https://doi.org/10.56093/JISAS.V80I1.13
Available online at https://isas.live/newsite
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 80(1) 2026 135–148

Improved Crop Yield Forecasting using Wavelet-Based Denoising and Artificial Neural Networks Framework

Md Yeasin, Ranjit Kumar Paul, Pushkar Bora, Himadri Shekhar Roy, Prakash Kumar and A.K. Paul

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 01 April 2026; Revised 09 April 2026; Accepted 15 April 2026

SUMMARY

Agricultural crop yields are often non-stationary, noisy and influenced by multiple environmental factors, which limits the effectiveness of traditional statistical models. Advanced hybrid approaches that can capture both linear and non-linear patterns, along with multi-scale characteristics, are required for reliable forecasting. This study aimed to develop and evaluate a hybrid wavelet-based forecasting framework by integrating the Maximal Overlap Discrete Wavelet Transform (MODWT) with Artificial Neural Network (ANN) and Autoregressive Integrated Moving Average (ARIMA) models for predicting crop yield series of Total Pulses, Gram, and Tur in India. Annual yield data from 1950 to 2023 were analysed using a hybrid modelling framework. The MODWT was applied to decompose the original time series into multi-resolution components, followed by denoising through thresholding techniques. The decomposed components were then modelled using ANN and ARIMA. Optimal wavelet filters and decomposition levels were obtained based on predictive performance. The results showed that hybrid wavelet-based models significantly outperformed standalone ARIMA and ANN models. This study contributes to the literature by systematically optimizing wavelet filters and decomposition levels within a hybrid modelling framework and demonstrates the effectiveness of combining wavelet-based denoising with machine learning and statistical models for non-stationary agricultural time series data.

Keywords: Hybrid model; Neural network; Pulse yield;Wavelets and volatility.

1. INTRODUCTION

Accurate crop yield forecasting is crucial for ensuring food security, effective agricultural planning, market stability and informed policy decision-making. By analysing the historical data, early predictions become feasible which is helpful for farmers and agronomists for strategic planning in resource allocation, supply chain management, policy formulation and mitigating risks (Anjoy *et al.,*2017). Among the traditional statistical models, Autoregressive Integrated Moving Average (ARIMA) model is one the most preferred methods for crop yield forecasting (Mila and Parwin, 2019). ARIMA effectively capture linear trends, autocorrelation and seasonality under assumptions of stationarity and linearity. However, ARIMA model struggle with the inherent non-stationarity, non-linearity, noisy and high volatility data (Paidipati and Banik, 2019). Agricultural data often exhibit time varying volatility and heteroscedasticity arising from erratic weather, pest outbreaks and market shocks. Paul *et al.* (2009, 2014) and Musunuru *et al.* (2013) addressed these limitations using extensions such as Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized ARCH (GARCH) family of models for modelling volatility in agricultural commodity prices and yield data. These models are parametric which means they rely on predefined equations and error patterns that may not match reality. As the crop yields involve tricky non-

Corresponding author: Ranjit Kumar Paul
E-mail address: ranjitstat@gmail.com

linear changes, shifting trends over time and influenced by multiple factors, these specific parametric structures are not suitable to capture these complex patterns (Huang *et al.* 2012). Machine learning models address these shortcomings by capturing complex non-linear relationships without any strict assumptions about data distribution (Shawon *et al.,* 2025). It leads to higher predictive accuracy than traditional approaches (Lionel *et al.* 2025). Artificial Neural Networks (ANN) is type of machine learning and widely used in crop yield prediction. However, the performance of machine learning models may be affected by noise present in agricultural time series data (Alhnaity *et al.* 2021). Agricultural time series data often contain noise due to measurement errors, environmental variability and external disturbances. Noise can reduce forecasting accuracy and make it difficult for prediction models to learn meaningful patterns. Therefore, denoising is an important preprocessing step that helps in removing noise while preserving useful information (Rani, 2025). The decomposition techniques are used to decompose the original time series into simpler, multi-resolution components. Decomposition allows models to analyse different frequency components separately and improve prediction accuracy (Shao *et al.,* 2017). The Wavelet transform has emerged as a powerful tool for denoising and decomposition of non-stationary time series data. Unlike traditional methods, wavelet transform provides both time and frequency information simultaneously. Wavelet transform decomposes the original time series into low-frequency approximations, which capture smooth underlying trends, and high-frequency details which highlight short-term noise (Paul and Garai 2021). Paul *et al.* (2022) developed a wavelets-based ANN technique for forecasting nonlinear, non-stationary agricultural price series like tomato wholesale prices. This approach preprocesses data via wavelet decomposition (haar and d4 filters at levels 3 and 6) to extract clean signals, models components with ANN, then reconstructs via inverse transform. Paul *et al.* (2019) applied wavelet decomposition with 'haar' wavelet filter to India's sub divisional rainfall series followed by ARIMA and ANN modelling on denoised components, with inverse Maximal Overlap Discrete Wavelet Transform (MODWT) reconstruction outperforming standalone ARIMA in forecast accuracy. Hybrid machine learning approaches were also used for time series prediction in the literature (Mitra and Paul, 2017; Oikonomidis *et al.,* 2022;Das *et al.* 2023; Rakshit and Paul, 2024; Bakulin *et al.,* 2025; Naik, 2025; Khosravi *et al.,*2026). Paul *et al.* (2025) proposed Wavelet SVR hybrid using wavelet denoising with Daubechies/Haar filters on potato wholesale price series which is followed by SVR modelling of sub series and ensemble aggregation with outperforming standalone ARIMA, ANN and RF benchmarks. Paul ***et al.*** (2025) introduced WaveDeepCrop, a wavelet enhanced deep learning framework for crop yield forecasting that employs diverse filters such as haar,d4,la8,bl14 with d4 proving most effective for superior noise resilience and accuracy over basic haar expansions.

Despite the growing use of hybrid wavelet machine learning approaches in agricultural forecasting, several important limitations remain in the existing literature. Most previous studies have primarily focused on a limited set of mother wavelet filters, particularly Haar or Daubechies wavelets and have generally used fixed or predetermined levels of decomposition. However, agricultural time series data are often highly non-stationary, noisy and influenced by multiple environmental factors which may require different wavelet structures and decomposition levels to effectively capture their underlying patterns. As a result, the sensitivity of forecasting performance to the selection of mother wavelet filters and decomposition levels has not been systematically examined in the context of crop yield forecasting.

Therefore, this study aimed to address this gap by systematically evaluating a wide range of mother wavelet filters and decomposition levels for improving forecasting performance. The study employed the MODWT to decompose crop yield time series into multi-resolution components. Hard thresholding is applied to the detail coefficients to remove noise and the denoised series is reconstructed using inverse MODWT. The denoise series is then used to develop a WaveletANN forecasting framework. Through this approach, the study seeks to identify the optimal wavelet filter and decomposition level that enhance the overall forecasting accuracy of crop yield predictions.

2. METHODOLOGY

2.1 Autoregressive Integrated Moving Average (ARIMA) Model

The ARIMA model is the widely used statistical technique for time series analysis and forecasting. This modelling framework provides a systematic procedure for identifying, estimating and validating time series models.

Let y_t denotes the original time series data. Mathematically the ARIMA model can be expressed as

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \tag{1}$$

where B is the backshift operator, defined as $By_t = y_{t-1}$, p denotes the autoregressive (AR) order, d denotes the differencing order, q denotes the order of moving average (MA) and ε_t represents the white noise error term with $\varepsilon_t \sim iidN(0, \sigma^2)$. The autoregressive and moving average polynomials are given by $\phi(B) = 1 - \phi_1 B - \phi_2 (B)^2 - \ldots \phi_p (B)^p$ and $\theta(B) = 1 - \theta_1 B - \theta_2 (B)^2 - \ldots \theta_p (B)^q$ respectively. Here, ϕ_i and θ_j are the parameters of the AR and MA components respectively.

