

Dependable Flow Modeling In Upper Basin Citarum Using Multilayer Perceptron Backpropagation

Ika Sari Damayanthi Sebayang^{a,1,*}, Muhammad Fahmi^{b,2}

^a Mercu Buana University, Jl. Meruya Selatan No 1, Jakarta 11650, Indonesia

¹ikasari.damayanthi@mercubuana.ac.id*; ²ffmuhammadfahmi@gmail.com; ³rizalbasir03@gmail.com

* corresponding author

ARTICLE INFO (8 pt)

Article history:

Received 20 June 2020

Revised 14 Oct 2020

Accepted 06 Nov 2020

Keywords:

Dependable flow

Perceptron

Back propagation

Citarum

Rainfall-runoff

ABSTRACT

To determine the amount of dependable flow, a hydrological approach is needed where changes in rainfall become runoff. This diversification is a very complex hydrological phenomenon. Where this is a nonlinear process, with time changing and distributed separately. To approach this phenomenon, an analysis of the hydrological system has been developed using a model which is a simplification of the actual natural variables. The model is formed by a set of mathematical equations that reflect the behavior of parameters in hydrology. Modeling in this case uses artificial neural networks, multilayer perceptron combined with the backpropagation method is used to study the rainfall-runoff relationship and verify the model statistically based on the mean square error (MSE), Nash-Sutcliffe Efficiency (NSE) and correlation coefficient value (R2). Of the three models formed, model 3 provides optimum results with correlation levels using NSE per month as follows, in Cikapundung Sub-Basin NSE = 0,990703, R2 = 0,995008, and MSE = 0,00014443, while in Citarik Sub-Basin NSE = 0.9500, R2 = 0.97592, and MSE = 0.0010804 . From these results it can be seen that ANN has a fairly good ability to replicate random discharge fluctuations in the form of artificial models that have almost the same fluctuations and can also be applied in rainfall runoff modelization even though the results of the test results are not very accurate because there are still irregularities

Copyright © 2017 International Journal of Artificial Intelligence Research.

All rights reserved.

I. Introduction

Dependable Flow is the amount of discharge available to meet water needs with calculated risk of failure. In planning the provider of water projects must first look for the dependable flow, the purpose of which is to determine the planning discharge which is expected to always be available in the river [1]. The hydrological process can be simply described by the relationship between elements of input namely rain, process and output, namely in the form of flow [2]. The relationship of rainfall-runoff is an important problem in hydrology and is the most fundamental component in the process of evaluating water resources [3].

The relationship of rainfall runoff, especially the diversification of rainfall to runoff, in a watershed is a very complex hydrological phenomenon. Where, this process is non-linear with time changing and spatially distributed [4]. To approach this phenomenon, an analysis of the hydrological system has been developed by using a model which is a simplification of actual natural reality [5].

In this research, To calculate the level of debit reliability, the flow duration curve was used [6]. The flow duration curve is a graph showing the river flow for a certain period of time in a year. Its value was determined from the arc at flow duration graph plotted from discharge data recorded from the model.

Along with the development of the digital world (computers), a number of models have been developed to imitate this process. Both empirical models (black box models), conceptual models

(physical process based), continuous models (continuous events), lumped models, distribution models and single models [7].

These models are formed by a set of mathematical equations that reflect the behavior of parameters in hydrology, so that the parameters contained in the equation have physical meaning [8]. ANN is one form of artificial intelligence that has the ability to learn from data and does not require a long time in making models [7]. This model uses a set of linear and non-linear mathematical equations that do not take into account the physical process at all, the most important in this model is that the output produced approaches the real one [8]. In addition, ANN is also able to identify the structure of the model and is also effective in connecting simulation inputs and outputs and forecasting models [7]. Benefits of using ANN is his ability to study relationships pre-existing unknown between the input and output data of each system. Besides the modeling with ANN has that attribute desirable and learning abilities from examples without explicit physical data.

The application of artificial neural networks has been widely used by previous researchers, such as, [9] using the multilayer perceptron backpropagation method as an optimization of runoff rainfall on the island of Java, namely the Cijung River, Banten, [10] who used artificial nerves to predict discharge at The Jinsha River Basin, China. [11] made a prediction model of river water salinity using artificial neural networks, and [12], using artificial neural networks to predict rainfall and climatology. Based on the above studies, according to researchers the method of artificial neural networks needs to be applied to predict an increase in the number of new students, because by using this method to obtain good accuracy and fast process. [13] used an ANN to model the relationship between rainfall and daily runoff in the Leaf River, Collins, Mississippi, USA. In order to evaluate the ANN performance, they compared the results obtained to the results of the SACRAMENTO model and autoregressive models of the ARMAX type. In this situation, ANNs presented the smallest errors and the smallest differences in volume. Good results were also obtained by applied ANN to flow forecasting for flood control in the Red River, Canada, and compared the results to linear and nonlinear multiple regression models; [14] applied ANN to three basins in Kansas, USA, to model the mean monthly flow, and also compared the results with empirical models; and [15] applied the ANN to three different USA basins: the Fraser River, Colorado, Raccoon Creek River, Iowa and the Little Patuxent River in Maryland, for monthly and daily forecasting horizons. Modarres [16] tested multilayer perceptron models in the Zayandehrud watershed in Iran. In all cases, the ANNs presented very good results, superior to the models to which they were compared.

