Prediction of daily and monthly rainfall using a backpropagation neural network

Huu Nam Nguyen[1], Thuy-Anh Nguyen[2], Hai-Bang Ly[2], Van Quan Tran[2*], Long Khanh Nguyen[2], Minh Viet Nguyen[1], and Canh Tung Ngo[3]

[1] *Institute for Hydropower and Renewable Energy, Hanoi, Vietnam*
[2] *University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Hanoi, 100000, Vietnam*
[3] *Hydraulic Construction Institute, Hanoi, Vietnam*
[*] *Corresponding author. E-mail: quantv@utt.edu.vn*

Received: Sept. 08, 2020; Accepted: Dec. 23, 2020

In this study, the main goal is to develop a model using artificial intelligence (AI) based on the artificial neural network (ANN) for the prediction of daily and monthly rainfall. The authors compare the prediction accuracy of between daily and monthly rainfall, using meteorological parameters as input information (temperature, dew point, humidity, pressure, visibility, and wind speed). Validation of the developed model is achieved using various quantitative evaluation criteria such as correlation coefficient (R), Root-Mean-Square Error (RMSE), Mean Absolute Error (MAE), which are respectively 0.8063, 0.2487, and 0.0932 for the daily rainfall, and 0.8012, 0.0731 and 0.0578 for monthly rainfall. A comparison is then performed, which shows a higher prediction accuracy of monthly than daily rainfall. These reliable results could help in constructing a soft computing tool to predict accurately and quickly the daily and monthly rainfall.

Keywords: rainfall prediction, artificial intelligence, artificial neural network, daily rainfall, monthly rainfall.

http://dx.doi.org/10.6180/jase.202106_24(3).0012

1. Introduction

The rainfall is an important factor that directly affects many social-economic sectors such as agriculture, energy, and water resources of all countries in the world [1, 2]. In addition, extreme weather events, natural disasters such as floods, droughts, landslides, and local sea-level rise are closely related to the intensity and duration of rain [3–5]. Therefore, precisely and quickly predicting and quantifying rainfall intensity and duration is essential for every area [6][6]. Accurate and timely forecasting will help in planning economic development, preventing and minimizing extreme weather events [3, 7][7].

The current forecast technologies often use operational numerical modeling in conjunction with meteorological data [8]. Usually, these models use regression algorithms and averaging methods in weather forecasting [9, 10]. Other approaches use numerical methods [11] or experimental formulas [12]. These models provide predictions from long-term rainfall measurements with other relevant meteorological parameters such as air temperature, solar radiation, cloud information, wind speed, sunny time, relative humidity, coordinates, and altitude [13–15]. From the relationship of those experimental parameters, different experimental equations are proposed to predict and estimate the rainfall [16, 17]. In addition, numerous investigations use linear regression models and statistical methods for the prediction based on experimental data [18, 19]. However, the relationships of experimental parameters are strongly nonlinear. Therefore, the accuracy of empirical formulas is low with rainfall prediction including large number of inputs. The new approach Artificial Intelligence (AI) has strongly developed since a few decades to solve numerous problems with large number of inputs which have complex relations such as structural engineering [20, 21], material science [22],[23] as well as analysis and design structure [24].

Recently, the methods of rainfall prediction based on Artificial Intelligence (AI) such as Artificial Neural Network (ANN) or Adaptive Neural Fuzzy Inference System (ANFIS) have been successfully applied in many investigations [9, 25, 26]. These methods will help analyze data and make forecasting of rainfall over a certain period more accurately and efficiently than other conventional methods [27–29]. The selection of input parameters of the analytical process and calculation methods is also quite diverse. Dabhi and Chaudhary used various meteorological variables such as maximum temperature, minimum temperature, evaporation index, and relative humidity to predict rainfall in a region in India [30]. Meanwhile, Partal et al. [31] used various AI techniques (such as feed-forward backpropagation, radius basis function, and g-generalized regression neural network) with input data containing the daily parameters: average temperature, maximum temperature, minimum temperature, total specific humidity, total evaporation to predict total daily rainfall. The Support Vector Machine (SVM) has been used in rainfall prediction by OrtizGarcía [32], who used meteorological variables such as temperature, wind speed, direction wind, and humidity as input for the forecast process. Although there is an important number of studies in rainfall prediction in the literature, there are still some issues that deserved to be addressed, such as the use of different groups of input variables, rainfall prediction at various locations worldwide, or the applicability of other training algorithms to find the optimum weights and bias values of ANN model.

In the present study, the main focus is to evaluate the feasibility of using ANN with conjugate gradient algorithm for forecasting, and quantifying the total precipitation. The prediction algorithm is trained and validated based on a combination of 6 input parameters: temperature, dew point, humidity, pressure, visibility, and wind speed, in which the dew point and visibility are rarely used variables in the literature. The rainfall database of Austin, Texas is chosen in this study, which is split into the training dataset (accounts for 70 % of the total data), and the testing dataset (accounts for the remaining 30 % of the database). In parallel with the daily rainfall original dataset, the monthly rainfall dataset is constructed and used to compare with the corresponding actual results to evaluate the prediction performance of ANN model through common statistical measurements, i.e., correlation coefficient (R), root mean square error (RMSE) and mean absolute error (MAE).

2. Database collection

The analysis and forecasting process use a dataset consisting of 6 meteorological input variables (each variable includes the maximum, minimum, and average value) listed in table 1 and 2, respectively, including temperature, dew point, humidity, pressure, visibility, and wind speed. There are two sets of variables that describe the daily and monthly rainfall, considered as the output variables in this study. The values of inputs and outputs are on the same day/month and those have not been measured from the original dataset are excluded from the dataset. It is worth noticing that the monthly rainfall dataset is constructed based on the average values of daily data collected in the selected month, except for February 2013, which included only 8 days from February 21 to February 28. The dataset was obtained from WeatherUnderground.com, at the Austin KATT station [33]. The Austin KATT station is found in Texas. The daily rainfall and monthly rainfall are collected from February 21, 2013 to July 31, 2017.

Fig. 1 illustrates the input variables that affect rainfall in the training and prediction process. Fig. 1 and 2 show the input data distribution of daily rainfall and monthly rainfall over 6 meteorological input variables, resp ectively. The distribution of points is random. A correlation analysis is performed, and no strong correlation between inputs variable and output has been found. Indeed, for daily precipitation dataset, the strongest correlation with total precipitation is associated with the visibility (i.e., R=-0.45), following by humidity (R=0.37), dew point (R=0.14), pressure (R=-0.13), wind speed (R=0.04) and temperature (R=-0.02). Regarding the monthly dataset, the strongest correlation with precipitation is associated with the humidity (i.e., R=0.47), following by visibility (R=-0.44), dew point (R=0.23), pressure (R=-0.18), temperature (R=0.13) and wind speed (R=0.00).

