Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 25 (1) : 151 - 157 (March- 2023)
DOI: https://doi.org/10.54386/jam.v25i1.1895
https://journal.agrimetassociation.org/index.php/jam

Research Paper

Rainfall prediction using time-delay wavelet neural network (TDWNN) model for assessing agrometeorological risk

MRINMOY RAY [1], K. N. SINGH [1], SOUMEN PAL[1], AMIT SAHA[2], KANCHAN SINHA[1] and RAJEEV RANJAN KUMAR[1*]

[1]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110012, India*
[2]*Ministry of Statistics and Programme Implementation, Govt. of India, New Delhi-110001, India*
[*]*Correspondence author Email: rrk.uasd@gmail.com*

ABSTRACT

In an agriculturally dependent nation like India, accurate and effective rainfall forecasting methods are crucial for assessing agrometeorological risk. Forecasting rainfall is perhaps one amongst the most arduous tasks in this context due to the prevalence of a non-linear pattern. One of the most promising and frequently employed approaches for forecasting rainfall data is the Time-Delay Neural Network (TDNN) model. TDNN's non-parametric, data-driven, and self-adaptive characteristics make it increasingly attractive for modelling nonlinear dynamics and generating nonlinear forecasts. Nevertheless, since the conventional TDNN uses the sigmoid activation function, there is always a chance that the training process may converge to local minima. This study addresses the usage of a Time-Delay Wavelet Neural Network (TDWNN), which employs a TDNN architecture with a hidden layer activation function derived from the orthonormal wavelet family, in order to circumvent this issue. TDWNN has been empirically demonstrated using annual rainfall data from two districts in the Indian state of West Bengal. According to the findings of this study, the TDWNN model is superior to the conventional TDNN method to assess agrometeorological risk.

Keywords: Rainfall**,** Forecasting, TDWNN, TDNN

In developing nations like India, systematic and scientific rainfall forecasting may be crucial for improved agrometeorological risk assessment. Rainfall, owing to its irregular character, is regarded as the most significant crop yield limiting factor for rainfed agriculture. India's important crops, including rice, wheat, and maize, are reliant on rainfall. Therefore, for assessing agrometeorological risk, rainfall forecasting has emerged as a pertinent topic within the context of time series forecasting. Due to the presence of a nonlinear pattern, rainfall forecasting is one amongst the most arduous tasks in this context. Various statistical models may be used to model and forecast a particular series. The Auto Regressive Integrated Moving Average (ARIMA) (Box *et al.*, 1995) is the prevalent time series model. In certain instances, the ARIMA model has been used to forecast hydrological and meteorological phenomena (Leite and Peixoto, 1996; Valipour, 2016). As ARIMA is a linear model, it cannot represent the nonlinear temporal pattern of rainfall data. Statistical models such as bilinear (Granger and Anderson, 1978), Autoregressive Conditional Heteroscedasticity (ARCH) model (Engle, 1982), Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model (Bollerslev, 1986) and Threshold Autoregressive (TAR) model (Tong and Lim, 1980) are utilised to deal with non-linear patterns. However, these models include assumptions that are not always discernible when working with real data.

Consequently, TDNN or Artificial Neural Network (ANN) have gained significant traction in modelling non-linear patterns and generating non-linear forecasts in the arena of time series forecasting (Coban and Tezcan, 2022; Zhong *et al.*, 2022). The major advantage of this model is that it does not need any presumption about the considered time series data; rather, the pattern of the data is quite important to it, usually referred as the data-driven approach. TDNN has been extensively used for rainfall forecasting up to this point. Several fascinating research studies have linked the TDNN, including soil temperature prediction (George *et al.*, 2001), typhoon rainfall in Taiwan (Lin and Chen, 2005) weekly rainfall probability analysis (Kulshrestha *et al.*, 2007), monsoon rainfall in Kerala (Guhathakurta, 2006), monsoon rainfall in India

Article info - DOI: *https://doi.org/10.54386/jam.v25i1.1895*
Received: 27 September 2022; Accepted: 11 January 2023; Published online : 17 February 2023
This work is licenced under a Creative Common Attribution 4.0 International licence @ Author(s), Publishing right @ Association of Agrometeorologists

