relies on two fundamental functions: the scaling function and the wavelet function, defined as follows [26,27]:

$$\emptyset(t) = \sum_{b=0}^{N-1} a_b \, \emptyset\,(2t-b) \qquad (10)$$

$$\psi(t) = \sum_{b=0}^{N-1} c_b \, \emptyset\,(2t-b) \qquad (11)$$

where $c_b = (-1)^b a_{N-1-b}$

The scaling coefficients in this context represent the wavelet coefficients whereas N represents the total number of coefficients which directly depends on the number of vanishing moments. The formula computes the wavelet coefficients from the scaling coefficients.

2.3.2.2 Coif lets wavelet

Ingrid Daubechies developed Coiflet wavelets in 1989, which were a major improvement to wavelet theory for use in signal processing. Coiflet wavelets have two vanishing moments in both the wavelet function (ψ) and the scaling function (ϕ). Due to this characteristic, Coiflet wavelets can represent polynomial trends in signals better than other types of wavelets. These wavelets are typically called Coif N, where "Coif" represents Ronald Coifman, who was involved with Daubechies' work, and "N" represents the order of the wavelet [22,28]. A direct relationship exists between the wavelet's order and its mathematical properties.

Wavelet length: 6N, Number of vanishing moments for the wavelet function ψ:

L= 2N, Number of vanishing moments for the scaling function ϕ: L1=2N−1.

By using wavelets, the scaling function ϕ(x) is derived from the dilation equation,

$$\phi(\mathrm{x}) = \sqrt{2}\sum_{\mathrm{k}\in\mathrm{z}} \mathrm{h_k}\phi(2\mathrm{x}-\mathrm{k}) \qquad (12)$$

And the other wavelet functions ψ(x) is obtained as

$$\psi(\mathrm{x}) = \sqrt{2}\sum_{\mathrm{k}\in\mathrm{z}} \mathrm{g_k}\psi(2\mathrm{x}-\mathrm{k}) \qquad (13)$$

Where $h_k = -1^k h(N-1-k)$ and N is the number of scaling and wavelet coefficients. The sets of scaling (h_k) and wavelet (g_k) function coefficients vary depending on their corresponding wavelet bases [27]

2.3.2.3 Symlets Wavelet

Symlets are a particular category of wavelet, which are a derivative of the Daubechies family (which was developed by Ingrid Daubechies in 1992), but improves their symmetry, while still providing the same main characteristics of Daubechies wavelet (i.e., Compact Support and the Number of Vanishing Moments). The Symlet mathematical formulation uses the same mathematical definition as that of the Daubechies wavelet: both are defined by a scaling function φ (t), as well as a wavelet function ψ (t) and have the same basic characteristics: they must be orthogonal to one another and have compact support. Additionally, Symlets use improved scaling coefficients, which allows them to achieve greater symmetry and a linear phase response, which is advantageous for the applications of edge detection and accurate signal reconstruction. The symbol N is typically used to represent the wavelet order, although some sources use 2N instead [26,27].

2.3.3 Thresholding Wavelet

Thresholding is a fundamental nonlinear technique in wavelet-based denoising. It involves comparing the discrete wavelet transform (DWT) coefficients against a specified threshold, where coefficients below the threshold are either set to zero or modified accordingly. This approach is widely used in data segmentation and noise reduction, where the coefficients are divided into two categories [29]: one representing noise and the other representing the true signal. Threshold estimation techniques vary depending on the nature of the data, and the most important of these techniques have been adopted in this study.

2.3.3.1Universal Threshold

Donoho and Johnstone (1994) proposed universal threshold, which is given by

$$\delta = \hat{\sigma}_{(MAD)}\sqrt{2\log(N)} \qquad (14)$$

Where N is the data length series, and $\delta = \hat{\sigma}_{(MAD)}$ is the estimator of standard deviation of details coefficients, which is estimated as:

$$\hat{\sigma}_{(MAD)} = \frac{MAD}{0.6745} \qquad (15)$$

Where MAD is the median absolute deviation of the wavelet coefficients at the finest scale [30].

2.3.3.2 Minimax Threshold

The optimal minimax threshold method, introduced by Donoho and Johnstone in 1995 as an advancement over the universal threshold method, relies on an estimator designed to achieve the minimax risk, expressed as:

$$\tilde{R}(F) = in_{\tilde{f}} f \, sup_{f\in\tilde{R}(F)} R\big(\tilde{f}, f\big) \qquad (16)$$

Where

$$R(\tilde{f}, f) = \frac{1}{N}\sum_{i=1}^{N} E\left[\tilde{f}_i - f_i\right]^2 \qquad (17)$$

Where and denote the vectors of true and estimated sample values. Unlike universal thresholding methods, the minimax threshold estimator is specifically designed to minimize the overall mean square error (MSE) without leading to excessive smoothing of the signal. When applied within soft-thresholding frameworks, this approach effectively reduces the associated estimation risk [31].

