Bulletin of Electrical Engineering and Informatics
Vol. 14, No. 4, August 2025, pp. 3075∼3090
ISSN: 2302-9285, DOI: 10.11591/eei.v14i4.8441

Rainfall forecasting by utilizing adaptive neuro-fuzzy inference system in Aceh Besar District

Hizir Sofyan[1], Nidya Tatsara[1], Yolanda[2], Tarmizi Usman[3], Irvanizam[4]
[1]Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh, Indonesia
[2]Graduate School of Mathematics and Applied Sciences, School of Postgraduate, Universitas Syiah Kuala, Banda Aceh, Indonesia
[3]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh, Indonesia
[4]Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh, Indonesia

Article Info

Article history:

Received Mar 6, 2024
Revised Jun 2, 2025
Accepted Jul 5, 2025

Keywords:

Adaptive neuro-fuzzy inference system
Forecasting
Fuzzy inference system
Mean absolute percentage error
Rainfall

ABSTRACT

Forecasting is a common thing to capture events in future based on previous information. However, some classical time-series methods, including moving average (MA), autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA), and simple exponential smoothing (SES), have limitations in predicting nonlinear time-series data. Therefore, this paper aims to utilize the adaptive neuro-fuzzy inference system (ANFIS) model, a combination of the fuzzy inference system (FIS) and neural network architecture to forecast a nonlinear rainfall problem. This model can capture the non-linear data, adaptation capability, and speedy learning capacity. We used the data consisting of temperature (ºC), humidity (%), and wind speed (km/hour) as input variables and rainfall (millimeter) as an output variable at two stations and one rain post in Aceh Besar District, from January 2009 to December 2019. The results demonstrated that ANFIS with generalized Bell (gBell) membership function on epoch 10 can successfully conduct rainfall forecasting in Aceh Besar District with the best-predicted value. The mean absolute percentage error (MAPE) of the prediction at the Meteorology, Climatology, and Geophysics Agency (MCGA) Station or *Badan Meteorologi, Klimatologi dan Geofisika* (BMKG) Indrapuri is 6.73% for 80% of the training dataset and 20% of the testing dataset.

Corresponding Author:

Irvanizam
Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala
Banda Aceh, Indonesia
Email: irvanizam.zamanhuri@usk.ac.id

1. INTRODUCTION

Rainfall is the amount of water that falls on a flat surface during a specific period. The water droplets or ice crystals that fall from the clouds when they reach the earth's surface are called rain, and if they come out from the bottom of the clouds but do not contact the earth's surface, they are called virga. According to the Meteorology, Climatology, and Geophysics Agency (MCGA) Station or *Badan Meteorologi, Klimatologi dan Geofisika* (BMKG) published on its official website, rainfall consists of three groups. Moderate rainfall ranges from 20 mm to 50 mm, Heavy rainfall ranges from 51 mm to 100 mm, and Strong heavy rainfall greater than 100 mm. The factors that correlate with high and low rainfall are climatic elements, including temperature, humidity, and wind speed [1].

Modeling of rainfall dataset has become essential approach in daily life. It can be used to estimate the possible impacts of climate change [2], analyze changes in rainfall intensity and variability [3], evaluate suitable indices at some tropical observing sites [4], and design processes for urban infrastructure [5]. On the other hand, high rainfall intensity causes many hydro-logical problems and effects on widely ranging some sectors. In agriculture, rainfall intensity can divide rainfall into infiltration and runoff [6]. It impacts the movement of soil that can break the soil structure and burn protective vegetation so that it damages crop growth [7].

Indonesia is a densely populated country with agriculture as one of its primary economic sectors. Among its provinces, Aceh plays a significant role in agricultural production despite covering only 3.02% of Indonesia's total land area [8]. Aceh Besar District, in particular, practices intensive rice cultivation, with up to three harvests per year, producing approximately 159,929 tons of rice from a total harvest area of 34,205 hectares. However, farmers in this region frequently face crop failures due to extreme weather conditions, including droughts and floods. Notably, nearly seventy of the rice fields in Aceh Besar still rely solely on rainfall as their primary water source, making them highly vulnerable to climate variability.

Based on BMKG, December, January, and February are months that have potentially heavy rain, even sometimes effect floods in some areas in Aceh Besar, a district of the province of Aceh. This district is located in northern part of the province or borders with Banda Aceh city about 30.21 km at coordinate of 5° 03' 1.2"- 5° 045' 9.007" N and 95° 55' 43.6" - 94° 59' 50.13" E. It has an area of 290,350.73 hectares. Most of its territory is mainland and a small part in the archipelago. The altitude of 35% areas of this territory is between 10 and 100 meters above sea level (mdpl). According to the climate situation, the strong heavy rainfall (December-January) in the territory can reach until up to 300 mm. This condition impacted river overflows and flooding at 15 sub-districts in Aceh Besar and Banda Aceh in May 2020 [9].

Given the topographical and climatic conditions mentioned above, it is imperative for the government of Aceh Besar and BMKG to establish an advanced warning system capable of delivering accurate, reliable, and timely decisions. The conventional approach to flood prediction relies on monitoring river water levels and categorizing them based on predefined danger thresholds. However, this method is reactive rather than predictive, as it only provides warnings once rainfall has already occurred, leaving limited time for residents to take precautionary measures or evacuate if necessary.

Traditional rainfall prediction methods, such as autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA), have been widely used due to their ability to model time-series data. However, these models struggle to handle the nonlinearity and stochastic nature of rainfall patterns, leading to reduced forecasting accuracy [10]. Machine learning techniques, such as artificial neural networks (ANN), have shown improvements in prediction capabilities, but they often require large datasets and extensive computational resources [11]. Furthermore, numerical weather prediction models, such as the general circulation model (GCM), lack the resolution needed for localized forecasting. These limitations highlight the necessity of an alternative approach that can effectively capture complex rainfall dynamics.

Nowadays, rainfall prediction has become a critical issue in tropical countries, where precise forecasting is essential for preventing crop failures and improving food security. Consequently, numerous studies have explored various forecasting models to enhance predictive accuracy. Swain *et al.* [12] developed an ARIMA-based model for monthly rainfall prediction in Odisha, India, integrating the Bayesian information criterion (BIC) and Akaike information criterion (AIC) to optimize model selection. Mahmud *et al.* [13] applied the SARIMA method for monthly rainfall forecasting in Bangladesh, demonstrating satisfactory predictive capabilities. Ridwan *et al.* [14] employed machine learning techniques to estimate rainfall intensity in Tasik Kenyir Terengganu, Malaysia. Hung *et al.* [15] enhanced rainfall forecasting performance in Bangkok, Thailand, using an ANN approach, which exhibited superior accuracy compared to traditional models. Thoeun [16] combined the GCM with the regional climate model (PRECIS) to analyze rainfall and temperature variations in highland areas of Cambodia. In Indonesia, Caraka *et al.* [17] applied a generalized space-time autoregressive model to forecast rainfall intensity in East Java, effectively illustrating spatial rainfall patterns using geospatial information system (GIS) software.

The trend of rainfall is hard to model and learn due to the uncertainty of climate change and the complexity of evaporation processes that generate rainfall. Most of scientists have considered as a linear data problem and have used time series (TS) methods to predict future events based on its past data patterns. The TS forecasting models, including moving average (MA), ARIMA, SARIMA, vector autoregression moving average (VARMA), and simple exponential smoothing (SES), have limitations, especially in nonlinear forecasting data [18].

Nevertheless, in reality, rainfall dataset is very potential in non-linear pattern. Hence, we need to select suitable forecasting techniques to overcome the limitations of linear predictive models. There are recently artificial intelligent (AI) models that have been proposed by some researchers as an alternative instead of the classical TS models in predicting time-series datasets. The AI models can predict the future dataset through the stochastic dependencies that can learn from the historical dataset with nonparametric nonlinear models [19]. One of the AI models is adaptive neuro fuzzy inference system (ANFIS) method. Many applications of this method build in for the prediction of cement strength [20], rainfall [21], sand concrete strength [22], nanofluids thermophysical properties [23], and cell-characteristic performance [24]. Those methods involved statistical methods and assisted by computer systems, statistical computation methods, fuzzy logic, and neural networks [25].