3. Artificial neural network

An artificial neural network (ANN) is a model of information processing simulated based on the activity of the biological nervous system, consisting of a large number of neurons attached to process information [34]. ANN can store information and use such information in the prediction of new data. The general architecture of an ANN consists of 3 components, which are the input layer, the hidden layer, and the output layer. The relationship between the input data and the result is shown by equation (1).

$$Y = f(X_1, X_2, \ldots, X_n) \tag{1}$$

in which Y is the result to be forecasted; X1, X2, ..., Xn are the number of input data.

These data classes are linked through weights, bias, and transformation functions. The basic structure of an ANN is shown in Fig. 3.

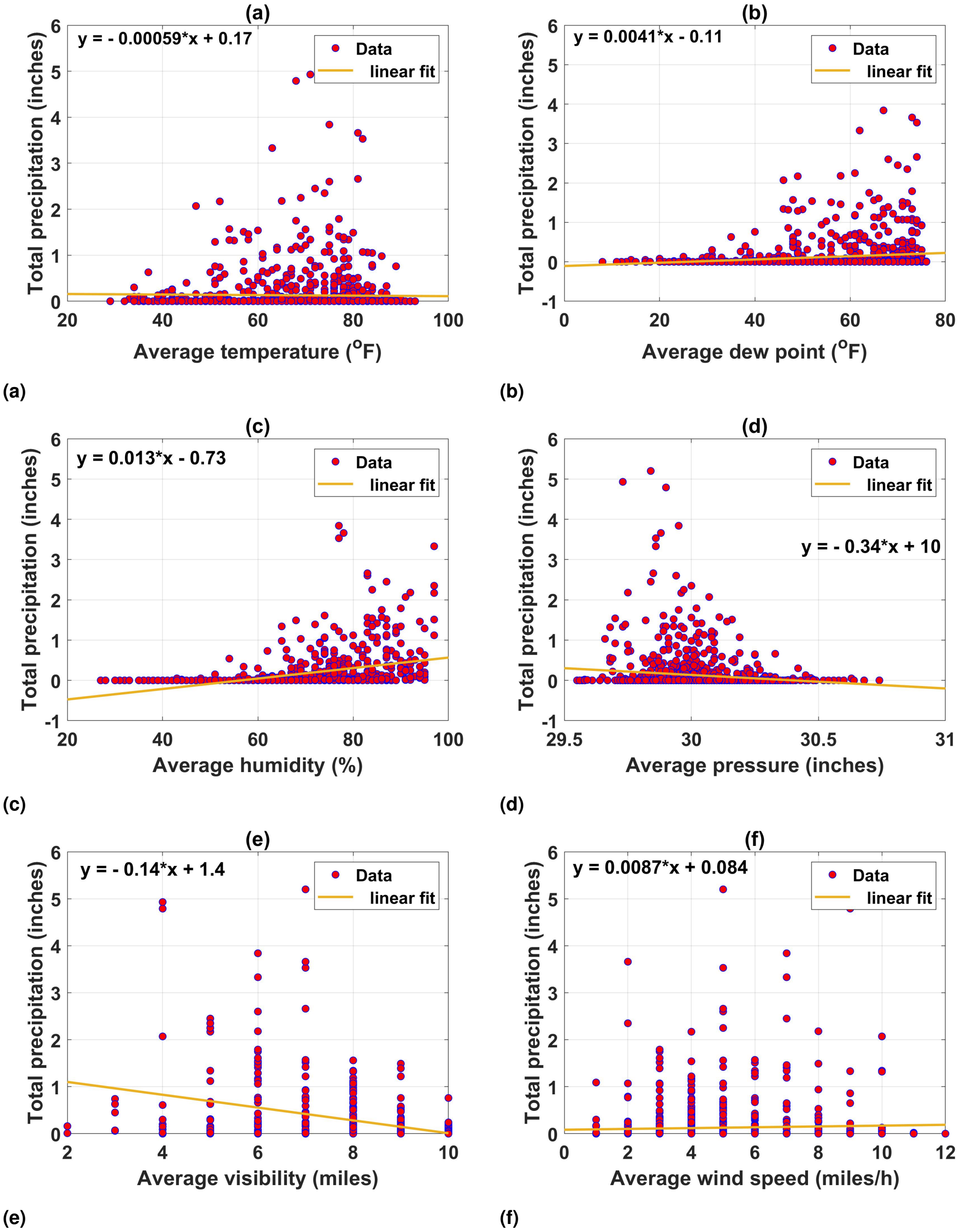


Fig. 1. Histograms of the input and output variables for daily rainfall.

An ANN can have one or more hidden layers. Each hidden layer is a coefficient matrix, the size of the coefficient matrix is the number of neurons in the hidden layer. The number of hidden layers depends on the shape and size of input dataset, as well as requirements for the accuracy of output data. In general, there is no specific regulation to determine the number of hidden layers nor the number of neurons in hidden layers [35]. The determination of these parameters is based on a trial and error assessment process.

In the framework of this paper, the authors have proposed a number of networks, and after checking the predictive errors of the networks, it is found that a one-way multi-layer network gives the best results. The structure of the selected network consists of 5 layers, including an input layer, three hidden layers, and an output layer (Fig. 4). The output layer consists of one neuron, and the number of neurons in the input layer is correlated with the number of input variables in the network. The number of neurons