2.3.3.3 SURE Thresholding

Donoho and Johnstone's 1994 minimum recommendation for the Stein estimator of risk has been lowered to 40. The sure thresholding technique calculates the threshold estimate (δk) for each level k based on the wavelet coefficients and the supplied soft threshold estimator [32].

$$SURE(\eta, W) = N - 2 \neq \{j: |W_j| \leq \eta\} - \sum_{j=0}^{d} min\left(|W_j|, \eta\right) \qquad (18)$$

Where [Wk: k = 1,2, …, d] be a wavelet coefficient in the k^{th} level, and then, select δ^s that minimizes $SURE$ (δ, W)

$$\eta^s = \mathrm{argmin}_{\eta} \mathrm{SURE}(\eta, \mathrm{W}) \qquad (19)$$

2.3.4 Thresholding Rules

The primary thresholding rules encompass both soft and hard thresholding. The type used in this research is Soft Thresholding

Soft thresholding involves setting to zero all signal values smaller than δ and then subtracting δ from the values larger than δ. This process is defined as follow

$$W_n^{(s)} = sign\{Wn\}(|Wn| - \delta) \qquad (20)$$

Where

$$W_n = \begin{bmatrix} +1 & if\, w_n > 0 \\ 0 & if\, w_n = 0 \\ -1 & if\, w_n < 0 \end{bmatrix} \qquad (21)$$

And

$$(|Wn| - \delta) = \begin{bmatrix} (|Wn| - \delta) & if\, (|Wn| - \delta) \geq 0 \\ 0 & if\, (|Wn| - \delta) < 0 \end{bmatrix} \qquad (22)$$

Coefficients smaller than the threshold are set to zero, and additionally, all coefficients greater than the threshold are reduced by the amount of the threshold. Thus, soft thresholding is a continuous mapping [33].

3 Application aspect

We apply the proposed hybrid approach in this research, which combines the SARIMA model, neural networks, and wavelet transform, to model time series. We use real-world data to evaluate the effectiveness of wavelet integration in both the SARIMA model and neural networks. We also use simulation data with different samples for the same purpose. The results are then compared to determine the best model.

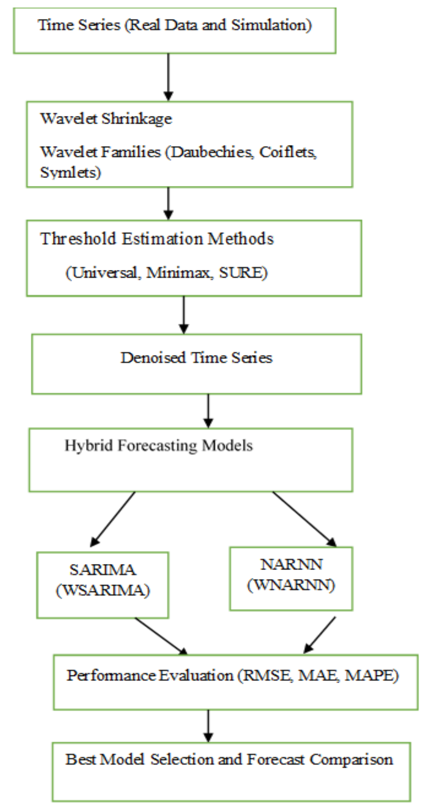

Figure 1: Flowchart of the proposed hybrid forecasting framework.

3.1 Real Data

The data were analysed using the Box-Jenkins methodology on a time series representing the monthly peak electricity demand over the period 2014 to 2024 (132 observations), and the best-fit SARIMA model was found. This was done through three main steps: identification, estimation, and diagnostic validation. Figures 2 and 3 illustrate the time series data in their original form and after applying the first difference to achieve stationarity. The augmented Dickey-Fuller (ADF) test was performed after the first difference, and the results indicated the absence of a unit root, confirming that the series was stationary. Based on the lowest values of AIC, SBIC, and RMSE, and the highest value of the log-likelihood model, the best-fit model was determined as follows: SARIMA (1,0, **1**) (0,**1**,**1**)$_{12}$.

The parameters of the optimal SARIMA model (Table1) were estimated using Maximum Likelihood Estimation (MLE). Utilizing the Ljung-Box test to analyses the autocorrelation in the RC has enabled us to verify that our results are model-adequate. Using the p-value associated with the Ljung-Box test, we compared it with α = 0.05 to confirm that our residuals were randomly distributed.

