A Novel ANN-based Model for Short-term Rain Prediction from ST Radar Moment Data in and around the Radar Site

Nishanta Choudhury
Department of Physics, B Borooah College, Guwahati, Assam, India

Diganta Kumar Sarma
Department of Physics, B Borooah College, Guwahati, Assam, India

Parag Kalita
Department of Physics, B Borooah College, Guwahati, Assam, India

Samiran Patgiri
ST Radar station, Gauhati University, Guwahati, Assam, India

Utpal Sarma
ST Radar station, Gauhati University, Guwahati, Assam, India

Mahen Konwar
Indian Institute of Tropical Meteorology, Pune, India

ABSTRACT
This study deals with the development of an Artificial Neural Network (ANN) based model for short-term rain prediction using moment data from Stratosphere-Troposphere (ST) radar located at 26.14^o N, 91.73^o E, ~50 m above MSL. The ST radar that operates at 212.5MHz scans in five directions East, West, North, South and Zenith with a tilting angle of ~12^o and collects data in various height resolutions viz. 75m, 150m & 300m. For the present study, 150m and 300m height resolution data are considered. The radar data for each of the beam directions are considered separately between the heights of 1500m to 3000m. The neural network is trained with the collocated rain rate from the Automatic Weather Station (AWS) with a 30-minute time lag between radar data and rain rate. The four moments data viz. Return power (P), Doppler shift (DS), Spectral Width (SW) and Signal-to-Noise Ratio (SNR) from ST radar are considered as input parameters to the ANN and the ~30-minute time lag of rain rate data are considered as target values. The developed model will be able to predict rain rate for a diameter of 3 km with the radar as the centre. For that purpose, the neural network is trained separately for each of the beams (East, West, North, South and Zenith). Thus, five separate models are developed for predicting rain rates in and around the radar station. Further, the rain rate from these five models can be clubbed together to get an overall picture of the rain rate over a diameter of 3km. It is observed that the correlation coefficient for the training and validation of the MLPs is more or less similar (~90 %) for all the beams. But the other errors viz. mse, mae, sse & rmse are lowest for MLPs of the zenith beam compared to the other beams. The performance of the derived model was tested with three months of independent dataset data viz. May, June & July 2024 which are not part of the training dataset. It is observed that out of five MLPs for the five beams, the zenith beam performed better than the other MLPs. Though the performance of the MLPs of other beams is less compared to the MLP of the zenith beam, they could predict the rainy situation.

Keywords
Rain prediction, ST radar, Artificial Neural Network.

1. INTRODUCTION
Water is one of the planet's most critical and life-sustaining resources. Water management is a significant concern due to rising water consumption and pollution. Accurate and timely rainfall forecasting is crucial for water resource management, irrigation, planning, and reservoir operation. Predicting when and how much rain will fall is a crucial aspect of weather forecasting that can benefit a variety of fields, including agriculture, disaster management and emergency response etc. [1], [2], [3]. Various methods are in use for the prediction of short-term as well as long term rainfall prediction viz. conventional ground-based measurements, remote sensing, and numerical weather models. Machine learning is used extensively in weather monitoring in terms of estimation and prediction of rainfall [4], [5]. [6], [7]. Rainfall is a highly discontinuous process in space and time and occurrences of rainfall is a non-linear process. Because of the highly non-linear behavior of rainfall, it is very difficult to estimate and predict rainfall from in-situ data and the situation becomes more complicated for remote sensing observation. As technology has advanced, the methods of measuring rainfall have improved a lot, offering more accurate, reliable, and timely information [8]. These advancements are crucial for enhancing our understanding of weather patterns, managing water resources, and preparing for natural disasters. Further, regional models work better than global models. Sarma et al. 2008 showed the better functioning of the regional model for rainfall estimation over the Global model.

