Application of Feed Forward Backpropagation Neural Network in Monthly Rainfall Prediction

Putu Doddy Heka Ardana[1], I Gusti Made Sudika[2], Ni Kadek Astariani[3], Gede Sumarda[4]

[1]Ngurah Rai University, Indonesia, doddyhekaardana@unr.ac.id
[2]Ngurah Rai University, Indonesia, made.sudika@unr.ac.id
[3]Ngurah Rai University, Indonesia, kadek.astariani@unr.ac.id
[4]Ngurah Rai University, Indonesia, gede.sumarda@unr.ac.id

ABSTRACT

Rainfall is one of the most critical parameters in a hydrological model. A few models have been created to investigate and predict the rainfall conjecture. in recent years, soft computing models like Artificial Neural Network (ANN) have been widely used to model a complex hydrological processes. In this paper an attempt has been made to find an alternative feed forward backpropagation neural network (FFBNN) architecture for rainfall prediction. The FFBNN with 12-10-5-1 architecture have been attempted and demonstrated with a case study on Tukad Mati watershed. The developed model on rainfall prediction pattern have been calibrated and validated with appropriate statistical methods. For model performance criteria used mean square error (MSE) and coefisient correlation (R) as an indicator. It was found that the FFBNN are capable in modelling monthly rainfall prediction in the Tukad Mati watershed with 0,011913 for average of MSE and 0,85769 for average of R testing value from three rain gauge station. The FFBNN approach could provide a very effective to solve problems in monthly rainfall prediction.

Key words : Feed forward backpropagation neural network, Prediction, Rainfall, Input, Output.

1. INTRODUCTION

Rainfall is a form of the precipitation process. Rainfall is the amount of water that falls on a flat surface during a certain period measured in units of height (mm) above the horizontal surface if there is no evaporation, runoff and infiltration (1). Rainfall intensity is the magnitude of the amount of rainfall that falls which is expressed in high rainfall or rain volume per unit of time. The amount of rainfall intensity varies, depending on the duration of rainfall and the frequency of occurrence (2). The intensity of high rainfall is one of the factors causing the disaster, especially floods because it can increase the capacity of the river so that the river cannot accommodate large water capacity. As a result, river runoff overflowed the surrounding area. Like a fairly large flood on January 2018 in the area of Dewi Sri Kuta - Badung which included the middle and downstream part of the *Tukad Mati* watershed (3) and (4).

The amount of rainfall that occurs cannot be determined with certainty but can be predicted or estimated. By using historical data on the magnitude of past rainfall data, it can be predicted how much rainfall will occur in the future. In addition, detailed rainfall forecasts up to the regions are still less. Research on rainfall forecasts for a region that is very important in various sectors, for example in water resources management and agriculture as well as in disaster mitigation, in particular, is flooding. There are three ways that are usually used to analyze the amount of rainfall in a region, namely the average arithmetic method, the polygon Thiessen and isohyet which have deficiencies and the advantages of each.

The ability of artificial neural network (ANN) in forecasting, especially rain forecasting has been developed by researchers as an effective and efficient predictor. The research conducted by (5) provides very promising results for ANN in rain predicting. ANN has been used in forecasting monthly rainfall events in Udaipur City, Rajasthan, India using meteorological variables and testing their performance with multiple linear regression (6). Rainfall prediction using data mining with the Bayesian approach gives quite accurate results when using large training data (7). The use of data mining with the application of artificial neural network is also used in rainfall forecasting in Queensland, Australia where the RMSE value of this model produces similar values to observational data for the next three months' rain data (8). Artificial neural network models can also be combined with other models in the analysis process. As in rainfall forecasting by using a combination of perceptron and Wavelet multi-layer models called the hybrid model provides good results in rainfall predicting for the next day in the Turkish state (9). Accurate results were also contributed by the artificial neural network model in predicting rainfall compared to the Box-Jenkins model (10). A scientific model of precipitation was created by researcher which was used as the original precipitation and dissected with the artificial neural network model. The neural

network fit for learning the unpredictable relationship depicts the space-time advancement of precipitation, for example, that natural intricate precipitation recreation shows (11). Besides being successful in predicting rainfall, the artificial neural network model has also been widely used in modelling the rainfall-runoff transformation (12–14); prediction of sedimentation and sediment transport in rivers (15,16), artificial neural networks in water quality analysis (17), development of water level forecasting (18), drought forecasting (19) etc. as stated by (20) that ANN is very good for hydrological modelling.