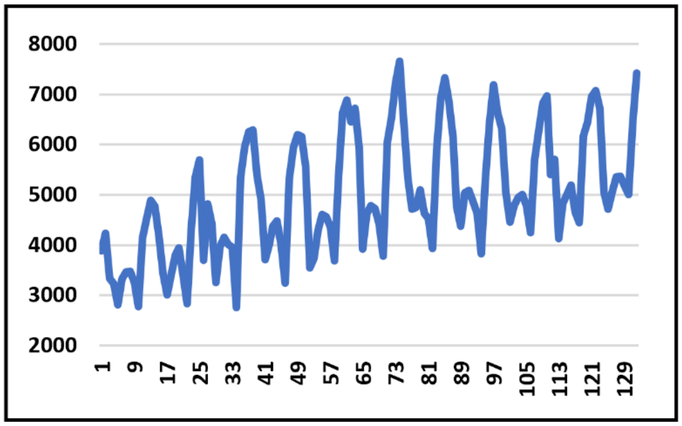

Figure 2: Original Data**.**

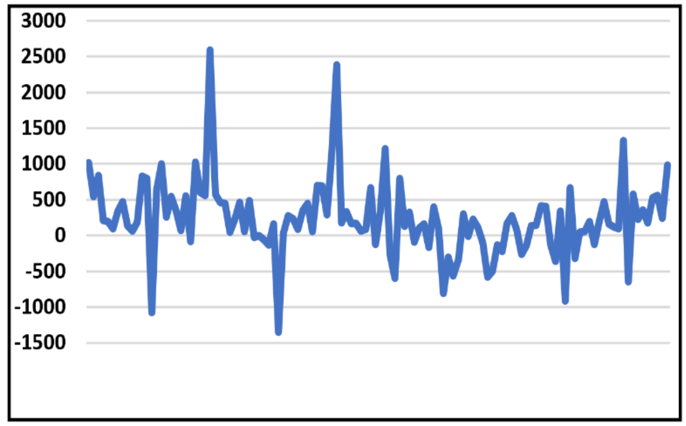

Figure 3: After first difference.

Table 1: Estimated Parameters Values of the SARIMA (1,0, **1**) (0,**1**,**1**)$_{12}$ Model.

Parameter	Estimate	Standard Error	t	P-value
AR (1)	0.99998	0.004	252.075	0
MA (1)	0.8058	0.0546	14.7636	0
SMA (1)	0.756	0.0593	12.7579	0

This study proposes a methodology that uses wavelet transforms to address the noise problem in time series forecasting by removing noise from time series data using wavelet transforms. The proposed methodology involves applying Wavelet Shrinkage Filters to wavelet functions from the Daubechies, Symlets, and Coiflets families, utilizing threshold estimation methods including Universal, Minimax, and SURE, with the Soft Thresholding rule, following these steps:

1. **Determination of the wavelet decomposition level:** This level is defined as a positive integer not exceeding $\leq log_2 n$, where n is the number of time series observations.
2. **Selection of the optimal wavelet order:** The optimal order for each wavelet function was determined based on the performance of denoising techniques using statistical evaluation metrics (RMSE, MAE, MAPE). The order that yielded the lowest values for these metrics was selected (Table 2).

Table 2: The best wavelet filter for (Daubechies, coiflets, and symlets) for real data.

Wavelet	Threshold Method	RMSE	MAE	MAPE
db2	Universal	846.08	701.591	14.17%
db18	Minimax	3293.5	3147.68	63.12%
db40	SURE	0.042	0.0363	0.00%
Coif1	Universal	3167.2	3007.8	60.64%
Coif5	MiniMax	952.97	753.93	14.74%
Coif2	SURE	0.289	0.245	0.01%
Sym3	Universal	2772.3	2601.6	0.5201%
Sym3	MinMax	5111.5	4973.9	1%
Sym5	SURE	0.3962	0.329	0.0001%

3.1.1 Application of wavelets to SARIMA and NARNN

After transforming the monthly electricity demand time series data using the proposed wavelet transformations (Daubechies, Coiflets, and Symlets) as shown in Table2, we applied these wavelet transformations to the SARIMA model, and we have the proposed model (Wavelet-SARIMA). We compared the results to see whether the proposed method improved the model and reduced noise. The results are shown in Table 3.

We also applied a nonlinear autoregressive neural network (NAR) to the same time series, its structure consists of three main layers: the input layer, the hidden layer, and the output layer. The input layer represents a time series of monthly peak electricity demand, denoted by (yet). We conducted multiple experiments on the number of neurons in the hidden layer from 1 to 10 in each experiment, and the best was 10 neurons, and changing the number of delay periods from 1 to 15, and 12 steps was the best. The output layer consists of a single neuron representing the predicted value. Experimental results indicated that the optimal network structure was (1:12:10). The model was developed employing the Levenberg-Marquardt algorithm, a well-known and efficient numerical optimization method. A dataset of 132 observations was used and randomly divided into three subsets: 70% for training, 15% for testing, and 15% for validation. The results are shown in Table 3 before and after applying the proposed method.

Then we will apply the different wavelet transformations in Table 2 to the NARNN network to obtain the second proposed models, which is the proposed Wavelet-NARNN model, and the results are shown in Table 3.

Table 3: Comparison of Results Between the Proposed Methods, SARIMA, and NARNN.

