The 7th International Conference on Engineering Mechanics and Automation (ICEMA - 2023)
Hanoi, November 11, 2023

DOI: 10.15625/vap.2023.0135

Applying artificial neural network model in calculating rainfall rate from radar reflectivity at Thua Thien Hue, Dong Van meteorological station

Chinh Kien Nguyen
Hydrodynamic and Disaster Mitigation
Institute of Mechanics - VAST
Hanoi, Vietnam
nckien@imech.vast.vn

Hang Nguyen Thi
Hydrodynamic and Disaster Mitigation
Institute of Mechanics - VAST
Hanoi, Vietnam
hangnguyen.imech@gmail.com

Thanh Huong Duong Thi
Hydrodynamic and Disaster Mitigation
Institute of Mechanics - VAST
Hanoi, Vietnam
dthuong.imech@gmail.com

Thanh Hang Do
National Centre for Hydrometeorological Forecasting - NCHMF
Hanoi, Vietnam
dothanhhang234@gmail.com

***Abstract* — In Vietnam and the world, the rainfall rate estimated from radar reflectivity has been used to monitor and forecast heavy rain for many years because of its advantages in spatial distribution compared to the observed rainfall at measuring stations. However, rainfall estimates from radar reflectivity often use simple empirical formulas, which do not accurately reflect the nonlinear relationship between rainfall rate and radar reflectivity. In this study, the authors experimented with using an artificial neural network to estimate rainfall rate from radar reflectivity for Thua Thien Hue and Dong Van meteorological station. The results are compared with the use of empirical formulas, showing the good applicability of the artificial neural network model in calculating rainfall rate from radar reflectivity.**

Keywords— Artificial neural network, radar reflectivity, rainfall rate.

I. INTRODUCTION

Rainfall data is an important input to hydrological models for streamflow forecasting. However, the most common source of rain data currently used is from ground monitoring stations - which are highly reliable, but the current distribution of measuring stations is still sparse, scattered and uneven across regions, leading to does not meet the detailed spatial requirements of some hydrological models.

There have been many studies in Vietnam and the world using weather radar to calculate rainfall because of the advantages of space (uniformly distributed over a wide area with a radius of hundreds of kilometers) and time (data updated continuously - about every 10 minutes).

Estimating rainfall from radar reflectivity is mainly done based on a number of empirical formulas, mostly using the "Marshall-Palmer form" formula $Z=a.R^b$, the results obtained are is still limited by the simplicity of the formula as well as the difficulty in determining the empirical coefficient for each region.

There are many studies around the world on the use of radar image data sources in rain forecasting: Studies that build a quantitative relationship between rainfall and reflectivity with spatial distribution in large area and high time frequency by the authors Krajewski and Smith, 2002 [4]; Smith and J.A., 1991 [6]. In 2002, Marco Borga [5] pointed out that the accuracy of rainfall rate from radar reflectivity is also affected by terrain and vegetation because they can cover electromagnetic waves emitted by radar, so before using radar rain data for rainfall-runoff simulation, leading to the need to calibrate radar rain data to minimize errors; In 2021, Seng Wang et al [7] used rainfall estimates from radar and satellite for integrated prediction of flood and landslide risks.

In Vietnam, radar image data has been researched by scientists to effectively use image data sources to forecast heavy rainfall and storms. Nguyen Huong Dien 2015 [1] based on the response data of weather radar located at 4 rain measuring stations, has proposed an empirical formula to estimate rainfall for the Southeast region.

In recent years, the Artificial Neural Network (ANN) model has been effectively applied by the authors in many hydrological and hydraulic problems, etc. for a number of basins in Vietnam. In this study, the authors tested using an artificial neural network model to calculate rainfall from radar reflectivity at Thua Thien Hue meteorological station in Thua Thien Hue province and Dong Van meteorological station in Ha Giang province of Vietnam. This is a new research direction in Vietnam.