The artificial neural networks are contrasting to classical statistical techniques, the advantages of artificial neural networks as a forecasting method are their ability to recognize certain patterns using learning and training algorithms as the human brain works (21). Artificial neural networks provide various kinds of network architecture and training. Network architecture and training used can be chosen so that artificial neural networks can study and analyze past data patterns more precisely so that more accurate results are obtained (22).

Related to has been explained above, the aims of this study are to predicting the region monthly rainfall from three rain stations which have an influence on the Tukad Mati watershed using the Artificial Neural Network model with feed forward backpropagation neural network (FFBNN). The application of ANN was based on consideration of the characteristics of the watershed that is different from previous studies.

2. ULITERATURE REVIEW

2.1 Study Area and Rainfall Data

In the present study, object for investigation covering three rain stations which have an influence on the Tukad Mati watershed. The Aseman rain station, the Ngurah Rai rain station, and the Sanglah rain station are contibuting to Tukad Mati river watershed. Tukad Mati watershed has a drainage area of 38,42 km2 and length of the river is 18,52 km which cover the area of Denpasar city and Badung regency in Bali province. The rainfall monthly data from three stations stations were used for model investigation. The data contains information for a period of ten years (2009 to 2018). The ANNs model was trained using the monthly rainfall and the next month rainfall as the target. The database was collected by the BMKG Region III, Departement of Public Work in Bali Province, and River Basin Department of Bali Penida (BWS – Bali Penida). The area of Tukad Mati watershed is depicted in Figure 1 and Figure 2.

Figure 1 : The location of study area

2.2 ANN For Rainfall Forecasting

Artificial Neural Network (ANN) is a concept of processing data inspired by the natural neurons system, like the process of information in the human brain. The basic component of this concept is constructed from an information processing system consisting of larger number of component relating (neurons), working in unison for solve certain problems. How ANN works like the way humans work, that is learning through examples. The feed-forward backpropagation neural networks (FFBNN) is a form of supervised training. The network must be provided with both sample inputs and outputs when supervised training method is used. Backpropagation refers to a varied group of ANN, whose design include of various interrelated layers (23). The back-propagation algorithm is a supervised iterative training method for multilayer feed-forward nets with a differentiable nonlinear function. FFBNN often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range 0 to 1, then the output layer should use a sigmoid transfer function (such as logsig). Every neuron is completely associated with neurons in the following layer (Figure 3). All contribution to a neuron in a specific layer is from the procedure layer and these undirectional qualities are known as weights and biases (24).
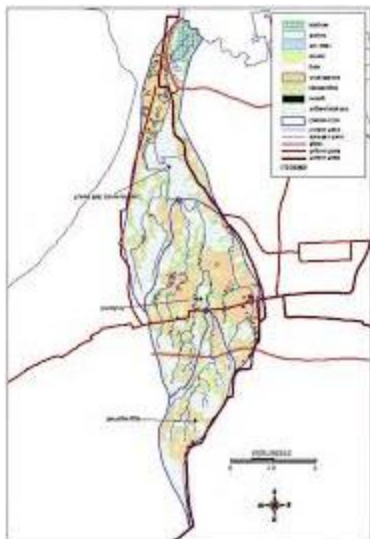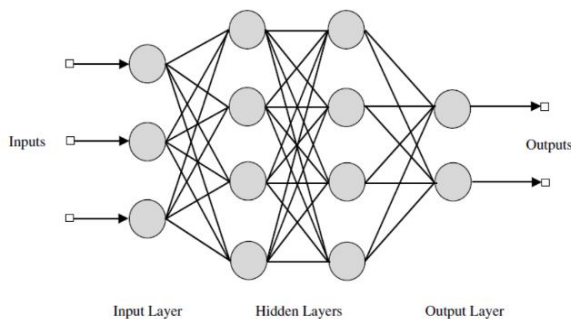
Figure 2 : Tukad Mati watershed

The equation of the calculated sigmoid is given by

$$y = f(x) = \frac{1}{1 + e^{-\alpha x}} \qquad (1)$$

The identity (purelin) function will be use to linked hidden layer and output layer. The equation of the logistic sigmoid is specified by

$$f(x) = x \qquad (2)$$

The architecture of feed-forward backpropagation neural network is depicted in Figure 3.