Technological developments and the complexity of getting the parameters needed by using empirical formulas, moreover it is quite difficult to get results from quite a lot of data. To simplify this, the authors conducted a reliable discharge modeling that would be based on artificial neuron networks with the help of MATLAB software. This artificial nervous system method is useful only for real time not for recalculation.

II. Artificial Neural Network

Artificial neural networks (ANN) are information processing systems that have characteristics similar to biological neural networks, which are artificial representations of the human brain, which contain millions of nerve cells (neurons) and function to process information. Neurons have the same characteristics in ANN, consisting in groups called layers. Neurons in one layer are connected in the other adjacent layers. The strength of the relationship between adjacent neurons is represented in the strength of the relationship or weight.

An ANN generally consists of three layers, namely input layer, hidden layer and output layer. The input layer consists of neurons that receive an input from the outside environment. The input entered is a description of a problem. The hidden layer consists of neurons that receive input from the input layer, and then bring the output to the next layer. The output layer is called the output units, consisting of neurons that receive output from the hidden layer and send it to the user.

Rosenblatt (1957) in Hung et al., (2009) introduced the first simplest form of artificial neural network called a perceptron, which consists of only one layer [17]. Inputs are given directly to the

output unit via a weighting connection. In 1960 a multi layer perceptron (MLP) was developed and gradually became a neural network topology that was widely used in various fields, and included in the feedforward network category. In this network, besides input and output units, there are other units (often called hidden layer).

For system modeling, there are several advantages to using artificial neural networks [18], namely:

- Nonlinear system. Artificial neural networks have the ability to map relationships that are not linear, so it is very promising to overcome the problem of nonlinear control.
- Learning and adapting. Artificial neural networks that have been trained to use a set of data will be able to issue output equal to or close to the input pair output given, even able to issue output with inputs that have never been trained. Artificial neural networks can also be trained online.
- Multivariable system. Applicable artificial neural network for multi-variable processes.

The common artificial neural network structure is a multi layer perceptron (MLP). Figure 1 illustrates MLP structure, which consists of input, hidden and output layers.

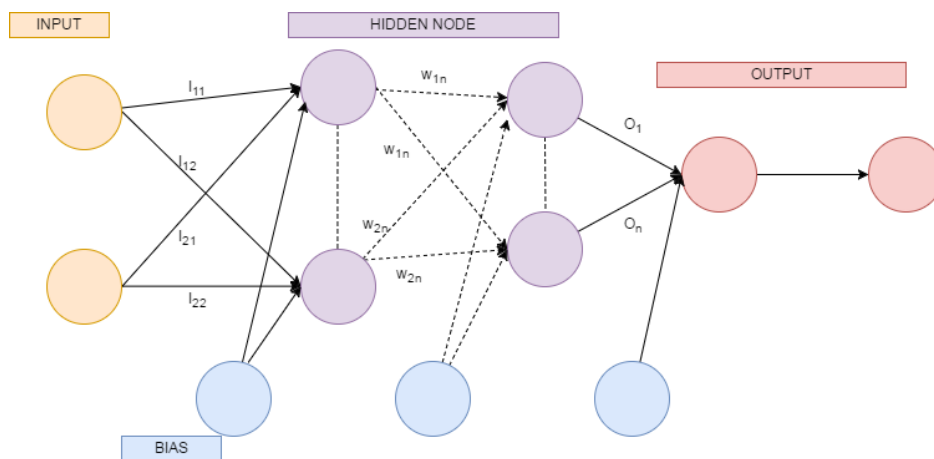


Fig. 1. Multilayer perceptron structure.

The output produced by ANN for a particular input pattern depends on the value (weight) of the relationship between neurons in the ANN. In ANN, for pattern recognition applications an initial training is needed so that it can be used in receiving input from outside. An ANN can solve a complicated problem if the appropriate weight values are used between neurons in different layers. The right weight value is obtained through the training process. The training process is a process to change the weight between neurons so that a network can solve a problem.

An important step in developing an ANN model is determining the weight of the matrix through training. There are two types of training mechanisms, namely supervised training and unsupervised training. Supervised training requires supervision from outside to guide the training process. This algorithm uses a number of pairs of input-output data that are used for example, where data used as an example should use data that is already known to be correct. The output from the network is then compared to the expected output data (sample output) to get the difference between the estimated output and the actual output. This difference is used to change the weight of the network to obtain the same output or close to the target. The mechanism of a supervised training called the back-propagation training algorithm [19] is generally widely used in engineering applications. Because ANN does not consider the physical problem, ANN is a blackbox model but can detect physical processes in ANN models that have been trained [20], [21], [22].