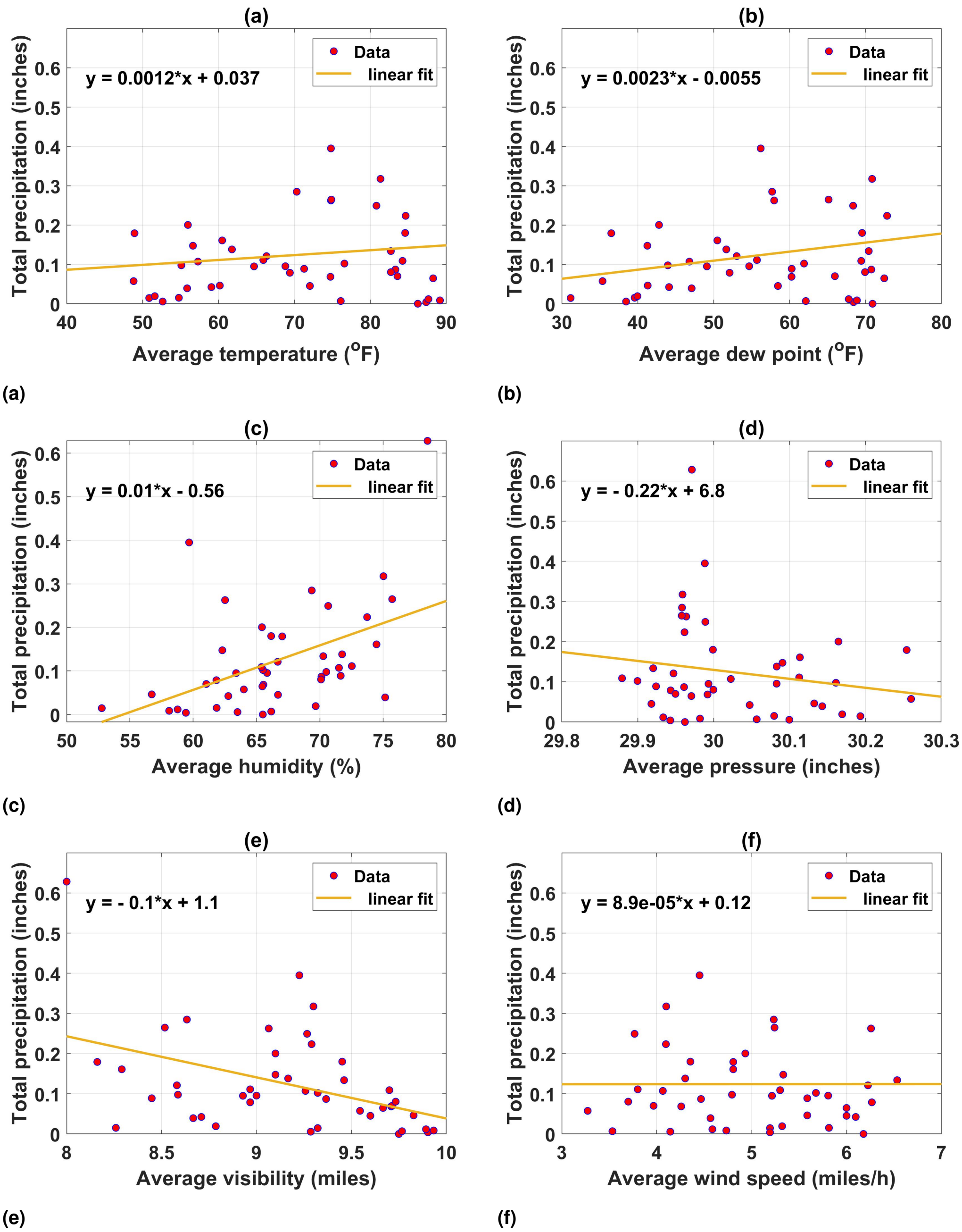


Fig. 2. Histograms of the input and output variables for monthly rainfall.

in the hidden layer is determined through the process of running a model with a small number of neurons and increasing until the predictive results and predictive error reach the allowed value. The hyperbolic tangent math function is used to transform information in the layers of the network, and the ANN network is set up and run in the MATLAB environment.

Network training is the process of determining the optimal weight of an ANN. During network training, weights and bias are evaluated in each iteration of the network. The backpropagation (BP) algorithm is the most widely used training technique for continuous function mapping, which has proven to be theoretically correct [36]. During training, calculations are carried out from the input layer directed to the output layer, and the errors are then propagated from the output layer to the preceding layers. These networks are well proficient in learning linear and nonlinear relationships between inputs and outputs. In order to achieve high

Table 1. Statistical analysis of the inputs and output of Dataset 1.

Variable	Task	Unit	Min	Median	Average	Max	St.D*
High temperature	Input	^{o}F	32.00	83.00	80.95	107.00	14.84
Average temperature	Input	^{o}F	29.00	73.00	70.53	93.00	14.14
Low temperature	Input	^{o}F	19.00	62.00	59.60	81.00	14.34
High dew point	Input	^{o}F	13.00	66.00	61.31	80.00	13.72
Average dew point	Input	^{o}F	8.00	61.00	56.38	76.00	14.98
Low dew point	Input	^{o}F	2.00	55.00	50.68	75.00	16.28
High humidity	Input	%	37.00	90.00	87.74	100.00	11.19
Average humidity	Input	%	27.00	66.00	66.27	97.00	12.74
Low humidity	Input	%	10.00	42.00	44.30	93.00	17.28
High pressure	Input	inches	29.63	30.08	30.11	30.83	0.18
Average pressure	Input	inches	29.55	30.00	30.02	30.74	0.17
Low sea level pressure	Input	inches	29.41	29.91	29.93	30.61	0.17
High visibility	Input	miles	5.00	10.00	9.99	10.00	0.17
Average visibility	Input	miles	2.00	10.00	9.14	10.00	1.48
Low visibility	Input	miles	0.00	9.00	6.81	10.00	3.76
High wind speed	Input	miles/h	6.00	13.00	13.19	29.00	3.44
Average wind speed	Input	miles/h	1.00	5.00	4.98	12.00	2.07
Highest wind speed	Input	miles/h	9.00	21.00	21.30	57.00	5.88
Total rainfall	Output	inches	0.00	0.00	0.13	5.20	6.21

*St.D. = Standard Deviation;

accuracy prediction results, Powell-Beale Conjugate Gradient with momentum weight and bias learning function has been employed.

The conjugate gradient method is a search algorithm that is performed in the direction of the conjugate gradient to minimize errors, which produces generally faster convergence than steepest descent directions. The conjugate gradient algorithms start by searching for the steepest direction (the negative of the gradient) in the first iteration.

$$P_0 = -G_0 \tag{2}$$

where P_0 is the initial search gradient and G_0 is the initial gradient

Then, given the current weight X_k, a line search is performed to determine the optimal distance from which a new estimate for the next weight vector

$$X_{k+1} = X_k + \alpha_k P_k \tag{3}$$

where α_k is the k_{th} learning rate. It is clear that the direction of the $(k+1)_{th}$ search is conjugate to the direction of the $(k)_{th}$ search as in equation 4

$$P_k = -G_k + \mu_k P_{k-1} \tag{4}$$

where P_{k-1} is the earlier search direction, and μ_k is a positive scalar. It should be noted that the computation manner of μ_k determines the characteristics of Conjugate Gradient Algorithms.

With conjugate gradient algorithms, the direction of the search is continuously restarted to ensure better performance. One such reset method was proposed by Powell

Table 2. Statistical analysis of the inputs and output of Dataset 2.