Figure 3 : Architecture of FFBNN

Forecasting is a natural that exactness is significant. The input paramaters for rainfall determining model are diverse kinds of information need and distinctive sorts of techniques, and should be dealt with appropriately. Measurable strategies are typically connected liniear data while Artificial Neural Network (ANN) are related with non linear data (25).

With ANN, rainfall forecasting is expected to be done quickly. To construct an ANN model for rainfall forecasting, determination of area for input data and parameters is important. The input data is to be taken from a particular region on which the model is prepared and tried with the goal that the model can create exact outcomes. The quantity input data given to show additionally improves exactness of the model by giving the outcomes with a high level of comparability among anticipated and real ouput data. The accessible data might be loud therefore, the data should be cleaned. Correspondingly, it must be standardized in light of the fact that, all of the parameters are of various units and standardization will support the input and output parameters to connect with one another (26). The data should be devided become to training and testing data in appropriate extent with the goal that the outcomes can be anticipated, tried and approved legitimately. Structure of the ANN model likewise greatly affects age of exact outcomes. The multilayer ANN helps in foreseeing nonlinear information all the more effectively. The initiation capacity will be diverse for various layers of ANN according to require.

3. METHOD

3.1 Analysis of Rainfall

In hydrological analysis process, precipitation data must be tested for consistency with the aim of the results are valid. Rescaled Adjusted Partial Sum (RAPS) were used in this study. The advantage of this method is to avoid the impacts of different data units and random errors on analysis. Time series data of rainfall during 10 years (2009-2018) have been analyzed for their consistency data trough standard deviation and cumulative deviation test. Formula of RAPS is expressed, as follows:

$$RAPS_k = \sum_{t=1}^{k} \frac{y_t - \overline{y}}{s_y} \qquad (3)$$

3.2 FFBNN Model Application

The training and testing data set collected during years 2009 to 2018 were selected to evaluate the performance of the neural network model for Tukad Mati watershed in Bali. Since the original recorded data contain area rainfall which prior hydrological analysis (Rescaled Adjusted Partial Sum method for rainfall data consistency). In this research, the rainfall data devided in two sets: the seven years of data (2009-2015) are used for training process and the three years of data (2016-2018) are used for testing process, which will be done for Tukad Mati watershed.

Normalized process will be done for the input and target rate to scale the unit interval in to 0 and 1. Following (27), the normalized value shown from equation below

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}(BA - BB) + (BB) \qquad (4)$$

In this paper, we used 12-10-5-1 architecture scenario (12 input layers; 10 first hidden; 5 second hidden layers, and 1 output layer). The coefficient of correlation (r) and mean square error (MSE) value will be used for model performance criteria In this study, MATLAB software version R2018b will be practiced to analysis process.

4. RESULTS AND DISCUSSION

In hydrology process, consistency data analysis is the first step to do. Rescaled Adjusted Partial Sum (RAPS) used to

calculate the consistency of rain data. The data of three rainfall station were analyzed with this method. Three rainfall gauge are Aseman, Ngurah Rai, and Sanglah with 10 years data (2009-2018). The result of RAPS analysis are presented in Table 1 respectively.

Table 1: The RAPS analysis result

Year	Q/√n			R/√n		
	ASM	NGR	SGL	ASM	NGR	SGL
10	0.75	0.71	0.61	1.19	1.14	0.92

FFBNN with 12-10-5- architecture were used for analysis. Monthly rainfall data used as input to ANN models. The length of the data used for calibration is 7 years (84 months) and for validation is 3 years (36 months). The model structures with least mean square error (MSE) are considered as the best result[28][29][30]. The Goodness-of-fit statistic of selected ANN models during calibration and validation is presented in Table 2.