III. Material and Method

A. Location

In this study, Citarum Upper Basin is the location of the study. Increasing the number of population which is increasing every year in the Citarum Upper Basin has an impact on increasing land needs. Changes in forest areas and agricultural land into residential land and various other uses have caused many negative impacts on land and water resources that occur in watershed areas.

In this study, the location of the observation was the Cikapundung sub-watershed and the Citarik sub-watershed. Cikapundung and Citarik are the two biggest sub-watershed in Citarum Upper Basin. The Cikapundung watershed covers an area of 15,386.5 hectares with the administrative area of West Bandung Regency, Bandung City, and Bandung Regency. The Cikapundung River starts at Mount Bukit Tunggul, flows through the city and district of Bandung and empties into the Citarum River. The length of the Cikapundung River reaches 28,000 meters with the width of the river upstream of 22 meters and downstream of 26 meters as shown in Figure 3. The Citarik watershed has an area of 4,315.41 hectares, which is spread into five villages namely Dampit Village, Tanjungwangi Village, Cimanggung Village, Sindulang Village, and Tegal Manggung Village.

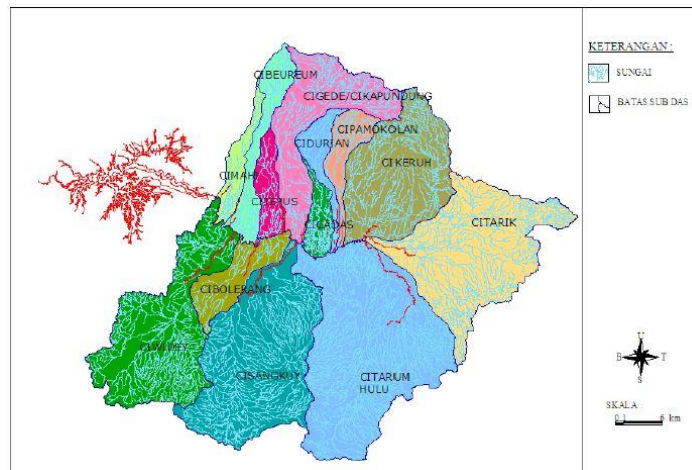


Fig. 2. Map of the Upper Citarum Basin.

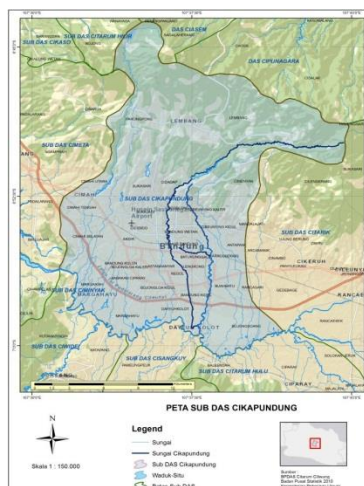


Fig. 3. Map of the Cikapundung sub-watershed.

B. Data and Software

The data used are daily rainfall data, discharge data and water level data for 10 years from 2008-2017. Rainfall data for Cikapundung sub-watershed is obtained from 3 rain observation stations

(Dagopakar, Margahayu, Mateo Lembang) and used as regional rainfall data for Citarik sub-watershed is obtained from 3 rain observation stations too (Jatiroke, Tanjungsari, Cibiru). Data sources were obtained from BBWS Citarum and PUSAIR. Topographic data is obtained from the Geospatial Information Agency (BIG). The software used to process numerical data is used by Microsoft Excel 2010 and to create a model used MATLAB R2015a which has several features grouped based on certain applications (Toolbox) related to Artificial Neural Networks namely Neural Network Tools.

C. Modeling Procedure

The modeling process with ANN is carried out in several stages, the first providing data to be used as input and data that will be targeted. The data is further divided into training data and testing data. Training data is used to train models that are made to approach the targeted results. While testing data is used to test the model whether the model functions to predict well.

In this case, a model with the MLP-BP (Multilayer Perceptron Back-Propagation) method was used. The correlation formed from the model will be presented using the Nash-Sutcliffe Efficiency (1) and R2 (2) as indicated by formulas, to ensure the model has a small error value, the approach with mean square error (MSE) is formulated (3).

$$NSE = 1 - \frac{\sum_{t=1}^T (Y_m^t - Y_o^t)^2}{\sum_{t=1}^T (Y_o^t - \bar{Y}_o)^2} \quad (1)$$

Where Y_m^t is modeled data, Y_o^t is the observation data and \bar{Y}_o is the average of the observation data.

$$R^2 = \left[\frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} \right] \quad (2)$$

Where Y is the data modeled and X is the observation data.

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - Y'_t)^2 \quad (3)$$

Where Y'_t is the data modeled and Y_t is the observation data.