Fuzzy logic is a system described by linguistic knowledge [26]. Neural networks can build structures such as neural networks in the human brain with computer programs to complete calculations during a learning process [27]. Fuzzy logic and neural networks have limitations on complex systems. Fuzzy logic also has difficulties in determining the proper membership rules and functions. Meanwhile, neural networks require very long and complicated stages and processes. Therefore, it needs a hybrid method that combines fuzzy logic and neural networks using soft computing to deal with those limitations [28].

ANFIS is a variance of ANN that combines neural networks and fuzzy logic principles with the Takagi–Sugeno fuzzy inference system (FIS). The advantages of ANFIS are that it can provide a solution to many kinds of nonlinear and complex problems effectively. ANFIS is also better for solving nonlinear forecasting problems. Therefore, many researchers utilize forecasting models to know about events in the future by looking at past activities [29]. Additionally, ANFIS is widely applied to predict TS data. Among them are to forecast the Technical Assistance and Information Exchange (TAIEX) stock [30], to predict unsaturated hydraulic conductivity [31], and to predict the compressive strength of cement [32]. Based on these studies' results, the ANFIS model tends to achieve a small error value in many forecasting problems.

Due to the advantages of the ANFIS, many researchers have focused on developing, improving, and applying this method in their studies. ANFIS has been widely utilized in various fields, including meteorology, engineering, and medical sciences, due to its ability to model complex nonlinear relationships effectively. Duranoğlu *et al.* [33] demonstrated that the optimal type and number of membership functions (MFs) for ANFIS-based adsorption prediction can be systematically determined using the Box-Behnken experimental design, leading to improved accuracy, with the triangular MF yielding the best performance. Mwaura and Liu [34] proposed an ANFIS approach to enhance the detection and diagnosis of incipient faults in steam generator tubes, demonstrating high sensitivity, superior local interpolation, and accurate time-series prediction, with model validation conducted through simulations in the Qinshan I NPP reactor using RELAP5/SCDAP Mod4.0. Jin *et al.* [35] proposed a novel approach to simplify ANFIS for high-dimensional systems using importance-confidence-similarity (ICS) measures, effectively reducing fuzzy rules while maintaining accuracy, as demonstrated in chaotic TS and industrial temperature prediction applications. Sada and Ikpeseni [36] compares the performance of ANN and ANFIS in predicting machining responses, revealing that while both models exhibit strong predictive capabilities, ANN outperforms ANFIS in accuracy, with lower prediction errors and a higher correlation coefficient. Kanwal and Jiriwibhakorn [37] demonstrated that ANFIS-based models outperform traditional and other AI-based methods in fault detection, classification, and localization in transmission lines, achieving lower error rates and higher accuracy, making them suitable for real-time protection system applications.

This study aims to develop an ANFIS-based rainfall forecasting model tailored to the climatic conditions of Aceh Besar District. By comparing the performance of different membership functions and evaluating the optimal number of training epochs, this study seeks to determine the most effective configuration for accurate rainfall prediction. The results of this research are expected to contribute to improved early warning systems for flood prevention and enhanced agricultural planning. This study will focus on a station that can cover almost half of this district. To determine the best prediction result, we will demonstrate the mean absolute percentage error (MAPE) values of the training and testing datasets with two different membership functions and look at the best epoch performance.

The leftover of the paper is engaged as follows. Section 2 describes the research methodology involving datasets, study location, and a proposed neural network procedure. Section 3 explains the results and discussions. Finally, section 4 summarizes some conclusions, limitations, and potential future studies.

2. METHOD

2.1. Datasets and location

The data used in this study are secondary data obtained from the MCGA Station or BMKG Indrapuri, Aceh Besar District. Dataset is monthly rainfall data from January 2008 to December 2020 and consists of three parameters: temperature, humidity, and wind speed. The data were collected using automatic meteorological sensors installed at the BMKG Indrapuri Station (geographically: 5° 21' 21.492" N - 95° 34' 44.3244" E). The location of BMKG station is presented in Figure 1 where the distance from Banda Aceh City is about 12 km.

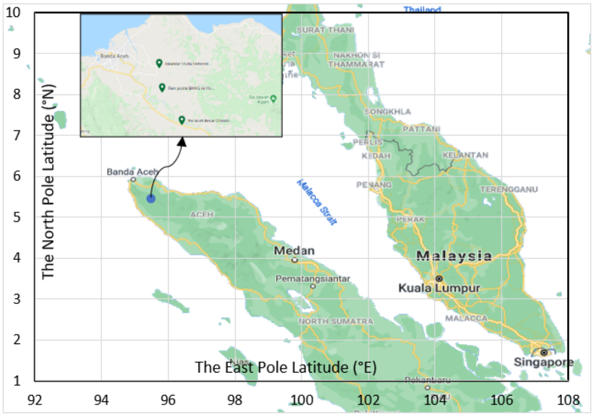

Figure 1. The location of study by Google Maps in Aceh Besar District

2.2. Neural network

A neural network, also known as an ANN, is a computational model inspired by the structure and functioning of the human brain. It is termed "artificial" because it is implemented through computer programs that perform various computational tasks during the learning process [38]. An ANN consists of interconnected processing units known as neurons, which are linked through communication pathways structured according to a specific network architecture. The process of determining the weights of these connections is governed by a learning algorithm. The architectural structure of an ANN is illustrated in Figure 2.

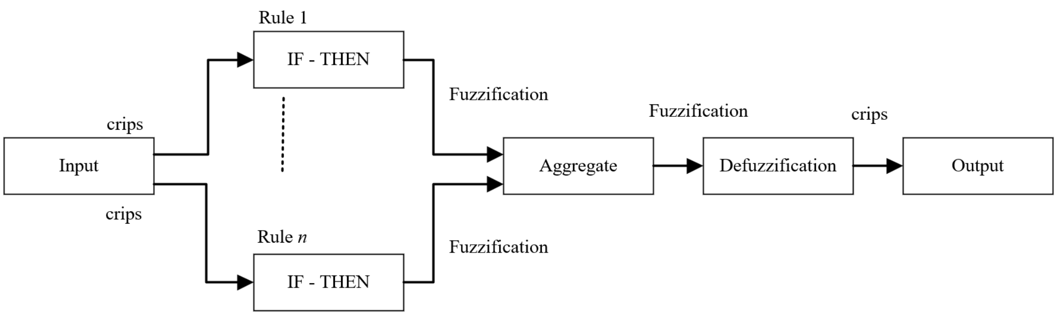

Figure 2. The diagram of FIS

An ANN has two stages of information processing which are training and testing stages. It also includes epoch and learning rate in the backpropagation training process [39]. Thus, it is very efficient in adapting and learning, but it has some negative attributes such as black box [40]. Meanwhile, fuzzy logic can capture the uncertainty of human thinking and interpret expert knowledge into calculable numerical data. Therefore, Zhang *et al.* [41] proposed ANFIS by integrating both NNs and fuzzy logic principles and compared them with the bilayered NN model. The ANFIS architecture uses both stipulated input-output data pairs and human knowledge to construct an input-output mapping by using a hybrid learning procedure. This architecture is constructed to model functions in nonlinear, identify nonlinear components on-line in a control system, and predict an uncontrolled TS. Similar to the Sugeno fuzzy rule, the architecture of ANFIS adopts the FIS and consists of five layers that can be seen in Figure 3.