II. METHOD AND DATA

A. *Research methods*

Artificial Neural Network (ANN), referred to as neural network, is an information processing model that simulates the information processing method of biological neural systems. It is made up of a large number of elements (called processing elements or neurons) connected together through links (called link weights) that work as a unified whole to solve a problem specific problem [3].

A neuron is an information processing unit and is the basic component of a neural network. The structure of a neuron is depicted in "Fig. 1".

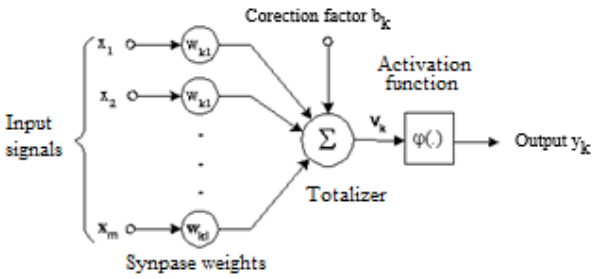

Fig. 1. The structure of an artificial neuron

In there:

x_i (i = 1, m): The set of input signals of the neuron;
w_{jk}: Set of links, each link between the j^{th} input signal and k neuron is represented by a weight, randomly initialized at the time of network initialization and continuously updated during the learning process;
$\sum$: The totalizer is used to calculate the sum of the product of the inputs with their associated weights.
b_k: Bias, included as a component of the transfer function;
φ(.): Transfer function is also called Activation function. Input data of this function are the result of the totalizer and the given bias;
y_k: The output signal of a neuron, with each neuron having at most one output.

Although each individual neuron can perform certain information processing functions, the computational power of an artificial neural network is mainly derived from the combination of neurons in a unified architecture. "Fig. 2". shows the most widely used multi-layer feedforward network model among current ANN models. A general MLP network is a network with n (n ≥ 2) layers, which includes an output layer (nth layer) and (n-1) hidden layers.

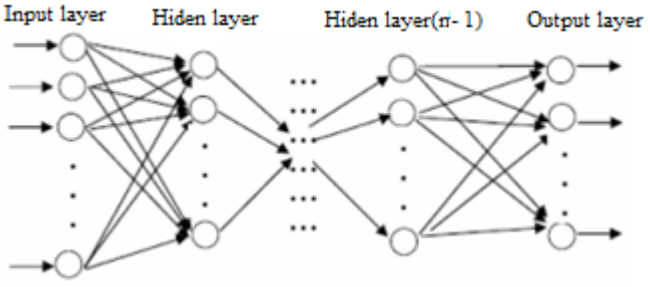

Fig. 2. Multi-layer feedforward network

The advantage of an artificial neural network is that it allows building a computational model with a very high ability to learn from data. The neural network training process is based on the regression error between the calculated value and the measured value. The training algorithm will adjust the connection weights of the neural network to minimize the regression error on the training samples. After the network is successfully trained, the weight matrix will be updated for use in the forecasting process.

B. *Data collection*

Thua Thien Hue is one of the provinces located in the east of the Truong Son Mountain range of the Central Coast, so the rain regime and rainfall here are both influenced by the Southeast Asian monsoon circulation mechanism and affected by strong influence of geographical location (topography) as well as terrain conditions. In 2022, Thua Thien Hue had many heavy rains, typically on October 14 - 15 and on December 2 - 3, that caused serious floods [10].

Ha Giang is a province in the Northeast region of Vietnam. Located in the tropical monsoon region and a high mountainous region, the rainfall regime in Ha Giang is quite diverse. The province's average annual rainfall is about 2,300 - 2,400 mm, this is one of the regions with the highest rainfall in Vietnam. In 2022, from May 27 to 29, heavy rain caused floods and landslides in some places, causing heavy damage in Dong Van district [9].

Therefore, the authors chose to use the time period of 2022 to calibrate and verify the method of calculating rainfall from radar reflectivity using the ANN model for both measuring stations.

Radar reflectivity data was extracted at points neighboring the location of Thua Thien Hue meteorological station and Dong Van meteorological station with a radius of about 8 km, taken from the free website of the High Air Meteorological Station - Ministry of Natural Resources and Environment [8].