D. Application

Training data and testing data use rainfall, water level and discharge data from 2008-2017. For training used data from 2008-2012 and testing used data from 2013-2017. The transformation process is carried out to change the input and output data values to a scale of 0 to 1 using the following equation:

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

Three artificial neural network structures were formed to be tested, each of which was as follows:

- Model 1

Model 1 has 2 input nodes, rainfall and discharge data, 8 hidden layer nodes and 1 layer output node as in Fig. 5 and Fig. 8

- Model 2

Model 2 has 3 input nodes, rainfall data, discharge and water level, 12 hidden layer nodes and 1 layer output node as in Fig. 6 and Fig. 9

- Model 3

Model 3 has 3 input nodes, rainfall data, discharge and water level, 12 hidden node layers and 1 output layer node. in model 3 is calculated by separating the model into per month as in Fig. 7 and Fig. 9

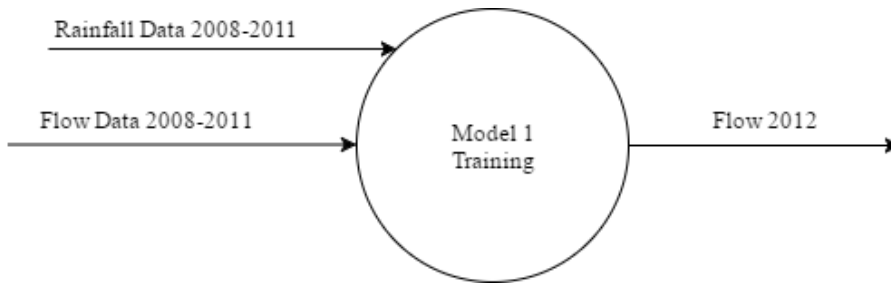


Fig. 5 Artificial Neural Net Model 1

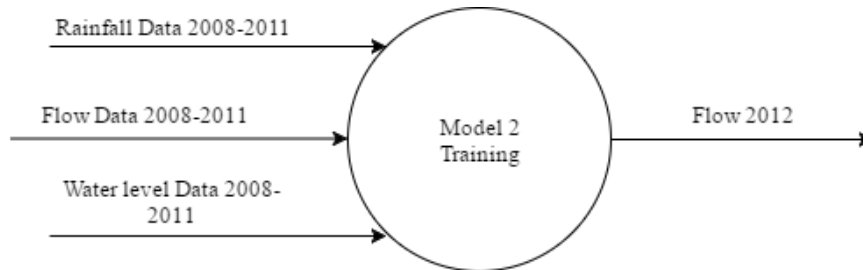


Fig. 6 Artificial Neural Net Model 2

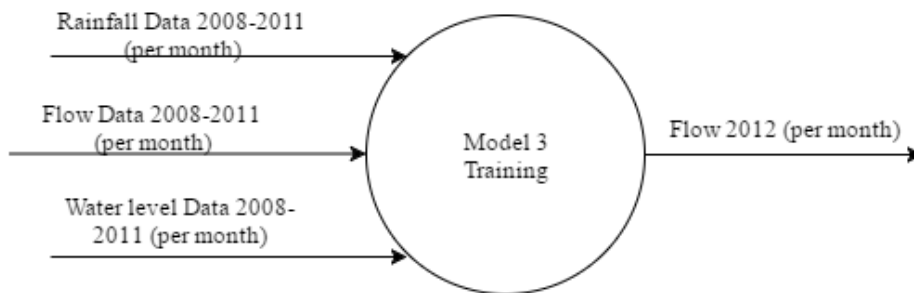


Fig. 7 Artificial Neural Net Model 3

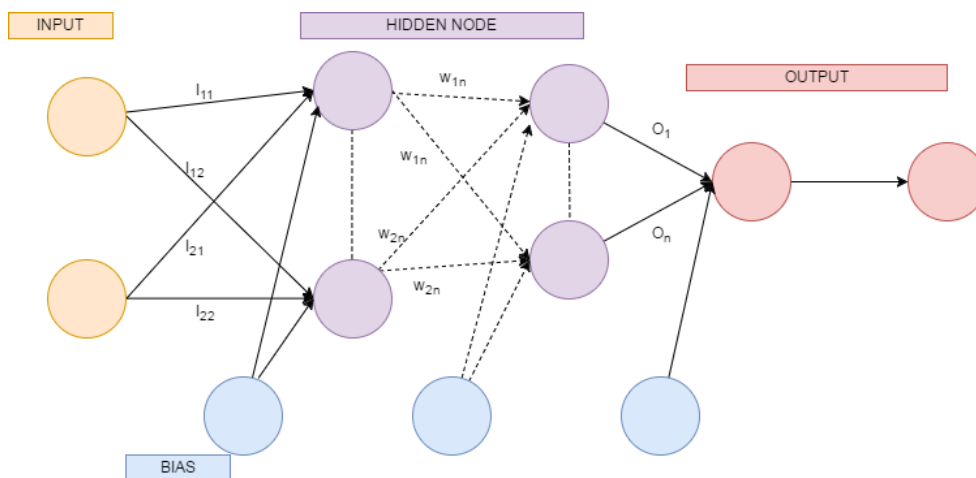


Fig. 8 Structure Artificial Neural Net Model 1

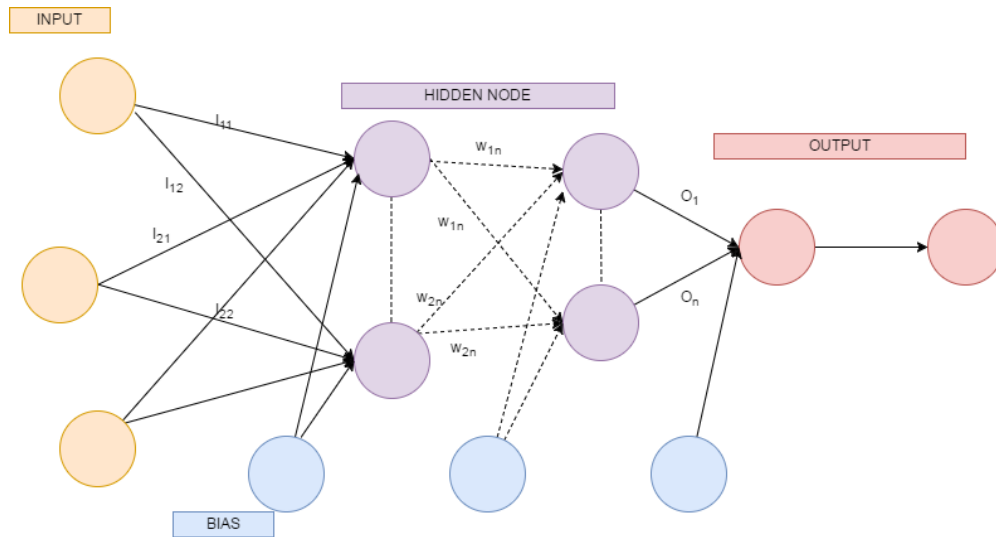