Artificial Neural Network is one of the most recent advances in machine learning [5]. Such soft computing models for rainfall analysis are being widely used by researchers all around the world [9], [10], [6], [7]. It has the capability to learn to correctly denote primary data, detect the important features, and increase the effectiveness of prediction compared to traditional models [11]. Many researchers prefer to use ANN for rainfall forecasting because it is a data-driven model do not require restrictive assumptions about the form of the basic model and also can implicitly detect complex nonlinear relationships between dependent and independent variables [12]. Aizansi et al. 2024 [13] have shown that Multi-Layer Perceptron (MLP) networks give better results in a case study conducted for monthly rainfall prediction in the Republic of Benin. A similar study by Mislan et al.2015 [14] showed that the use of the Backpropagation algorithm provided a good model to predict rainfall in Tenggarong, East Kalimantan- Indonesia. Short-term rainfall forecasting plays an important role in hydrologic modeling and water resource management problems such as flood warnings and real-time control of urban drainage systems [15]. Radar systems have emerged as a crucial tool for high-resolution rainfall prediction, offering data on precipitation structures through multiple beam orientations.

In this paper, a short-range rain prediction model is developed using the ANN technique from ST-radar data. The four moments data viz. Return power (P), Doppler shift (DS), spectral width (SW), and signal-to-noise ratio (SNR) with corresponding rain from rain gauge are utilized for the development of the model. These four moments data is able to depict a precipitating system. Rao et al., 1999 [16] studied the precipitating systems using the moments data of mesosphere-stratosphere-troposphere (MST) radar, for the present study, moment data within the height from 1500 m to 3000 m are considered for better representation of precipitation at ground. Thus, the objectives of the present study are to integrate and process data from the ST Radar and AWS observations to predict rain rates in mm/hr. To develop an ANN-based short-term rain prediction model from these processed data that will be able to predict rain in and around the radar site within a 3 km diameter with radar as the center. Finally, the performance of models for each radar beam (East, West, North, South, and Zenith) is analyzed by comparing radar-derived predictions with ground-based AWS measurements.

2. SYSTEM DESCRIPTION AND DATA PREPARATION

For the present study, to develop an ANN-based model for short-term rain prediction, moments data from the Stratosphere-Troposphere (ST) radar located at 26.14^{o} N, 91.73^{o} E, ~50 m above MSL is used along with the Rain Gauge data from AWS installed at the radar site. This ST radar is a coherent pulsed Doppler radar operating in Doppler Beam Swinging (DBS) mode consisting of 576 antenna elements arranged in a circle with a square grid. The radar operates at 212.5 MHz and scans the atmosphere in five directions East, West, North, South, and Zenith with a tilting angle of ~12^{o}. It collects data in various height resolutions viz. 75m, 150m & 300m for different heights starting from 750 m up to 20000 m. For the present study, 150 m and 300 m height resolution data are considered. Further, the moments data between the heights 1500 m and 3000 m are considered for a better representation of rain. On the other hand, the rain accumulation given by the AWS rain gauge is converted to rain rate (mm/hr) and considered for the present study.

Eight-month radar data for the years 2023 and 2024 viz. April, May, June, July, and August are considered for the present study. The ST radar collects data by scanning the atmosphere in five directions during a time frame. The five scans are divided into a number of frames viz. 5, 10, 20, etc. If the frame number is 5 then each beam will collect data only once. On the other hand, if the frame time is 10 each beam will collect data twice during the allotted time frame. The whole time frame is thus divided amongst the different beams depending on the number of frames and thus time is allotted for each beam data.

The corresponding rain rate (mm/hr) from rain gauge data is offset by 30 minutes with respect to ST radar moment data i.e. rain rate data lagging by 30 minutes with respect to the time of each frame of radar data is considered. While finding the 30-minute lagging rain gauge data, a 6-minute window (± 3 minutes) with respect to the time of each frame of radar data is applied. Thus, for radar data of a particular time frame, the rain gauge data within the 6-minute window is considered. The mean value of these rain gauge data over that window time is considered which represents the 30-minute lagging time data corresponding to the time frame radar data.