(Chattopadhyay and Chattopadhyay, 2008), Bangkok, Thailand (Hung *et al.*, 2009), and Chennai, India (Geetha and Selvaraj, 2011), evapotranspiration modelling (Naidu and Majhi, 2021) FAO-Penman-Monteith equation (FAO-PM, weather parameters based wheat yield prediction (Arvind *et al.*, 2022). Due to the usage of the sigmoid activation function in traditional TDNN, however, it is always possible for the training process to converge to a local minimum. TDWNN, or Wavelet Neural Network (WNN), was developed lately and uses a wavelet activation function in the hidden neuron. TDWNN has an excellent track record (Alexandridis and Zapranis, 2013; Banakar and Azeem, 2008; Ray *et al.*, 2016; Ray *et al.*, 2020). The TDWNN inherits wavelet function along with neural network properties like self-adapting, self-learning, resilience, time-frequency localization, and nonlinearity. Also, the theory of wavelets makes sure that TDWNN can represent a nonlinear system in an accurate and efficient way.

Continuing the idea of TDWNN with a wavelet-based activation function, the focus of this paper is to evaluate TDWNN with traditional TDNN for rainfall forecasting.

MATERIALS AND METHODS

Data description

Annual rainfall data for two districts, north 24 parganas and Bankura, from 1901 to 2018 were collected from the India Meteorological Department (https://mausam.imd.gov.in/). The main crop in both districts is rice. Consequently, the development of an appropriate model for rainfall forecasting will be beneficial for agrometeorological risk assessment. The time series data set for each district has a total of 118 observations. The first 106 observations were used for model development, whereas the last 12 observations were utilised for model validation.

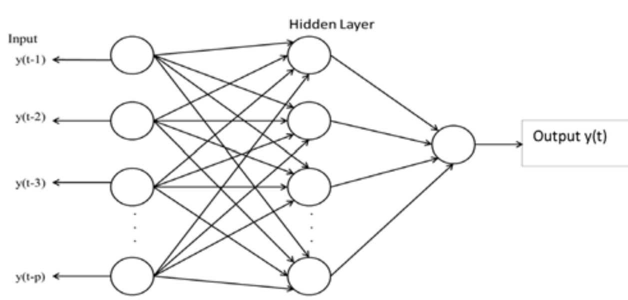

Fig. 1: TDNN Architecture

TDNN model

Time delay neural networks (TDNNs) are a kind of generalised nonlinear model capable of identifying several nonlinear data patterns. This approach does not involve any presumptions about the process of data generation; rather, it is pretty much entirely reliant on the characteristics of the data; this method is commonly referred to as the data-driven approach. The most prominent network for time series modelling and forecasting is a single hidden layer feed forward network. This model is referred to as multilayer TDNNs since it is characterised via use of a three-layer network of rudimentary computing components. The input layer comes first, then the hidden layer, and ultimately the output layer. Diagrammatic depiction of the TDNN model is shown in Fig. 1.

The following is a numerical representation of the association connecting the output (y_t) and the inputs ($y_{t-1}, y_{t-2}, \ldots, y_{t-p}$)

$$y_t = f\{\sum_{k=0}^{a} w_k\, h(\sum_{l=0}^{b} w_{lk}\, y_{t-l})\} \tag{1}$$

where, w_k $(k=0,1,2,\ldots,a)$ and w_{lk} $(l=0,1,2,\ldots,b, k=0,1,2,\ldots,a)$ are parameters of the model usually referred as connection weights, a symbolizes the number of the input layer's nodes and b symbolizes the number of the hidden layer's nodes, and h and f represent, respectively, the activation function at the hidden and output layers.. The activation function represents the nonlinearity of the interaction between the network's inputs and outputs. As shown in Eq. 2, the sigmoid function is the most extensively employed hidden layer activation function.