The ARIMA model is particularly effective in capturing linear temporal dependencies in time series data and has been widely applied in economic, agricultural and environmental time series forecasting studies.

2.2 Wavelet Transform

The MODWT is an extension of the traditional Discrete Wavelet Transform (DWT) that is particularly suitable for time series analysis. Unlike DWT, which applies downsampling at each decomposition level and therefore reduces the length of the coefficient series, MODWT eliminates the down sampling step and retains the original length of the signal at all decomposition levels. This property preserves the alignment between the transformed coefficients and the original observations, resulting in translation invariance and improved interpretability. These characteristics make MODWT well suited for analysing non-stationary time series, which are commonly observed in agricultural and economic data.

Let X_t, $t = 0, 1, \ldots, N-1$, denote a discrete time series. The initial scaling coefficients are defined as

$$s_{0,t} = X_t \tag{2}$$

At decomposition level j, the wavelet (detail) coefficients and scaling (approximation) coefficients are obtained with wavelet filter h_l and scaling filter g_l. The filters are up sampled by inserting $2^{j-1} - 1$ zeros between coefficients and allow the transform to operate at multiple resolutions while preserving the full sample size. This procedure produces a multiresolution representation of the series consisting of wavelet details $\widetilde{d}_1, \widetilde{d}_2 \ldots, \widetilde{d}_j$ and the final scaling component $\widetilde{s}_j$. Typically, the maximum decomposition level J is chosen such that $J \approx log_2 N$.

Several orthogonal wavelet filters can be used within the MODWT framework. The haar wavelet is the simplest orthogonal wavelet with filter length two and one vanishing moment. It provides computational simplicity and is effective for detecting abrupt changes in the time series, although it has limited ability to capture smooth variations. The Daubechies d4 wavelet has four vanishing moments and a compact support, enabling it to capture localized features and sharp transitions in non-stationary time series. However, it exhibits some asymmetry that may introduce minor phase distortion. The Least Asymmetric la8 wavelet provides improved symmetry and smoother approximations compared with d4, resulting in better time localization and interpretability in multi scale analysis. Another alternative is the Best Localized bl14 wavelet, which has a longer filter length and improved frequency localization. This property enhances its ability to isolate specific frequency components and reduce variance in the decomposed series.

Haar wavelet is a simple and base wavelet filter. It has the advantages of being fast and memory efficient. It is a discontinuous step function. The high-pass(g) and low-pass(h) filters is given by $g = \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}}\right]$ and

$h = \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \right]$ which are the shortest possible wavelet filters (Burrus *et al.,*1998).

$$\text{Mother Wavelet } (\psi(n)) = \begin{cases} 1 & 0 \le n \le 0.5 \\ -1 & 0.5 \le n \le 1 \\ 0 & otherwise \end{cases} \quad (3)$$

Daubechies wavelet filter is an invertible wavelet transforms that generalize the haar transform. It is more localized and smoother (Sharif and Khare, 2014). The Daubechies family wavelets are denotedas dbN or dN, where N is the order of the wavelet. The order N is the number of vanishing moments for the particular Daubechies wavelet. A wavelet has N vanishing moments if

$$\int_{-\infty}^{\infty} t^k \psi(t)\, dt = 0 \text{ for } 0 \leqslant k < N \quad (4)$$

The Daubechies wavelets are orthogonal but not symmetrical (Lindfield and Penny, 2019). The range over which they are non-zero is $[0, 2N-1]$.

In this study, Least Asymmetric wavelet of order 8 (la8) from the Symlet family of wavelets proposed by Ingrid Daubechies (1992) as a modification of the Daubechies wavelets has been used. The primary objective of constructing Symlet wavelets was to achieve near symmetry while preserving orthogonality and compact support. Similar to Daubechies wavelets, the la8 wavelet is implemented using a pair of quadrature mirror filters consisting of a low pass scaling filter h and a high pass wavelet filter g. The la8 wavelet has eight vanishing moments which means it can exactly represent polynomials up to degree seven. A wavelet has N vanishing moments if

$$\int_{-\infty}^{\infty} t^k \psi(t)\, dt = 0 \text{ for } 0 \leqslant k < N \quad (5)$$

Thus, the la8 wavelet satisfies the above condition for N=8. Due to la8's near symmetry, orthogonality and good time frequency localization, it is widely used in signal processing, image compression and time series analysis.

Biorthogonal Least asymmetric wavelet of length 14 (bl14) wavelet is employed for Wavelet-ARIMA decomposition due to its good smoothness and symmetry property. The bl14 belongs to the Biorthogonal wavelet family. Unlike orthogonal wavelets such as the Daubechies and Symlet families, biorthogonal wavelets employ two distinct sets of scaling and wavelet functions in which one for decomposition and another for reconstruction. This property relaxes the strict orthogonality condition and allows the construction of wavelets that are perfectly symmetric. Due to its symmetric filters and efficient reconstruction properties, the bl14 wavelet is frequently used in maximal overlap wavelet transforms for analysing non-stationary time series analysis.

For denoising, soft or hard thresholding is applied to the detail coefficient. High frequency noise dominated details are suppressed while approximation coefficients are retained. The Inverse Maximum Overlap Discrete Wavelet Transform (IMODWT) reconstruct the denoised sample via up sampling and synthesis filters ensuring perfect reconstruction for orthogonal wavelets.

After decomposition, the original series can be reconstructed using the Inverse MODWT (IMODWT). Reconstruction is achieved by combining the wavelet and scaling coefficients using the corresponding synthesis filters. For orthogonal wavelet families, the reconstruction filters are equivalent to the time-reversed analysis filters. By iteratively combining the coefficients from the highest decomposition level to the lowest level, the original signal can be recovered exactly. This perfect reconstruction property ensures that the MODWT provides a consistent and reversible transformation of the time series.

2.3 Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) follows human brain neuron connections to learn complex input-output relationships in data through layered processing. ANN method consists of three layers namely, input (raw features), hidden (nonlinear transformations) and output (predictions). It makes the connections among the layers by connecting every single layer with the all possible neuron in the other layers. Hyperparameters like learning rate (step size), activation functions, hidden nodes (model capacity), and input scaling critically determine accuracy; poor tuning leads to under/overfitting.

ANNs can be mathematically expressed as a composition of non-linear functions that transform input variables into an output through weighted connections. Among different ANN architectures, the Multi-Layer Perceptron (MLP) is the mostly used model for regression and forecasting problems (Hounmenou *et al.*. 2021).

Let an ANN model consists of p input neurons, h hidden neurons and one output neuron. The input signal received by the j^{th} hidden neuron is given by

$$z_j = \beta_{0j} + \sum_{i=1}^{p} \omega_{ij} x_i \tag{6}$$

where x_i is input variables where i = 1, 2,…p , and ω_{ij} is the weight connecting the ith input neuron to the jth hidden neuron, β_{0j} is the bias term associated with the jth hidden neuron.

Then the hidden neuron applies a nonlinear activation function to this weighted sum. In many applications, the logistic (sigmoid) activation function is used:

$$\phi\left(z_j\right) = \frac{1}{1+\mathrm{e}^{-z_j}} \tag{7}$$

Thus, the output of the j^{th} hidden neuron becomes

$$v_j = \phi\left(\beta_{0j} + \sum_{i=1}^{P} w_{ij} x_i\right) \tag{8}$$

where $\phi(.)$ is the activation function used in hidden neurons.