Method	**RMSE**		**MAE**		**MAPE**	
	SARIMA	**NARNN**	**SARIMA**	**NARNN**	**SARIMA**	**NARNN**
SARIMA	388.070	--------	274.710	-------	5.546	--------
NARNN	--------	511.210	--------	365.140	-------	7.789
W(db5-Universal)	112.540	163.670	78.230	111.990	1.534	2.255
W(db8-MiniMax)	162.490	105.860	125.630	76.409	2.555	1.544
W (db25-SURE)	107.460	38.494	76.699	26.018	1.543	0.529
W (Coif1-Universal)	97.296	109.900	66.848	74.561	1.264	1.437
W (Coif5-MiniMax)	241.422	286.950	182.473	206.280	3.563	4.282
W(Coif4-SURE)	241.422	363.970	182.473	250.190	3.563	5.107
W(Sym2-Universal)	88.385	109.990	61.318	73.763	0.205	1.447
W(Sym2-MiniMax)	190.810	227.150	141.090	162.600	2.815	3.275
W (Sym8 - SURE)	224.278	259.930	145.804	177.340	0.340	3.635

Table 3 shows the performance of the proposed wavelet-based methods compared to SARIMA and NARNN models using three evaluation metrics. The results clearly indicate that wavelet-based processing improves model performance, with the proposed models achieving the best performance. The wavelet (Sym2-Universal) results for SARIMA showed the lowest RMSE (88.385), MAE (61.318), and MAPE (0.2045), outperforming traditional SARIMA. Among the hybrid NARNN-based models, W(db25-SURE) for NARNN achieved the lowest RMSE (38.49), MAE (26.018), and MAPE (0.529), outperforming the standalone NARNN models on all metrics and the best overall performance. These results demonstrate the effectiveness of integrating wavelet transforms with time series forecasting models.

3.2 Simulation Data

This study presents a simulation-based evaluation of wavelet-based noise removal techniques aimed at improving the performance of the SARIMA (1,0,1) $(0,1,1)_{12}$ model. The simulation data structure is illustrated in Figure 4. The method involves identifying the optimal level of decomposition and wavelet order from three different wavelet families: Daubechies (orders 2 to 45), Coiflets (1 to 5), and Symlets (2 to 8). The algorithm is applied using a soft thresholding rule, along with threshold estimation methods including Universal, Minimax, and SURE. The simulation design includes varying sample sizes (150, 250, 350) and different values of noise variance ($\sigma^2 = 3, 4, 5$). Each scenario consists of 250 iterations.

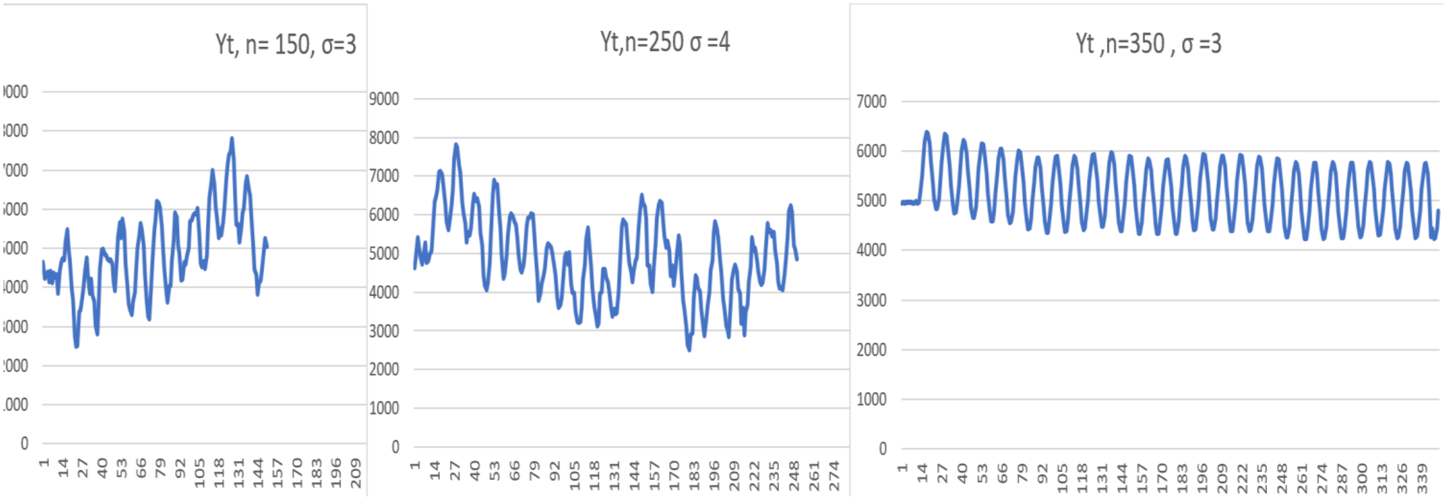

Figure 4: Data Simulation.