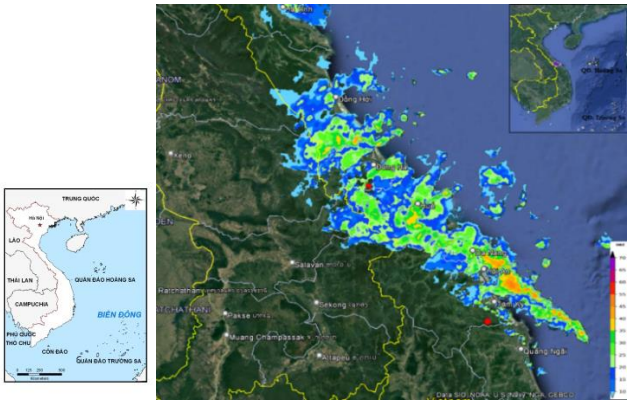

Fig. 3. Radar reflectivity of Thua Thien Hue area at 3:30 p.m. October 15, 2022

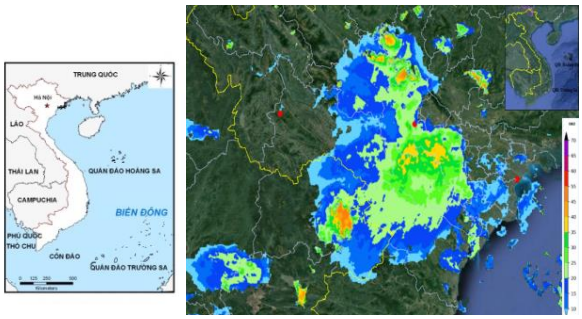

Fig. 4. Radar reflectivity of Ha Giang area at 12:30 p.m. May 28, 2022

Collecting rain data at Thua Thien Hue meteorological station is from January to December 2022 to calibrate and test the model. For Dong Van meteorological station, due to the time of data collection following heavy rains, the data series used to calibrate and test the model is May and June 2022.

The time used to calibrate and validate the model for each meteorological station is presented in "Table I".

TABLE I. CALIBRATION AND VERIFICATION OPTIONS

Station	Time	
	Calibration	*Verification*
Hue	7h 01/01/2022 - 19h 30/11/2022	7h 01/12/2022 - 19h 21/12/2022
Dong Van	7h 10/05/2022 - 19h 05/06/2022	7h 06/06/2022 - 19h 18/06/2022

After collecting rain data at meteorological stations, it is necessary to standardize the data and eliminate extreme values (negative data, objective errors,...) and synchronize time with radar reflectivity data to as input to the calculation model.

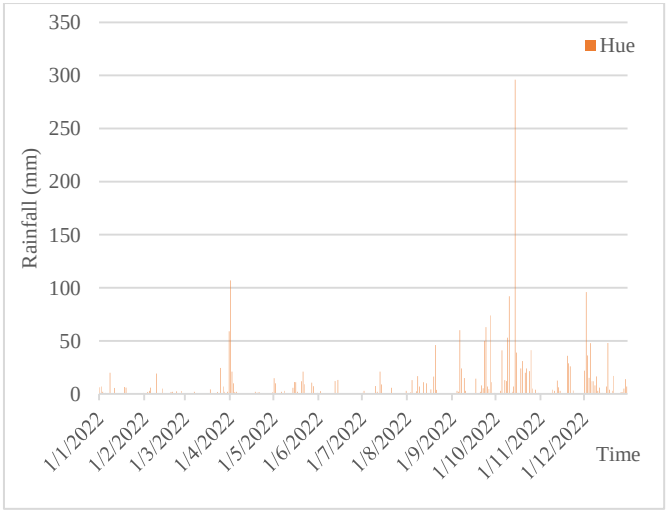

Fig. 5. Rainfall measured at Thua Thien Hue station.