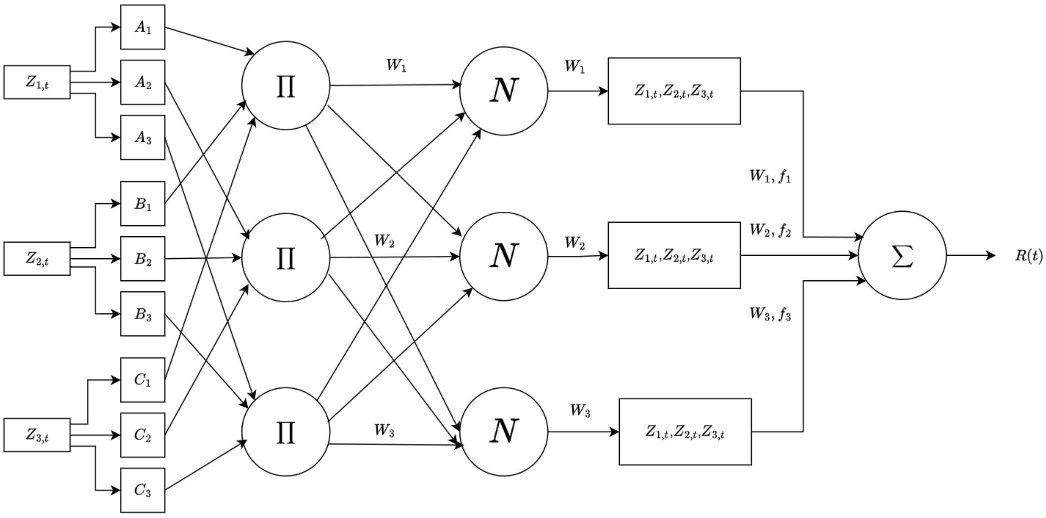

Figure 3. The architecture of ANFIS with two inputs, one output, and two rules

Figure 3 illustrates the architecture of the ANFIS model employed in this study, which consists of three inputs, $Z_{1,t}$, $Z_{2,t}$, and $Z_{3,t}$, and a single output, $R(t)$. The input variables represent the average monthly rainfall, which are utilized for rainfall prediction using the ANFIS model. The architecture comprises five layers. The first layer serves as the fuzzification layer, where each neuron is adaptive and parameterized by its activation function. The output of each neuron corresponds to the membership degree based on the input's membership function. The second layer consists of fixed (non-adaptive) neurons that compute the product of all incoming signals, effectively implementing the logical AND operator. The output of this layer represents the firing strength of each fuzzy rule, with each neuron corresponding to a specific rule as defined in (1).

$$w_i = \mu_{Ai}.\mu_{Bi} \tag{1}$$

However, in layer 3, each neuron is a non-adaptive neuron that calculates the ratio of the firing strength to $i(w_i)$ to the total firing strength in layer 2. The output of this layer is normalized firing strength by (2).

$$\bar{w}_t = \frac{w_i}{w_1 + w_2}, i = 1, 2 \tag{2}$$

Layer 4 is a neuron adaptive to an output where the normalized firing strength at layer 3 and p_i, q_i and r_i are the neuron's parameters or called consequent parameters. The result can be found by (3).

$$\bar{w}_i f_i = \bar{w}(p_i Z_{1,t} + q_i Z_{2,t} + r_i) \tag{3}$$

Layer 5 is a single neuron which is the sum of all outputs from layer 4. The single neuron can be obtained from (4).

$$\sum\nolimits_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{4}$$

The ANFIS architecture comprises five distinct layers, each performing a specific function. The first layer, classified as an adaptive layer, contains adaptive nodes, with each node capable of adjusting its parameters based on the learning process [42]. The output of this layer is derived by fuzzifying the input variables using membership functions, such as bell-shaped or Gaussian functions. To determine the premise parameters in this layer, clustering techniques are commonly employed as an initial step.

Currently, fuzzy C-means (FCM) is recognized as a straightforward yet effective clustering method applicable to various domains [43]. FCM is a clustering algorithm that assigns each data point a degree of membership to one or more clusters. The fundamental principle of FCM lies in the iterative determination of cluster centers, which represent the central locations of the respective clusters. This process is repeated iteratively, allowing the cluster centers to gradually converge toward their optimal positions. The iteration is guided by the minimization of an objective function, which quantifies the weighted distance between data points and cluster centers, based on the membership degrees. The final output of the FCM algorithm includes a set of cluster centers and the corresponding membership degrees for each data point. The procedural steps of the FCM algorithm are outlined as follows [44], and can be presented in a numbered list.

1) Let X_{kj} denote the data matrix to be clustered, where k represents the number of data points (objects) and j denotes the number of attributes (criteria) associated with each data point;
2) Define the parameters required for the FCM algorithm, including: the number of clusters to be formed ($c \geq 2$), the fuzziness weighting exponent ($w > 1$), the maximum number of iterations ($n_{\max}$), the convergence threshold or stopping criterion (ξ is small positive value), and the initial value of the objective function (P_0);
3) Initialize the membership degrees of each data point to the clusters (i.e., the degree of membership of data point k to cluster i) using a random method, where k denotes the number of data points and i represents the number of clusters;
4) Compute the center of cluster (V) for all clusters using (5);

$$V_{ij} = \frac{\sum_{k=1}^{n} (\mu_{ik})^{w} X_{kj}}{\sum_{k=1}^{n} (\mu_{ik})^{w}} \tag{5}$$

Here, V_{ij} denotes the center of the i-th cluster for the j-th attribute. The term μ_{ik} represents the membership degree of the k-th data point to the i-th cluster, as recorded in the partition matrix U. Meanwhile, X_{kj} refers to the value of the k-th data point with respect to the j-th attribute, also from the data matrix. The parameter w is the fuzziness (weighting) exponent used in the clustering process;
5) Calculate the objective values (P_n) using (6).

$$P_n = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^{w} d_{ik}^{2} \tag{6}$$

where P_n is objective value on n-th iteration and d_{ik} is the function of the Euclidean distance of the i-th cluster center and k-th data;
6) Fix the membership degree of data in each cluster (correction of partition matrix) using (7).

$$\mu_{ik} = \left[\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(w-1)} \right]^{-1} \tag{7}$$

with the value of d_{ik} can be calculated using (8).

$$d_{ik} = d(X_k - v_i) = \left[\sum_{j=1}^{m} (X_{kj} - v_{ij}) \right]^{1/2} \tag{8}$$

where d_{ik} is the Euclidean distance from the cluster center of the i-th and k-th data. X_{kj} is the data (at matrix U) on the j-th attribute and k-th data;
7) Terminate the iteration if the cluster centers V remain unchanged. Alternatively, the stopping criterion can be defined based on the change in the objective function, such that: $|P_n - P_{n-1}| < \xi \quad \text{and} \quad n > n_{\max}$, where ξ is a small positive threshold and $n_{\max}$ is the maximum number of iterations. If neither stopping condition is met, return to Step 4. Once the iteration terminates, each data point is assigned to a cluster. The final cluster assignment is based on the highest membership value in the partition matrix.

The data that have been clustered will be used to calculate the value of standard deviation (a) and mean (c) for each cluster. Then the values will be the initial premise parameters in ANFIS method. After constructing the ANFIS methodology, the datasets were selected in this study and divided into two groups. One is as training dataset, and other is as testing dataset. The composition of datasets is 80% for training and 20% for testing. Afterward, the system output will be evaluated to approximate the expected results to the system. If there is not exist, then the learning algorithm will be modified by taking new network architecture. We used root mean squared error (RMSE) to test the dataset. The RMSE is also used to know how much the forecast deviates from the period's actual value. The RMSE can be derived from (9).

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(Z_t - \widehat{Z}_t \right)^2} \tag{9}$$

After that, the rainfall prediction results are also evaluated with MAPE. The ANFIS will consider and decide if the prediction value of MAPE in the training process must be below 20%. This evaluation can be expressed using (10):

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \frac{\left| Z_t - \widehat{Z}_t \right|}{Z_t} \tag{10}$$

where Z_t is the prediction value at time t, $\widehat{Z}_t$ is the observation value at time t, and n is the number of datasets.

In summary, we define ten steps in developing the methodology. These steps are as follows: i) input the data of temperature, humidity, and wind speed from 2009 until 2019 in Aceh Besar District; ii) find out summary statistics of rainfall, temperature, humidity, and wind speed from 2009 until 2019 in Aceh Besar District and plotting the data; iii) transform the ANFIS network architecture using the Sugeno model order 1 consisting of 5 layers; iv) cluster classes using the FCM algorithm with the number of clusters used is 3. Then the standard deviation and average values are calculated for each cluster. The standard deviation (a) and average (c) values will be the initial premise parameters for the BMKG rainfall station; v) the degree of membership is calculated using the generalized bell (gBell) membership function with standard deviation (a), b = 1, and average (c) of each cluster at each station and rain post; vi) determine the amount of training data from January 2009 to December 2015, and the testing data used were from January 2016 to December 2018. At this stage, the number of epochs and learning rates was also tested until the value was obtained the smallest MAPE at each station and rain post. Besides, we also attempt to analyze the prediction result for the dataset in which the composition of training and testing is 80:20; vii) ANFIS training was carried out from layer 1 to layer 5 in order to obtain consequent parameter values. The recursive LSE is used to correct the consequent parameters (forward) while the error propagation model is used to correct the premise parameter (backward) at each station and rain post; viii) forecast rainfall in 2018 uses the ANFIS model with the smallest MAPE value at each station and rain post; ix) compare the results of the 2018 rainfall forecasting with actual rainfall data (test data) in 2018 by plotting data at each station and rain post; and x) made a conclusion.