RMSE, MAE, and MAPE were used to evaluate prediction accuracy for each of the configurations being evaluated by the proposed wavelet model. Using these evaluations resulted in the identification of three optimal levels of configuration for each wavelet family considered (see Table 4). The results indicate that the proposed method selects configurations that are more efficient than those identified using the other two methods for the three families of wavelets studied, leading to better noise reduction in forecasted values, as well as higher SARIMA model performance. These findings highlight the importance of identifying the optimal wavelet configurations when forecasting time series data.

Table 4: The best three wavelet filters: Daubechies, Coiflets, and Symlets, for each sample, based on the performance metrics of the proposed methods.

	N=150, $\sigma^2 = 3$		N=250, $\sigma^2 = 4$		N=350, $\sigma^2 = 3$	
SARIMA model parameters used in the simulation	wavelet	Threshold level	wavelet	Threshold level	wavelet	Threshold level
	db5	Universal	db5	Universal	db5	Universal
	db2	MiniMax	db8	MiniMax	db8	MiniMax
	db22	SURE	db27	SURE	db25	SURE
AR=0.98 MA=0.76 SMA=0.89	Coif1	Universal	Coif1	Universal	Coif1	Universal
	Coif5	MiniMax	Coif5	MiniMax	Coif5	MiniMax
	Coif4	SURE	Coif5	SURE	Coif4	SURE
	Sym2	Universal	Sym2	Universal	Sym2	Universal
	Sym2	MiniMax	Sym2	MiniMax	Sym5	MiniMax
	Sym6	SURE	Sym8	SURE	Sym8	SURE

3.2.1 Hybrid Model Results

Based on the results presented in Table 4, we evaluated the prediction accuracy of the two proposed methods.

1 The First Proposed Method

The first recommended approach includes a hybrid approach that utilizes the SARIMA model as well as Wavelet. It combines the two models into a Wavelet-SARIMA model with a range of different datasets. By employing the previous statistical criteria, we assessed the performance of the proposed hybrid model against the other models. Using Tables 8, 9, and 10, summarised the results from performance comparison.

Table 5: Performance metrics for the SARIM and the wavelet SARIMA model, n =150.

Model	RMSE	MAE	MAPE

SARIMA (2,0,2) (0,2,2)12	146.735	76.2843	1.42701
W (db5-universal -soft)	157.03	84.9483	1.61193
W (db2 – MiniMax- soft)	154.549	87.8783	1.66687
W (db22 – SURE- soft) *	25.8577*	17.9033*	0.343137*
W (Coif1 –Universal- soft)	160.471	108.315	2.1334
W (Coif5 – MiniMax- soft)	131.874	75.4508	1.39404
W (Coif4 – SURE- soft)	107.66	48.12	0.918855
W (Sym2- Universal- soft)	121.049	60.35	1.14912
W (Sym2-MiniMax -soft)	117.6	75.9576	1.43892
W (Sym6 – SURE- soft)	123.557	58.6928	1.13456

Table 6: Performance metrics for the SARIM Model and the wavelet SARIMA model, n=250.

Model	RMSE	MAE	MAPE
SARIMA (2,0,2) (0,2,2)12	53.7628	27.4919	0.558716
W (db5-universal- Soft)	39.4299	17.5853	0.356143
W (db8- MiniMax-Soft)	47.3705	22.6385	0.454862
W (db27- SURE- Soft)	26.5987	19.9239	0.391483
W (Coif1- Universal- Soft)	75.5411	32.3238	0.629575
W (Coif5- MiniMax - Soft)	26.8737	13.7475	0.276046
W (Coif5 – SURE- Soft)	28.3624	12.8633	0.25424
W (Sym2- Universal -Soft)	67.8833	39.9539	0.787502
W (Sym2-MiniMax -Soft)	66.8776	41.1785	0.813833
W (Sym6-SURE - Soft) *	16.2225*	11.5408*	0.228119*

Table 7: Performance metrics for the SARIM Model and the wavelet SARIMA model, n=350.

Model	RMSE	MAE	MAPE
SARIMA (2,0,2) (0,2,2)12	47.8502	23.784	0.478189
W (db5-universal- Soft) *	13.8559*	8.27347*	0.169182*
W (db8- MiniMax-Soft)	45.7119	19.1956	0.390527
W (db25- SURE- Soft)	45.5383	21.3105	0.402114
W (Coif1- Universal- Soft)	47.701	28.2666	0.554264
W (Coif5- MiniMax - Soft)	15.5748	7.41437	0.145376
W (Coif4 – SURE- Soft)	19.6276	10.1301	0.192057
W (Sym2- Universal -Soft)	50.3707	27.8481	0.542846
W (Sym2-MiniMax -Soft)	147.159	56.6136	1.09476
W (Sym8-SURE - Soft)	19.7928	9.89497	0.192273

The results reported in Tables (5–7) indicate that the new proposed model, Wavelet-SARIMA (WSARIMA), is much better than the other two methodologies, (Universal, Minimax), and (SURE) for estimating statistical thresholds, regardless of how large of a sample size you have observed (150, 250, and 350). The model clearly outperformed all of the others based on performance metrics (RMSE, MAE, MAPE). All three models had lower scores than the original WSARIMA algorithm.