The final output of the MLP model is obtained by combining the outputs of all hidden neurons through another weighted summation:

$$Y = I\left(a_0 + \sum_{j=1}^{h} w_j v_j\right) \tag{9}$$

where a_0 is the bias term of the output neuron, w_j is the weight connecting the jth hidden neuron to the output neuron.

Since the identity function I(.) is used in the output layer for regression problems, the output can be written as

$$Y = a_0 + \sum_{j=1}^{h} w_j \phi\left(\beta_{0j} + \sum_{i=1}^{p} w_{ij} x_i\right) \tag{10}$$

The equation shows that the ANN model represents a non-linear mapping between input variables and the output variable, where hidden neurons capture complex nonlinear interactions among inputs. The parameters of the ANN model, including the weights w_{ij}, w_j and biases β_{0j} and a_0, are estimated during the training phase using the backpropagation algorithm.

2.3.1 Wavelet based Hybrid Models

In order to capture the multi scale time series characteristics, the MODWT is applied. Let X_t denote the original time series of N length. The maximum decomposition level is chosen as $J \leq \log_2(N)$ where N is the total number of observations. In the hybrid Wavelet-ANN (WaveANN) framework the original series is decomposed into one approximation (V_j) and j detail sub-series (W_1…., W_j) using appropriate wavelet filters. Each sub-series is modelled independently using ANN model.

Mathematically, the reconstructed forecast can be written as

$$\widehat{X_t} = \sum_{j=1}^{J} \widehat{W_{j,t}} + \widehat{V_{J,t}} \tag{11}$$

where $\widehat{W_{j,t}}$ represents the predicted detail components, $\widehat{V_{J,t}}$ represents the predicted approximation component, $\widehat{X_t}$ represents the predicted series.

This hybrid wavelet-based modelling approach effectively captures both nonlinear patterns and multi scale temporal dynamics. This approach mitigates non-stationarity, reduce noise impact and improves overall accuracy. The schematic diagram of Wavelet-ANN/ARIMA model is shown in the Fig.1.

2.4 Performance Comparison Measures

The accuracy of above-described forecasting models and empirical comparison is carried out in terms of following criterion:

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i - \widehat{y_i}\right)^2}{N}} \tag{12}$$

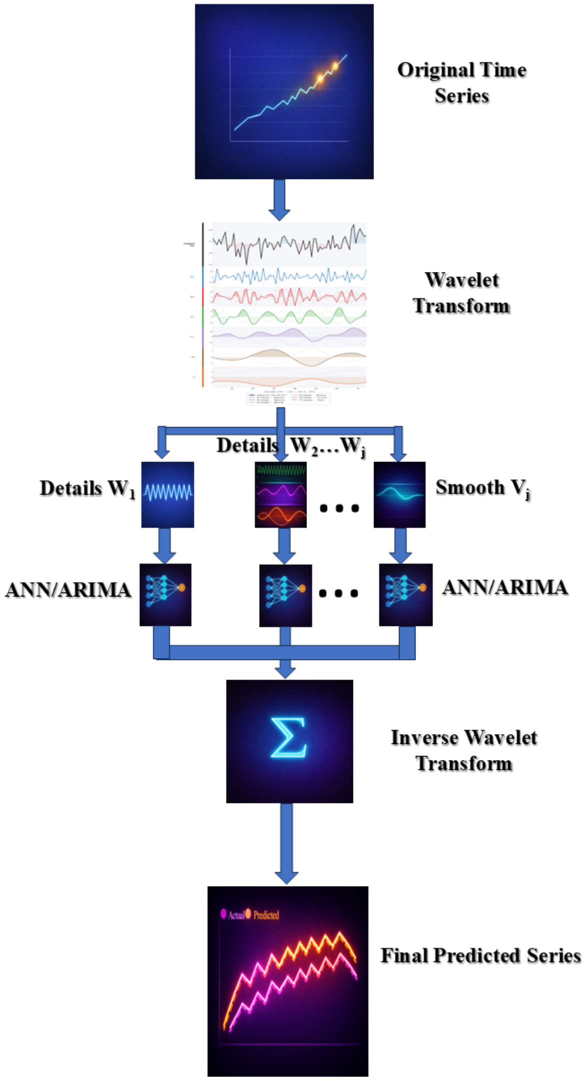


Fig. 1. Schematic diagram of Wavelet ANN/ARIMA model

Mean Absolute Error (MAE):

$$MAE = \frac{\sum_{i=1}^{N} |y_i - \hat{y}_i|}{N} \tag{13}$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = 100 * \frac{\sum_{1=1}^{n} |y_i - \hat{y}_i| / y_i}{N} \tag{14}$$

where y_i and $\hat{y}_i$ are the actual and predicted value respectively.

Lower value of RMSE, MAE and MAPE indicates more accuracy of the model for forecasting performance.

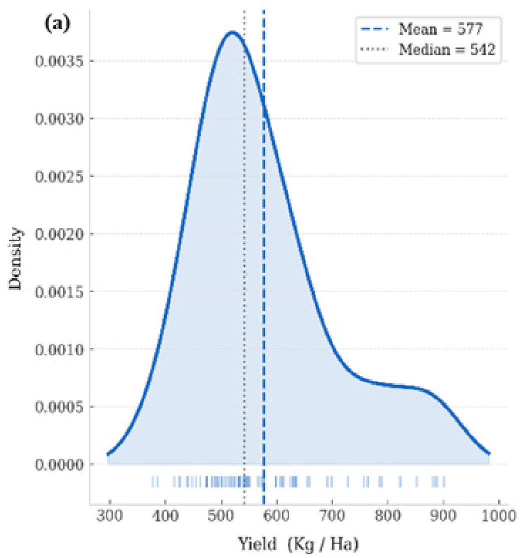
3. RESULTS AND DISCUSSION

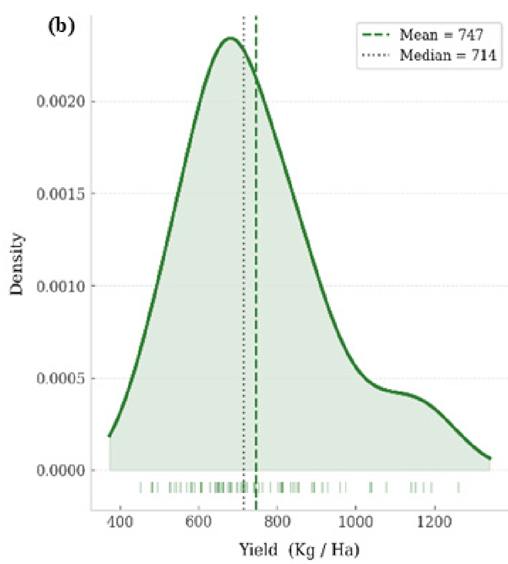
The yearly crop yield data of Pulses from 1950-2023 in India, namely, Total Pulses, Gram, Tur, were obtained from Unified Portal for Agricultural Statistics (UPAg) (https://www.upag.gov.in/). Total 74 observations are present in each series. Pulses are the world's most important protein-rich staple foods and provide essential nutrition across diverse cuisines. India leads global pulses production with contributing approximately 25% of worldwide output and produced 24.49 million tonnes in 2023-24. In Total Pulses, Kharif Pulse Crops have 33% production and Rabi Pulse Crops have 67% production. India produces Gram over 11.5 million tonnes annually in 2022-23 which accounts for more than 70% of Total Rabi Pulses production. India accounts for nearly 47% of Total Kharif Pulses production of Tur (https://dpd.gov.in/). Madhya Pradesh, Rajasthan, Maharashtra, Uttar Pradesh and Karnataka dominate production in Pulses.