Fig. 9 Structure Artificial Neural Net Model 2 and 3

IV. Results and Discussion

The performance of each network model in modeling the rainfall-runoff relationship for the Cikapung watershed was calculated by mean square error (MSE), Nash-Sutcliffe Efficiency (NSE) and correlation coefficient (R²). The training performance results of the model will be displayed in the Table 1 and will be represented in graphical form as in Fig. 10-15.

Table 1. The training model performance results.

Item	Training					
	Cikapung			Citarik		
	NSE	R ²	MSE	NSE	R ²	MSE
Model 1	0,9642	0,9907	6,21E-04	0,9500	0,97388	0.00038541
Model 2	0,9712	0,9743	0,0017	0,9500	0,97477	0.00072723
Jan	0,9954	0,9945	6,02E-04	0,9990	0,99949	0.00022048
Feb	0,9970	0,9971	1,64E-04	0,9990	0,99977	1.16E-06
Mar	0,9961	0,9966	7,68E-05	0,9998	0,99991	1.85E-06
Apr	0,9963	0,9966	1,75E-04	0,9998	0,99989	8.86E-07
Mei	0,9946	0,9962	2,72E-04	0,9998	0,99993	5.37E-08
Model 3	0,9967	0,9970	3,36E-05	0,9990	0,99962	5.81E-06
Jul	0,9927	0,9929	1,65E-05	0,9990	0,99959	5.31E-07
Agust	0,9913	0,9832	5,73E-06	0,9980	0,99989	2.43E-06
Sep	0,9951	0,9962	1,13E-05	0,9990	0,99917	7.89E-07
Okt	0,9679	0,9945	5,18E-05	0,9990	0,99971	8.87E-08
Nov	0,9794	0,9989	6,11E-05	0,9996	0,99979	1.21E-05
Des	0,9860	0,9964	2,63E-04	0,9990	0,99957	2.30E-06

The following are the results of the hydrograph and dependable flow graph of 3 models:

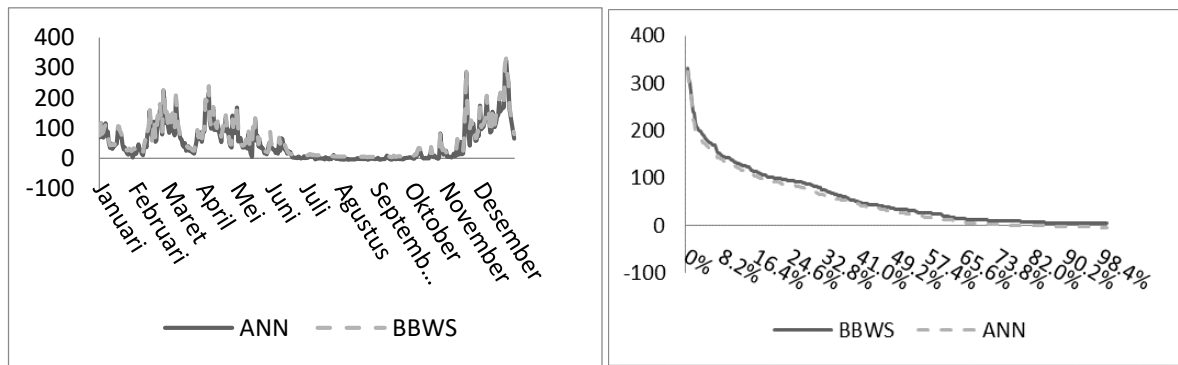


Fig. 10 Results Of Training Model 1 for Cikapungdung sub -watershed

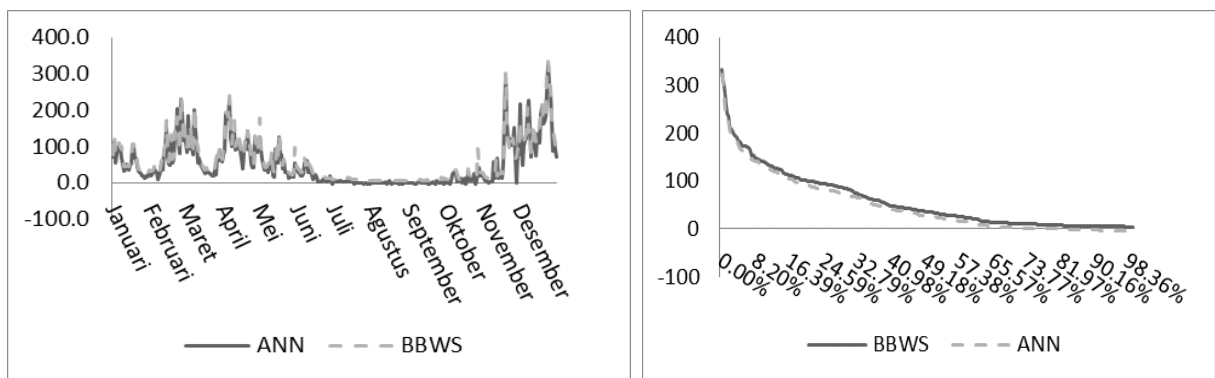


Fig. 11 Results Of Training Model 2 Cikapungdung sub -watershed

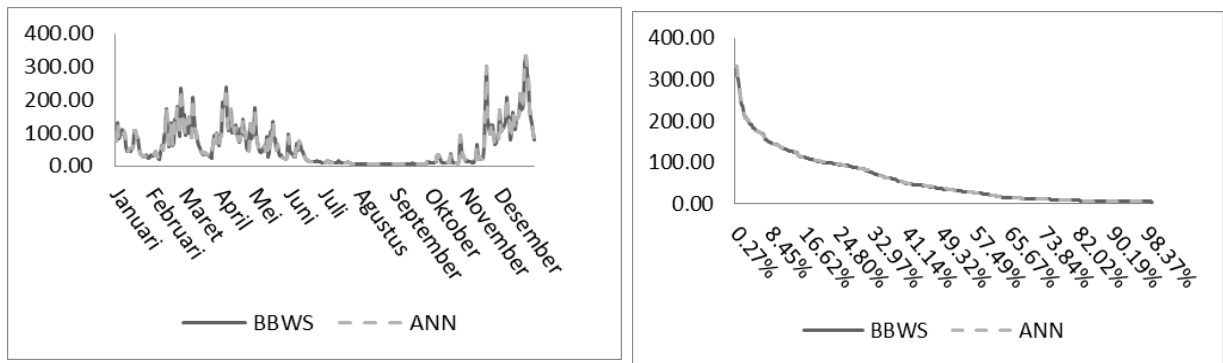


Fig. 12 Results Of Training Model 3 Cikapungdung sub -watershed

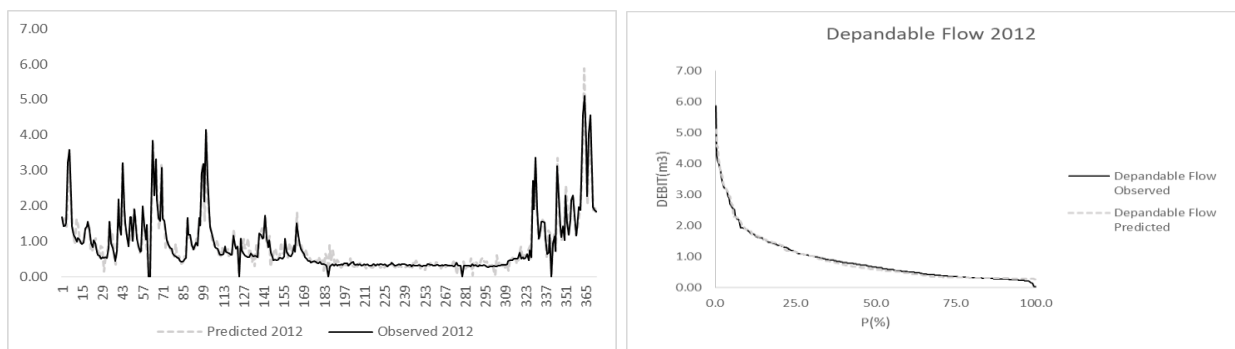


Fig. 13 Flood Duration Curved Model 1 Citarik sub -watershed

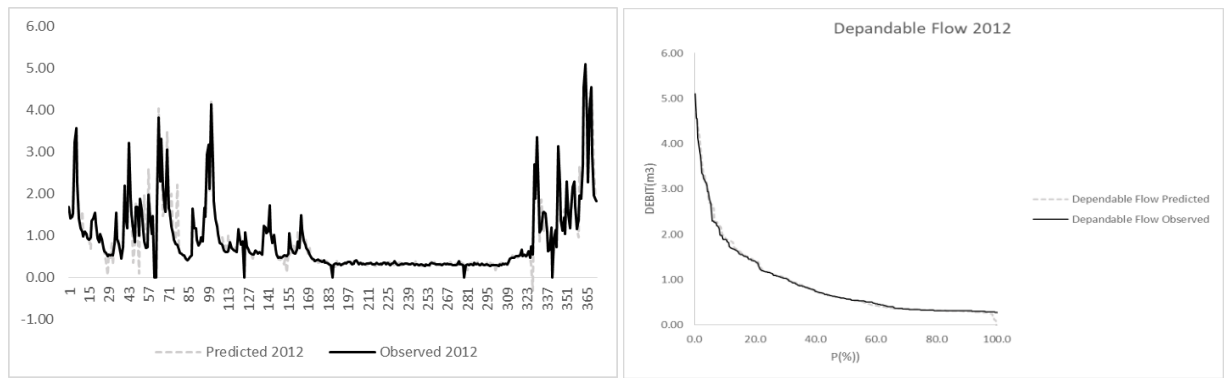


Fig. 14 Flood Duration Curved Model 2 Citarik sub -watershed

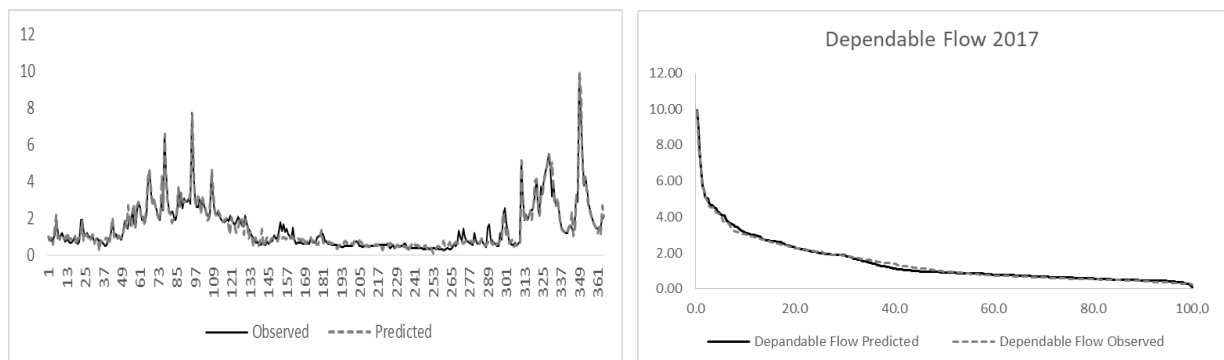


Fig. 15 Flood Duration Curved Model 3 Citarik sub –watershed

V. Conclusion

The ability of artificial neural network models to predict can be adjusted by changing existing parameters. In this case, the use of the number of neurons in the hidden layer, the more neurons used, the iteration process will be shorter and get faster results. But too much use of neurons can cause the network to become over fitting. Then it must be adjusted to the existing data input in order to avoid over fitting. In this study, the use of neurons was effective with 50 neurons for each hidden layer.