3. METHODOLOGY

The methodology for the present work is shown in the schematic diagram in Figure 1. Five different MLPs are trained using ANN for five different beam directions using the respective ST radar moment data. These MLPs are named as ANN_E, ANN_W, ANN_N, ANN_S & ANN_Z for the east, west, north, south & zenith beams.

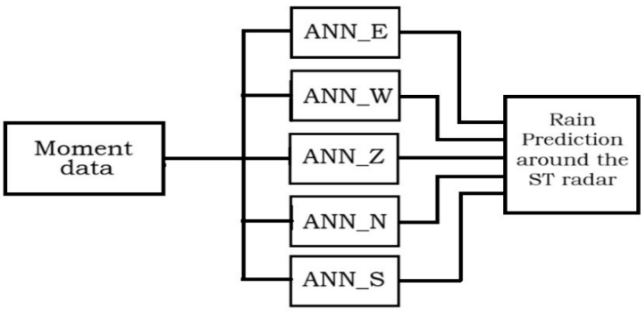

Figure 1. Schematic diagram of the methodology

The input consists of four moment data viz. returned power (P), Doppler shift (DS), Spectral width (SW), and Signal to Noise Ratio (SNR). All five MLPs were trained with these input parameters along with the rain rate with a 30-minute time lag. The trained ANN models are used to predict rain after 30 minutes. Once the model is developed the moments data are fed into the models and it predicts whether there will be rain or not after half an hour. As the models are developed for different beams rain can be predicted in all directions around the ST radar within an aerial distance of around 1.5 km.

4. TRAINING OF THE ANN AND RAIN PREDICTION

An ANN-based model is developed to predict rain from ST radar data. The moment data at different heights from 1500 m to 3000 m is considered as input data for training of the MLPs from five months: April, May June, July and August 2023. The output dataset consists of 30-minute time lag rain gauge data. The number of input dataset for the five beams are listed in Table I. Out of this dataset, 80% is utilized for training and the remaining for validation of the MLPs. Only those input data are considered for which the rain rate is between 0 to 30 mm/hr.

The high rain rate data are avoided in the training dataset. This is mainly because of the fact that the number of high rain rate data is very less which may affect the proper training of the MLPs. Further, the main thrust of the present study is to train the MLPs for predicting rain not specifically the amount of rain rate.

Table 1. Number of data points for the training dataset

Radar beam	No of data point
EAST	8940
WEST	9520
NORTH	8756
SOUTH	8462
ZENITH	7980

All the MLPs were trained in the MATLAB environment. For training purposes, the back-propagation gradient descent method was utilized using the 'traingdx' network training function that updates weight and bias values according to gradient descent momentum and an adaptive learning rate. Depending on the nature of the input dataset each of the MLPs is trained with two hidden layers. The training of the MLPs was tested with various nodes of the hidden layers. The number of nodes required for proper training in the case of the slanting beams (East, West, North, and South) is found to be more compared to the Zenith beam. This may be attributed to the fact that the dependency of the moment data with rain is less for the slanting beams compared to the Zenith. The architecture of the MLPs considered are 4-30-22-1, 4-30-25-1, 4-30-28-1, 4-32-24-1 & 4-25-16-1 for the east, west, north, south & zenith beams respectively. The error statistics for the training dataset are shown in Table II. The functioning of the MLPs is visualized in terms of the correlation coefficient, mean square error (mse), mean absolute error (mae) sum of square error (sse) and root mean square error (rmse).