To obtain a definitive understanding of the yield dynamics across Total Pulses, Gram and Tur, several descriptive statistics were presented in Table 1. In Table 1, Total Pulses exhibits a mean yield of 576.74 kg/ha with CV as 21.81%, indicating substantial yield fluctuations relative to its average. The distribution shows positive skewness, suggesting a right tailed pattern where higher yields occur less frequently but with greater magnitude, and leptokurtic kurtosis of 0.56 suggests peaked distribution with heavier tails than normal. The Jarque-Bera test confirms significant non-normality. Gram has the highest average yield of 746.97 kg/ha and displays the highest yield variability of 24.47% which reflects substantial production instability. Gram exhibits positive skewness of 0.87 and leptokurtic kurtosis of 0.46 indicate peaked distribution with extreme yield events. The Jarque-Bera test confirmed a non-normal yield distribution. Tur displayed mean yield of 721.32 kg/ha with variability of 12.85%. It shows near symmetric skewness of -0.04 and leptokurtic kurtosis of 0.64. The distribution Tur is also non-normal. The kernel density and line plot of all the three crops have been presented in Fig. 2 and 3 respectively.

Table 1. Descriptive Statistics of different Pulses yield (kg/ha)

Statistics		Total Pulses	Gram	Tur
Mean		576.74	746.97	721.32
Standard Error		14.62	21.25	10.78
Median		542.00	713.50	717.50
Mode		475.00	481.00	718.00
Standard Deviation		125.79	182.77	92.71
CV (%)		21.81	24.47	12.85
Kurtosis		0.56	0.46	0.64
Skewness		1.05	0.87	-0.04
Jarque Bera test	Test statistics	32.00	29.18	17.16
	p value	<0.0001	<0.0001	<0.001





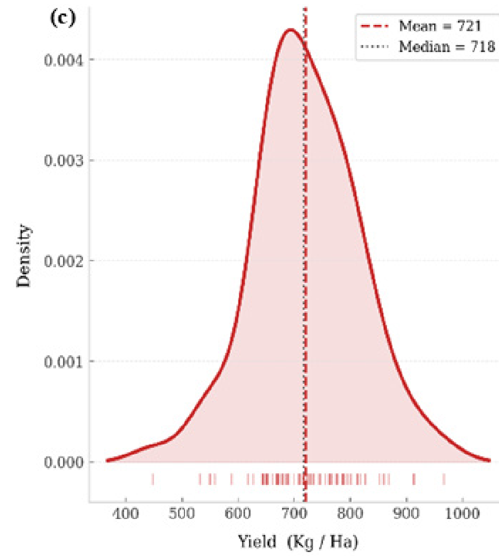


Fig. 2. Kernel density of (a) Total Pulse, (b) Gram and (c) Tur yield (kg/ha)

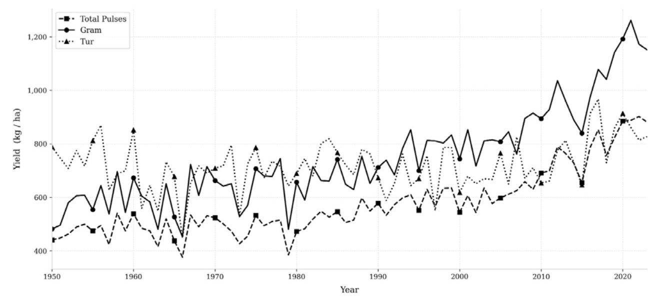


Fig. 3. Line plot of annual yield (kg/ha) for Total Pulses, Gram and Tur in India

To capture the multi-scale temporal dynamics ofyield data, MODWT is employed. The maximum decomposition level is determined by the formula $J = \text{int}\left(\log_2 N\right)$, where N represents the series length i.e.74. For the present analysis, though J_{max} has been found to be 6, but the optimal decomposition levels were set at 5, 4 and 5 for Total Pulses, Gram and Tur respectively in order to more precisely capture the variation present in the respective series.

The optimal lag for each model was determined for each MODWT decomposed sub-series. For the original yield series of Total Pulses, Gram, and Tur, significant autocorrelation was observed up to lag 4. The ACF plots for the Total Pulses, Gram and Tur are shown in the Fig. 4. Accordingly, lag 4 was selected as the input for the ANN models. For the ARIMA models, the appropriate models were identified as ARIMA (0,1,1) for Total Pulses, ARIMA (1,1,1) for Gram, and

ARIMA (1,0,0) for Tur based on AIC. In the wavelet-based models, the yield series were decomposed using MODWT. The detail components of Total Pulses, Gram, and Tur showed consistent significant autocorrelation at lag 4. This supported the use of a uniform input lag 4 for the WaveANN model.For the WaveARIMA model, the autoregressive (AR) order was determined separately for each component based on ACF analysis. Higher-order AR structures (p = 5) were used for high-frequency detail components. The smooth trend component was modelled using ARIMA (2,2,2) for Total Pulses, ARIMA (2,2,3) for Gram, and ARIMA (2,2,0) for Tur. The ACF plots for MODWT components for Total Pulses, Gram, and Tur are shown in the Fig. 5.The train–test split in the ratio of 80:20 was adopted to maintain the sequential integrity of the time series data.

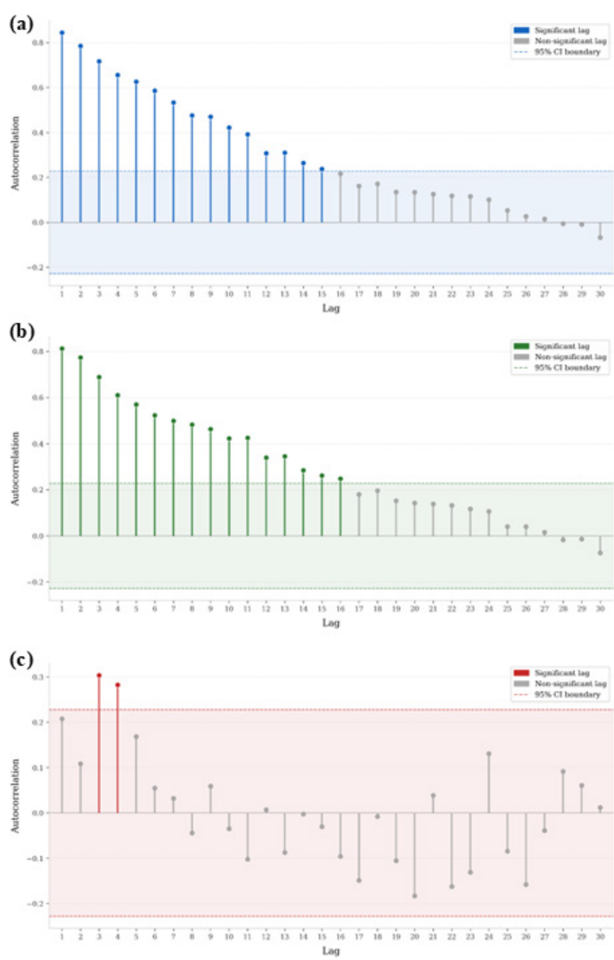
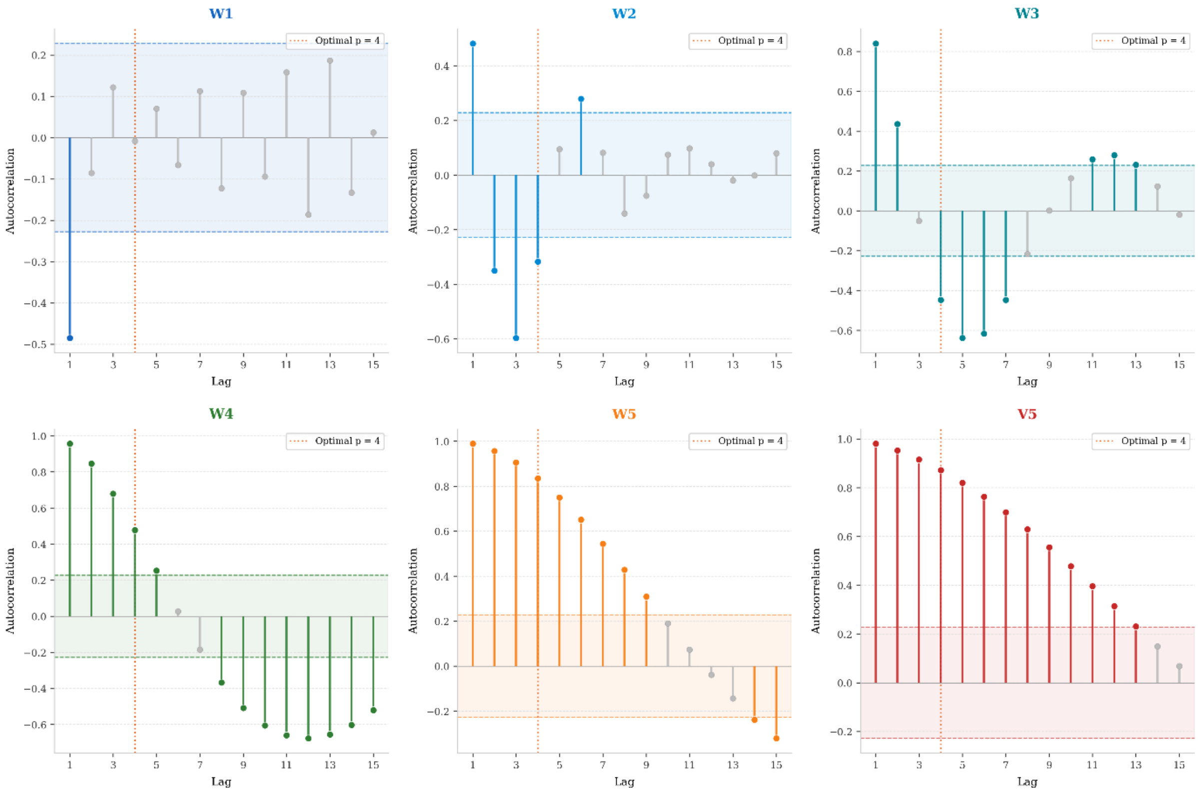