From the three models that have been formed it can be concluded that the third model has a correlation value and a good error value to be used. These results it can be seen that ANN has a fairly good ability to replicate random discharge fluctuations in the form of artificial models that have almost the same fluctuations and can also be applied in rainfall runoff modelization even though the results of the test results are not very accurate because there are still irregularities. With the dependable flow modeling this can be a parameter for decision making what will be the solution to water use in the Upper Basin Citarum. The results of modeling this artificial neural network can be used by the stakeholder to be used as decision makers adjusted to the real conditions in the Upper Basin Citarum, especially in Cikapundung and Citarik.

VI. References

- [1] Soemarto, C.D., 1995, Hidrologi Teknik, Penerbit Erlangga, Jakarta, vi+316 hlm.
- [2] Bessaih, N., Mah, Y. S., Muhammad, S.M., Kuok, K.K., and Rosmina, A.B., “Artificial Neural Networks for Daily Runoff Simulation”, Faculty of Engineering, Universiti Malaysia Sarawak, 2003.
- [3] Junsawang, P., J. Asavanant, C. Lursinsap. 2007. “Artificial Neural Network Model for Rainfall-Runoff Relationship”, ASIMMOD, Chiang Mai, Thailand.
- [4] Rajurkar, M. P., Kothyari, U. C. & Chaube, U. C. (2004). Modeling of daily rainfall–runoff relationship with artificial neural network.