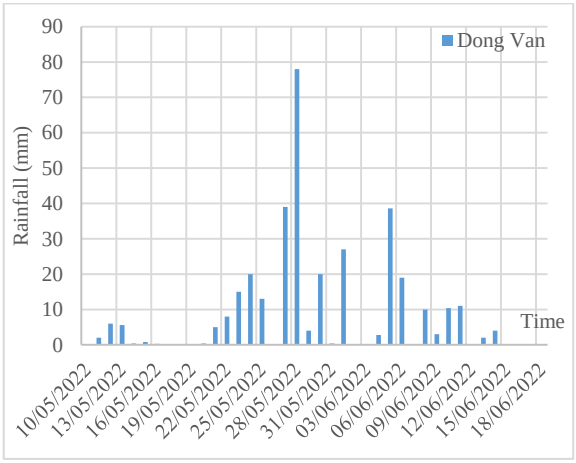

Fig. 6. Rainfall measured at Dong Van station

III. RESULTS AND DISCUSSIONS

Using the ANN model to calibrate and verify the Hue and Dong Van meteorological stations, in addition to comparing with measured values, the authors will compare with the results calculated from the empirical formula, which has the following Marshall-Palmer form:

$$Z = a.R^b \Rightarrow R = c.10^{d.Z'} \quad (*)$$

In there:

Z is radar reflectivity, $Z = 10^{Z'/10}$ (Z' in dbZ unit);
R is rainfall rate;
a, b, c, d are the experimental coefficients.

From the collected radar reflectivity and rainfall measured, the authors detected the experimental coefficients c and d for Thua Thien Hue and Dong Van meteorological stations using the Shuffled Complex Evolution method [2]. The coefficient detection results for Hue station are c = 4,837 and d = 0,055 and Dong Van station are c = 4,457 and d = 0,073.

A. *Calibration*

Use input data including radar reflectivity data and rainfall measured corresponding to the calibration period to calculate rainfall according to the empirical formula as well as to train the artificial neural network model. The results of rainfall calculated by artificial neural network (ANN) and Empirical Formula (EF) of Hue station are shown along with rainfall measured (Obs) in "Fig. 7".

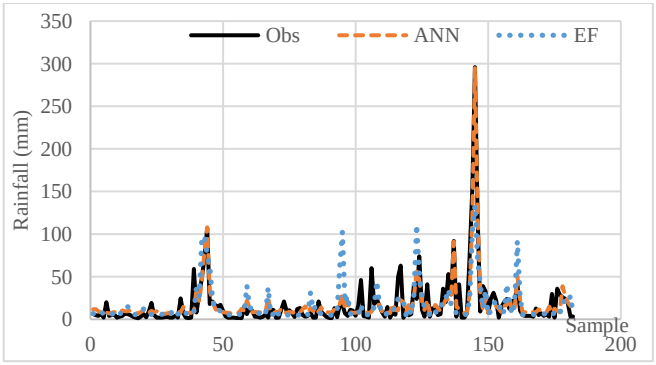

Fig. 7. Rainfall at Thua Thien Hue station is calculated using the EF and ANN model of the calibration plan

Some heavy rain values calculated from the ANN model are equivalent to the measured values, such as the largest rainfall in the measured values in Hue is 296 mm while the results calculated by ANN reach 294,7 mm. The error is 1,2 mm (equivalent to 0,4%), the calculation result from the empirical formula is 134 mm, the difference is 162 mm (equivalent to 55%) at 13:00 on October 14, 2022.

Using the Nash Sutcliffe Efficiency (NSE) index, the EF method at Hue station has NSE = 0,43, for the ANN method it has NSE = 0,82.

Similar to Thua Thien Hue meteorological station, in the calibration plan, the rainfall calculated from the empirical formula and ANN model along with actual measured data at Dong Van station is shown in "Fig. 8".