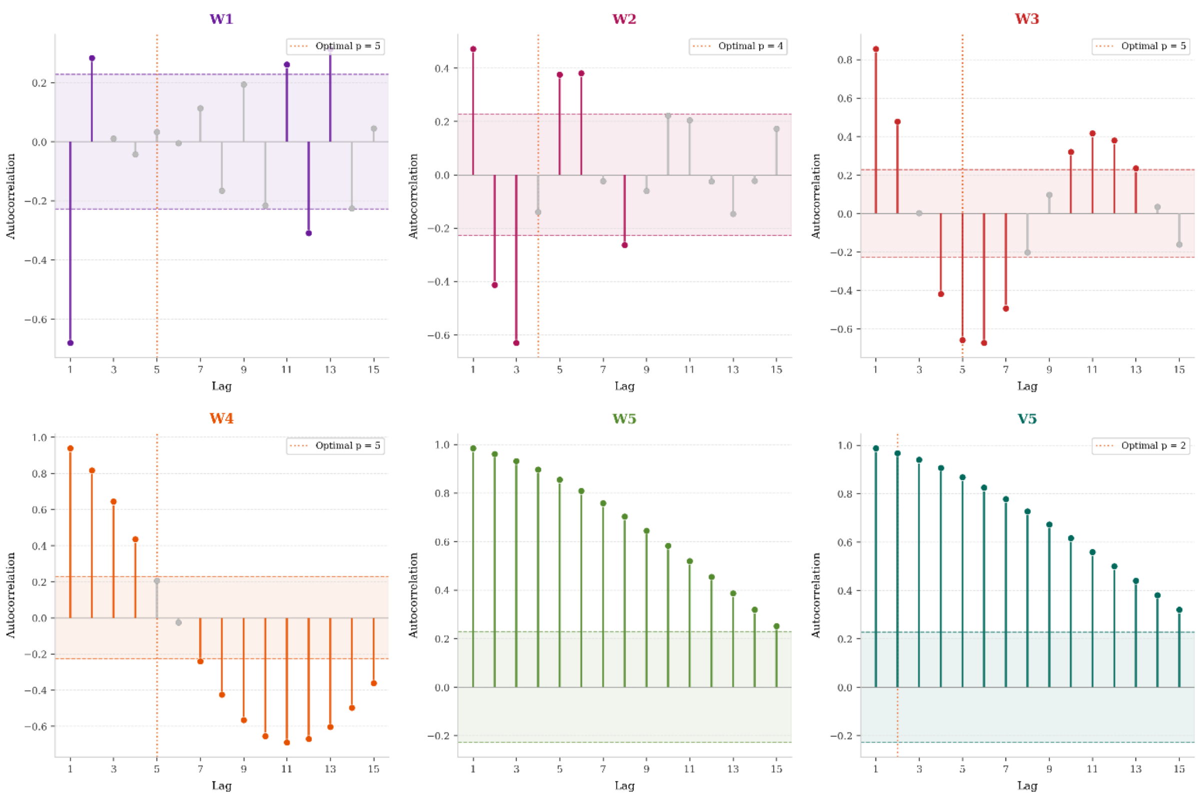
Fig. 4. ACF plots of yearly (a) Total Pulse, (b) Gram and (c) Tur Yield (kg/ha)

The decomposed components for Total Pulses, Gram can be visualized in Fig. 6. To evaluate the prediction performance at different levels of decomposition, the best decomposition level was selected. The optimal level was found to be 5 for WaveANN and WaveARIMA models in the case of Total Pulses, WaveARIMA for Gram, and both WaveANN and WaveARIMA for Tur. For WaveANN in the case of Gram, the best level was 4. Based on prediction performance, the most suitable wavelet filter was also identified. The la8 filter performed best for WaveANN in Total Pulses and Gram, while the bl14 filter was found to be optimal for WaveANN in Tur and for WaveARIMA in Total Pulses, Gram, and Tur. The wavelet decomposition in Fig. 6, effectively separates the time series into different frequency components.