- [5] Abdulla, F. and L.A. Badranih. 2000. "Application of a Rainfall-Runoff Model to Three Catchments in Iraq". *Journal of Hydrological Sciences*, 45: 13-25.
- [6] Ika Sari Damayanthi S. et. al, "Identification of Renewable Energy Potential in Ciberang River, Cisarua Village, Bogor, West Java", *IOP Conf. Ser.: Mater. Sci. Eng.* 343 . April. 2018.
- [7] Setiawan, B.I. dan Rudiyanto, 2004. "Aplikasi Neural Networks Untuk Prediksi Aliran Sungai", *Prosiding Semiloka Teknologi Simulasi dan Komputasi serta Aplikasi 2004 – BPPT, Jakarta*.
- [8] Kitanidis, P.K. and R.L. Bras. 1980a. "Adaptive Filtering Through Detection of Isolated Transient Errors In Rainfall-Runoff Models". *Water Resource Res.*, 16 (4): 740-748.
- [9] Ika Sari Damayanthi S. et. al, "OPTIMIZATION RAINFALL-RUNOFF MODELING FOR CIUJUNG RIVER USING BACK PROPAGATION METHOD", *SINERGI Vol. 22, No.3* : 193-204. October. 2018.
- [10] Hsu, K. L., Gupta, H. V. & Sorrosian, S. (1995) Artificial neural network modelling of the rainfall-runoff process. *Water Resour. Res.* 31(10), 2517–2530.
- [11] Jain, S. K., Das, A. & Srivastava, D. K. (1999) Application of ANN for reservoir inflow prediction and operation. *J. Water Resour. Plan. Manage. ASCE* 125(5), 263–271.
- [12] Lima, C. H. R. & Ferreira Filho, W. M. (2003) Análise de modelos de redes neurais aplicados ao processo chuva-deflúvio no semi-árido. *XV Simpósio Brasileiro de Recursos Hídricos*. Curitiba. Relação de trabalhos. Curitiba: Associação Brasileira de Recursos Hídricos, ABRH. CD ROM.
- [13] Machado, F. W. (2005) Modelagem chuva-vazão mensal utilizando redes neurais artificiais. MSc Thesis, Universidade Federal do Paraná, Curitiba, Paraná, Brazil.
- [14] Machado, F. W., Santos, I., Perreira Filho, D. L. B. & Mine, M. R. M. (2005) Avaliação do ajuste e extrapolação de curvas de descarga através de redes neurais. *XX Congresso Nacional del Agua – CONAGUA & III Simposio de Recursos Hidricos del Cono Sur*. Mendoza, Argentina.
- [15] Modarres R. (2009) Multi-criteria validation of artificial neural network rainfall-runoff modeling. *Hydrol. Earth System Sci.* 13(3), 411–421.
- [16] Muller, I. I. (1995) Métodos de avaliação da evaporação e evapotranspiração: análise comparativa para o Estado do Paraná. MSc Thesis, Universidade Federal do Paraná, Curitiba, Paraná, Brazil.
- [17] Neelakantan, T. R. & Pundarikanthan, N. V. (2000) Neural networkbased simulation-optimization model for reservoir operation. *J. Water Resour. Plan. Manage. ASCE* 126(2), 57–64.
- [18] Murray, R., Neumerkel, D., Sbarbaro, D., 1992, *Neural Networks for Modeling and Control of A Non-Linear Dynamic System*. Proceedings of the 1992 IEEE International Symposium on Intelligent Control, Glasgow, Scotland, pp. 404–409.
- [19] Rumelhart, E., G. Hinton and R. Williams, 1986. "Learning Internal Representations by Error Propagation". *Parallel Distributive Process*, 1: 218-362.
- [20] Wilby, R.L., R.J. Abrahart, and C.W. Dawson. 2003. "Detection Of Conceptual Model Rainfall-Runoff Processes Inside An Artificial Neural Network", *Hydrol. Sci. J.*, 48 (2): 163-181.
- [21] Jain, A., K.P. Sudheer and Srinivasulu, S. 2004. "Identification Of Physical Processes Inherent In Artificial Neural Network Rainfall-Runoff Models". *Hydro. Process*, 118 (3): 571-581.
- [22] Sudheer, K.P. and A. Jain. 2004. "Explaining The Internal Behavior of Artificial Neural Network River Flow Models". *Hydrol. Process*, 118 (4): 833-844.
- [23] N. J. de Vos and T. H. M. Rientjes, "Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation," *Hydrology and Earth System Sciences*. pp.111-126. July, 2005.
- [24] Asaad Y. Shamseldin, "Artificial neural network model for river flow forecasting in a developing country," *Journal of Hydroinformatics*. doi: 10.2166/hydro.2010.027. 2010.
- [25] Kuok King Kuok and Nabil Bessaih, "Artificial Neural Networks (ANNS) For Daily Rainfall Runoff Modeling," *The Institution of Engineers*. Vol. 68, No.3. September, 2007.
- [26] Yonas B. Dibike and Dimitri P. Solomatine, "River Flow Forecasting Using Artificial Neural Networks," *EGS journal of Physics and Chemistry of the Earth*. MS-No. EGS1.1-99002.