The traditional SARIMA model has demonstrated reasonable performance without the use of noise reduction techniques, but the performance of the traditional SARIMA model did not achieve the same level as that of the proposed WSARIMA. The proposed WSARIMA has very good performance and is very effective in processing difficult time series data.

Specifically, in Table (5), the SURE threshold estimating method and the db22 wavelet filter produced the best results using a sample size of 150, which resulted in the lowest metrics across all performance measures.

Table (6) shows that the Sym6 filter performed best with an efficiency rating of 0.71 when using the SURE method to analyze a sample size of 250. Similarly, Table (7) reveals that the db5 filter performed extremely well with the Universal threshold method when applied to a sample size of 350. The use of wavelet transformation reduces noise while isolating critical factors for optimal model fit and prediction development for complex real-valued time series data.

2 The Second Proposed Method

Based on the results presented in Table 4, the second proposed method is a hybrid model that combines the traditional nonlinear autoregressive neural network (NARNN) model with the Wavelet model to form a Wavelet-NARN model with different data sizes.

Their performance was evaluated and their effectiveness verified using performance metrics, and correlation coefficient. This comparison aims to determine which of the two approaches is more effective in prediction. Tables 8, 9, and 10 summarize the results and comparisons.

Table 8: Performance metrics for the NARNN and the wavelet NARNN model, n=150, $\sigma^2 = 3$.

Models	**RMSE**	**MAE**	**MAPE**	**Corr.**
NARNN	29.333	19.698	0.38425	0.99919
W(db5-Universal-NARNN)	127.39	60.563	1.1598	0.98153
W (db2 -MiniMax -NARNN)	142.6	94.903	1.80E+00	0.9661
W (db22 -SURE - NARNN) *	19.51*	12.424*	0.24035*	0.99953*
W (Coif1 -Universal - NARNN)	271.13	152.23	2.942	0.85219
W (Coif5 - MiniMax -NARNN)	208.53	110.34	2.1114	0.95479
W (Coif4 -SURE - NARNN)	167.02	91.118	1.7523	0.96621
W (Sym2 - Universal - NARNN)	155.6	101.79	1.943	0.94121
W (Sym2 - MiniMax - ANN)	1.53E+02	9.38E+01	1.7909	0.95986
W (Sym6 - SURE - ANN)	144.29	71.602	1.396	0.97895

Table 9: Performance metrics for the NARNN and the wavelet NARNN model, n=250, $\sigma^2 = 3$.

Models	**RMSE**	**MAE**	**MAPE**	**Corr.**
NARNN	20.954	14.821	0.29343	0.99942
W(db5-Universal-NARNN)	22.35	17.667	0.35251	0.96429
W(db8-MiniMax-NARNN)	15.755	9.6211	0.1891	0.99965
W(db27-SURE-NARNN) *	11.553*	7.8451*	0.15514*	0.99981*
W(Coif1-universal-NARNN)	28.27	17.461	0.34009	0.9973
W(Coif5-MiniMax-NARNN)	12.842	8.0632	0.15944	0.99978

W(Coif5-SURE-NARNN)	14.641	8.5765	0.16883	0.99965
W(Sym2-universal-NARNN)	39.082	19.14	19.14	0.99316
W(Sym2-MiniMax-NARNN)	37.982	19.951	0.38243	0.99655
W(Sym8-SURE-NARNN)	13.77	9.7541	0.19411	0.99974

Table 10: Performance metrics for the NARNN and the wavelet NARNN model, n=350, $\sigma^2 = 3$.

Models	RMSE	MAE	MAPE	Corr.
NARNN	18.361	13.238	0.25869	0.99952
W(db5-Universal-NARNN) *	16.384*	9.6961*	0.18961*	0.99956*
W (db8-MiniMax NARNN)	63.089	28.68	0.57475	0.99289
W (db25-SURE - NARNN)	80.441	21.479	0.41638	0.97741
W (Coif1-Universal NARNN)	37.64	19.373	0.38335	0.99413
W (Coif5-MiniMax NARNN)	38.854	20.537	0.39767	0.99709
W (Coif4-SURE NARNN)	40.07	17.315	0.33032	0.99713
W (Sym2-Universal NARNN)	47.631	25.981	0.50812	0.9911
W (Sym2-MiniMax NARNN)	30.784	18.815	0.36296	0.9973
W (Sym8 - SURE NARNN)	18.842	9.9836	0.1968	0.99944

The results in Tables 8–10 clearly show that the proposed Wavelet-NARNN (WNARNN) model consistently outperformed the traditional NARNN model at all simulation sample sizes (150, 250, and 350) when used in combination with the statistical thresholding methods—Universal, Minimax, and SURE. This is confirmed by lower RMSE, MAE, and MAPE values, as well as higher correlation coefficients. Incorporating wavelet-based noise removal significantly improved the prediction accuracy.