3. RESULT AND DISCUSSION

3.1. Meteorological parameters

Figure 4 presents the monthly average of rainfall in Aceh Besar from January 2009 until December 2019. As we can see in the figure, in the end of each year, rainfall pattern showed a higher rainfall from October to December for each year. During these eleven years, the highest was experienced in December 2018 and reached more than 500 mm.

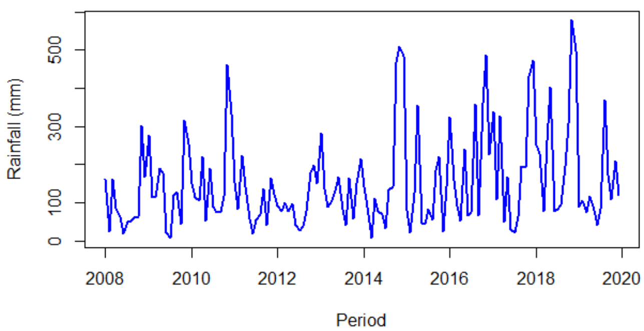

Figure 4. Monthly average variations of rainfall in Aceh Besar District for eleven years (2009-2019)

3.2. Rainfall forecasting

The monthly dataset for eleven years presented in Figure 4 can be directly used to predict monthly rainfall due to it will not take a long-time computation. We then normalized the dataset into the range from 0 to 1 using (8) to avoid large computation steps in the pre-processing.

We used 11-year data to predict this rainfall case where the data from January 2009 to December 2015 (7 years) are as the training data. In the training process, we used gBell and Gaussian Membership Functions for defuzzification. This step evaluated a FIS that can produce RMSE and MAPE. The rest of the 4-year dataset (from January 2016 to December 2019) will be the testing data. In other words, we evaluated by trial and error, in this case, is 70%:30%. It was also involved the epoch from 10 to 50 to obtain the optimal results. Figure 5 demonstrates the result of testing for 4-year data.

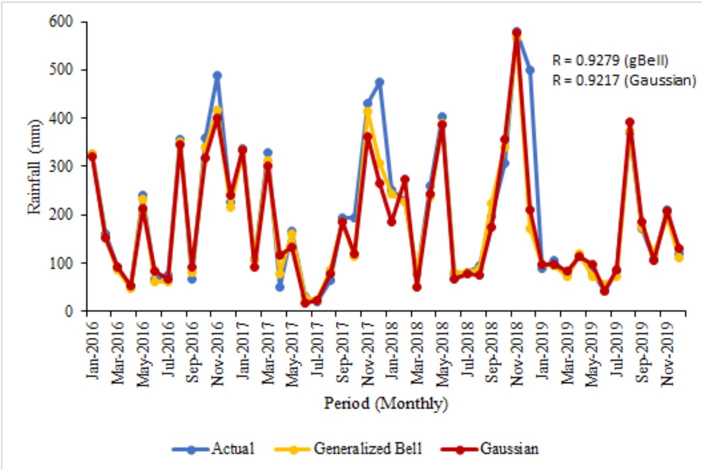

Figure 5. Rainfall prediction result for testing in Aceh Besar District based on eleven years data (2009-2019) where the composition of training and datasets are 70%:30%

Figure 6 and Table 1 show the performance for testing results through evaluating some different epoch values where the comparison of MAPE and RMSE values for gBell and Gaussian membership functions used 70:30 dataset. Figures 6(a) to (f) illustrate the comparison between actual monthly rainfall data and the predicted results obtained using two different membership functions, namely gBell and Gaussian, during the testing phases at epoch 10, epoch 20, epoch 30, epoch 40, and epoch 50. The x-axis represents the monthly period from January 2018 to December 2019, while the y-axis shows the rainfall amount measured in mm. In five figures, the black line represents the actual observed rainfall, the blue line represents the predicted rainfall using the gBell membership function, and the red line represents the predicted rainfall using the Gaussian membership function.

Table 1. The performance for training and testing with epoch where the composition of training and testing datasets are 70%:30%

Performance indicator	Epoch=10	Epoch=20	Epoch=30	Epoch=40	Epoch=50
gBell membership function					
MAPE training	0.3132	0.4326	0.5272	0.6761	0.6237
MAPE testing	0.1085	0.1220	0.1735	0.1956	0.2299
RMSE training	52.1930	58.6299	60.2769	62.3214	86.8944
RMSE testing	56.3857	60.4027	68.1765	70.8584	78.0764
Gaussian membership function					
MAPE training	0.3622	0.4159	0.5879	0.5909	0.7300
MAPE testing	0.1251	0.1456	0.1973	0.2047	0.2371
RMSE training	56.1624	59.9002	59.0007	59.6853	72.3810
RMSE testing	63.4532	56.1980	66.8511	71.0093	68.0432

As the training epochs increase from 10 to 50, there is a clear trend of improvement in the predictive performance of both the gBell and Gaussian membership functions. In the testing epoch 10 figure, both models are still in the early learning stage. The predictions exhibit significant deviations from the actual rainfall values, especially during extreme rainfall events. While the gBell function shows a slightly better alignment with the actual data compared to the Gaussian function, both models struggle to fully capture the peaks and sharp fluctuations.