Variable	Task	Unit	Min	Median	Average	Max	St.D*
High temperature	Input	^{o}F	58.32	81.65	80.54	101.53	12.73
Average temperature	Input	^{o}F	48.82	71.65	70.29	89.17	12.63
Low temperature	Input	^{o}F	37.61	61.13	59.51	77.07	12.66
High dew point	Input	^{o}F	41.36	62.46	61.24	75.45	10.81
Average dew point	Input	^{o}F	31.16	57.83	56.34	72.84	12.25
Low dew point	Input	^{o}F	22.19	52.57	50.65	69.58	13.45
High humidity	Input	%	74.32	88.03	87.84	95.17	4.53
Average humidity	Input	%	52.77	66.00	66.61	78.52	5.63
Low humidity	Input	%	29.07	43.03	44.88	62.13	7.57
High pressure	Input	inches	29.95	30.07	30.12	30.39	0.11
Average pressure	Input	inches	29.88	29.99	30.03	30.26	0.10
Low sea level pressure	Input	inches	29.80	29.91	29.93	30.15	0.09
High visibility	Input	miles	9.77	10.00	9.99	10.00	0.04
Average visibility	Input	miles	8.00	9.26	9.16	9.93	0.52
Low visibility	Input	miles	3.58	6.74	6.84	9.36	1.29
High wind speed	Input	miles/h	10.90	13.29	13.22	15.39	1.17
Average wind speed	Input	miles/h	3.27	5.06	4.98	6.53	0.86
Highest wind speed	Input	miles/h	17.48	21.42	21.33	25.93	2.22
Total rainfall	Output	inches	0.00	0.01	0.12	0.63	0.12

*St.D. = Standard Deviation;

[37], based on an earlier version proposed by Beale [38]. The algorithm is restarted, and the direction of the search is reset to the negative of the gradient if the following equation 5 is satisfied [?].

$$|G_{k-1}^{T}G_k| \leq 0.2||G_k||^2 \tag{5}$$

4. Forecast Verification

4.1. Performance indicators

In order to validate developed ANN models, three statistical criteria, namely correlation coefficient (R), root mean square error (RMSE), and absolute mean error (MAE) are used in this research. The R value used to investigate the linear correlation between the target and the predicted values, which lies in the range [-1; 1]. Both RMSE and MAE measure the average intensity of the error [39], where MAE represents the residual error between the target and predicted values for each dataset, and RMSE denotes the square root of the average residual error between target values and predicted values for each data set. For these two indicators, the smaller value signifies the better performance of the ANN model. The closer the absolute value of R is to 1, the stronger the ANN model is in predicting. R, RMSE, and MAE are defined by the following equations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(r_{0,j} - r_{t,j})^2} \tag{6}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|r_{0,j} - r_{t,j}| \tag{7}$$

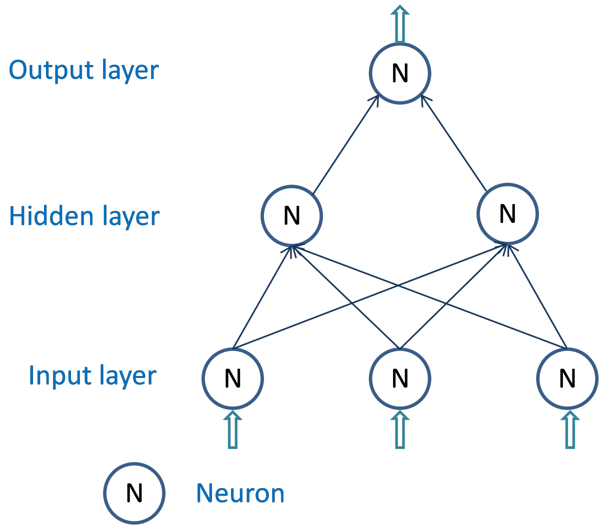


Fig. 3. The general structure of an ANN.

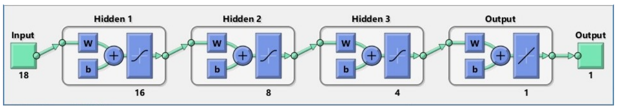


Fig. 4. ANN structure used in this study.

$$R = \frac{\sum_{j=1}^{n}(r_{0,j} - r_0)(r_{t,j} - r_t)}{\sqrt{\sum_{j=1}^{n}(r_{t,j} - r_t)^2 \sum_{j=1}^{n}(r_{0,j} - r_0)^2}} \quad (8)$$

where: n is the number of database, r_0 and r_0 is the actual experimental value and the average real experimental value, r_t and r_t is the predicted value and the average predicted value, calculated according to the model forecast.

4.2. Methodology

The present study is carried out based on the proposed methodology that comprises four main steps as follows:

(1) Step 1: Data preparation: in this first step, sample data are collected. Then, the dataset is divided into two sets, one for constructing the model, called the training dataset, and the other one for evaluating the estimated model, called the testing dataset. The training dataset was generated from 70 % of total data, whereas the testing dataset was built from the 30 % remaining data.

(2) Step 2: Determination of input variables for each model is performed. The selection of input variables is experimental or based on the trial and error method. A correlation analysis is performed, and no correlation between inputs has been found. Therefore, the original inputs collected from the literature are used.

(3) Step3: Construction of the model: The architecture of ANN is constructed and developped in this step.

(4) Step 4: Validation of the proposed models: in this final step, the predicting ability for the proposed model is evaluated. Furthermore, the predicted output are compared with the actual data, and error is calculated. Statistical indicators, including R, RMSE, and MAE are employed to validate the models.

5. Results and discussion

5.1. Daily rainfall

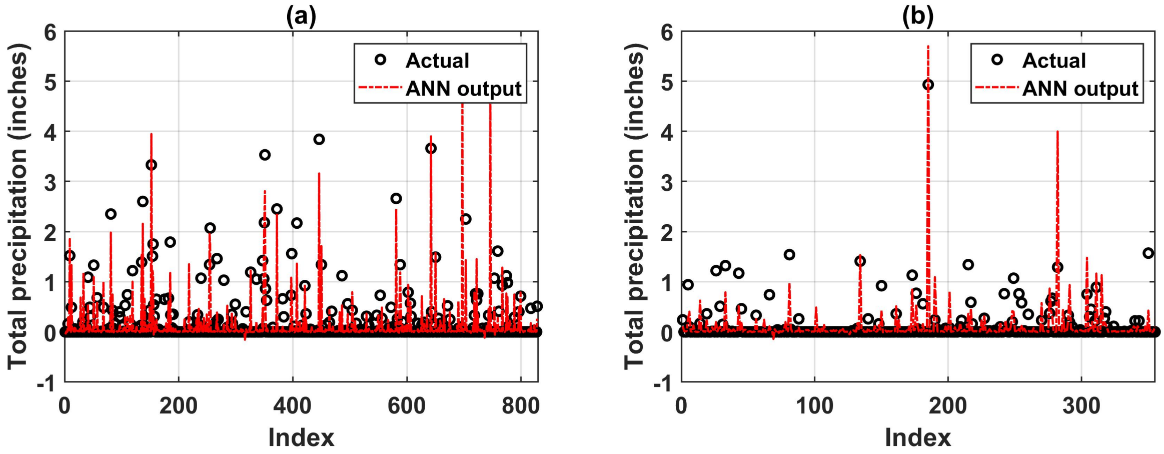
The ANN model uses the performance indicators to validate the prediction: R, RMSE, MAE, mean of error (denoted as Error_mean), and standard deviation error (denoted as Error_std). Considering the daily rainfall forecast, the performance indicators are shown in table 3 for both training and testing datasets. This is the best result obtained over 50 simulations, selected based on the highest values of R and lowest values of RMSE and MAE.