Table 2. Error statistics for the training and validation of the MLPs

	MLPs	Correlation	MSE	MAE	SSE	RMSE
TRAINING	ZENITH	0.87	0.003	2.48	15.23	3.90
	EAST	0.89	0.004	2.74	22.05	4.70
	WEST	0.88	0.004	2.72	21.14	4.60
	NORTH	0.86	0.004	2.69	20.78	4.56
	SOUTH	0.89	0.004	2.71	20.47	4.52
VALIDATION	ZENITH	0.86	0.011	2.43	14.78	3.84
	EAST	0.88	0.016	2.54	22.96	4.79
	WEST	0.89	0.017	2.80	23.34	4.83
	NORTH	0.85	0.019	2.92	24.96	4.99
	SOUTH	0.88	0.015	2.61	20.67	4.55

It can be noticed that the correlation coefficient for the training and validation of the MLPs is more or less similar for all the beams. But the other errors viz. mse, mae, sse & rmse are lowest for MLPs of the zenith beam compared to the other beams. Figure 2 represents the scatter plot of the training and validation data for all five beams.

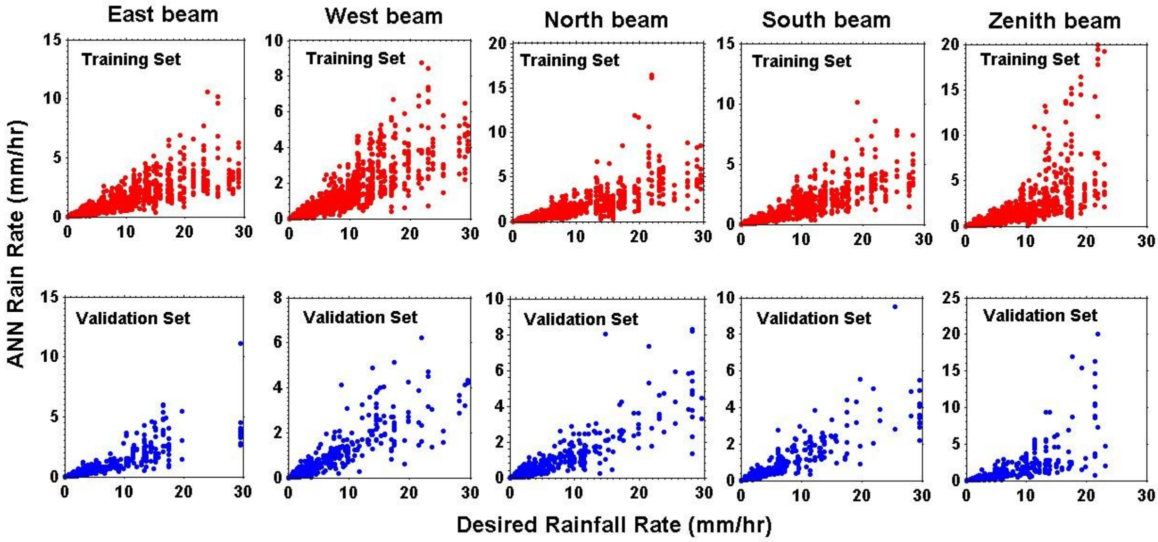

Figure 2. Scattered plot between the desired rain rate and ANN rain rate for training and validation of the MLPs of five beams

4.1 Independent validation of the ANN model:

The trained MLPs for different beams are further validated with independent validation datasets which are not the part of training and validation dataset. The independent dataset comprises three months ST radar moment dataset of May, June and July 2024. To test the functioning of the MLPs the moment dataset is matched with 30-minute time lag rain gauge data. After that, only those radar moment data are considered for which the rain rate is less than or equal to 30 mm/hr. As discussed in the methodology section these radar datasets for each of the beams are fed to corresponding MLPs that give output as the rain after 30 minutes. The output of the MLPs for each of the beams is then compared with the corresponding target value of the rain gauge. The error statistics are shown in Table III.