Table 2: Performance of FFBNN (12-10-5-1)

Rainfall Station	Performance of ANN Architecture (12-10-5-1)		
	MSE	R - training	R - testing
Aseman	0.016163	0.90499	0.92309
Ngurah Rai	0.01503	0.82932	0.81556
Sanglah	0.0045471	0.90149	0.83441

The performance of the developed ANN models in terms of MSE is depicted in Figure 4, Figure 5 and Figure 6 for calibration and validation respectively.
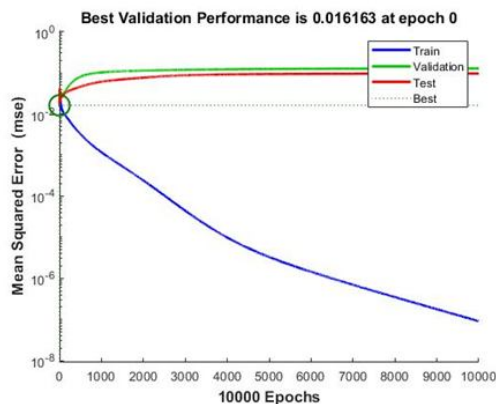

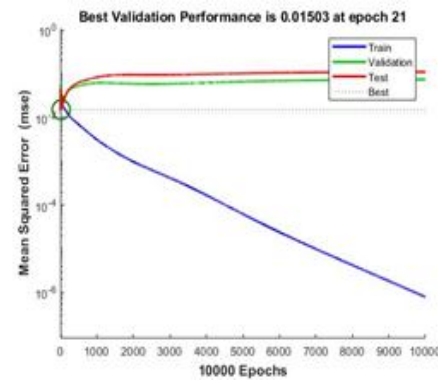**Figure 4 :** Best validation performance (Aseman Rainfall Station)

Figure 5 : Best validation performance (Ngurah Rai Rainfall Station)
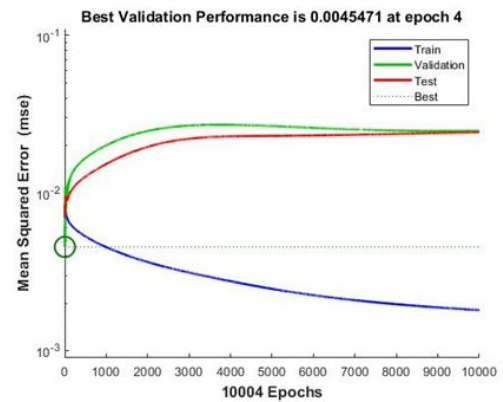

Figure 6 : Best validation performance (Sanglah Rainfall Station)
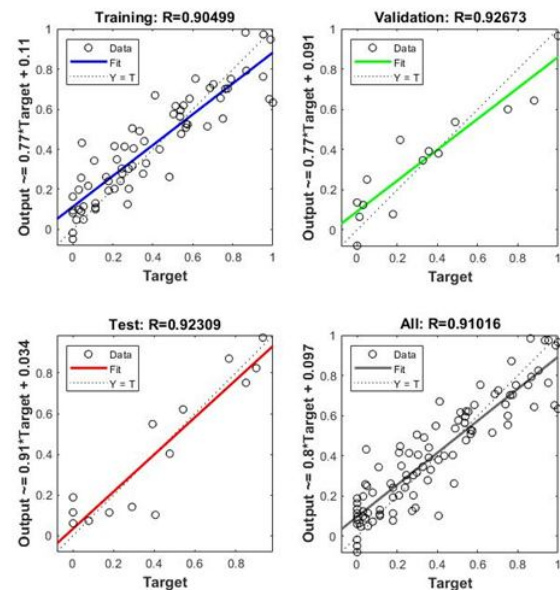

Figure 7 : "R" value (Aseman Rainfall Station)

From the Figure 7, it can be seen that the correlation coefficient between the training process (0.90499) and testing process (0.92309) at Aseman rainfall station shows a positive linear relationship which means the actual value can be approached by predictive value.

For the Ngurah Rai rainfall station, the value of coeficient correlation between training dan testing shown a strong relationship too. It seen the "R" value for training is 0.82932 and the value of "R" testing is 0.81556.
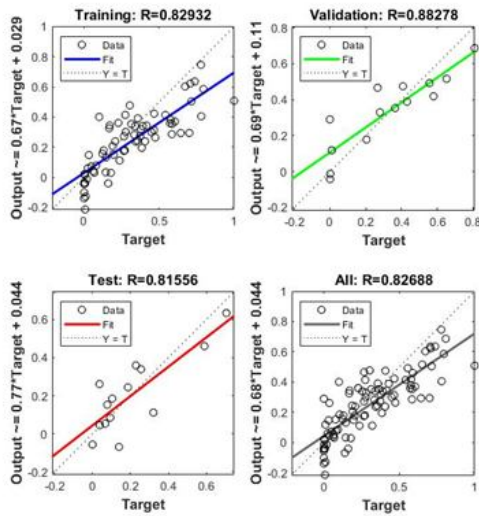

