Implementation of Rainfall Forecasting using ANN

N. Z. A. Nordin, M. M. Yunus* and A. N. Dahalan

Centre of Telecommunication Research and Innovation (CeTRI), Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya 76100 Durian Tunggal, Melaka, Malaysia

*mawar@utem.edu.my

Abstract – Prediction of rainfall is essential for many sectors, including flood mitigation, disaster prevention, agriculture production, and water resource management. Rain forecasts are used to warn of natural disasters such as floods or to plan planting activities to improve the quality of crop yields. An accurate rainfall prediction remains a challenging task due to the uncertainty of natural phenomena considering that rainfall is a non-linear and dynamic process. Various climatic variables such as temperature, relative humidity, wind speed and direction, are among those that affect the dynamic process of rainfall. This study suggests a machine learning technique based on Artificial Neural Network (ANN) model to forecast rainfall by using two-years historical local rain rate and others meteorological data. Evaluation metrics such as mean absolute error, root mean square error, and correlation coefficient, are used to evaluate model's performance. The results show good performance and sensible prediction accuracy. The comparison with existing model clearly depicted that ANN model predicts rainfall better than statistical models particularly at lower and moderate rain rate.

Keywords – Rainfall Forecasting, Machine Learning, Artificial Neural Network (ANN), Rain Rate

I. INTRODUCTION

The prediction of rainfall is one of the most crucial and challenging concerns in today's environment. Extensive simulations and complicated machine models are needed for accurate forecasting since rainfall is frequently dynamics and non-linear variation [1]. Rainfall has become a significant source of concern due to these unexpected changes in rainfall patterns. Temperature, humidity, wind speed, wind direction and precipitation have all impacted rainfall. These are the main influencing elements for rain. Rainfall forecasting is a complex but necessary operational responsibility for meteorological departments worldwide. A significant collection of previous information from past data records is needed for prediction [2,3].

Machine learning methods have recently been effectively used to simulate various complicated phenomena that are difficult to find a closed-form mathematical solution. Machine learning can tackle complex issues in dynamic environments without explicit programming, which would require highly complex lines of code [4]. Artificial Neural Networks (ANN) which are subset of machine learning are always an option to accurately predict the rainfall. ANN is self-adaptive method where predictions are made based on functional relationships between historical multiple inputs data captured in the training process [5,6]. This work utilized ANN as prediction tool based on NARX (Nonlinear Autoregressive with Exogenous Inputs) Neural Network. The NARX network is also a type of feedforward neural network that is trained using supervised learning. It connects the output layer to one or more hidden layers through feedback connections. To forecast future values, the network can use the previous output values as inputs [7]. For network training, external inputs like weather data are also fed in. The network weights and biases are adjusted iteratively using an optimization algorithm to reduce the gap values between the predicted and

the actual output. This process keeps continued until the network's accuracy level is adequate. Once trained, the network can forecast future output values based on future input information.

II. MATERIALS AND METHOD

A. Data Collection

The inputs used in this work mainly been collected from a weather station which is installed at FKEKK Laboratory roof top inside the campus of UTeM, Melaka. The weather station consists of a tipping bucket rain gauge, wind speed, wind direction, temperature, and relative humidity sensors. Two-years of meteorological data (2019-2021) was collected from the weather station for the prediction purpose.

B. Data pre-processing

This process was done to filter an unorganized and inaccurate input data into a format which the model can utilize and understood. Raw data has many inconsistency and missing features that will affect the analysis in later process. The number of null values need to be updated with the mean values and eliminated unnecessary columns or rows for the missing values. Since the model is based on mathematical calculations and equations, encoding categorical data into numeric form is important. The feature selection process, which is part of pre-processing, involves choosing only those features that will assist the model in accurately predicting rainfall. This is to reduce training time [8]. The dataset was first encoded, and then prepared to undergo the experiment.

C. Data Normalization

Data is required to be normalized because the units between the input and output are different. For normalization, the mean or average of these data rainfall is calculated [9].