$$h(v) = \frac{1}{1+e^{-v}} \tag{2}$$

In this research, the gradient decent back propagation technique was utilised to train the network. Training focuses on minimizing the error function, which assesses the gap between predicted and actual values. The error function commonest is mean squared error, which is denoted as:

$$E = \frac{1}{M}\sum_{m=1}^{M}(\varepsilon_l)^2 = \frac{1}{M}\sum_{m=1}^{M}\left[y_t - f\{\sum_{k=0}^{a} w_k\, h(\sum_{l=0}^{b} w_{lk}\, y_{t-l})\}\right]^2 \tag{3}$$

where, M denotes the number of error terms. Iteration is used to estimate the neural network's parameters w_k and w_{lk}. Initially network weights are arbitrarily chosen from a uniform distribution. In each iteration, the weights of the connections altered by a certain amount Δw_{lk}

$$\Delta w_{lk}(t) = -\alpha \frac{\partial E}{\partial w_{lk}} + \beta \Delta w_{lk}(t-1) \tag{4}$$

where, α symbolizes the learning rate and $\frac{\partial E}{\partial w_k}$ symbolizes the partial derivative of the function E with respect to the weight w_k. β is the momentum rate. The $\frac{\partial E}{\partial w_k}$ can be represented as follows-

$$\frac{\partial E}{\partial w_k} = -\varepsilon_l(m) \times f'(v) \times y_l(m) \tag{5}$$

where, $\varepsilon l\ (m))$ is the residual at mth iteration

$f'(v) =$ derivative of the output layer's activation function. As the activation function of the output layer is the identical function in time series forecasting, hence $f'(x) = 1$. $y_l(m)$ is the expected output. Now, network connection weights from input to hidden nodes have altered by a certain amount Δw_{lk}

$$\Delta w_k(t) = -\alpha \frac{\partial E}{\partial w_k} + \beta \Delta w_k(t-1) \tag{6}$$

where

$$\frac{\partial E}{\partial w_{lk}} = h'(v) \times \sum_{k=0}^{a} \varepsilon_l(m) \tag{7}$$

If *h'(v)* is the activation function, then for sigmoid function *h'(v)* is as follows,

$$h'(v) = \frac{\exp(-v)}{(1+\exp(-v))^2} \quad (8)$$

The learning rate is a user-defined neural network's tuning parameter that determines how slowly or quickly the optimal weight is achieved. The learning rate is set low enough to prevent divergence. The momentum term prevents a local minimum from forming throughout the learning process. Regardless of the fact that there exist no accepted theories for the selection of learning rate and momentum, learning rate and momentum are often determined via experimentation. Typically, learning rate and momentum values range from 0 to 1. After computing the final weights using Equation 1, the final output is produced.

TDWNN model

Time delay wavelet neural networks (TDWNNs) used a wavelet-based activation function as opposed to a fundamental logistic function. Various wavelet functions may be utilised for approximating functions or estimating outputs from inputs in a neural network structure. Utilizing a wavelet activation function in the hidden layer generates the concept of a wavelet neural network (WNN). Consider a WNN with one or more inputs, one output layer, and one hidden layer, all of which have orthonormal wavelet basis activation functions. In this context, the hidden nodes are referred to as wavelons. The output of a network with a single input is denoted as follows:

$$\psi_{\gamma,p}(x) = \psi\left(\frac{x-p}{\gamma}\right) \quad (9)$$

where and *p* are parameters for dilation and translation, respectively. If the hidden layer of a single-input, single-output wavelet neural network is formed of wavelons, then the final output is the weighted sum of wavelon outputs.

$$y = \sum_{t=1}^{\lambda} w_t\, \psi_{\gamma_t,p_t}(x) + \bar{y} \quad (10)$$

where $\bar{y}$ is added to accommodate functions having a mean other than zero (since the wavelet function Ψ(X) has a zero mean). Fig. 2 displays a schematic representation of WNN. In this study, the Morlet wavelet function for a wavelet network was constructed. The representation of the Morelet function is as follows:

$$h(v) = \exp(-v^2)\cos(5v) \quad (11)$$

This wavelet is generated from a function which is proportional to the cosine and normal probability density function. If is the activation function, then for Morlet wavelet function appears as follows,

$$h'(v) = -5\sin(5v)\exp(-v^2) - 2v\cos(5v)\exp(-v^2) \quad (12)$$

Non-linearity test

The BDS test developed by Brock, Dechert, and Scheinkman (Broock *et al.*, 2010) was employed to determine the existence of nonlinear dynamics in this study. It has uses in weather forecasting and temperature forecasting and is thoroughly documented (Kim *et al.*, 2003).