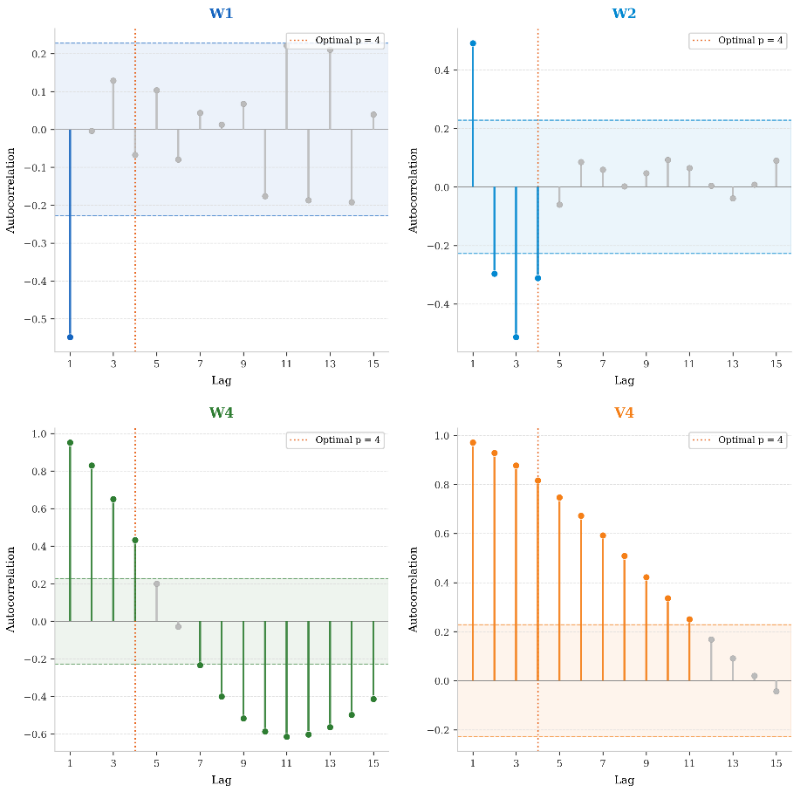
The forecasting performance of four models, namely ARIMA, ANN, WaveANN, and WaveARIMA, was evaluated using datasets of the Total Pulses, Gram, and Tur series. The comparative results are reported in Table 2.The results presented in Table 2 indicate that the hybrid wavelet-based models consistently outperform the standalone models. In particular, the WaveANN and WaveARIMA models exhibit notable improvements in forecasting accuracy over the conventional ANN and ARIMA approaches. For the Total Pulses series, the WaveANN model achieved the lowest forecasting errors on the test dataset, with RMSE, MAE, and MAPE values of 23.88, 18.04, and 2.44, respectively. In case of the Gram series, the WaveARIMA model demonstrated superior predictive performance, yielding RMSE, MAE, and MAPE values of 31.18, 24.26, and 2.26, respectively. For the Tur series, the WaveANN model again outperformed other models, with RMSE, MAE, and MAPE values of 42.62, 26.59, and 3.64, respectively, on the test dataset. The findings suggest that the incorporation of wavelet decomposition significantly enhances forecasting performance. This improvement can be attributed to the ability of wavelet-based methods to capture the multi-scale characteristics inherent in time series data, thereby improving the predictive capability of the models. The out of sample forecasts for each of the series for the period 2024 to 2030 has been computed based on the best fitted model and the same is reported in Fig. 7.


W1
Optimal p = 4
W2
Optimal p = 4
W3
Optimal p = 4
W4
Optimal p = 4
W5
Optimal p = 4
V5
Optimal p = 4
Autocorrelation
Lag

(a)


W1
Optimal p = 5
W2
Optimal p = 4
W3
Optimal p = 5
W4
Optimal p = 5
W5
V5
Optimal p = 2
Autocorrelation
Lag

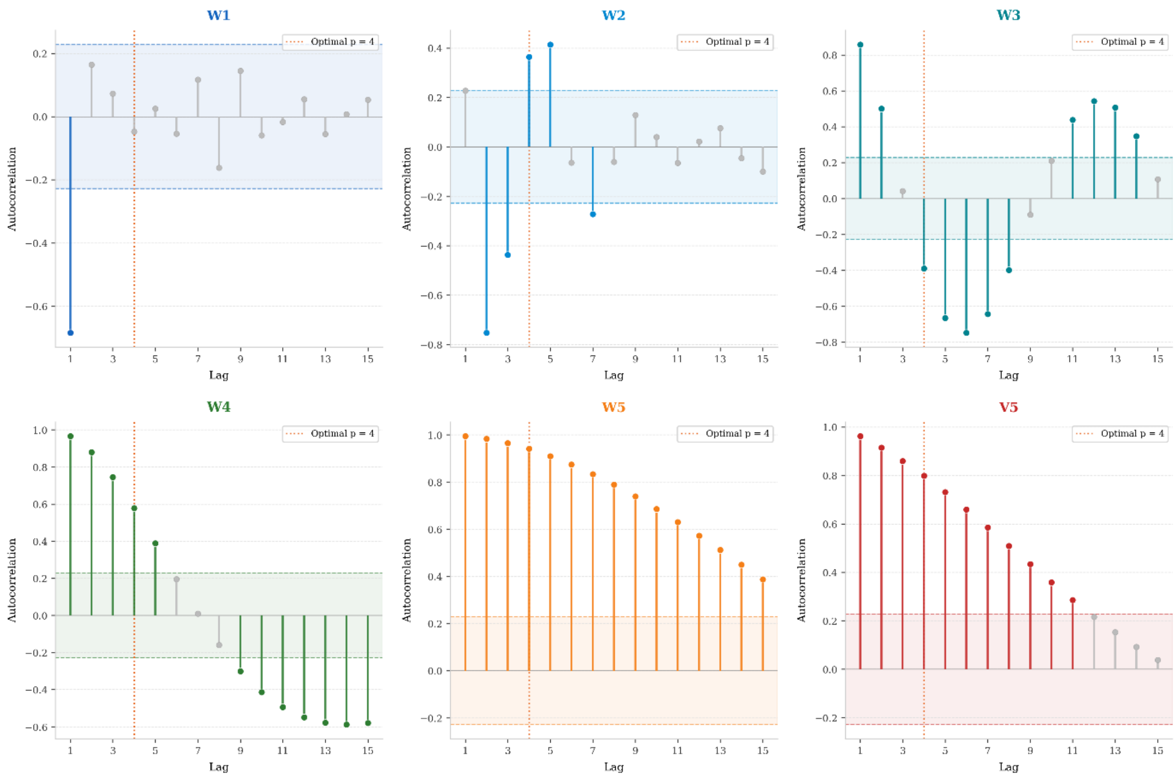
(b)


W1
W2
W3
W4
V4
Optimal p = 4
Autocorrelation
Lag

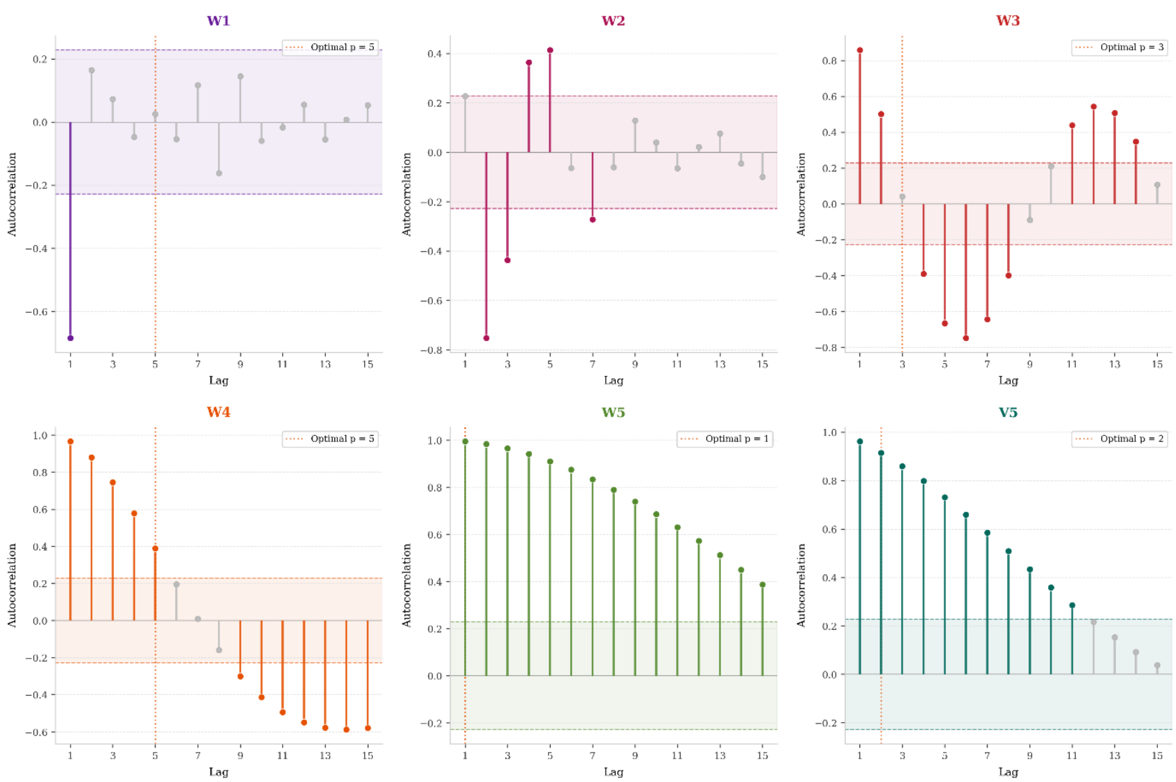
(c)