Figure 8 : "R" value (Ngurah Rai Rainfall Station)

For the Sanglah rainfall station, the value of coeficient correlation between training dan testing shown a strong relationship too. It seen the "R" value for training is 0.90149 and the value of "R" testing is 0.83441.
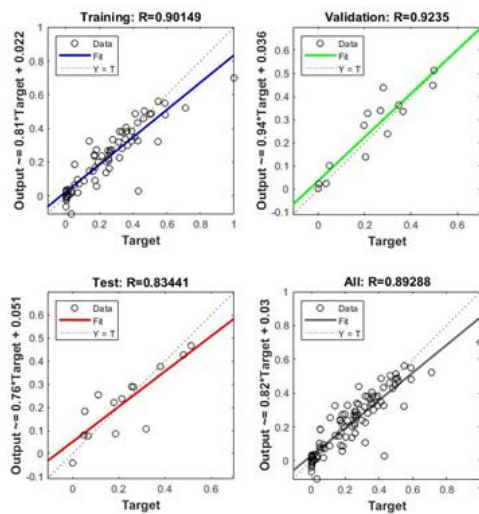

Figure 9 : "R" value (Sanglah Rainfall Station)

The relation among the observed and predicted monthly rainfall depend from the value of coefisient correlation (R). The high R values indicate a close relationship between them. It was observed from the distribute plots of the observed and predicted monthly rainfall for three rainfall station during calibration and validation for 12-10-5-1 FFBPNN architecture in Figure 10, 11 and 12 respectively.
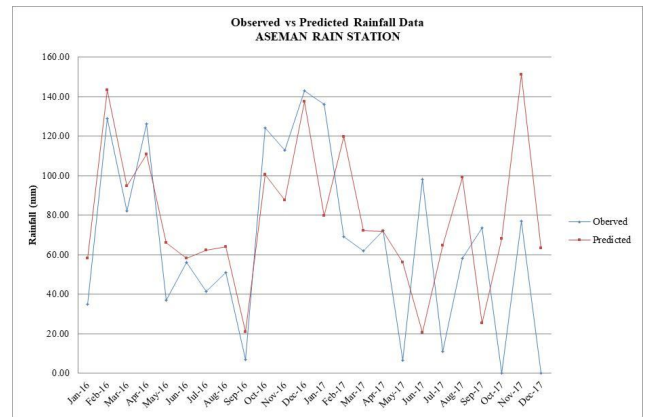

Figure 10 : Comparison of Rainfall Values between Observed and ANN Model during Validation Period (Aseman)
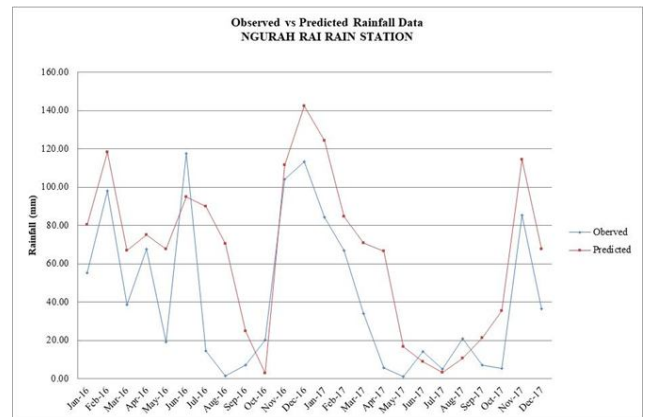

Figure 11 : Comparison of Rainfall Values between Observed and ANN Model during Validation Period (Ngurah Rai)