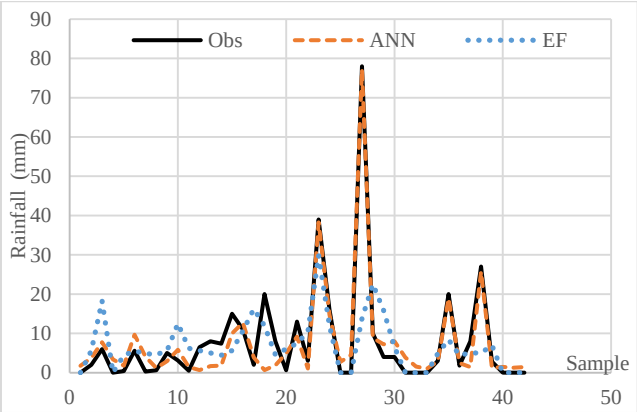

Fig. 8. Rainfall at Dong Van station is calculated using the empirical formula and ANN model of the calibration plan

With the calibrated data at Dong Van station, the data is collected according to the flood, the number of samples to calibrate is less than at Hue station, leading to better calibrating results. The largest actual measured rainfall in the calibration value series at Dong Van station is 78 mm at 7:00 pm on May 27, 2022, the rainfall calculated from EF only reaches 13,7 mm with an error of up to 64,3 mm (equivalent to 82,4%) while the rainfall calculated in the ANN model is 77,2 mm, the error is only 0,81 mm (equivalent to 1,03%).

The effectiveness index for the calibration plan at Dong Van station with the EF method is NSE = 0,23 and with the ANN model, NSE = 0,9.

It is noticed that, for the calibration plan, the rainfall calculated by the ANN model gives the coefficient NSE = 0,82 at Hue station and NSE = 0,9 at Dong Van station at a very good level (0,75 ≤ NSE ≤ 1,0), showing enough reliability to use the trained weight matrix for verification step.

B. Verification

Use the experimental coefficients c and d that were found in the calibration plan to estimate the rainfall according to the empirical formula method for the verification step. Similarly, use the trained weight matrix from the ANN model calibration step to calculate for the verification step. The results of rainfall at Thua Thien Hue station in the verification plan are shown in "Fig. 9".

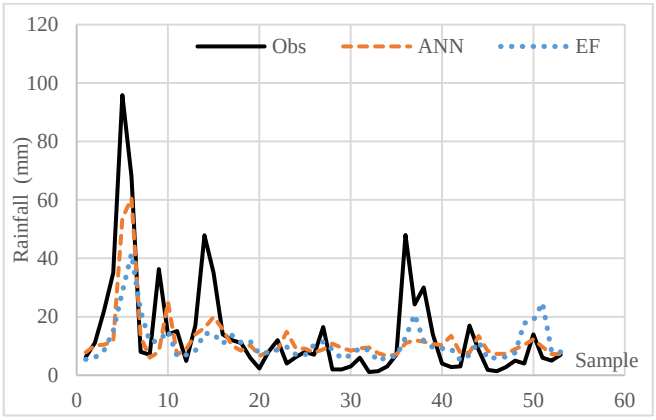

Fig. 9. Rainfall at Thua Thien Hue station is calculated using the empirical formula and ANN model of the verification plan

In the verification data series, at 13:00 on December 2, 2022, the highest rainfall at Hue station was 68 mm, the result of the ANN method was 60,6 mm, the error was 7,4 mm (equivalent to 10,8%), while the result of empirical formula method is 42 mm (error is 38,8%).

The NSE index of the verification plan with the method of estimating rainfall using the EF at Thua Thien Hue station is NSE = 0,33 and NSE = 0,54 with the ANN method.

Similar to the Thua Thien Hue meteorological station, the rainfall results calculated from the EF and the ANN model in the verification plan are shown along with the actual measured data at Dong Van station in "Fig. 10".