W1
W2
W3
W4
W5
V5
Optimal p = 5
Optimal p = 2
Autocorrelation
Lag

(d)



(e)



(f)

Fig. 5. ACF Plots of MODWT components for Total Pulses with (a) WaveANN la8 and (b) WaveARIMA bl14, Gram with (c) WaveANN la8 and (d) WaveARIMA bl14, Tur with (e) WaveANN bl14 and (f) WaveARIMA bl14

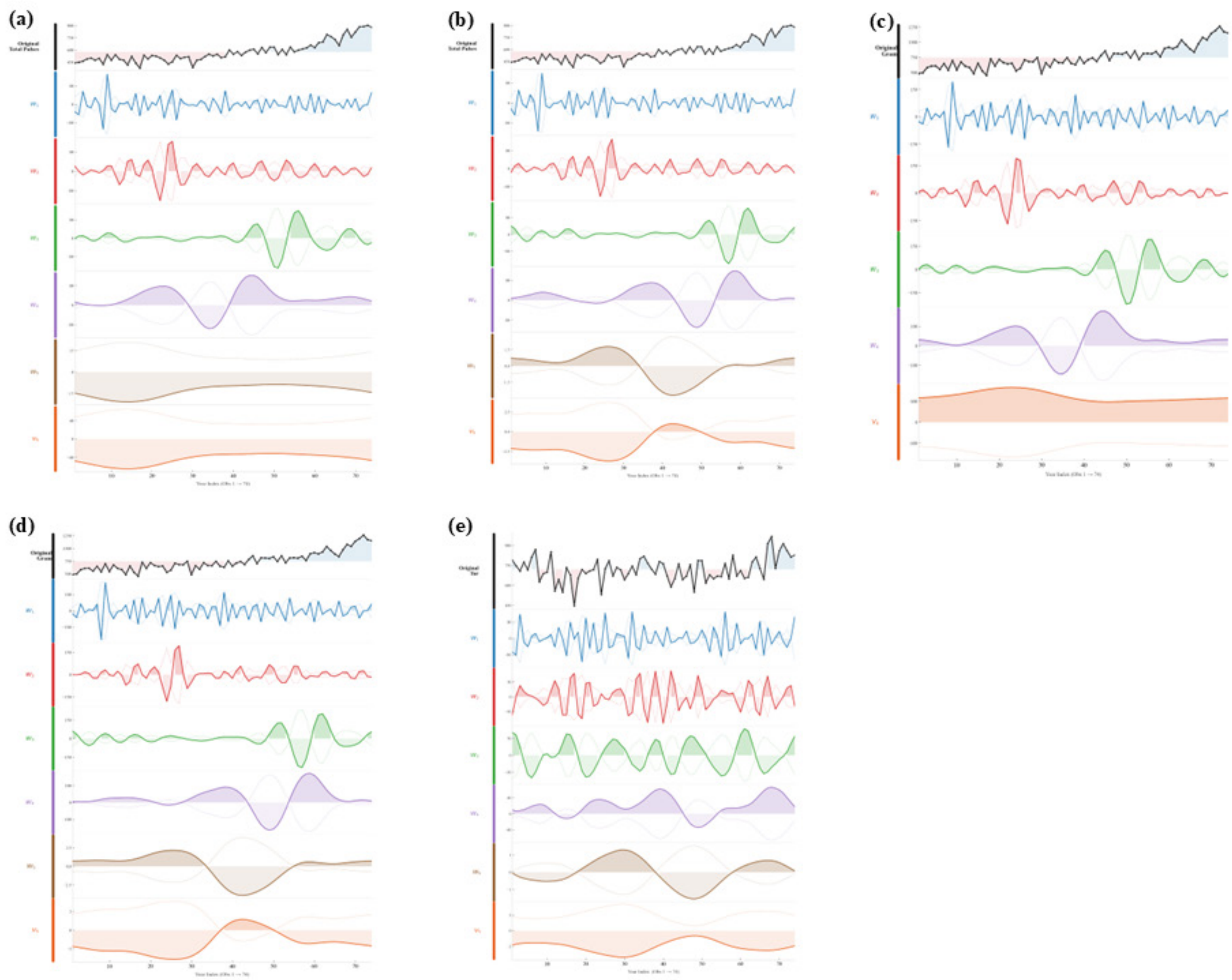
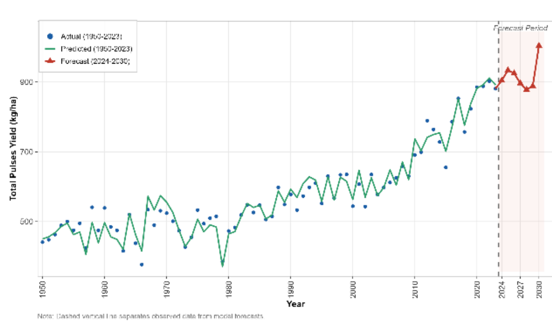
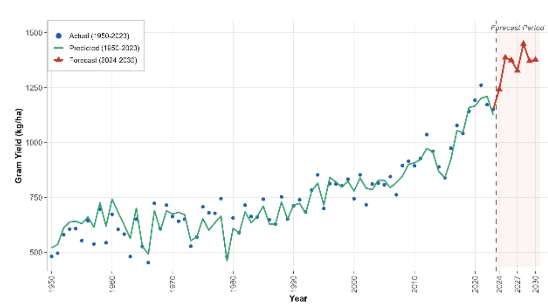


Fig. 6. Wavelet decomposition level for Total Pulses with (a) la8 and (b) bl14 filter, Gram with (c) la8 and (d) bl14 filter, Tur with (e) bl14 filter

Table 2. Performance Measures of models for different Pulses

	Total Pulses			Gram			Tur		
Models	**RMSE**	**MAE**	**MAPE (%)**	**RMSE**	**MAE**	**MAPE (%)**	**RMSE**	**MAE**	**MAPE (%)**
ANN	183.02	162.44	19.82	266.86	225.70	20.59	133.15	108.92	13.14
ARIMA	156.76	137.26	16.66	198.61	169.48	15.46	131.75	108.06	12.80
WaveANN	23.88	18.04	2.44	37.36	26.60	2.76	42.62	26.59	3.64
WaveARIMA	25.38	19.87	2.49	31.18	24.26	2.26	55.66	45.48	5.97





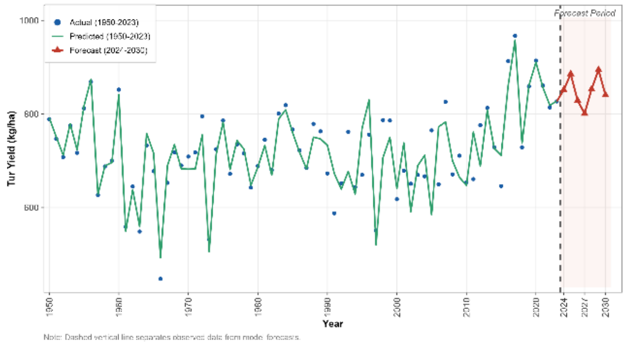


Fig. 7. Actual, predicted and out-of-sample forecast of Total Pulse, Gram and Tur yield