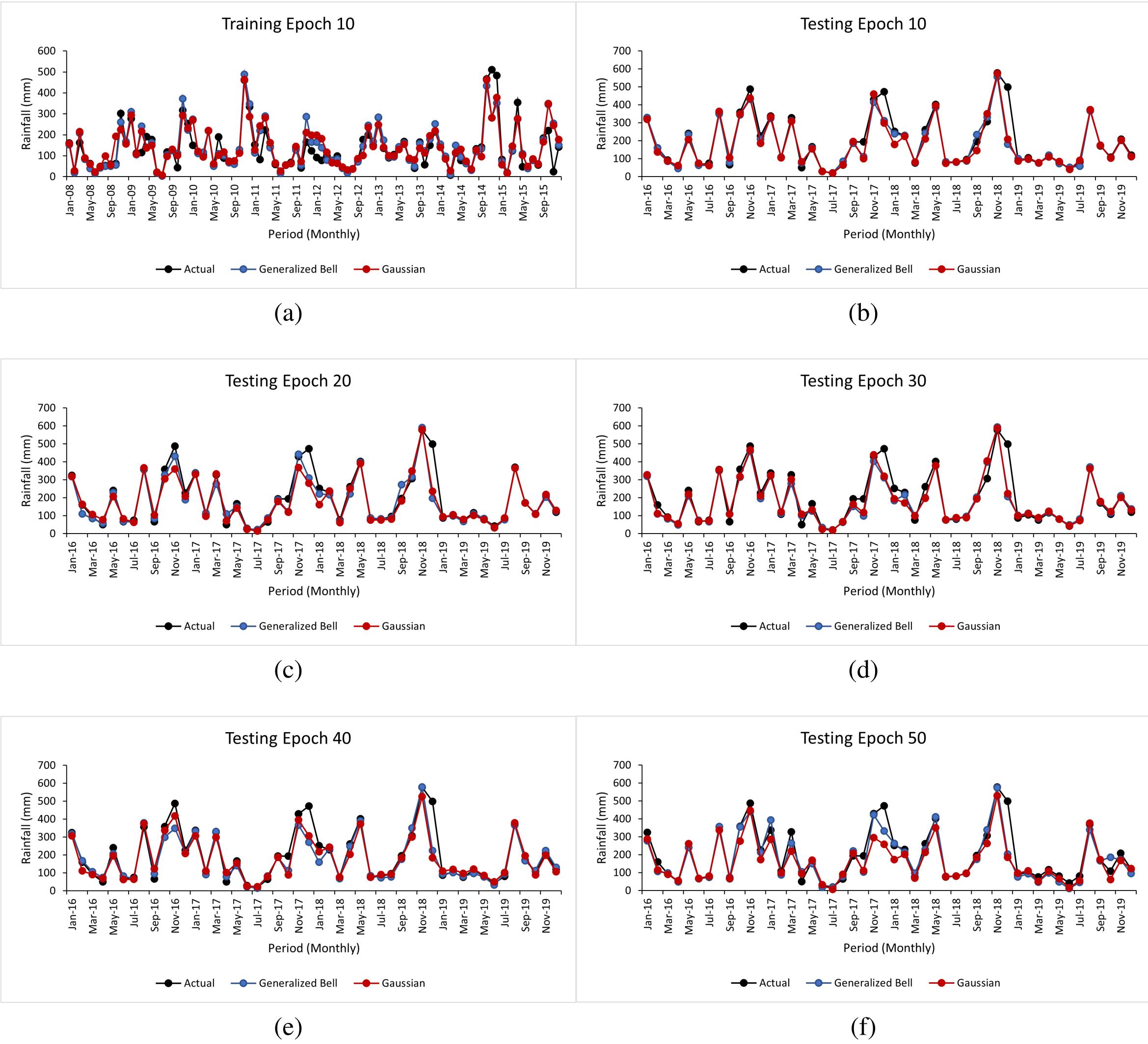

Figure 6. The results according to eleven years data (2009-2019) to forecast rainfall in Aceh Besar using gBell and Gaussian where data comparison for training and testing is 70:30; (a) with training at 10-th epoch, (b) with testing at 10-th epoch, (c) with testing at 20-th epoch, (d) with testing at 30-th epoch, (e) with testing at 40-th epoch, and (f) with testing at 50-th epoch

By epoch 20 and epoch 30, the models begin to show notable improvements. The predicted rainfall patterns start to resemble the actual values more closely, particularly in capturing the timing and magnitude of rainfall peaks. The gBell function consistently demonstrates smoother and more stable predictions, with a better fit to the real data. The Gaussian function, while improved, still occasionally underestimates or overshoots the rainfall peaks in certain months. The improvement indicates that both models benefit from additional training, reducing prediction errors and increasing reliability.

At epoch 40, the models reach a level of prediction quality where the lines almost perfectly overlap with the actual data in many time periods. The accuracy is particularly evident in the clear detection of rainfall spikes and troughs. Both membership functions perform well, but the Gaussian function seems to catch up with the gBell in terms of trend-following and magnitude approximation. This epoch marks a stage of model maturity, where both functions generalize well on unseen data.

However, at epoch 50, some signs of overfitting or diminishing returns become apparent—especially in the Gaussian model. While the gBell function maintains relatively stable and accurate predictions, the Gaussian function begins to exhibit slight inconsistencies, particularly in the over- or underestimation of extreme values. The fluctuation in prediction accuracy suggests that too many epochs may lead the model to memorize the training data rather than generalize well to testing data.

Overall, increasing the number of training epochs improves the model's ability to replicate the actual rainfall patterns, with the gBell membership function consistently showing more reliable and accurate performance. Epoch 40 appears to offer the best balance between accuracy and generalization, while epoch 50 may indicate the onset of overfitting. Therefore, careful selection of the optimal number of training epochs is crucial to ensure robust and generalizable predictions.

Evaluating the prediction rain-fall value using the gBell membership function can impact the MAPE or RMSE values on the epoch value. If we increase the epoch value, the MAPE or RMSE value also increases slightly. It indicates that ANFIS can predict rainfall.

Assume that the gBell is the best membership function. It then shows that the RMSE value for both testing and training increases continually according to increasing the epoch value. The same thing occurred for the MAPE value for training and testing. The smallest MAPE and RMSE values were reached when the given at 10-th epoch. Meanwhile, the Gaussian membership function yielded to fluctuate the RMSE value for training and testing. At 10-th epoch, the RMSE training reached the smallest value while the MAPE value for either testing or training appears to increase continually with the increase in the epoch value.

Compared with the results in Figure 7 and Table 2 used the composition of 80%:20%. Figures 7(a) to (f) illustrate the comparison between actual monthly rainfall data and the predicted results obtained using two different membership functions, namely gBell and Gaussian, during the testing phases at epoch 10, epoch 20, epoch 30, epoch 40, and epoch 50. The x-axis represents the monthly period from January 2018 to December 2019, while the y-axis shows the rainfall amount measured in mm. In five figures, the black line represents the actual observed rainfall, the blue line represents the predicted rainfall using the gBell membership function, and the red line represents the predicted rainfall using the Gaussian membership function.

Table 2. The performance for training and testing with epoch where the composition of training and testing datasets are 80%:20%

Performance indicator	Epoch=10	Epoch=20	Epoch=30	Epoch=40	Epoch=50
gBell membership function					
MAPE training	0.3772	0.3753	0.3896	0.6041	0.6422
MAPE testing	0.0673	0.0994	0.1001	0.1293	0.1583
RMSE training	62.1890	58.4584	58.7092	62.0837	71.8061
RMSE testing	55.7271	57.6898	66.8699	61.7495	74.3214
Gaussian membership function					
MAPE training	0.3717	0.4019	0.5252	0.6221	0.7090
MAPE testing	0.0960	0.1352	0.1419	0.1644	0.2081
RMSE training	58.8623	56.5308	61.5801	61.2534	62.6755
RMSE testing	57.8335	71.0524	57.7439	59.2218	56.4862

As the number of training epochs increases from 10 to 50, there is a visible improvement in the prediction performance of both the gBell and Gaussian models. At epoch 10, the predicted values still show some noticeable deviations from the actual rainfall, especially during peak periods such as March 2018, October–December 2018, and August 2019. The Gaussian model performs better than the gBell model in approximating those peaks, although both models exhibit underestimations in some months.

By epoch 20, the accuracy of both models improves. The Gaussian model continues to capture the trends more precisely, especially in months with high rainfall like May 2018 and November 2018, where it closely aligns with the actual data. The gBell model also shows reduced error compared to epoch 10 but still tends to slightly underestimate during high rainfall months. This suggests that additional training allows the models to better generalize the underlying patterns in the data.

At epoch 30, the Gaussian model maintains its strong performance and continues to align well with the actual rainfall pattern, showing consistent accuracy even in months with significant rainfall spikes, such as November 2018 and August 2019. In contrast, the gBell model's predictions appear to plateau or slightly decline in quality during extreme rainfall events, indicating potential limitations in its ability to capture more complex nonlinear patterns, despite longer training.

Figure 7. The results according to eleven years data (2009-2019) to forecast rainfall in Aceh Besar using gBell and Gaussian where data comparison for training and testing is 80:20: (a) with training epoch=10; (b) with testing at 10-th epoch; (c) with testing at 20-th epoch; (d) with testing at 30-th epoch; (e) with testing at 40-th epoch; and (f) with testing at 50-th epoch

At epoch 40 and epoch 50, it can be observed that both models are capable of following the general pattern of the actual rainfall data. The predicted values from both membership functions successfully capture the major peaks and troughs throughout the observed period. However, the Gaussian function tends to produce slightly higher peak predictions compared to the gBell function, especially at higher rainfall values. The prediction results show minor improvements in fitting the actual data pattern, indicating enhanced model performance.

Increasing the number of training epochs enhances the prediction performance, particularly for the Gaussian model, which demonstrates greater accuracy and stability across all epochs. The gBell model, while improving between epoch 10 and 20, shows less consistent performance by epoch 30, especially in high rainfall scenarios. This comparison highlights the superiority of the Gaussian membership function in this context, especially as the training process matures. In general, both membership functions demonstrate good reliability in forecasting monthly rainfall during the testing phases.

Similarly, the MAPE value for testing using the gBell is slightly smaller than the results in Table 1. In addition, the RMSE value for testing in Table 2 also showed the increasing the epoch val-ue in which at the epoch = 10, the smallest RMSE and MAPE values obtained 55.7271 and 0.0673. In other words, the prediction results of the evaluation of these two membership functions showed that gBell membership function demonstrated in Table 2 produced the best results, as shown in Figure 8.