It is the non-linear variant of the well-known Box-Pierce Q statistic used to determine the non-linearity of data. Following are the statements of the alternative and null hypotheses:

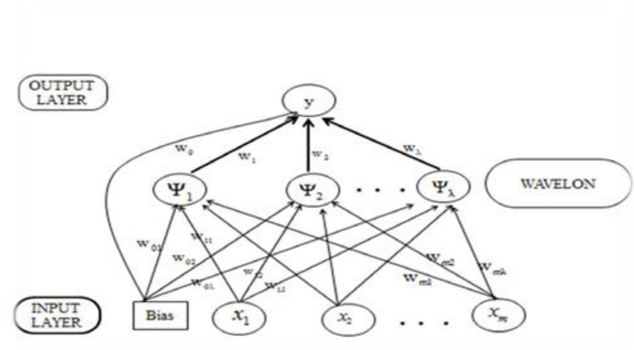

Fig 2: TDWNN Architecture

Null hypothesis (H_0): Data are *i.i.d* (independently and identically distributed). It indicates that the data is linear.

Alternative hypothesis(H_1): Data do not conform to *i.i.d.* It deduces the data's nonlinearity.

Test Statistic: BDS test statistic is defined as:

$$B_{e,\epsilon} = \sqrt{T}\,\frac{A_{e,\epsilon} - A_{1,\epsilon}}{s_{e,\epsilon}} \quad (13)$$

where, represents the embedding dimension, represents the time series data at time *t*, *t*= 1,2,…,T and

$$A_{e,\epsilon} = \frac{2}{T_e(T_e-1)}\sum\sum_{e\le s<t\le T} I(y_t^e, y_s^e, \epsilon)$$
$$A_{1,\epsilon} = P(|y_t - y_s| < \epsilon)^e\ ;\ T_e = T - e + 1 \quad (14)$$

$$I(y_t^e y_s^e, \epsilon) = \begin{cases} 1, & if\ |y_{t-i} - y_{s-i}| < \epsilon,\ i = 0,1,2,\dots,e-1 \\ 0, & Otherwise \end{cases} \quad (15)$$

Assuming the null hypothesis, converges to a standard normal distribution.

Model evaluation criterion

The TDWNN model's modelling and forecasting ability was compared with that of the conventional TDNN model employing Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE). The following is the formula for computing MSE and MAPE.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_t - \hat{y}_t)^2 \quad (16)$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_t - \hat{y}_t}{y}\right| \times 100 \quad (17)$$

where, n is the number of values that have been forecasted or modelled., yt depicts the actual value at the time t and yt estimate corresponds to the modelled or predicted value. In addition, the modelling and forecasting accuracy was evaluated using the DM test developed by Diebold and Mariano (1995). Using the residuals of the fitted models, it compares two forecasting methods.

Additionally, it is important to note that the DM test does not rely on model. Assume that the time series data are y_t (t =1,2,3...T) Forecast error for forecasting method 1: $(e_{t1}) = y_t - F_{t1}$.. Forecast errors for forecasting method 2: $(e_{t2}) = y_t - F_{t2}$

F_{t1} and F_{t2} are the forecasted values of methods 1 and 2, respectively.

Now, a function of error, denoted as (e_{ti}); $i=1,2$. must be defined. This is an illustration of a loss function. Now, the differential loss is characterised as follows:

$$L_t = g(e_{t1}) - g(e_{t2}) \quad (18)$$

Null hypothesis: $E(L_t)=0$ or the forecast accuracy is same for two methods.

Alternative hypothesis: $E(L_t)\neq 0$, or the forecast accuracy is different for two methods.