4. CONCLUSION

This study applied a wavelet-based decomposition framework integrated with forecasting techniques, namely ANN and ARIMA, to model and forecast the yield series of Total Pulses, Gram, and Tur. The proposed hybrid approach was evaluated using historical time series data, wherein the wavelet transform was employed to decompose the original series into multiple components corresponding to different frequency levels. This decomposition effectively captured the underlying structure of the series and reduced noise present in the raw data.The forecasting performance of both standalone models and hybrid wavelet-based models was assessed using three standard accuracy measures. The empirical results indicated that the hybrid models consistently outperformed the standalone models across both training and testing datasets. In particular, the WaveANN and Wave ARIMA models showed significant improvements in forecasting accuracy compared to conventional ANN and ARIMA models. The findings suggested that incorporating wavelet decomposition enhanced the ability to capture the multi-scale characteristics inherent in agricultural time series data, thereby improving predictive performance. Consequently, the proposed hybrid wavelet-based framework provided a robust and reliable approach for modelling and forecasting agricultural yield series. For future research, alternative wavelet filters and advanced machine learning techniques could be explored to further improve the predictive performance of the models.

ACKNOWLEDGEMENTS

The study is carried out under the project "Development of AI enabled models and web solution for prediction of crop yield". The authors are thankful to the Anusandhan National Research Foundation (ANRF) for funding the project under Core Research Grant Scheme.

REFERENCES

Alhnaity, B., Kollias, S., Leontidis, G., Jiang, S., Schamp, B. and Pearson, S. (2021). An autoencoder wavelet based deep neural network with attention mechanism for multi-step prediction of plant growth. *Information Sciences*, **560**, 35-50

Anjoy, P., Paul, R.K., Sinha, K., Paul, A.K. and Ray, M. (2017). A hybrid wavelet based neural networks model for predicting monthly WPI of pulses in India. *Indian Journal of Agricultural Sciences,* **87**(6), 834-839.

Bakulin, D.V., Beletskaya, N.V., Gorshkov, N.N., Djioeva, M.I., Mokhnatkina, U.S., Petrusevich, D.A. and Polyakov, A.A. (2025). Hybrid Models of Time Series Using Wavelet Transform. *Journal of Communications Technology and Electronics*, **70**(7), 309-318

Burrus S., Gopinath R., and Guo H. (1998). Introduction to Wavelets and Wavelet Transforms: A Primer. *Recherche*, **67**.

Das, P., Jha, G.K., Lama, A., and Parsad, R. (2023). Crop yield prediction using hybrid machine learning approach: a case study of lentil (Lens culinaris Medik.). *Agriculture*, **13**(3), 596.

Daubechies, I. (1992) Ten lectures on wavelets. *Society for industrial and applied mathematics*.

Hounmenou, C.G., Gneyou, K.E. and Kakai, R.G. (2021). A Formalism of the General Mathematical Expression of Multilayer Perceptron Neural Networks.

Huang, W., Huang, Z., Matei, M. and Wang, T. (2012). Price volatility forecast for agricultural commodity futures: the role of high frequency data. *Romanian Journal of Economic Forecasting.*

Khosravi, S., Nikoo, M.R., Zamani, M.G., Barzegar, R., Najafipour, P. and Gandomi, A.H., (2026). Enhancing long-term water quality forecasting with a hybrid deep-learning approach integrating MODWT, CNN, and GRU. *Hydrological Sciences Journal*, **71**(2), 270-293.

Lindfield, G. and Penny, J. (2019). Analysing data using discrete transforms. *Numerical Methods*, 383-431

Lionel, B.M., Musabe, R., Gatera, O. and Twizere, C., (2025). A comparative study of machine learning models in predicting crop yield. *Discover Agriculture*, **3**(1),151.

Mila, F.A. and Parvin, M.T., (2019). Forecasting area, production and yield of onion in Bangladesh by using ARIMA model. *Asian Journal of Agricultural Extension, Economics & Sociology*, **37**(4),1-12.

Mitra, D. and Paul, R.K. (2017). Hybrid time-series models for forecasting agricultural commodity prices. *Model Assisted Statistics and Applications*, **12**,255–264

Musunuru, N., Yu, M. and Larson, A. (2013). Forecasting volatility of returns for corn using GARCH Models. *Texas Journal of Agriculture and Natural Resources*, **26**:42-55.

Naik, R.T. (2025). Comparative Study of Wavelet–ANN and Wavelet–ARIMA Models for Groundwater Level Forecasting. *International Journal on Science and Technology*, **16**(4).

Oikonomidis, A., Catal, C. and Kassahun, A. (2022). Hybrid deep learning-based models for crop yield prediction. *Applied artificial intelligence*, **36**(1), 2031822.

Paidipati, K. and Banik, A. (2020). Forecasting of rice cultivation in India–a comparative analysis with ARIMA and LSTM-NN models. *EAI Endorsed Transactions on Scalable Information Systems*, **7**(24).

Paul, R. K., Prajneshu, and Ghosh, H. (2009). GARCH Nonlinear Time Series Analysis for Modeling and Forecasting of India's Volatile Spices Export Data. *Journal of the Indian Society of Agricultural Statistics,* **63**(2), 123-131.

Paul, R.K., Ghosh, H. and Prajneshu (2014). Development of out-of-sample forecast formulae for ARIMAX-GARCH model and their application. *Journal of the Indian Society of Agricultural Statistics,* **68**(1), 85-92.

Paul, R.K., Paul, A.K., and Bhar, L.M. (2019). Wavelet-based combination approach for modelling sub-divisional rainfall in India. *Theoretical and Applied Climatology,* **139**(3-4), 949-964.

Paul, R.K. and Garai, S. (2021). Performance comparison of wavelets-based machine learning technique for forecasting agricultural commodity prices. *Soft Computing,* **25**(20), 12857-12873

Paul, R.K. and Garai, S. (2022). Wavelets based artificial neural network technique for forecasting agricultural prices. *Journal of the Indian Society for Probability and Statistics*, **23**(1), 47-61.

Paul, R.K., Garai, S., and Yeasin, M. (2025). Wavelet-SVR: A novel wavelet-based support vector regression algorithm for time series forecasting. *International Journal of Agricultural Sustainability,* **23**(1), 1-18.

Paul, R.K., Yeasin, M. and Manjunatha, B. (2025). WaveDeepCrop: wavelet enhanced deep learning framework for noise-resilient crop yield forecasting. *Proceedings of the Indian National Science Academy*, **91**(4), 1310-1323.

Rakshit, D. and Paul, R.K. (2024). Hybrid ARFIMA-LRNN Model for Forecasting Commodity Prices. *Journal of the Indian Society of Agricultural Statistics*, **78**(3), 221–229.

Rani M. (2025). Application of wavelet transform in agricultural sciences*. International Journal of Statistics and Applied Mathematics*,**10**(8), 33-36.

Shao, Z., Chao, F., Yang, S.L. and Zhou, K.L. (2017). A review of the decomposition methodology for extracting and identifying the fluctuation characteristics in electricity demand forecasting. *Renewable and Sustainable Energy Reviews,* **75**, 123-136.

Sharif, I. and Khare, S. (2014). Comparative analysis of Haar and Daubechies wavelet for hyper spectral image classification. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **40**, 937-941.

Shawon, S.M., Ema, F.B., Mahi, A.K., Niha, F.L. and Zubair, H.T. (2025). Crop yield prediction using machine learning: An extensive and systematic literature review. *Smart Agricultural Technology*, **10**, 100718.