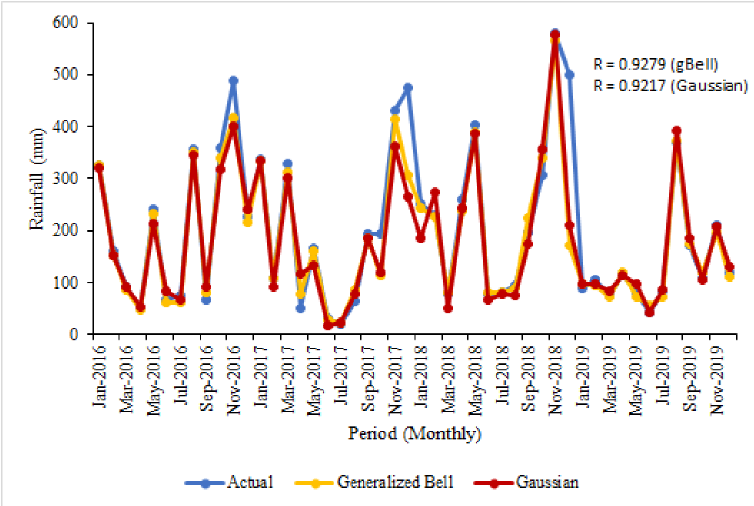

Figure 8. Rainfall prediction result for testing in Aceh Besar District based on eleven years data (2009-2019) where the composition of training and datasets are 80%:20%

These findings highlight that an increase in data variability does not necessarily lead to decreased performance in rainfall forecasting using ANFIS. Instead, the proposed approach can benefit from a larger training dataset and refined membership function optimization, ensuring enhanced predictive accuracy and model stability. In addition, we found that the choice of membership function correlates with the accuracy and stability of rainfall forecasting using ANFIS. The proposed method in this study tended to have an inordinately higher proportion of accurate predictions as the gBell membership function, combined with an optimized epoch value, resulted in lower MAPE and RMSE values.

The findings of this study have several important implications for the field of rainfall forecasting and the broader application of AI in meteorology. First, the results indicate that an increase in data variability does not necessarily degrade the performance of ANFIS models. This challenges traditional assumptions that highly variable datasets lead to less reliable predictions. Instead, it highlights the importance of training data diversity and the role of data-driven optimization in improving forecasting accuracy. As a result, researchers and practitioners can focus on collecting and utilizing richer datasets rather than merely filtering out variability, ensuring that ANFIS models remain robust in different climatic conditions.

Second, the study emphasizes the critical impact of membership function selection in ANFIS-based rainfall forecasting. The use of the gBell membership function, combined with an optimized epoch value, resulted in lower MAPE and RMSE values, suggesting that model performance is highly sensitive to the choice of fuzzy parameters. This finding has implications for future AI-driven forecasting models, as it underscores the need for careful selection and fine-tuning of fuzzy logic components to enhance both accuracy and stability. The insights gained from this research can guide future developments in FIS, helping meteorologists and climate scientists to build more precise and adaptive forecasting models.

Finally, these results suggest that ANFIS can be effectively applied to rainfall prediction in regions with complex weather patterns, such as Aceh Besar District. The ability of ANFIS to adapt to varying environmental conditions makes it a valuable tool for improving early warning systems and climate resilience strategies. By integrating optimized ANFIS models into meteorological forecasting platforms, decision-makers can enhance disaster preparedness, optimize water resource management, and mitigate the socio-economic impacts of extreme weather events.

Furthermore, the study underscores the crucial role of membership function selection in ANFIS-based forecasting models. The gBell membership function, when combined with an optimized epoch value, led to significantly lower MAPE and RMSE values. This finding provides a strong foundation for future studies aiming to refine membership function design to further improve accuracy and computational efficiency. Additionally, it offers practical guidance for meteorologists and data scientists in selecting the most suitable fuzzy logic structures for their forecasting models, ensuring more reliable predictions in diverse climatic conditions.

In the future, these insights could contribute to the development of more adaptive and robust rainfall prediction models, particularly in the context of climate change, where weather patterns are becoming increasingly unpredictable. The integration of ANFIS with real-time data assimilation techniques, hybrid optimization algorithms, and ensemble learning approaches could further enhance the reliability and precision of rainfall forecasts. Additionally, policymakers and disaster management authorities can benefit from these advancements by incorporating ANFIS-based forecasting models into early warning systems, improving preparedness for extreme weather events and mitigating potential socio-economic impacts.

By building upon the findings of this study, future research can explore novel ways to optimize ANFIS models, such as employing deep learning techniques for feature extraction, utilizing high-resolution satellite data, or integrating internet of things (IoT) sensor networks to enhance real-time rainfall forecasting. These advancements will ensure that rainfall prediction models continue to evolve, providing more accurate and actionable insights for climate resilience, water resource management, and disaster mitigation strategies.

4. CONCLUSION

The ANFIS method has successfully carried out the prediction of rainfall in Aceh Besar District. The average performance of the testing process is about 6.73%. This method has predicted the rainfall dataset from 2009 to 2019 where the gBell membership function on epoch 10 for training and testing using the composition of 80:20 provided the best predictive value. Recent observations indicate that the ANFIS model with the gBell membership function provides more accurate and resilient rainfall forecasting compared to other tested configurations. Our findings offer definitive proof that this phenomenon is linked to the model's ability to adaptively learn complex rainfall patterns and capture nonlinear relationships, rather than being caused by increased quantities of historical rainfall data alone. However, we realize that the rainfall data used in this research is not large enough so this prediction model has limitations in obtaining more accurate results. In addition, this research conducted an extensive analysis of the gBell membership function and its influence on predictive performance. Nonetheless, further in-depth studies may be necessary to validate its stability and applicability, especially in relation to different datasets and practical implementations.

In a tropical country, climate change is erratic and climatic conditions are very fluctuating. It needs further rainfall forecasting using big data. Our research demonstrates that the MAPE value for testing using the gBell with an 80:20 ratio is more resilient than testing using the gBell with a 70:30 ratio. Future research may look into optimizing the selection of membership functions and parameter tuning techniques and practical methods for producing more adaptive and generalized FIS for diverse datasets. Additionally, future studies can also investigate and determine other potential categories of rainfall that can potentially affect floods, landslides, and disasters with integrating the decision-making methods, deep learning technologies, early warning systems, and IoT technologies to determine suitable forecasting parameters.

ACKNOWLEDGEMENT

The authors would like to thank all academic staff of the Department of Statistics, the Department of Mathematics, and the Department of Informatics, Universitas Syiah Kuala, who supported this study. Additionally, the authors thank anonymous reviewers and editors for their constructive suggestions and valuable comments.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Hizir Sofyan		✓		✓	✓	✓		✓	✓			✓	✓	✓
Nidya Tatsara	✓	✓	✓	✓		✓	✓		✓		✓			
Yolanda		✓	✓	✓		✓		✓	✓	✓			✓	
Tarmizi Usman		✓		✓	✓	✓		✓	✓		✓			✓
Irvanizam	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the fndings of this study are available from the corresponding author, [I], upon reasonable request.

REFERENCES

[1] C. Corradini, R. Morbidelli, C. Saltalippi, and A. Flammini, “Meteorological systems producing rainfall,” *Rainfall*, Elsevier, Chapter 2, pp. 27–48, 2022, doi: 10.1016/b978-0-12-822544-8.00011-1.

[2] R. M. Santos and R. Bakhshoodeh, “Climate change/global warming/climate emergency versus general climate research: comparative bibliometric trends of publications,” *Heliyon*, vol. 7, no. 11, pp. 1-15, Nov. 2021, doi: 10.1016/j.heliyon.2021.e08219.

[3] D. Dunkerley, “Sub-Daily Rainfall Intensity Extremes: Evaluating Suitable Indices at Australian Arid and Wet Tropical Observing Sites,” *Water*, vol. 11, no. 12, pp. 1-18, 2019, doi: 10.3390/w11122616.

[4] R. Maier, G. Krebs, M. Pichler, D. Muschalla, and G. Gruber, “Spatial Rainfall Variability in Urban Environments—High-Density Precipitation Measurements on a City-Scale,” *Water*, vol. 12, no. 4, pp. 1-16, 2020, doi: 10.3390/w12041157.