Table 3. Performance indicator of daily rainfall prediction using ANN model.

	Training dataset	Testing dataset	All dataset
RMSE	0.1831	0.2487	0.2050
MAE	0.0762	0.0932	0.0813
Error_mean	0.0005	-0.0080	-0.0021
Error_std	0.1832	0.2489	0.2051
R	0.9249	0.8063	0.8944

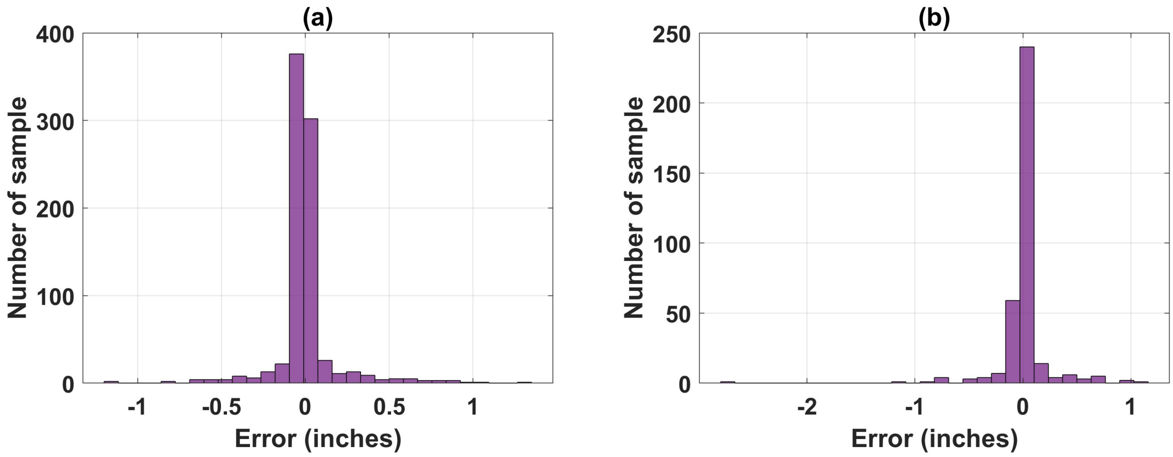
Figs. 5a, 5b shows the comparison between measured and predicted daily rainfall for the training and testing parts, respectively. Fig. 6 shows the error histograms between the actual and predicted values of daily rainfall prediction with respect to the training part (a) and the testing part (b). The range of error value is between [-1.5; 0.9] and [-2.8; 1.0] for the training and testing parts, respectively. However, most of the error between measured and predicted daily rainfall are close to 0, with about 380 of samples and 245 of samples for the training and testing parts, respectively.

The correlation of predicted and measured daily rainfall for (a) training data, (b) testing data, and (c) all data are shown in Fig. 7a, 7b, and 7c, respectively. The regression lines show that the prediction results are relatively close to reality, proving the effectiveness of ANN in predicting daily rainfall. Overall, the correlation coefficient R and the error histograms show the prediction rainfall capability of the proposed ANN model is high, with R=0.925, R=0.806 for the training, and testing parts, respectively.



(a) (b)

Fig. 5. Comparison between measured and predicted daily rainfall (a) training, (b) testing process.



(a) (b)

Fig. 6. Error statistics of daily rainfall prediction: (a) training part; and (b) testing part.

5.2. Monthly rainfall

The performance indicators for ANN prediction are shown in table 4. It is found that the performance of monthly rainfall prediction is better than that of daily rainfall prediction. That is clearly proved by the R, RMSE, and MAE values. The predicted monthly rainfall is close to the measured monthly rainfall (Fig. 8).

The better accuracy of ANN in predicting the monthly rainfall than that of daily is also shown by the error values between predicted and measured rainfall (cf. Fig. 9). The range of error values is [-0.1; 0.1] and [-0.15; 0.12] for the training and testing parts, respectively. Regarding monthly rainfall prediction, the magnitude of the error value is smaller than the daily prediction (i.e., 0.15 for monthly rainfall compared with 1.5 for daily rainfall). The most considerable magnitude of errors is 0.1 inches, whereas that of daily prediction is about 1 inch. It can be concluded that monthly rainfall prediction using ANN gives more accurate results than daily rainfall prediction.

Table 4. Performance indicator of daily rainfall prediction using ANN model.

	Training dataset	Testing dataset	All dataset
RMSE	0.0359	0.0731	0.04980
MAE	0.0283	0.0578	0.0371
Error_mean	-0.0053	-0.0144	-0.0080
Error_std	0.0360	0.0746	0.0498
R	0.9589	0.8012	0.9155