Table 3. Error statistics for the independent validation of the MLPs for all five beams

	MAY	JUNE	JULY
MSE	0.0107	0.0041	0.0041
MAE	0.6091	1.1349	0.6798
SSE	4.6553	10.6798	6.3536
RMSE	2.1576	3.268	2.5206

It is anticipated that the moment data for the zenith beam depicts the vertically falling rain more compared to the other beams which collect data at a slanting angle of ~12^0. It can clearly be noticed from Table III that the correlation and the other errors are less for the zenith beam compared to the other beams. Therefore, the MLP for the zenith beam ANN_Z is tested separately for the months of three months: May, June and July 2024. Table IV shows the improved error statistics for the ANN_Z for these months.

Table 4. Error statistics for the independent validation of ANN_Z

ERROR	RADAR BEAM				
	ZENITH	EAST	WEST	NORTH	SOUTH
Correlation coefficient	0.72	0.54	0.56	0.40	0.44
MSE	0.0025	0.0023	0.0022	0.0026	0.0029
MAE	0.9539	1.4354	1.2925	1.3899	1.2685
SSE	9.1911	13.6816	11.8413	13.5727	12.2631
RMSE	3.0317	2.9614	2.8051	3.0641	3.2314

For further visualization of the performance of the developed models, the time variation of the model output is observed in comparison to the 30-minute lagging rain gauge data. Figure 3 shows the time series plot of the MLPs ANN_E, ANN_W, ANN_N & ANN_S.

For more clarity, only a portion of the data is shown in the time series plot. It can clearly be noticed from these time series plots that the model could able to predict rain, particularly the non-rainy situations. Though the prediction during rainy situation is not satisfactory in terms of amount of rain rate but still the model could able to predict the existence of rain. On the other hand, the time series plot of the ANN_Z predicted rain shows better results. Figure 4 shows the time variation of the predicted rain from ANN_Z for the month of May, June and July 2024.

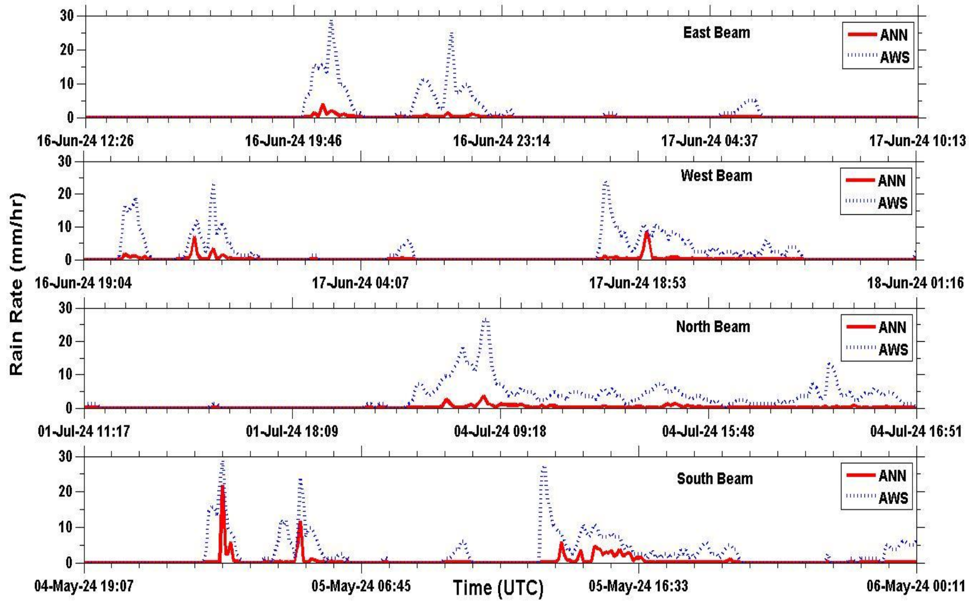

Figure 3. Time series plot of the Predicted rain rate and desired rain rate for east, west, north & south beams

The situation is far better for ANN_Z predicted rain in comparison to the prediction by the other four models for the east, west, north & south beams. ANN_Z could predict rainy situations as well as the amount of rain. The important point to be noted is that all the models are able to predict the rainy and non-rainy situations though the performance of ANN_Z is better compared to the other four models. Nonetheless, using the five models for the five radar beam data it will be possible to predict rain around the ST-radar site within a diameter of 3 km as for the tilting angle of ~12^{o} and at the height of 3000m the horizontal range of the radar beams is ~ 1.5km.