Test statistic:

$$DM = \frac{\bar{L}}{\sqrt{\frac{2\pi \hat{s}_d(0)}{T}}} \quad (19)$$

Table 1: Summary statistics of rainfall data of two districts of West Bengal

Summary Statistics	Districts	
	North 24 Parganas	Bankura
Mean	1179.11	1075.70
Median	1142.00	1060.20
Minimum	615.00	664.10
Maximum	2141.80	1700.20
Standard Deviation	275.91	226.70
Skewness	0.65	0.42
Kurtosis	0.20	-0.36
Coefficient of Variation (%)	0.23	0.21

Table 2: BDS test result of nonlinearity for North 24 Parganas district

Dimensions	Epsilon		Test statistic	Probability
2	eps(1)	137.0998	6. 0148	<0.0001
	eps(2)	274.1997	6. 6224	<0.0001
	eps(3)	411.2995	8.6342	<0.0001
	eps(4)	548.3994	7.0334	<0.0001
3	eps(1)	137.0998	9. 6707	<0.0001
	eps(2)	274.1997	7.5504	<0.0001
	eps(3)	411.2995	8.2524	<0.0001
	eps(4)	548.3994	8.7971	<0.0001
4	eps(1)	137.0998	8.1355	<0.0001
	eps(2)	274.1997	7.6780	<0.0001
	eps(3)	411.2995	7.3090	<0.0001
	eps(4)	548.3994	6.7909	<0.0001

$$\bar{L} = \sum_{t=1}^{T} L_t$$

Population mean, $\mu=E(L_t)$, Spectral density at frequency 0;

$$s_d(0) = \frac{1}{2\pi}\sum_{k=-\propto}^{\propto} \propto_L (k) \quad (20)$$

$\sum_{k=-\propto}^{\propto} \propto_L (k)$ is the autocovariance of loss differential at lag

Using annual rainfall data, we present empirical examples of the aforementioned models in the next section. This research aims to examine the usage of two distinct activation functions (Logistic and Morlet) in neural network-based architecture and to determine the most reliable model.

RESULTS AND DISCUSSION

This section presents the details of forecasting rainfall data from two districts in West Bengal, India, employing the TDWNN as well as TDNN models. Table 1 provides a summary statistic of the datasets. The first and most important step prior to modelling is to check the linearity of the data. As a result, BDS test was employed to determine whether or not the data under consideration had a nonlinearity pattern. The outcomes of the BDS tests for Bankura and the north 24 parganas are shown in Tables 2 and 3, respectively.

According to the tables, both districts have a nonlinear trend. This indicates that the ARIMA model cannot be used to predict the data, since it is incapable of describing the nonlinear temporal pattern. However, TDNN and TDWNN may be used to forecast the considered data because to their data-driven modelling techniques.

Having proving the existence of nonlinearity employing the BDS test, appropriate TDNN and TDWNN models must be fitted to predict the aforementioned rainfall data. The optimal amount of lagged observations and number of hidden nodes is critical in both TDNN as well as TDWNN modelling. Although there are no documented theories for determining the appropriate number of lagged observations and number of hidden nodes. Hence, in order to determine the optimal values, it is important to carry out trials.

Table 3: BDS test result of nonlinearity for Bankura district

Dimensions	Epsilon		Test statistic	Probability
2	eps(1)	113.3509	7.6718	<0.0001
	eps(2)	226.7019	8. 5766	<0.0001
	eps(3)	340.0528	6.4059	<0.0001
	eps(4)	453.4037	6.9886	<0.0001
3	eps(1)	113.3509	9.4864	<0.0001
	eps(2)	226.7019	8.1227	<0.0001
	eps(3)	340.0528	6.2895	<0.0001
	eps(4)	453.4037	6.9101	<0.0001
4	eps(1)	113.3509	7.4489	<0.0001
	eps(2)	226.7019	6.0847	<0.0001
	eps(3)	340.0528	6.8289	<0.0001
	eps(4)	453.4037	6.2772	<0.0001

Table 4: Summary of model fitting

Districts	Models	Learning rate	Momentum	No. of input lag	No. of hidden unit	Total parameters
North 24 Parganas	TDNN	0.03	0.004	3	2	11
	TDWNN	0.02	0.004	3	1	6
Bankura	TDNN	0.03	0.002	3	2	11
	TDWNN	0.03	0.001	3	1	6