[5] J. Chen, X. Liu, and H. Huang, “Exploration of Urban Green Infrastructure Design Practice Methodology under Multidisciplinary Collaboration,” in *2022 International Conference on Computer Technologies (ICCTech)*, Melaka, Malaysia, 2022, pp. 87-95, doi: 10.1109/icctech55650.2022.00022.

[6] C. Langhans, G. Govers, J. Diels, W. Clymans, and A. Van den Putte, “Dependence of effective hydraulic conductivity on rainfall intensity: loamy agricultural soils,” *Hydrological Processes*, vol. 24, no. 16, pp. 2257–2268, 2010, doi: 10.1002/hyp.7800.

[7] G. R. Kishore, C. K. Roopa, and B. S. Harish, “Predicting the cause for Crop Damage using Machine Learning Approaches,” in *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, Mysuru, India, 2022, pp. 1-4, doi: 10.1109/mysurucon55714.2022.9972399.

[8] A. Wijaya, R. Firmansyah, Z. Said, and B. Nathania, “Monitoring of Indonesia Tropical Rainforests and Land Cover Change Using Hybrid Approach of Time Series Landsat Data,” in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 5980-5983, doi: 10.1109/igarss.2019.8900121.

[9] M. Al’ala and Syamsidik, “Coastal Flooding Impacts Induced Sea Level Rise on Banda Aceh Coastal Morphology,” in *IOP Conference Series: Earth and Environmental Science, The 11th Aceh International Workshop and Expo on Sustainable Tsunami Disaster*, 2019, vol. 273, doi: 10.1088/1755-1315/273/1/012007.

[10] B. Siregar, E. B. Nababan, A. Yap, U. Andayani, and Fahmi, ”Forecasting of raw material needed for plastic products based in income data using ARIMA method,” in *2017 5th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*, Malang, Indonesia, 2017, pp. 135-139, doi: 10.1109/ICEEIE.2017.8328777.

[11] A. Maulana, T. R. Noviandy, R. Idroes, N. R. Sasmita, R. Suhendra, and I. Irvanizam, ”Prediction of Kovats Retention Indices for Fragrance and Flavor using Artificial Neural Network,” in *2020 International Conference on Electrical Engineering and Informatics (ICELTICs)*, Aceh, Indonesia, 2020, pp. 1-5, doi: 10.1109/ICELTICs50595.2020.9315391.

[12] S. Swain, S. Nandi, and P. Patel, “Development of an ARIMA Model for Monthly Rainfall Forecasting over Khordha District Odisha India,” *Recent Findings in Intelligent Computing Techniques, Proceedings of the 5th ICACNI*, 2018, pp. 325–331, doi: 10.1007/978-981-10-8636_634.

[13] I. Mahmud, S. H. Bari, and M. T. U. Rahman, “Monthly rainfall forecast of Bangladesh using autoregressive integrated moving average method,” *Environmental Engineering Research*, vol. 22, no. 2, pp. 162–168, 2016, doi: 10.4491/eer.2016.075.

[14] W. M. Ridwan, M. Sapitang, A. Aziz, K. F. Kushiar, A. N. Ahmed, and A. El-Shafie, “Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia,” *Ain Shams Engineering Journal*, vol. 12, no. 2, pp. 1651–1663, 2021, doi: 10.1016/j.asej.2020.09.011.

[15] 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*, vol. 13, no. 8, pp. 1413–1425, 2009, doi: 10.5194/hess-13-1413-2009.

[16] H. C. Thoeun, “Observed and projected changes in temperature and rainfall in Cambodia,” *Weather and Climate Extremes*, vol. 7, pp. 61–71, 2015, doi: 10.1016/j.wace.2015.02.001.

[17] R. E. Caraka *et al.*, “Generalized Spatio Temporal Autoregressive Rainfall-Enso Pattern In East Java Indonesia,” in *2018 Indonesian Association for Pattern Recognition International Conference (INAPR)*, Jakarta, Indonesia, 2018, pp. 75-79, doi: 10.1109/inapr.2018.8627042.

[18] H. Hermansah, D. Rosadi, A. Abdurakhman, and H. Utami, “Automatic time series forecasting using nonlinear autoregressive neural network model with exogenous input,” *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 5, pp. 2836–2844, 2021, doi: 10.11591/eei.v10i5.2862.

[19] I. J. R. Filho, R. M. Marcacini, and S. O. Rezende, “On the enrichment of time series with textual data for forecasting agricultural commodity prices,” *MethodsX*, vol. 9, 2022, doi: 10.1016/j.mex.2022.101758.

[20] S. Elhishi, A. M. Elashry, and S. El-Metwally, “Unboxing machine learning models for concrete strength prediction using XAI,” *Scientific Reports*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-47169-7.

[21] S.-M. Chen and S.-W. Chen, “Fuzzy Forecasting Based on Two-Factors Second-Order Fuzzy-Trend Logical Relationship Groups and the Probabilities of Trends of Fuzzy Logical Relationships,” *IEEE Transactions on Cybernetics*, vol. 45, no. 3, pp. 391-403, 2015, doi: 10.1109/tcyb.2014.2326888.

[22] P. Sihag, N. K. Tiwari, and S. Ranjan, “Prediction of unsaturated hydraulic conductivity using adaptive neuro-fuzzy inference system (ANFIS),” *ISH Journal of Hydraulic Engineering*, vol. 25, no. 2, pp. 132–142, 2017, doi: 10.1080/09715010.2017.1381861.

[23] Z. Said, L. S. Sundar, H. Rezk, A. M. Nassef, S. Chakraborty, and C. Li, “Thermophysical properties using ND/water nanofluids: An experimental study, ANFIS-based model and optimization,” *Journal of Molecular Liquids*, vol. 330, 2021, doi: 10.1016/j.molliq.2021.115659.

[24] Q. Zhang *et al.*, “Computational prediction and characterization of cell-type-specific and shared binding sites,” *Bioinformatics*, vol. 39, no. 1, 2022, doi: 10.1093/bioinformatics/btac798.

[25] M. Ishaque, M. G. M. Johar, A. Khatibi, and M. Yamin, “A novel hybrid technique using fuzzy logic, neural networks and genetic algorithm for intrusion detection system,” *Measurement: Sensors*, vol. 30, 2023, doi: 10.1016/j.measen.2023.100933.

[26] H. Hasanzadeh and J. Majidpour, “Application of fuzzy logic in order to plan and create software for converting sound to text,” *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 3, pp. 1127–1136, Jun. 2020, doi: 10.11591/eei.v9i3.1723.

[27] I. Irvanizam, I. Syahrini, R. P. F. Afidh, M. R. Andika, and H. Sofyan, ”Applying Fuzzy Multiple-Attribute Decision Making Based on Set-pair Analysis with Triangular Fuzzy Number for Decent Homes Distribution Problem,” in *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, Parapat, Indonesia, 2018, pp. 1-7, doi: 10.1109/CITSM.2018.8674290.

[28] D.-A. Pham and S.-H. Han, “Design of Combined Neural Network and Fuzzy Logic Controller for Marine Rescue Drone Trajectory-Tracking,” *Journal of Marine Science and Engineering*, vol. 10, no. 11, pp. 1-23, 2022, doi: 10.3390/jmse10111716.

[29] F. Petropoulos *et al.*, “Forecasting: theory and practice,” *International Journal of Forecasting*, vol. 38, no. 3, pp. 705–871, 2022, doi: 10.1016/j.ijforecast.2021.11.001.

[30] C.-H. Cheng, L.-Y. Wei, J.-W. Liu, and T.-L. Chen, “OWA-based ANFIS model for TAIEX forecasting,” *Economic Modelling*, vol. 30, pp. 442–448, 2013, doi: 10.1016/j.econmod.2012.09.047.

[31] A. Peters, T. L. Hohenbrink, S. C. Iden, M. Th. van Genuchten, and W. Durner, “Prediction of the absolute hydraulic conductivity function from soil water retention data,” *Hydrology and Earth System Sciences*, vol. 27, no. 7, pp. 1565–1582, 2023, doi: 10.5194/hess-27-1565-2023.

[32] H. A. Al-Jamimi, W. A. Al-Kutti, S. Alwahaishi, and K. S. Alotaibi, “Prediction of compressive strength in plain and blended cement concretes using a hybrid artificial intelligence model,” *Case Studies in Construction Materials*, vol. 17, 2022, doi: 10.1016/j.cscm.2022.e01238.