The formula used in Normalization:

$$Mean(M) = \frac{Sum\ of\ all\ entries}{Total\ number\ of\ entries} \qquad (1)$$

$$Standard\ Deviation(SD) = \sqrt{\frac{\sum(x - \mu)^2}{N}} \qquad (2)$$

Each parameter's mean (M) and standard deviation (SD) are computed. Each weather index values used as a parameter are normalized using

formula (3) after computing their Mean and Standard Deviation. It assists in maintaining the connection between the values of the actual data [10].

$$Normalized\ value = \frac{(x - M)}{SD} \qquad (3)$$

D. Training, Validation and Testing

The inputs data are then been imported into MATLAB workspace to train the networks utilizing NTSTOOL features. The datasets were then divided into three parts, consisting of 70% of data for training, 15% of data for validation, and another 15% of data for testing. The ANN was trained using the Levenberg-Marquardt algorithm based on the feed-forward backpropagation neural network learning algorithm and additional network parameters. The model continues to study the correlation between the rain data values and other meteorological parameters until the performance goal is met. The training ceases automatically when the generalization stops improving, through an increase in the mean square error of the validation samples and the target values. The relationship between the output and the target is then analyzed to examine whether the current output can be used to estimate future rainfall at a given time.

E. Data Evaluation

After completing the testing and validation step, all the results are kept in the workspace. Fig. 1 shows a plotted graph between the target and outputs to compare the training, validation and test data. The results clearly depicted that proposed model has a minimum error and great accuracy.
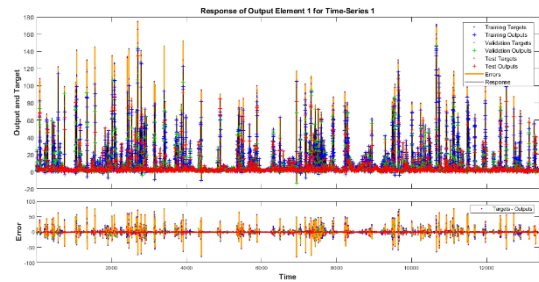

Fig. 1 Response of Output for Time Series Data

III. RESULTS AND DISCUSSION

A. Neural Network Performances

From Fig. 2, the NARX network shows the final Neural Network after completing the whole training, in which the inputs are 4 and the output is 1. In

contrast, 1:2 shows the feedback delay for each input feeding up to the network, W shows the weights of the network, and b is the delay in this or sometimes shows the bias Neuron. The closed-loop network shows the first step of training the network in which the inputs are taken and trained. According to NARX, that model should be feedback, so the output, then feedback to the inputs. The Predict one step ahead view is the next step after the closed-loop network in which the feedback delay is 0:1, and from the previous step, the output was feedback, so in this, the output again feedback to that network. The Neural Network Training View shows what parameters the Neural Network selects to show the best results. According to this, it takes 14 epochs and 00:00:02 seconds to train the model with the given performance and Gradient Descent values for minimal Mean Square Error (MSE).
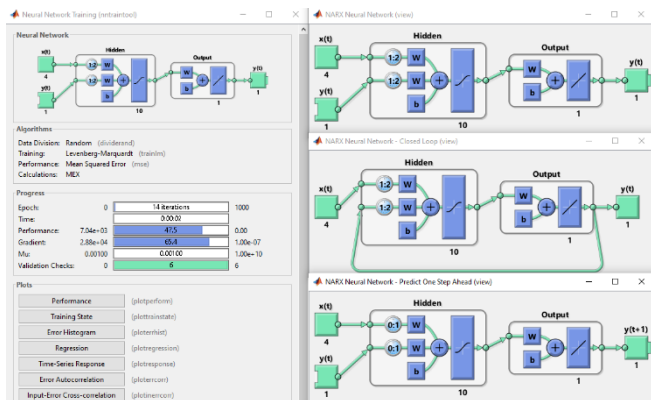