At a sample size of 150, the combination of the db22 wavelet and the SURE threshold achieved the best results (Table 8). The db27 model achieved its highest performance at a sample size of 250 according to SURE results shown in Table 9, while at a sample size of 350 the db5 model achieved equivalent performance with the global threshold and Sym8-SURE element (Table 10). The wavelet transformation improved the model's capability to identify essential time series components because it removed undesired sounds while it permitted critical signal patterns to emerge, which resulted in greater predictive accuracy than the NARNN model.

Conclusions

The research study established its findings through actual data analysis and simulation testing that used sample groups of 150 and 250 and 350 which researchers tested 250 times each. The main findings of the study can be described through the following summary.

- The results showed that specific wavelet combinations including **Sym2 with Universal thresholding** and **db25 with SURE method** achieved the best performance by reducing all statistical evaluation metrics. This demonstrates how choosing suitable wavelet functions together with thresholding methods can improve forecasting accuracy.

- The hybrid models (WSARIMA and WNARNN) achieved better results than their traditional counterparts (SARIMA and NARNN) because wavelet transformation enhanced signal quality while minimizing signal interference. The results produced higher correlation coefficients together with diminished prediction errors. The proposed hybrid models serve as dependable instruments for both electricity demand prediction and peak electricity demand modelling.

Limitations and Future Work

Despite the positive results obtained in this study, several limitations that should be acknowledged. The proposed models were evaluated using a single dataset related to the maximum electricity demand in the Kurdistan region of Iraq. Consequently, applying this methodology on alternative datasets and different application domains may enhance the assessment of its generalization capability. This study also focused on integrating wavelet transformation with SARIMA and NARNN models only, without addressing modern deep learning models, which could potentially offer improved predictive performance and adaptability to complex datasets.

References

[1] Gross, G. & Galiana, F. D. Short-term load forecasting. *Proc. IEEE*, **75**, 1558–1573 doi :10.1109/PROC.1987.13927 (2017).

[2] Fan, S. & Hyndman, R. J. Short-term load forecasting based on a semi-parametric additive model. *IEEE Trans. Power Syst.*, **27**, 134–141 doi :10.1109/TPWRS.2011.2162082 (2011).

[3] Zhang, G., Patuwo, B. E. & Hu, M. Y. Forecasting with artificial neural networks. *Int. J. Forecast.*, **14**, 35–62 doi :10.1016/S0169-2070(97)00044-7 (2018).

[4] Box, G. E., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M. Time Series Analysis: Forecasting and Control (John Wiley & Sons, 2015).

[5] Brockwell, P. J. & Davis, R. A. Introduction to Time Series and Forecasting (Springer, New York, 2002). doi :10.1007/978-3-319-29854-2

[6] Percival, D. B. & Walden, A. T. Wavelet Methods for Time Series Analysis (Cambridge Univ. Press, 2000).

[7] Amjad, N. & Keynia, F. Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. *IEEE Trans. Power Syst.*, **24**, 306–318 doi :10.1109/TPWRS.2008.2006997 (2008).

[8] Oonsivilai, A. & El Hawary, M. E. Wavelet neural network based short-term load forecasting of electric power system commercial load. In Proc. Can. Conf. Electr. Comput. Eng. (CCECE), **3**, 1223–1228 (IEEE, 1999). doi :10.1109/CCECE.1999.804865

[9] Pindoriya, N. M., Singh, S. N. & Singh, S. K. Adaptive wavelet neural network-based energy price forecasting in electricity markets. *IEEE Trans. Power Syst.*, **23**, 1423–1432 doi :10.1109/TPWRS.2008.92225 (2008).

[10] Nury, A. H., Hasan, K. & Allam, M. J. B. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *J. King Saud Univ. Sci.*, **29**, 47–61 doi :10.1016/j.jksus.2015.12.002 (2017).

[11] Benrhmach, G., Namir, K., Ouyaghroumni, J. & Namir, A. Financial time series prediction using wavelet and artificial neural network. *J. Math. Comput. Sci.*, **11**, 5487–5500 doi :10.28919/jmcs/5978 (2021).

[12] Zheng, X., Li, Y., Wu, J., Wang, M. & Zhang, H. Short-time wind speed prediction based on Legendre multi-wavelet neural network. *CAAI Trans. Intell. Technol.*, **8**, 946–962 doi :10.1049/cit2.12157 (2023).

[13] Wang, Y., Sun, S. H., Fathi, G. & Eslami, M. Improving the method of short-term forecasting of electric load in distribution networks using wavelet transform combined with ridgelet neural network optimized by self-adapted Kho-Kho optimization algorithm. *Heliyon*, **10**, e28381 doi :10.1016/j.heliyon.2024.e28381 (2024).

[14] Sheppard, J. A., González, M. Z., Ledezma, A. & Gómez, J. A. Significance of wavelet transform and data mining in forecasting sea level time series at the Pacific entrance of the Panama Canal. *Rev. Esp. Ing. Cienc. Tierra*, **4**, 31–49 doi :10.48204/reict.v4n2.6747 (2025).