Table 5: Comparison of TDNN and TDWNN model under training and testing data set

Districts	Models	Training data		Testing data	
		MAPE	MSE	MAPE	MSE
North 24 Parganas	TDNN	12.14	28032.80	13.99	28601.57
	TDWNN	10.21	22470.01	12.22	25763.46
Bankura	TDNN	13.27	30649.50	20.52	89958.01
	TDWNN	9.57	16080.78	13.83	36343.61

Table 6: DM test for the statistical significance of TDNN Vs TDWNN model

Districts	Models	DM Statistic	P-value
North 24 Parganas	TDNN vs TDWNN	3.4496	<0.001
Bankura		5.5902	<0.001

The range of input lags from 1 to 15 and the range of hidden nodes from 1 to 10 have been explored to determine the optimal TDNN and TDWNN from single hidden-layered networks. The sigmoid activation function in the hidden nodes and the identity function in the output nodes, together with the gradient descent back-propagation technique, have been used in the process of training TDNNs. For training TDWNNs, we used the Morlet wavelet activation function in the hidden nodes and the identity function in the output node, in conjunction to the Gradient descent backpropagation approach. Monitoring Mean Square Error (MSE) and Mean Absolute Percent Error (MAPE) allows the identification of the optimal model at various learning rates and momentum after screening a significant number of initial models. The Table 4 provides a summary of the best TDNN and TDWNN models for each district.

Table 5 details the comparative prediction abilities of TDNN and TDWNN for both districts. Table 5 demonstrates that the MSE and MAPE of the TDWNN methodology are less than those of the conventional TDNN method in both districts. In conclusion, the TDWNN model outperformed the conventional TDNN model in both districts. As seen in Fig. 3, the predicted values of the TDWNN model are more in line with the observed values than those of the TDNN model. However, the aforementioned findings simply demonstrate the observed model discrepancies. Therefore, the DM test was used to examine if there was a statistically significant distinction between the TDWNN and TDNN models. Table 6's DM test result demonstrates unequivocally that there is a large difference in predictive capacity between TDWNN and TDNN models, which substantially corroborates Table 5's conclusion. As a conclusion, the TDWNN model is superior to the conventional TDNN model for assessing agrometeorological risk.

CONCLUSIONS

Empirical and theoretical research on time-series modelling and forecasting is rapidly expanding, with rainfall forecasting being one of the most complex topics in this domain. Several rainfall data sets exhibit complex nonlinearity in their observable properties. The capacity of standard statistical models to predict such nonlinear patterns is usually exceeded. The introduction of TDNN has offered a suitable solution. However, since classic TDNN utilises the sigmoid activation function, there is always the potential that the training process may converge to local minima. On the other hand, wavelet-based models are gaining popularity because to their relative ability to handle complex data sets and deliver more accurate predictions. This paper provides a comparison between TDWNN and conventional TDNN for rainfall forecasting. The results of this investigation indicate that the TDWNN approach outperformed the conventional TDNN technique. This TDWNN model may be evaluated further using other climatic variables, such as temperature and relative humidity, which are essential for agrometeorological risk assessment. Additional study on the implementation of various wavelet-based activation functions and a comparative analysis are conceivable.

Conflict of Interest Statement: The author(s) declare(s) that there is no conflict of interest.

Disclaimer: The contents, opinions, and views expressed in the research article published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

Publisher's Note: The periodical remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