Fig. 2 Neural Network Training Tool

The primary indicators to demonstrate the performance ANN training is the Mean Squared Error (MSE). The MSE helps in tracking the training data to ensure that the error gradient decreases at each iteration up to a certain error function limit. Figure 4.2 shows the performance plot based on MSE. The best validation performance was attained at epoch 8 out of 14 epochs with MSE of 46.4716 as shown in Fig. 3.
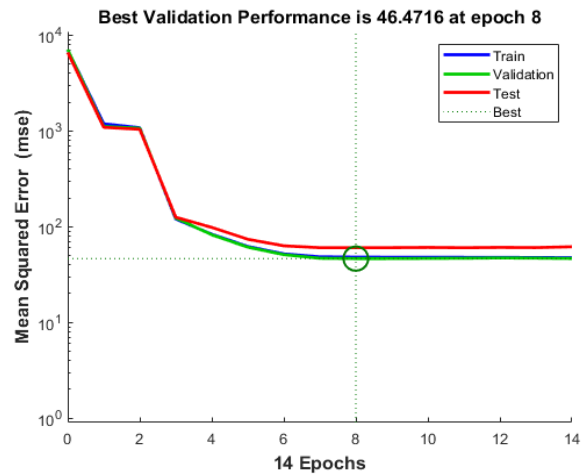

Fig. 3. Performance plot based on MSE

Fig. 4 indicates the performance results of the training, validation and testing for the model. The dotted line represents the condition where the entire neural network and the results are equal to the expected values, whereas the solid lines show the best fit linear regression between the actual and targeted output. The performance response shows a good regression coefficient of greater than 0.85 for all processes.
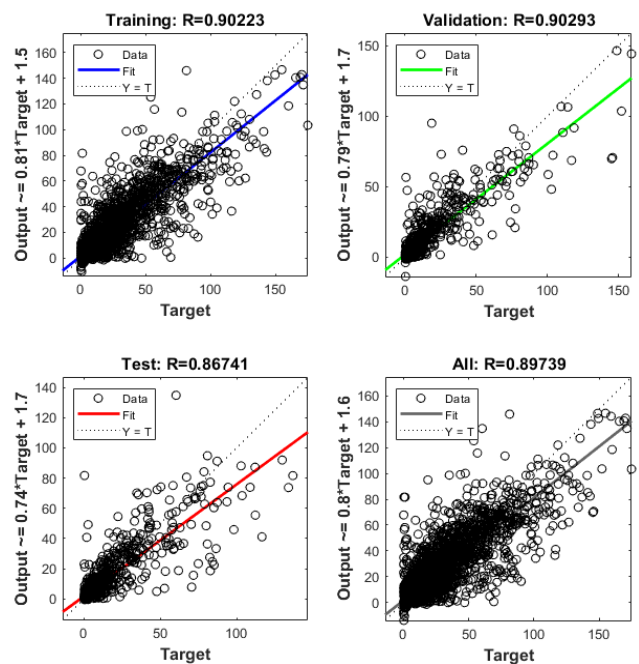

Fig. 4. The Training, Validation and Test Performance

B. Comparison Analysis

Fig. 5 shows the time series of rainfall rate data, which shows a good fit and similarity between actual and predicted output values. In this respect, it

is reasonable to state that the ANN model has a high degree of accuracy in forecasting.
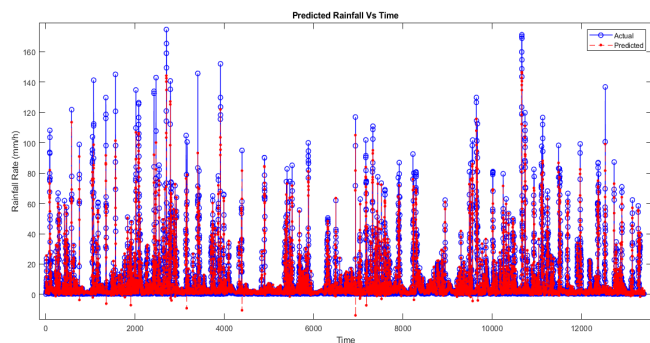