The monthly rainfall prediction is compared with mea-

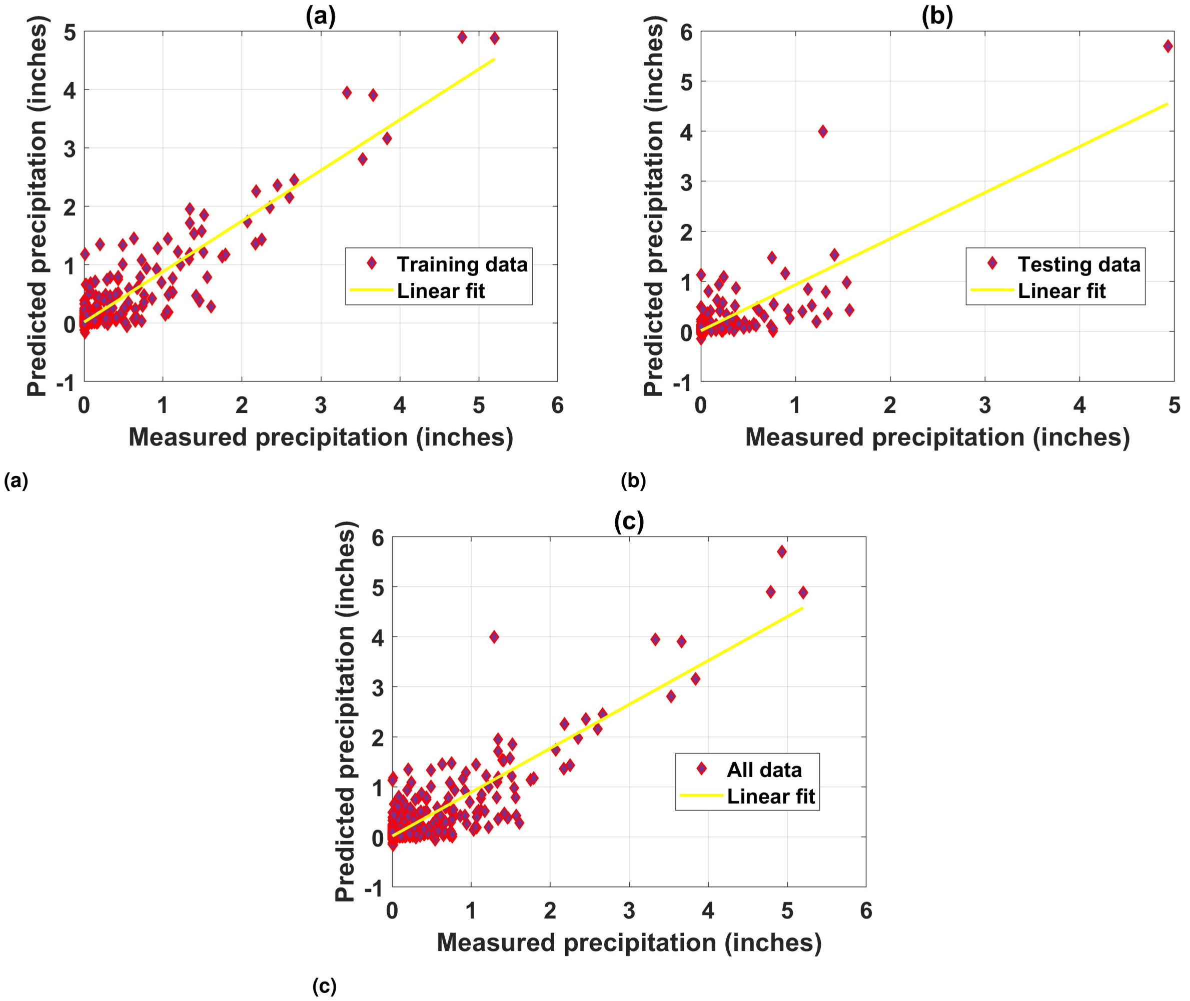


Fig. 7. Correlation of predicted and measured daily rainfall (a) training data, (b) testing data and (c) all data.

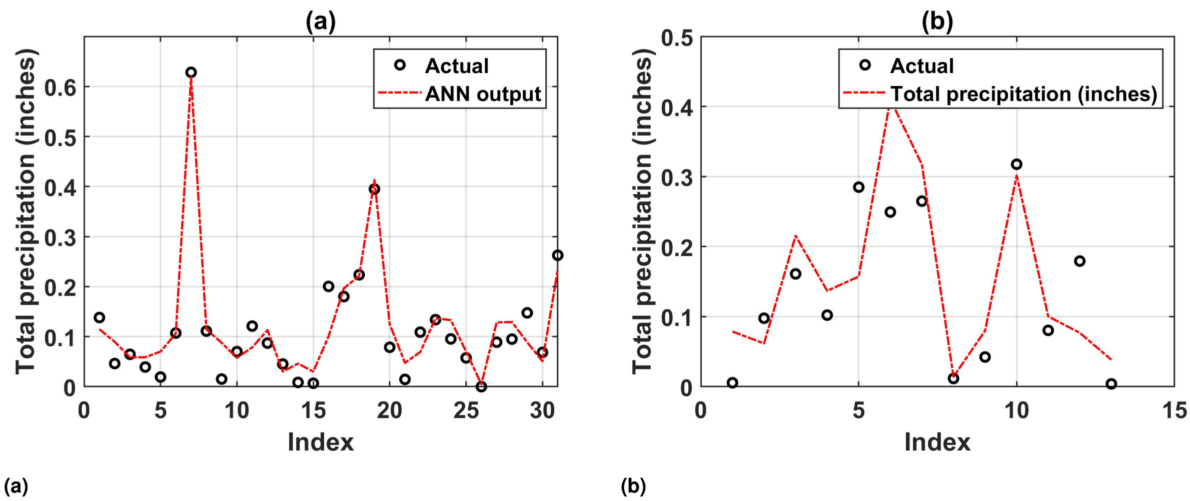


Fig. 8. Output data and actual parameters of monthly rainfall prediction.

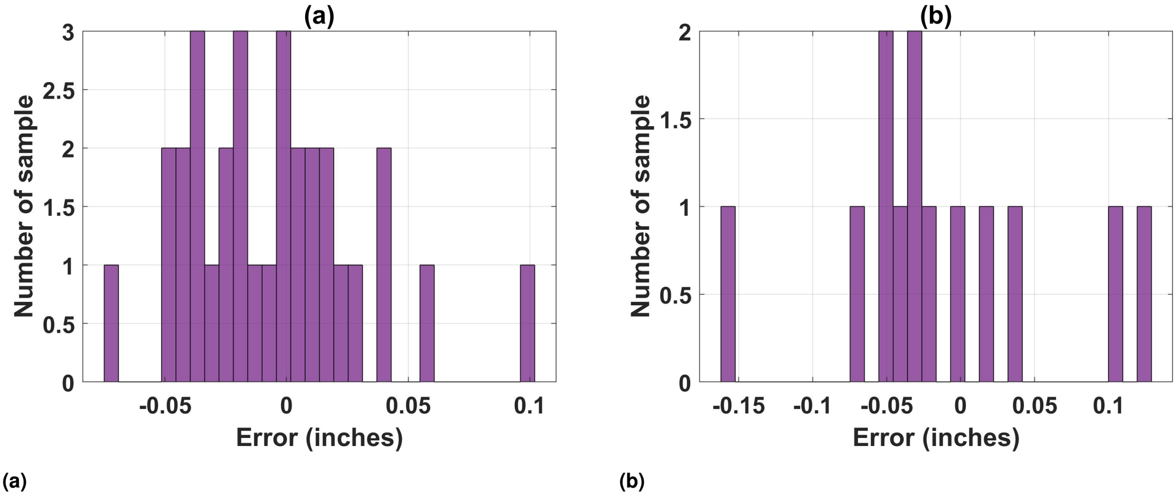


Fig. 9. Error statistics of daily rainfall prediction.