[15] Ali, T. H., Hamad, A. A., Mahmood, S. H. & Ahmed, K. H. ARIMAX time series analysis for a general budget in the Kurdistan Region of Iraq using wavelet shrinkage. *Commun. Stat. Case Stud. Data Anal. Appl.*, **11**, 164–188 doi :10.1080/23737484.2025.2486984 (2025).

[16] Shumway, R. H. & Stoffer, D. S. Time Series Analysis and Its Applications (Springer, 2006).

[17] Mahmood, S. H. & Ahmed, M. M. Application of Elman neural network and SARIMA model to modeling road traffic accident in the Kurdistan Region of Iraq. *Iraqi J. Stat. Sci.*, **20**, 43–56 (2023).

[18] Fausett, L. V. Fundamentals of Neural Networks: Architectures, Algorithms, and Applications (Prentice-Hall, 1994).

[19] Pamain, A., Rao, P. K. & Tilya, F. N. Prediction of photovoltaic power output based on different non-linear autoregressive artificial neural network algorithms. *Glob. Energy Interconnect.*, **5**, 226–235 doi :10.1016/j.gloei.2022.04.019 (2022).

[20] Rahim, S. A. & Botani, D. S. I. Hybrid neural network and time series approaches for forecasting CO_2 emissions: Iraq as a case study. *Passer J. Basic Appl. Sci.*, **8(1)**, 186–194 doi :10.24271/PSR.2025.550660.2392 (2026).

[21] Beermat, G., Namir, K., Namir, A. & Bouyaghroumni, J. Nonlinear autoregressive neural network and extended Kalman filters for prediction of financial time series. *J. Appl. Math.*, **2020**, 1–6 doi :10.1155/2020/5057801 (2020).

[22] Athermanous, R., Kim, D., Hossain, M. S., Alazab, M. & Jolfaei, A. Nonlinear autoregressive neural networks to predict hydraulic fracturing fluid leakage into shallow groundwater. *Water*, **12**, 1–14 doi :10.3390/w12030841 (2020).

[23] Qader, H. M., Birdawod, H. Q., Qader, H. M. & Sedeeq, B. S. Capability process to optimize specification by impact of wavelet analysis. *Cihan Univ.-Erbil J. Humanit. Soc. Sci.*, **9**, 53–60 doi :10.24086/cuejhss.v9n1y2025.pp53-60 (2025).

[24] Hasan, M. T., Ali, T. H. & Kareem, N. S. Multivariate CUSUM Daubechies discrete wavelet transformation coefficients charts for quality control. *Passer J. Basic Appl. Sci.*, **7(1)**, 533–546 doi :10.24271/PSR.2025.491277.1834 (2025).

[25] Abramovich, F., Bailey, C. T. & Captains, T. Wavelet analysis and its statistical applications. *J. R. Stat. Soc. Ser. D (The Statistician)*, **49**, 1–29 (2000).

[26] Elias, I. I. & Ali, T. H. Optimal level and order of the Coiflets wavelet in the VAR time series denoise analysis. *Front. Appl. Math. Stat.*, **11**, 1526540 doi :10.3389/fams.2025.1526540 (2025).

[27] Ali, T., Barware, N. H. S. & Ramadhan, D. L. Using the hybrid proposed method for quantile regression and multivariate wavelet in estimating the linear model parameters. *Iraqi J. Stat. Sci.*, **20**, 9–24 doi :10.33899/iqjoss.2023.178679 (2023).

[28] Elias, I. I. VARMA time series model analysis using discrete wavelet transformation coefficients for Coiflets wavelet. *Passer J. Basic Appl. Sci.*, **7(2)**, 657–677 doi :10.24271/PSR.2025.512121.2011 (2025).

[29] Guo, T., Zhang, T. & Lim, E. A review of wavelet analysis and its applications: Challenges and opportunities. *IEEE Access*, **10**, 58869–58903 doi :10.1109/ACCESS.2022.3179517 (2022).

[30] Jain, L. & Singh, P. A novel wavelet thresholding rule for speckle reduction from ultrasound images. *J. King Saud Univ. Comput. Inf. Sci.*, **34**, 4461–4471 doi :10.1016/j.jksuci.2020.10.009 (2022).

[31] Schluter, S. & Deuschle, C. Using wavelets for time series forecasting: Does it pay off. *IWQW Discuss. Pap. Ser.*, **2010**, 8–9 (2010).

[32] Yan, Z., Zhang, Y., Liu, Q., Wang, M. & Chen, L. Application of ultrasonic Doppler technology based on wavelet threshold denoising algorithm in fetal heart rate and central nervous system malformation detection. *World Neurosurg.*, **149**, 380–387 doi :10.1016/j.wneu.2020.10.030 (2020).

[33] Joy, J., Peter, S. & John, N. Denoising using soft thresholding. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, **2**, 1027–1032 (2013).