Fig. 5. Measured and Predicted Rainfall versus Time

Consequently, the comparison of measured and predicted rain rates were plotted together with the ITU-R P.837 model as depicted in Fig. 6. It can be seen that measured and predicted results have similar distribution for lower rain rates but slightly underestimated at rain rates higher than 40 mm/h. This might be due to the short duration of data recorded as well as year-to-year variablity in Melaka. On the other hand, in comparison with measured data, the distribution of ITU-R rain rate is overestimated for almost exceeded time.
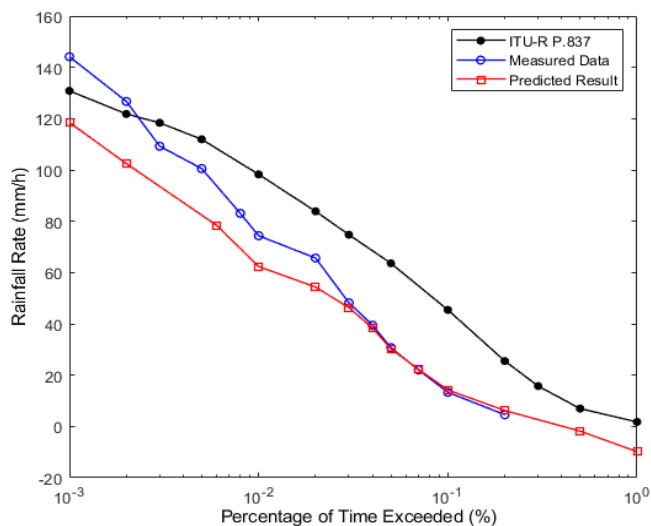

Fig. 6. Comparison of rainfall rate versus percentage of time exceeded between measured, predicted, and ITU-R P.837 model.

IV. CONCLUSION

This study presents the ANN model to predict the rainfall rate based on 4 local meteorological inputs including rainfall, temperature, humidity, and wind. The results show good performance and sensible prediction accuracy with regression coefficient 0.89739. Based on the cumulative distribution of rain rate models over UTeM, the predicted result shows significant same distribution with measurement. However, the predicted result is underestimated at the higher rain rate might cause by the short duration of testing data inputs. In contrast, the distribution of ITU-R rain rate is overestimated compared to the proposed model for almost entire time exceeded. This prediction model might help regional water resource management, agricultural planning, and disaster prevention.

REFERENCES

[1] Traxl, D., Boers, N., Rheinwalt, A. *et al.*, "The role of cyclonic activity in tropical temperature-rainfall scaling," *Nat Commun*, vol. 12, 2021.

[2] Alozie E, Abdulkarim A, Abdullahi I, Usman AD, Faruk N, Olayinka I-FY, Adewole KS, Oloyede AA, Chiroma H, Sowande OA, *et al*., "A Review on Rain Signal Attenuation Modeling, Analysis and Validation Techniques: Advances, Challenges and Future Direction," *Sustainability*, vol. 14, 2022.

[3] Mishra, N. et al., "Development and analysis of artificial neural network models for rainfall prediction by using time-series data," *International Journal of Intelligent Systems and Applications*, 10(1), pp. 16–23, 2018.

[4] Liyew, C.M. and Melese, H.A., "Machine learning techniques to predict daily rainfall amount," *Journal of Big Data*, vol. 8:153, 2021.

[5] Abdolrasol, Maher & Hussain, *et al*., "Artificial Neural Networks Based Optimization Techniques: A Review," *Electronics*, 2021.

[6] U. Kaimkhani, B. Naz, and S. Narejo, "Rainfall Prediction Using Time Series Nonlinear Autoregressive Neural Network," *International Journal of Computer Science and Engineering*, vol. 8, 2021.

[7] Gupta, A., Mall, H.K. and Janarthanan., S., "Rainfall prediction using machine learning," *First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR)*, 2022.

[8] A. Kala and S. G. Vaidyanathan, "Prediction of Rainfall Using Artificial Neural Network," *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2018.

[9] Wakure, G. et al., "Rainfall prediction model using machine learning techniques," *International Journal for Research in Applied Science and Engineering Technology*, 10(5), pp. 939–944, 2022.

[10] Ahuna, M.N., Afullo, T.J. and Alonge, A.A, "Rain attenuation prediction using artificial neural network for Dynamic Rain Fade Mitigation," *SAIEE Africa Research Journal*, 110(1), pp. 11–18, 2019.