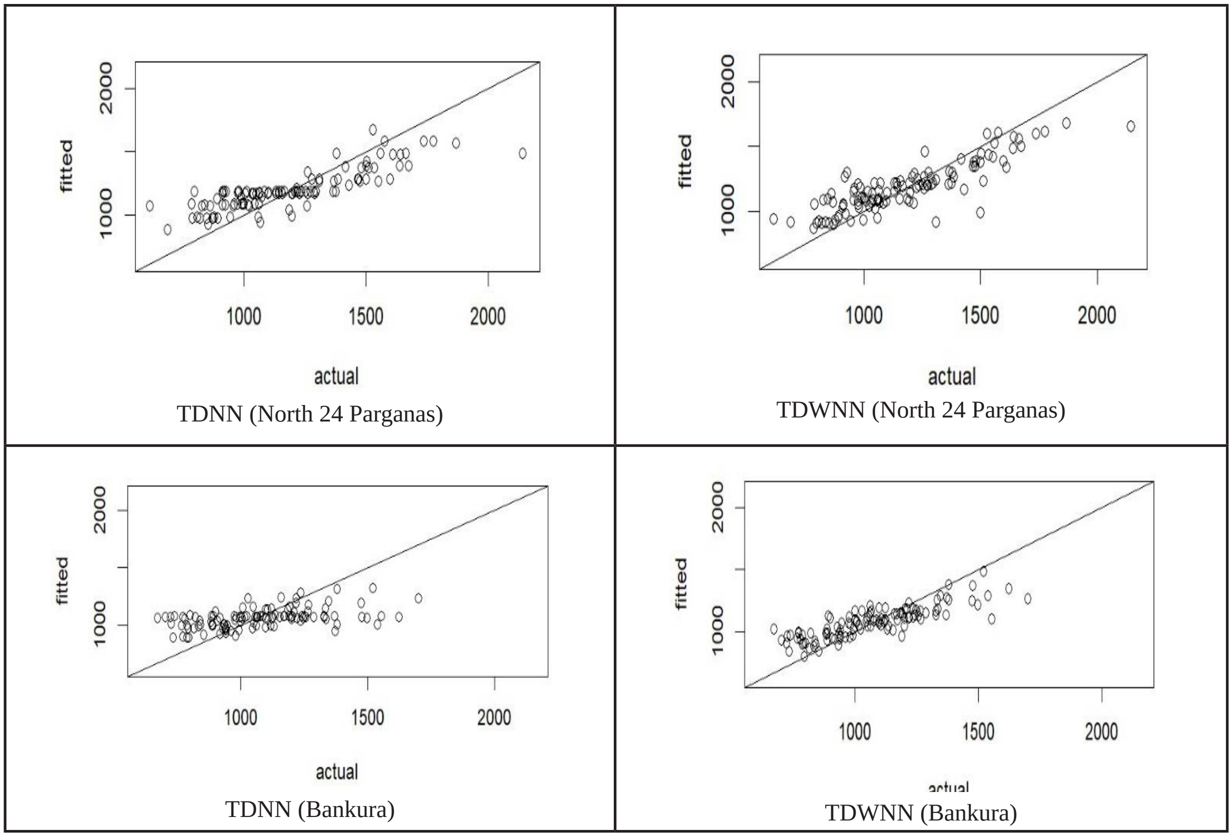

Fig 3: Graphical representation of the performance of models

REFERENCES

Alexandridis, A.K. and Zapranis, A.D. (2013). Wavelet neural networks: A practical guide. *Neural Networks*. 42:1–27. DOI: https://doi.org/10.1016/J.NEUNET.2013.01.008

Arvind, K.S., Vasisth, A., Krishanan, P. and Das, B. (2022). Wheat yield prediction based on weather parameters using multiple linear, neural network and penalised regression models. *J. Agrometeorol.*, 24(1):18-25. DOI: https://doi.org/10.54386/jam.v24i1.1002

Banakar, A. and Azeem, M.F. (2008). Artificial wavelet neural network and its application in neuro-fuzzy models. *Appl. Soft Comput.*, 8:1463–1485. DOI: https://doi.org/10.1016/J.ASOC.2007.10.020

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econom.*, 31:307–327. DOI: https://doi.org/10.1016/0304-4076(86)90063-1

Box, G., Jenkins, G., Reinsel, G. and Ljung, G. (1995). Time Series Analysis: Forecasting and Control. John Wiley & Sons.