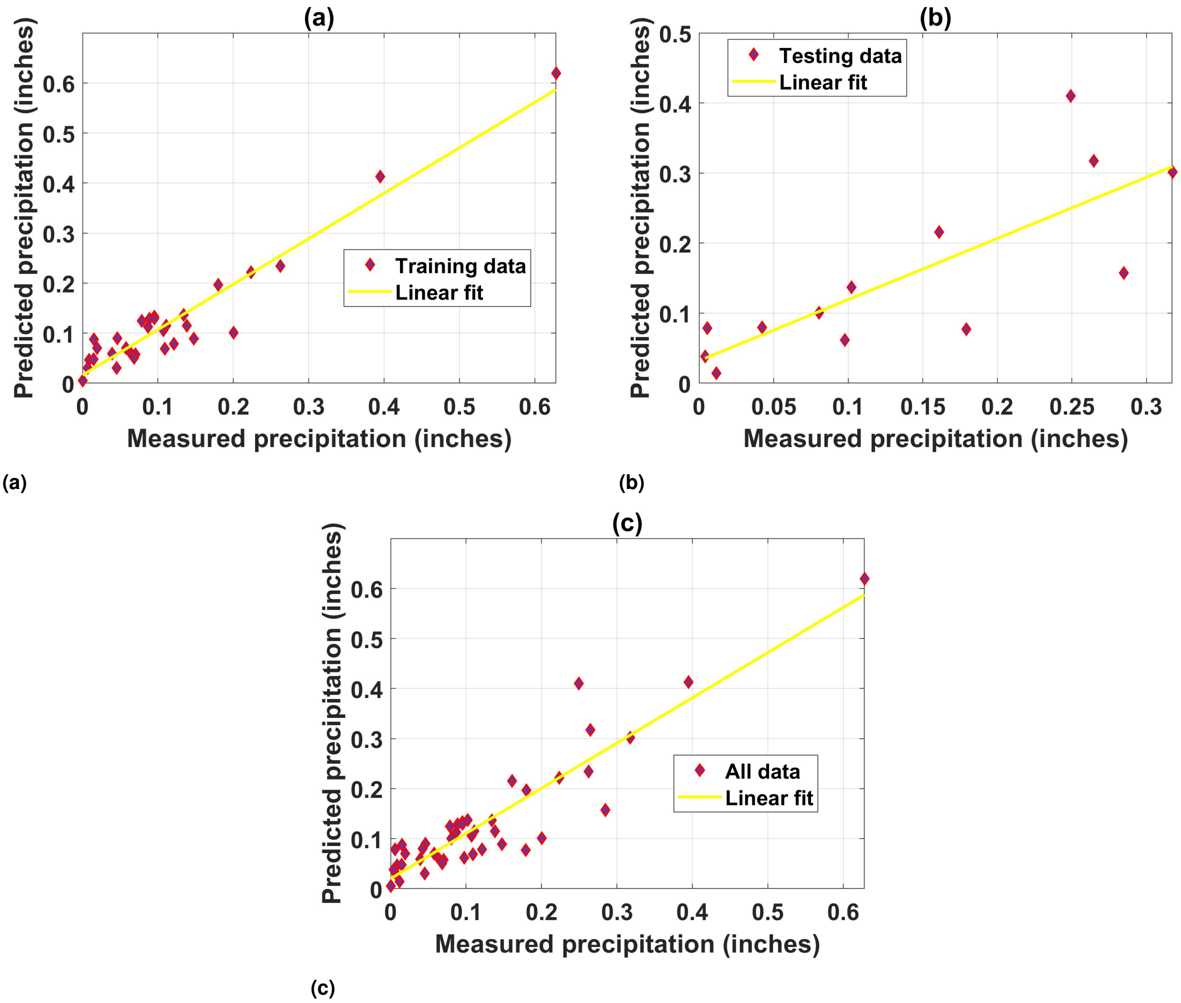


Fig. 10. Correlation of predicted and measured monthly rainfall.

sured monthly rainfall through correlation graphs, shown in Figs. 10a, 10b, and 10c for the training, testing, and all data, respectively. The correlation of monthly prediction case is better than the correlation of daily prediction.

Rainfall prediction is considered as a challenging task, which has been addressed in many studies in the literature. The values of performance metric of the proposed ANN model is this study is significant compared with the relevant literature. As for example, Kuligowski and Barros [40] used back-propagation neutal network for rainfall prediction and obtained a value of R of 0.549. Basides, Hung et al. [41] used ANN and obtained a maximum value of R of 0.88 for hourly rainfall prediction. Overall, the results in this study show that the ANN model can be applied to predict the daily and monthly rainfall with high accuracy.

6. Conclusion

In this investigation, the ANN model has shown its efficiency in predicting and quantifying daily/monthly rainfall through 6 input parameters: temperature, dew point, humidity, pressure, visibility, and wind speed. For the development of the ANN model, 70 % of data is randomly chosen for the training process, and the remaining 30 % of data is chosen for the validation phase. The results show that the ANN model is a promising algorithm to predict daily rainfall. The values of the testing phase are 0.8063, 0.2487, and 0.0932 for R, RMSE, and MAE, respectively, for the daily rainfall. The performance indicators of the testing part are 0.8012, 0.0731, and 0.0578 for R, RMSE, and MAE, respectively, for monthly rainfall. The results also show a higher accuracy prediction of monthly rainfall than daily rainfall. The results could help in constructing a reliable soft computing tool to predict accurately and quickly the daily and monthly rainfall.

7. Acknowledgements

This dataset was obtained from WeatherUnderground.com, at the Austin KATT station.

References

[1] Trong Anh Trinh. The Impact of Climate Change on Agriculture: Findings from Households in Vietnam. *Environmental and Resource Economics*, 71(4):897–921, dec 2018.

[2] Mohamed S. Haddad. Capacity choice and water management in hydroelectricity systems. *Energy Economics*, 33(2):168–177, mar 2011.

[3] Nejc Bezak, Mojca Šraj, and Matjaž Mikoš. Copula-based IDF curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides. *Journal of Hydrology*, 541:272–284, oct 2016.

[4] Saeid Janizadeh, Mohammadtaghi Avand, Abolfazl Jaafari, Tran Van Phong, Mahmoud Bayat, Ebrahim Ahmadisharaf, Indra Prakash, Binh Thai Pham, and Saro Lee. Prediction success of machine learning methods for flash flood susceptibility mapping in the Tafresh watershed, Iran. *Sustainability (Switzerland)*, 11(19), 2019.

[5] Soukayna Mouatadid, Nawin Raj, Ravinesh C. Deo, and Jan F. Adamowski. Input selection and data-driven model performance optimization to predict the Standardized Precipitation and Evaporation Index in a drought-prone region. *Atmospheric Research*, 212:130–149, 2018.

[6] Mislan, Haviluddin, Sigit Hardwinarto, Sumaryono, and Marlon Aipassa. Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggarong Station, East Kalimantan - Indonesia. In *Procedia Computer Science*, volume 59, pages 142–151, 2015.

[7] John Abbot and Jennifer Marohasy. Skilful rainfall forecasts from artificial neural networks with long duration series and single-month optimization. *Atmospheric Research*, 197:289–299, nov 2017.

[8] Shan He, Srivatsan V. Raghavan, Ngoc Son Nguyen, and Shie Yui Liong. Ensemble rainfall forecasting with numerical weather prediction and radar-based nowcasting models. *Hydrological Processes*, 27(11):1560–1571, may 2013.

[9] John Abbot and Jennifer Marohasy. Input selection and optimisation for monthly rainfall forecasting in queensland, australia, using artificial neural networks. *Atmospheric Research*, 138:166–178, mar 2014.