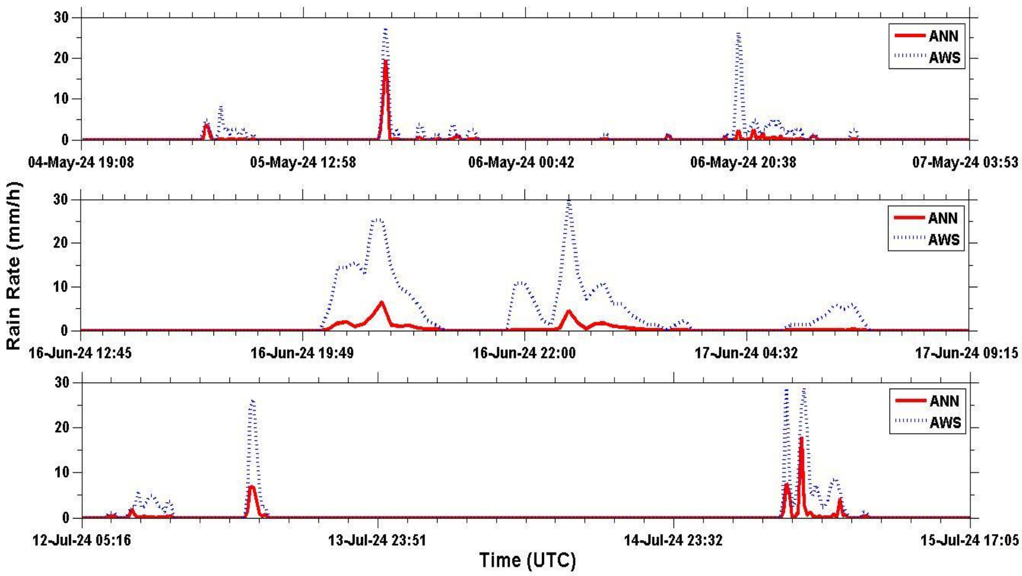

Figure 4. Time series plot of predicted rain rate with desired rain rate for the zenith beams (ANN_Z)

5. SUMMARY AND CONCLUSION

With the help of four moment data from ST- radar, an ANN-based model is developed to predict rain. Five MLPs were trained for the five-beam direction: east, west, north, south and zenith, named as ANN_E, ANN_W, ANN_N, ANN_S & ANN_Z. The main feature of the developed model is that it can predict rain at the radar site as well as in a region of 3 km diameter with the radar as center. If the four-moment data viz. return power, Doppler shift, Spectral width, and signal-to-noise ratio are fed then the model gives the predicted rain 30 minutes in advance. The performance of the derived model was tested with independent datasets that were not part of the training dataset. For that purpose, three months of data, May, June & July 2024 were considered. It is observed that out of five MLPs for the five beams, the zenith beam performed better than the other MLPs. However, the MLPs with off-zenith beams were not able to predict desired rain rate values but still able to predict the occurrence of rain. Further, the model is trained with a maximum rain rate of 30 mm/hr. This study will be extended to upgrade the developed model incorporating moment as well as the wind data for stratiform and convective rainy situations.

6. ACKNOWLEDGMENTS

The authors would like to thank IQAC, B. Borooah College (Autonomous), Assam, India, for the financial support under the Student Project Scheme. Support and data provided by ST Radar Centre, Gauhati University, Assam, India, for the present work are highly acknowledged. Thanks also go to the Principal, B. Borooah College (Autonomous), and the Department of Physics, B. Borooah College (Autonomous) for providing the necessary facilities and support to carry out the research work.