From the Figure 10, 11, and 12, the predicted rainfall by ANN model has fairly matched well and the trend of the observed rainfall and sometimes showing slightly higher value and lower values were observed during validation period (2016-2017), but it is an acceptable deviation range.
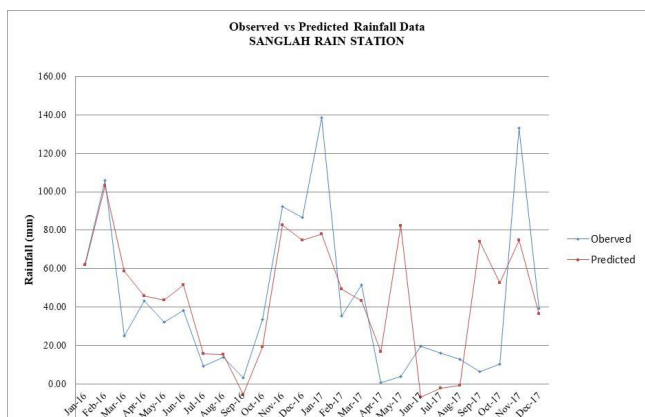

Figure 12 : Comparison of Rainfall Values between Observed and ANN Model during Validation Period (Sanglah)

5. CONCLUSION

The ANN model simulated monthly rainfall for three rainfall station has fairly matched with the observed values. Statistical analyses have been also performed to compare the simulated monthly rainfall with its measured counterpart. The high coefficient of correlation (R) values between the training process (0.90499) and testing process (92.309%) at Aseman rainfall station, "R" value for training is 0.82932 and the value of "R" testing is 0.81556 for Ngurah Rai stastion rainfall, and for the Sanglah rainfall station, the value of coeficient correlation for training is 90.149% and the value of "R" testing is 0.83441. That shows the close agreement between the measured and simulated monthly rainfall value during the calibration and the validation period. The low MSE value for 12-10-5-1 architecture of FFBPNN model during calibration and validation also shows better prediction too. The outcome shows obviously that the FFBNN is capable to forecast the rainfall in the Tukad Mati watershed.

ACKNOWLEDGEMENT

REFERENCES

1. Setiawan O. **Analisis variabilitas curah hujan dan suhu di Bali**. *J Anal Kebijak Kehutan.* 2012;9(1):66–79.
2. Triatmodjo B. *Hidrologi Terapan*. Beta Offset. Beta Offset Yogyakarta; 2013. 2 p.
3. **Tim SAR Evakuasi Warga Yang Terjebak Banjir**. 2018; Available from: http://basarnas.go.id/artikel/tim-sar-evakuasi-warga-yang-terjebak-banjir
4. **Banjir di Dewi Sri Badung Ternyata Kiriman Tukad Mati Denpasar** [Internet]. 2018. Available from: https://baliberkarya.com/index.php/read/2018/01/23/2018 01230004/Banjir-di-Dewi-Sri-Badung-Ternyata-Kiriman -Tukad-Mati-Denpasar.html
5. Mishra N, Soni HK, Sharma S, Upadhyay AK. A **Comprehensive Survey of Data Mining Techniques on Time Series Data for Rainfall Prediction**. 2017;11(2):168–84. https://doi.org/10.5614/itbj.ict.res.appl.2017.11.2.4
6. Sethi N, Garg K. **Exploiting Data Mining Technique for Rainfall Prediction**. 2014;5(3):3982–4.
7. Janbandhu CC. **Modelling Rainfall Prediction Using Data Mining Method - A Bayesian Approach**. *Int J Futur Revolut Comput Sci Commun Eng*. 2017;3(11):472–4.
8. Abbot J, Marohasy J. **Application of artificial neural networks to rainfall forecasting in Queensland, Australia**. *Adv Atmos Sci*. 2012;29(4):717–30. https://doi.org/10.1007/s00376-012-1259-9