[33] D. Duranoğlu, E. S. Altın, and İ. Küçük, “Optimization of adaptive neuro–fuzzy inference system (ANFIS) parameters via Box-Behnken experimental design approach: The prediction of chromium adsorption,” *Heliyon*, vol. 10, no. 3, 2024, doi: 10.1016/j.heliyon.2024.e25813.

[34] A. M. Mwaura and Y.-K. Liu, “Adaptive Neuro-Fuzzy Inference System (ANFIS) based modelling of incipient steam generator tube rupture diagnosis,” *Annals of Nuclear Energy*, vol. 157, 2021, doi: 10.1016/j.anucene.2021.108262.

[35] Y. Jin, W. Cao, M. Wu, Y. Yuan, and Y. Shi, “Simplification of ANFIS based on importance-confidence-similarity measures,” *Fuzzy Sets and Systems*, vol. 481, 2024, doi: 10.1016/j.fss.2024.108887.

[36] S. O. Sada and S. C. Ikpeseni, “Evaluation of ANN and ANFIS modeling ability in the prediction of AISI 1050 steel machining performance,” *Heliyon*, vol. 7, no. 2, 2021, doi: 10.1016/j.heliyon.2021.e06136.

[37] S. Kanwal and S. Jiriwibhakorn, “Advanced Fault Detection, Classification, and Localization in Transmission Lines: A Comparative Study of ANFIS, Neural Networks, and Hybrid Methods,” *IEEE Access*, vol. 12, pp. 49017-49033, 2024, doi: 10.1109/ACCESS.2024.3384761.

[38] S. Alvarez-Rodríguez and F. G. Peña-Lecona, “Artificial Neural Networks with Machine Learning Design for a Polyphasic Encoder,” *Sensors*, vol. 23, no. 20, pp. 1-24, 2023, doi: 10.3390/s23208347.

[39] A. Singh, S. Kushwaha, M. Alarfaj, and M. Singh, “Comprehensive Overview of Backpropagation Algorithm for Digital Image Denoising,” *Electronics*, vol. 11, no. 10, pp. 1-15, 2022, doi: 10.3390/electronics11101590.

[40] M. Almada, “Governing the Black Box of Artificial Intelligence,” *SSRN Electronic Journal*, 2023, doi: 10.2139/ssrn.4587609.

[41] G. Zhang, S. S. Band, S. Ardabili, K.-W. Chau, and A. Mosavi, “Integration of neural network and fuzzy logic decision making compared with bilayered neural network in the simulation of daily dew point temperature,” *Engineering Applications of Computational Fluid Mechanics*, vol. 16, no. 1, pp. 713–723, 2022, doi: 10.1080/19942060.2022.2043187.

[42] H. Moayedi *et al.*, “Novel hybrids of adaptive neuro-fuzzy inference system (ANFIS) with several metaheuristic algorithms for spatial susceptibility assessment of seismic-induced landslide,” *Geomatics, Natural Hazards and Risk*, vol. 10, no. 1, pp. 1879–1911, Jan. 01, 2019, doi: 10.1080/19475705.2019.1650126.

[43] S. E. Hashemi, F. Gholian-Jouybari, and M. Hajiaghaei-Keshteli, “A fuzzy C-means algorithm for optimizing data clustering,” *Expert Systems with Applications*, vol. 227, 2023, doi: 10.1016/j.eswa.2023.120377.

[44] J. C. Bezdek, R. Ehrlich, and W. Full, “FCM: The fuzzy c-means clustering algorithm,” *Computers & Geosciences*, vol. 10, no. 2–3, pp. 191–203, 1984, doi: 10.1016/0098-3004(84)90020-7.

BIOGRAPHIES OF AUTHORS

Hizir Sofyan is a professor in applied statistics at Universitas Syiah Kuala, Banda Aceh, Indonesia. Moreover, he serves as the chairman of the International Statistical Institute (ISI) South East Asia Outreach Committee. In 1992, he graduated with a B.Sc. degree in statistics from Indonesia's Sepuluh Nopember Institute of Technology in Surabaya. In 2003, he obtained his Ph.D. degree from the Institute of Statistics and Econometrics, School of Business and Economics, Humboldt-Universitat zu Berlin, Germany. His supervisors were Prof. Dr. Wolfgang Haerdle from Statistics Department, Humboldt Universitaet zu Berlin, Germany, and Prof. Oliver Guenther, Ph.D. (President of Potsdam University, Germany). He is currently the Head of Graduate School, Universitas Syiah Kuala, Banda Aceh, Indonesia. His current research interests include computational statistics, soft computing, fuzzy modelling, artificial-intelligence, advanced data-analysis, statistical geo-spatial, and deep learning. He can be contacted at email: hizir@usk.ac.id.

Nidya Tatsara received a B.Sc. degree in statistics from Universitas Syiah Kuala, Indonesia. She was a laboratory assistant of Bio statistics Lab and a former member of Computational and Applied Statistics Research Group (CASRG), the Department of Statistics, Universitas Syiah Kuala. She obtained a Digital Talent Scholarship of Fresh Graduate Academy for Data Science in 2020. Additionally, she was a data scientist at PT. eBdesk Teknologi in 2021. Currently, she is a data engineer of Bank Rakyat Indonesia (BRI). Her research interests include computational statistics, deep learning, applied statistics, and fuzzy logic. She is an experienced analyst with a demonstrated history of working using statistics tools. She has also programming skills such as Python and R programming languages. She can be contacted at email: nidya.tatsara@gmail.com.

Yolanda received a B.Sc. degree in statistics from Universitas Syiah Kuala, Indonesia. He is a laboratory assistant of Computational Statistics Laboratory and a member of Computational and Applied Statistics Research Group (CASRG), the Department of Statistics, Universitas Syiah Kuala. He is currently a M.Sc. student in disaster science at Universitas Syiah Kuala, under supervision of Prof. Dr. Hizir Sofyan. His thesis focuses on disaster statistics to develop sendai framework for disaster risk reduction (SFDRR) and suistainable development goals (SDGs) on Disaster Data in Aceh. In addition, he has an experience in big data analysis using statistics tools, including Python and R programming languages. He can be contacted at email: mr.yolanda24@gmail.com.

Tarmizi Usman received a B.Sc. degree in mathematics from Indonesia's Sepuluh Nopember Institute of Technology in Surabaya. He obtained his M.Sc. degree in mathematics science from Tohoku University, Japan in 1997. Later on, he received a Ph.D. degree in applied mathematics at system engineering from Tokushima University, Japan in 2000. He is currently an associate professor in applied mathematics with the Department of Mathematics at Universitas Syiah Kuala, Banda Aceh, Indonesia. His current research interests include applied mathematics in engineering applications, computational mathematics, mathematical modeling, advanced data-analysis, and optimization. He can be contacted at email: tarmiziusman@usk.ac.id.

Irvanizam is an associate professor in applied informatics or computer science at the Department of Informatics, Universitas Syiah Kuala, Banda Aceh, Indonesia. He obtained a B.Sc. degree in mathematics science majoring in informatics from Universitas Syiah Kuala, Banda Aceh, Indonesia, in 2002, an M.Sc. degree in computer science from Free University of Bozen-Bolzano, Bolzano, Italy, in 2010, and a Ph.D. degree in computer science from Universitas Sumatera Utara, Medan, Indonesia, in 2025. He is also the author of publications in prestigious international journals, many of which are indexed by Scopus and Web of Science, including Expert Systems with Applications, IEEE Access, Journal of Supercomputing, Applied Computational Intelligence and Soft Computing, Axioms, and Advances in Fuzzy Systems. He regularly serves as a reviewer for some reputable science and engineering journals such as Applied Soft Computing, Expert Systems with Applications, Artificial Intelligence Review, Soft Computing, International Journal of Computing and Digital Systems (IJCDS), Computers and Electrical Engineering, Alexandria Engineering Journal, etc. His research focuses on neutrosophic and fuzzy sets, decision support systems, soft computing, group decision-making, and artificial intelligence. He can be contacted at email: irvanizam.zamanhuri@usk.ac.id.