7. REFERENCES

[1] C. Chen, Q. Zhang, M. H. Kashani, C. Jun, S. M. Bateni, S. S. Band, S. S. Dash, and K. W. Chaou, "Forecast of rainfall distribution based on fixed sliding window long short-term memory," *Engineering Applications of Computational Fluid Mechanics*, vol. 16, pp. 248–261, 2022.

[2] U. Grote, "Can we improve global food security? A socio-economic and political perspective," *Food Security*, vol. 6, pp. 187–200, 2014.

[3] J. W. Jones, J. W. Hansen, F. S. Royce, and C. D. Messina, "Potential benefits of climate forecasting to agriculture," *Agricultural Ecosystems and Environment*, vol. 82, pp. 169–184, 2000.

[4] S. S. Patil and B. Vidyavathi, "A machine learning approach to weather prediction in wireless sensor networks," *International Journal of Advanced Computer Science and Applications*, vol. 13, 2022.

[5] E. G Dada, H. J Yakubu, DO Oyewola, "Artificial neural network models for rainfall prediction", *European Journal of Electrical Engineering and Computer Science*, vol. 5, pp. 30-35, 2021.

[6] D. K. Sarma, M. Konwar, J. Das, S. Pal, and S. Sharma, "A Soft Computing Approach for Rainfall Retrieval from the TRMM Microwave Imager," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 2879-2885, 2005.

[7] D. K. Sarma, M. Konwar, S. Sharma, J. Das, S. Pal, U. De, and G. Viswanathan, "An Artificial-Neural-Network-Based Integrated Regional Model for Rain Retrieval Over Land and Ocean," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 1689–1696, 2008.

[8] S. D. Latif, N. A. Hazrim, C. K. Chai-Hoon, J. L. Ng, B. Chaplot, Y. F. Huang, A. Ef-Shafie, and A. N. Ahmed, "Accessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches," *Alexandria Engineering Journal*, vol. 82, pp. 16-25, 2023.

[9] S. Kumar, K. K. Pandey, S. Kumar, S. Supriya, "Estimating rainfall–runoff modeling using the rainfall prognostic model-based artificial framework with a well-ordered selective genetic algorithm", *Journal of Hydroinformatics*, vol. 24, pp. 1066-1090, 2022

[10] Q. Liu, Y. Zou, X. Liu, and N. Linge, "A survey on rainfall forecasting using Artificial Neural Networks," *International Journal of Embedded Systems*, vol. 11, pp. 240-249, 2019.

[11] S. D. Nandakumar, R. Valarmathi, P. S. Juliet, and G. Brindha, "Artificial Neural Network for Rainfall Analysis Using Deep Learning Techniques," *Journal of Physics: Conference Series*, vol. 1964, 2021.

[12] O. R. Nayak, A. Mahapatra, and P. Mishra, "A survey on rainfall prediction using Artificial Neural Networks," *International Journal of Computer Applications*, vol. 72, no. 16, pp. 32-40, 2013.

[13] A.N. Aizansi, K.O. Ogunjobi, and F.K. Ogou, "Monthly rainfall prediction using artificial neural network (case study: Republic of Benin)," *Environmental Data Science*, vol. 3, pp. 1-25, 2024.

[14] H. Mislan, Haviluddin, S. Hardwinarto, Sumaryono, and M. Apiassa, "Rainfall monthly prediction based on artificial neural network: A case study in Tenggarong Station, East Kalimantan–Indonesia," *Procedia Computer Science*, vol. 59, pp. 142-151, 2015.

[15] F.R. Aderyani, S.J. Mouravi, and F. Jafari, "Short-term rainfall forecasting using machine learning-based approaches of PSO-SVR, LSTM and CNN," *Journal of Hydrology*, vol. 614, 2022.

[16] T. N. Rao, D. N. Rao, and S. Raghavan, "Tropical precipitating systems observed with Indian MST radar," *Radio Science*, vol. 34, pp. 1125-1139, 1999.