9. Altunkaynak A, Nigussie TA. **Prediction of daily rainfall by a hybrid wavelet-season-neuro technique**. *J Hydrol*. 2015;529(P1):287–301. https://doi.org/10.1016/j.jhydrol.2015.07.046
10. Rani BK, Govardhan A. **Rainfall Prediction Using Data Mining Techniques - A Survey**. 2013;23–30.
11. French MN, Krajewski WF, Cuykendall RR. **Rainfall forecasting in space and time using a neural network**. *J Hydrol*. 1992;137(1–4):1–31.
12. Dharma IS, Putera IA, Putu Doddy Heka Ardana. **Artificial Neural Networks Untuk Pemodelan Curah Hujan-Limpasan Pada Daerah Aliran Sungai (DAS) Di Pulau Bali**. *Bumi Lestari J Environ*. 2011;(May):9–22.
13. Kumar PS, Praveen T V., Prasad MA. **Artificial Neural Network Model for Rainfall-Runoff -A Case Study**. *Int J Hybrid Inf Technol*. 2016;9(3):263–72. https://doi.org/10.14257/ijhit.2016.9.3.24
14. Chen SM, Wang YM, Tsou I. **Using artificial neural network approach for modelling rainfall-runoff due to typhoon**. *J Earth Syst Sci*. 2013;122(2):399–405. https://doi.org/10.1007/s12040-013-0289-8
15. Nourani V. **Using artificial neural networks (ANNs) for sediment load forecasting of Talkherood river mouth**. *J Urban Environ Eng*. 2009;3(1):1–6.
16. Wang Y, Traore S, Kerh T. **Using ANNs for Modeling Suspended Sediment Concentration**.Pdf. 2008;(October 2002).
17. Zaheer I. **Application of Artificial Neural Network for Water Quality Management**. *Lowl Technol Int*. 2004;5(2):10–5.
18. Teng YH, Kuok KK, Imteaz M, Lai WY, Ling KX. **Development of Whale Optimization Neural Network for Daily Water Level Forecasting**. *Int J Adv Trends Comput Sci Eng*. 2019;8(3):354–62. https://doi.org/10.30534/ijatcse/2019/04832019
19. Doddy Heka Ardana P. **Drought Disaster Forecasting Based On Rainfall Runoff Transformation Modelling**. In 2015. Available from: https://www.researchgate.net/publication/316780800_Dr ought_Disaster_Forecasting_Based_On_Rainfall_Runoff _Transformation_Modelling_Case_Studies_In_Tukad_Pe tanu_Watershed
20. Dawson CW, Wilby RL. **Hydrological modelling using artificial neural networks**. *Prog Phys Geogr*. 2001;25(1):80–108.
21. Minarti, Imam SB. **Prediksi Terjadinya Hujan Harian dengan Metode Jaringan Syaraf Tiruan di Stasiun Meteorologi Bandara Minangkabau**. *Poli Rekayasa*. 2011;6(2).
22. Oktaviani C, Afdal. **Jaringan Syaraf Tiruan Dengan Beberapa Fungsi Pelatihan (Studi Kasus : Stasiun Meteorologi Tabing Padang , Tahun 2001-2012)**. 2013;2(4):228–37.
23. Kumar M, Singla MK, Gupta J, Nijhawan P. **Comparative Study on Backpropagation and Levenberg Marquardt Algorithm on Short Term Load Forecasting**. *Int J Adv Trends Comput Sci Eng*. 2019;8(2):194–202. https://doi.org/10.30534/ijatcse/2019/14822019

24. Tombul M, Ersin O. **Modeling of Rainfall-Runoff Relationship at the Semi-arid Small Catchments** [Internet]. *Springer*; 2006. 309–310 p. Available from: https://link.springer.com/bookseries/642

25. Sallehuddin R, Siti SM. **Hybrid grey relational artificial neural network and auto regressive integrated moving average model for forecasting time-series data**. *Appl Artif Intell*. 2009;23(5):443–86.

26. Abhishek K, Kumar A, Ranjan R, Kumar S. **A rainfall prediction model using artificial neural network**. *Proc - 2012 IEEE Control Syst Grad Res Colloquium, ICSGRC 2012*. 2012;(July):82–7.

27. Santosa B. ***Data Mining Terapan Dengan MATLAB***. Graha Ilmu; 2007.

28. U. Rahardja, T. Hariguna, & Q. Aini, "Understanding the Impact of Determinants in Game Learning Acceptance: An Empirical Study", International Journal of Education and Practice, 7(3), 136–145, 2019. http://doi.org/10.18488/journal.61.2019.73.136.145

29. I. W. K. Wijaya, A. A. G. Rai, & T. Hariguna, "The impact of customer experience on customer behavior intention use in social media commerce, an extended expectation confirmation model: An empirical study", Management Science Letters, 9(2019), 2009–2020, 2019. http://doi.org/10.5267/j.msl.2019.7.005

30. T. Hariguna, & Akmal,"Assessing students' continuance intention in using multimedia online learning", TELKOMNIKA (Telecommunication Computing Electronics and Control), 17(1), 187–193, 2019. http://doi.org/10.12928/TELKOMNIKA.v17i1.10328