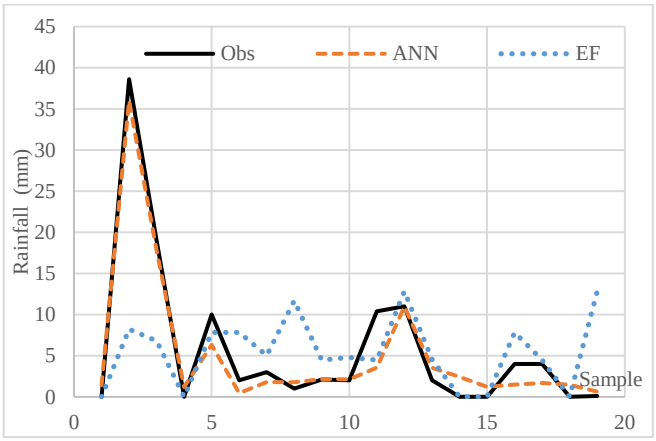

Fig. 10. Rainfall at Dong Van station is calculated using the empirical formula and ANN model of the verification plan

The largest measured rainfall in the series of data tested at Dong Van station was 38,6mm at 13:00 on June 5, 2022. The rainfall calculated from the empirical formula only reached 8,21 mm with an error of up to 30,39 mm (equivalent to 78,74%) while the rainfall calculated in the ANN model is 36 mm, the error is only 2,6 mm (equivalent to 6,74%).

The NSE index of the verification plan with the EF method at Dong Van station is NSE = 0,09 and achieves a poor grade (NSE < 0,5); in the ANN method NSE = 0,75 has a good level of reliability (0,75 < NSE ≤ 1,0).

C. Summary of results

Some error statistics: average error ME, average absolute error MAE, square error RMSE and efficiency coefficient NSE to evaluate the accuracy between the actual measured value and the calculated results of the empirical formula method and the ANN method. Each method has two results of calibration (CAL) and verification (VER) to compare at stations.