Broock, W.A., Scheinkman, J.A., Dechert, W.D. and LeBaron, B. (2010). A test for independence based on the correlation dimension. *Econom. Rev.*, 15:197–235. DOI: https://doi.org/10.1080/07474939608800353

Chattopadhyay, S. and Chattopadhyay, G. (2008). Identification of the best hidden layer size for three-layered neural net in predicting monsoon rainfall in India. *J. Hydroinformatics*, 10:181–188. DOI: https://doi.org/10.2166/HYDRO.2008.017

Coban, M. and Tezcan, S.S. (2022). Feed-Forward Neural Networks Training with Hybrid Taguchi Vortex Search Algorithm for Transmission Line Fault Classification. *Mathematics*, 10:3263. DOI: https://doi.org/10.3390/math10183263

Engle, R.F., (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50: 987. DOI: https://doi.org/10.2307/1912773

Geetha, G. and Selvaraj, R.S. (2011). Prediction of monthly rainfall in Chennai using back propagation neural network model. *Int. J. Eng. Sci. Technol.*, 3:211–213.

George, R.K., Rammohan, S., Kulshretha, M.S., Shekh, A.M. and Jaita, H., (2001). Prediction of soil temperature using artificial neural network. *J. Agrometeorol.*, 3:169–173. DOI: https://doi.org/10.54386/jam.v3i1-2.402

Granger, C.W.J. and Anderson, A.P. (1978). Introduction to bilinear time series models. Vandenhoeck and Ruprecht, Gottingen.

Guhathakurta, P. (2006). Long-Range Monsoon Rainfall Prediction of 2005 for the Districts and Sub-Division Kerala with

Artificial Neural Network. *Curr. Sci.*, 90:773–779.

Hung, N.Q., and Babel, M.S., Weesakul, S. and Tripathi, N.K. (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrol. Earth Syst. Sci.* 13:1413–1425. DOI: https://doi.org/10.5194/HESS-13-1413-2009

Kim, H.S., Kang, D.S. and Kim, J.H. (2003). The BDS statistic and residual test. Stoch. Environ. *Res. Risk Assess.*, 17:104–115. DOI: https://doi.org/10.1007/S00477-002-0118-0

Kulshrestha, M.S., George, R.K. and Shekh, A.M. (2007). Weekly rainfall probability analysis by gamma distribution and artificial neural network. *J. Agrometeorol.* 9:196–202. DOI: https://doi.org/10.54386/jam.v9i2.1128

Leite, S.M. and Peixoto, J.P. (1996). The autoregressive model of climatological time series: an application to the longest time series in Portugal. *Int. J. Climatol.*, 16:1165–1173.

Lin, G.F. and Chen, L.H. (2005). Application of an artificial neural network to typhoon rainfall forecasting. *Hydrol. Process.*, 19:1825–1837. DOI: https://doi.org/10.1002/HYP.5638

Naidu, D. and Majhi, B. (2021). Reference evapotranspiration modeling using radial basis function neural network in different agro-climatic zones of Chhattisgarh. *J. Agrometeorol.* 21:316–326. https://doi.org/10.54386/jam.v21i3.254

Ray, M., Rai, A., Ramasubramanian, V. and Singh, K.N. (2016). ARIMA-WNN hybrid model for forecasting wheat yield time series data. *J. Indian Soc. Agric. Stat.* 70:63–70.

Ray, M., Singh, K.N., Ramasubramanian, V., Paul, R.K., Mukherjee, A. and Rathod, S. (2020). Integration of Wavelet Transform with ANN and WNN for Time Series Forecasting: an Application to Indian Monsoon Rainfall. *Natl. Acad. Sci. Lett.*, 43:509–513. DOI: https://doi.org/10.1007/S40009-020-00887-2

Tong, H. and Lim, K.S. (1980). Threshold autoregression, limit cycles and cyclical data- with discussion. *J. R. Stat. Soc. Ser. B Methodol.* 42:245–292.

Valipour, M. (2016). How Much Meteorological Information Is Necessary to Achieve Reliable Accuracy for Rainfall Estimations? *Agric.*, 6(4):53. DOI: https://doi.org/10.3390/AGRICULTURE6040053

Zhong, C., Lou, W. and Wang, C. (2022). Neural Network-Based Modeling for Risk Evaluation and Early Warning for Large-Scale Sports Events. *Mathematics,* 10:3228. https://doi.org/10.3390/math10183228