[10] Roméo S Tanessong, P Moudi Igri, Derbetini A Vondou, P. H.Kamsu Tamo, and F Mkankam Kamga. Evaluation of probabilistic precipitation forecast determined from WRF forecasted amounts. *Theoretical and Applied Climatology*, 116(3-4):649–659, 2014.

[11] David R Novak, Christopher Bailey, Keith F Brill, Patrick Burke, Wallace A Hogsett, Robert Rausch, and Michael Schichtel. Precipitation and temperature forecast performance at the weather prediction center. *Weather and Forecasting*, 29(3):489–504, 2014.

[12] Francesco Silvestro and Nicola Rebora. Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain. *Journal of Hydrology*, 519(PA):1052–1067, 2014.

[13] Shaban G Gouda, Zakia Hussein, Shuai Luo, and Qiaoxia Yuan. Model selection for accurate daily global solar radiation prediction in China. *Journal of Cleaner*

Production, 221:132–144, 2019.

[14] Seyyed Mohammad Mousavi, Elham S. Mostafavi, and Pengcheng Jiao. Next generation prediction model for daily solar radiation on horizontal surface using a hybrid neural network and simulated annealing method. *Energy Conversion and Management*, 153:671–682, 2017.

[15] M. Chegaar and A. Chibani. Global solar radiation estimation in Algeria. *Energy Conversion and Management*, 42(8):967–973, 2001.

[16] Saleh A. AlHassoun. Developing an empirical formulae to estimate rainfall intensity in Riyadh region. *Journal of King Saud University - Engineering Sciences*, 23(2):81–88, 2011.

[17] Jinghao Niu and Wei Zhang. Comparative analysis of statistical models in rainfall prediction. In *2015 IEEE International Conference on Information and Automation, ICIA 2015 - In conjunction with 2015 IEEE International Conference on Automation and Logistics*, pages 2187–2190, 2015.

[18] M Rajeevan, D. S. Pai, R. Anil Kumar, and B. Lal. New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dynamics*, 28(7-8):813–828, 2007.

[19] Soyini A. Ashby, Michael A. Taylor, and Anthony A. Chen. Statistical models for predicting rainfall in the Caribbean. *Theoretical and Applied Climatology*, 82(1-2):65–80, 2005.

[20] Hung Quang Nguyen, Hai Bang Ly, Van Quan Tran, Thuy Anh Nguyen, Tien Thinh Le, and Binh Thai Pham. Optimization of artificial intelligence system by evolutionary algorithm for prediction of axial capacity of rectangular concrete filled steel tubes under compression. *Materials*, 13(5), 2020.

[21] Quang Hung Nguyen, Hai Bang Ly, Tien Thinh Le, Thuy Anh Nguyen, Viet Hung Phan, Van Quan Tran, and Binh Thai Pham. Parametric investigation of particle swarm optimization to improve the performance of the adaptive neuro-fuzzy inference system in determining the buckling capacity of circular opening steel beams. *Materials*, 13(10), 2020.

[22] Dong Van Dao, Hojjat Adeli, Hai Bang Ly, Lu Minh Le, Vuong Minh Le, Tien Thinh Le, and Binh Thai Pham. A sensitivity and robustness analysis of GPR and ANN for high-performance concrete compressive strength prediction using a monte carlo simulation. *Sustainability (Switzerland)*, 12(3), 2020.

[23] Hai Bang Ly, Binh Thai Pham, Dong Van Dao, Vuong Minh Le, Lu Minh Le, and Tien Thinh Le. Improvement of ANFIS model for prediction of compressive strength of manufactured sand concrete. *Applied Sciences (Switzerland)*, 9(18), 2019.

[24] Ahmad Siam, Mohamed Ezzeldin, and Wael El-Dakhakhni. Machine learning algorithms for structural performance classifications and predictions: Application to reinforced masonry shear walls. *Structures*, 22:252–265, 2019.

[25] P. T. Nastos, K. P. Moustris, I. K. Larissi, and A. G. Paliatsos. Rain intensity forecast using Artificial Neural Networks in Athens, Greece. *Atmospheric Research*, 119:153–160, jan 2013.

[26] Francesco Faccini, Fabio Luino, Guido Paliaga, Alessandro Sacchini, Laura Turconi, and Carmen de Jong. Role of rainfall intensity and urban sprawl in the 2014 flash flood in Genoa City, Bisagno catchment (Liguria, Italy). *Applied Geography*, 98:224–241, 2018.

[27] Budi Warsito, Rahmat Gernowo, and Aris Sugiharto. Rainfall Prediction by Using Wavelet General Regression Neural Network. *International Journal of Applied Mathematics and Statistics*™, 54(3):32–41, mar 2016.

[28] Mislan, Haviluddin, Sigit Hardwinarto, Sumaryono, and Marlon Aipassa. Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggarong Station, East Kalimantan - Indonesia. In *Procedia Computer Science*, volume 59, pages 142–151, 2015.

[29] Ninan Sajeeth Philip and K. Babu Joseph. A neural network tool for analyzing trends in rainfall. *Computers and Geosciences*, 29(2):215–223, 2003.

[30] Vipul K. Dabhi and Sanjay Chaudhary. Hybrid Wavelet-Postfix-GP Model for Rainfall Prediction of Anand Region of India. *Advances in Artificial Intelligence*, 2014:1–11, 2014.

[31] Turgay Partal, H. Kerem Cigizoglu, and Ercan Kahya. Daily precipitation predictions using three different wavelet neural network algorithms by meteorological data. *Stochastic Environmental Research and Risk Assessment*, 29(5):1317–1329, jul 2015.

[32] E. G. Ortiz-García, S. Salcedo-Sanz, and C. Casanova-Mateo. Accurate precipitation prediction with support vector classifiers: A study including novel predictive variables and observational data. *Atmospheric Research*, 139:128–136, 2014.

[33] Weather Underground. Local Weather Forecast, News and Conditions | Weather Underground.

[34] Jack V. Tu. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, 49(11):1225–1231, 1996.

[35] Mehdi Khashei and Mehdi Bijari. A novel hybridiza-

tion of artificial neural networks and ARIMA models for time series forecasting. In *Applied Soft Computing Journal*, volume 11, pages 2664–2675, 2011.

[36] David E. Rumelhart and James L. McClelland. *Parallel Distributed Processing*. The MIT Press, aug 2019.

[37] M. J.D. Powell. Restart procedures for the conjugate gradient method. *Mathematical Programming*, 12(1):241–254, dec 1977.

[38] E. Beale. *A derivation of conjugate-gradients*. 1972.

[39] T Chai and R R Draxler. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3):1247–1250, 2014.

[40] Robert J Kuligowski and Ana P Barros. Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Weather and Forecasting*, 13(4):1194–1204, 1998.

[41] N Q Hung, M S Babel, S Weesakul, and N K Tripathi. An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Sciences*, 13(8):1413–1425, 2009.