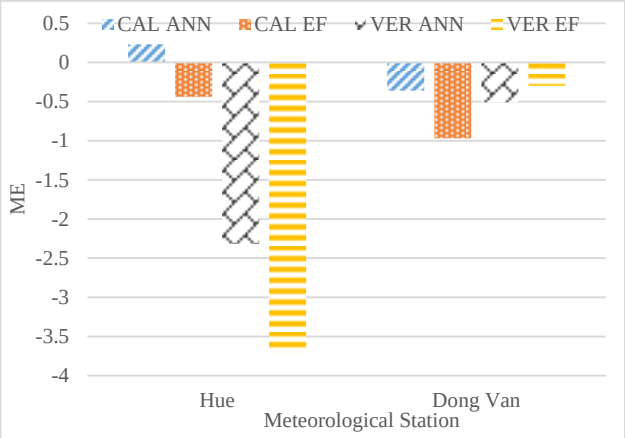

Fig. 11. ME error of calculation options

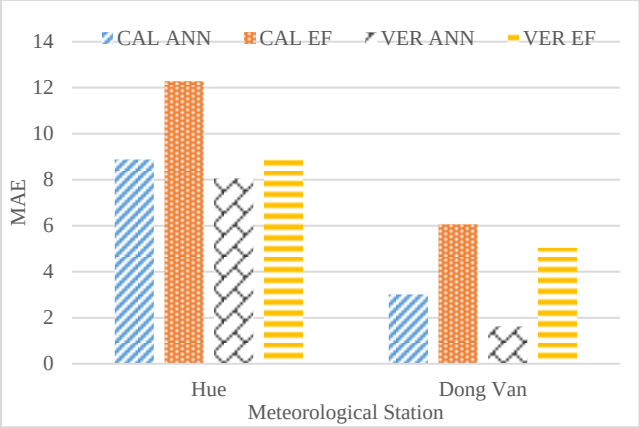

Fig. 12. MAE error of calculation options

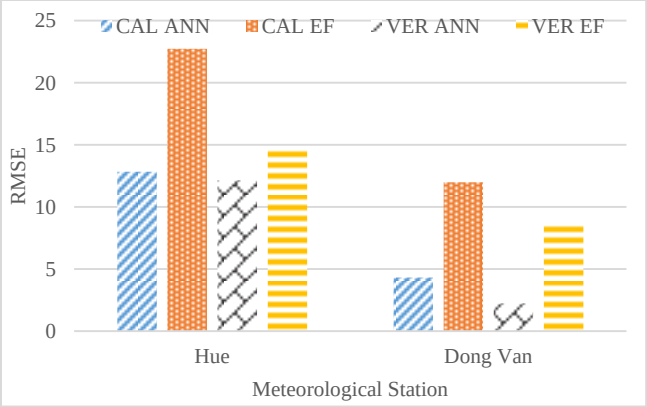

Fig. 13. RMSE error of calculation options

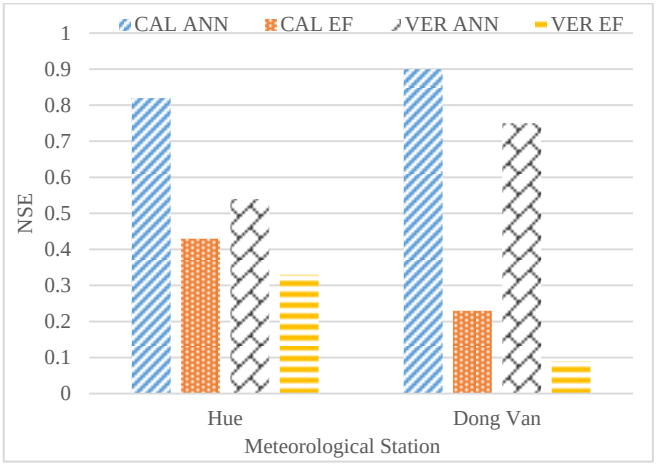

Fig. 14. NSE index of calculation options

IV. CONCLUSION

Calculating rainfall from radar reflectivity at Thua Thien Hue and Dong Van meteorological stations has been tested with the empirical formula method and artificial neural network model. The results show the better ability of the artificial neural network model compared to the method using empirical formulas. Because the radar reflectivity data is re-analyzed from radar images collected according to the color value table, the accuracy is not high (it is not the original data), leading to poor calculation results. In addition, using only short data series of 1 year or a few months to train the model may not accurately reflect the relationship between rainfall and radar reflectivity, it is necessary to continue collecting past data. as well as update current and future data to supplement the analysis and training process.

In summary, the study shows that the applicability of the ANN method is initially more promising than the method of estimating rainfall from radar reflectivity according to empirical formulas, and the results can be used for reference in current rain forecasting operations.

ACKNOWLEDGMENT

This work is funded by the research project from the Institute of Mechanics, Vietnam Academy of Science and Technology (VAST), code CSCL03.01/22-23.

REFERENCES

[1] Nguyen Huong Dien, “Empirical formula to calculate rain intensity from radar response for the Southeast region”, National University’s Science Magarine, Volume 31, No. 3S, 2015, pp. 66 -70.

[2] Nguyen Thi Hang, Nguyen Chinh Kien, Duong Thi Thanh Huong, “Applying optimal parameter methods for the distributed hydrological model IMECHTV2”, Jounal of Science and Natural Resouces and Environment , ISSN: 0866-7608, No. 42, 2022, pp. 113-123.

[3] Nguyen Chinh Kien, “Testing the application of artificial neural networks in hydrological and hydraulic forecasting”, The 10th National Conference on Mechanics, Hanoi, 12/2017.

[4] Krajewski, W.F. and Smith, J.A, “Radar Hydrology: Rainfall Estimation”, Advances in Water Resources, 25, 2002, pp. 1387-1394.

[5] Marco Borga, “Accuracy of radar rainfall estimates for streamflow simulation”, Journal of Hydrology, Volume 267, Issues 12, 1 October 2002, pp. 26-39.

[6] Smith, J.A. and Krajewski, W.F, "Estimation of Mean Field Bias of Radar Rainfall Estimates", Journal of Applied Meteorology, 30, 1991, pp. 397-412.

[7] Sheng Wang, "Exploring the utility of radar and satellite-sensed precipitation and their dynamic bias correction for integrated prediction of flood and landslide hazards", Journal of Hydrology, Volume 603, Part B, December 2021.

[8] Website of the High Air Meteorological Station - Ministry of Natural Resources and environment: http://amo.gov.vn/radar/

[9] Ha Giang portal: https://hagiang.gov.vn/

[10] Hue portal: https://huecity.gov.vn/