Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 125 (2015) 150 – 157

Procedia Engineering

www.elsevier.com/locate/procedia

The 5th International Conference of Euro Asia Civil Engineering Forum (EACEF-5)

Analysis of rainfall-runoff neuron input model with artificial neural network for simulation for availability of discharge at Bah Bolon Watershed

Setiono[a,*], Rintis Hadiani[a]

[a]Depatrment of Civil Engineering Sebelas Maet University, Jl. Ir. Sutami 36A, Surakarta 57126, Indonesia

Abstract

Indonesia is a tropical country with two seasons (wet and dry) which play the main role in water cycle process. Occurrence of rain continues into the flow of the discharge in the river with a huge energy potential that can be exploited for the life of the surrounding community. The occurrence and intensity of rain is random and difficult to predict in a certain period of time so that discharge is also difficult to be estimated although it is measured in the field in time of rainfall occurrence. The amount of runoff produced by the same depth of precipitation in a watershed will result a different magnitude with another watershed because it is influenced by land use in the watershed. This paper discusses the modeling of rainfall-runoff in the Watershed of Bolon in Simalungun district of North Sumatra Province using Artificial Neural Network (ANN) to determine the potential of the available discharge in the long term for the purpose of Micro Hydro Power (MHP). The software/program is developed with Scilab mathematical open source software (www.scilab.org) based on ANN algorithm. The data are record of monthly rainfall and discharge for 12 years (2001 to 2012). The models developed are 12 monthly neurons, 4 year neurons and series neuron (48 neurons) for input (rainfall) - output (runoff) neurons. The result shows that reliability the 12 monthly neurons is 99% (the best) followed by series neuron with 78% and 4 year neuron 77%. The chosen model (12 monthly neurons) then to be used for predicting the monthly discharge availability at Bah Bolon Site. Dependable discharges predicted with this software for year 2013 to 2016 consecutively are as follows: 0.678246 m^3/s, 0.655288 m^3/s, 0.678475 m^3/s and 0.678135 m^3/s.

Peer-review under responsibility of organizing committee of The 5th International Conference of Euro Asia Civil Engineering Forum (EACEF-5)

Keywords: Artificial Neural Network; Discharge; Hidrology Modeling; Micro Hydro Power

* Corresponding author. Tel.: +62-81-393-287-807; fax: +62-271-634-524.
*E-mail address:*setiono@ft.uns.ac.id

Peer-review under responsibility of organizing committee of The 5th International Conference of Euro Asia Civil Engineering Forum (EACEF-5)
doi:10.1016/j.proeng.2015.11.022

1. Introduction

Simulation of rainfall-runoff is a discharge based approach to predict rain that enters the watershed. Rainfall data is uncertain- random data. In the prediction of rainfall-discharge transformation, this uncertainty (random data) can be adapted in the model Monte Carlo [1]. Another method that is able to adapt this uncertainty is the expert system inference method [2 & 3].

The process of transformation of rainfall into discharge can be replicated and simplified in the form of models, commonly called rain-flow models. Many models have been developed in the rain-flow analysis. One of the conventional methods in Hydrology, the hydrology which simulate actual events is IHACRES [4]. In modern Hydrology, by inference, the models which were developed using Artificial Neural Network (ANN) for rain-flow analysis has been done by [5], [6], [7], [8] and [9].

ANN technology is a computational algorithm that is part of the Artificial Intelligence that simulates the works of the nervous system in the processing of biological stimuli/incoming information and then determines the response to that information. One of the applications that match the ANN is a matter of pattern recognition based on historical data that is fed to the ANN to perform the learning process. Variables are processed in the ANN is the number of inputs, number of layers, and number of outputs with the reliability of the model to 95% confidence.

This method is able to adopt the hydrologic process with all its uncertainty. Under these conditions, the ANN was used as a model in the analysis of rainfall-runoff [6]. The accuracy of determining the discharge can ease the analysis of discharge available in watersheds [10]. Potential discharge generated by rain-flow analysis in a watershed can be used as a reference to its use, for example for the calculation of the capacity of Micro Hydro Power (MHP) at Bah Bolon Simalungun District of North Sumatra.

2. Literature Review

- Hydrologic cycle
  The hydrologic cycle is the circulation of water that never stops, occurs when water evaporates into the air from the ground, sea or from plants (transpiration). The evaporated water then condenses and solidifies then drops as rain or snow. Rain water is not all straight up to the ground surface, but some have stuck by the plant canopy interception as the retained portion that will be evaporated back and some will go into the ground through percolation process, some will continue to penetrate the soil layer which will eventually become saturated groundwater. Most of the water on the surface will be a run off that fills hollows, get into rivers that flow rate is referred to as discharge [11].
- Rainfall
  Rainfall is one of hydrological variables which has a large variation in both its distribution, time and place. Rainfall which occurs in an area is called rainfall region and measured in milimeter [11]. Estimated rainfall data is used to estimate precipitation that falls in the surrounding area. Accuracy of rainfall measurement results in an area will depend on the spatial variability of rainfall throughout the area, thus, needed more equipment to measure rainfall, especially in areas with large steep slope and areas that receive heavy rainfall (thunderstorms) than frontal precipitation type [12]. The intensity of rainfall is rainfall depth per unit time (mm/h). Rainfall in an area is the average rainfall throughout the observation area, instead of 1 point rainfall observations. One point rainfall measurements cannot represent the depth of precipitation that falls on a place. Calculation method of rainfall in some regions of rainfall observation point is divided into 5, i.e Arithmetic Mean, Thiessen polygons, isohyets lines, intersection lines method and depth elevation method [11].
- Model of Artificial Neural Network (ANN)
  In general, ANN models have network structures with three or more layers (layer) consisting of input nodes, hidden nodes, and the output nodes. Node is the amount (value) of the layer (variable). [5], [6], [7], [8] and [9] have attempted to investigate the relationship between rainfall and runoff with this model in some areas and resulted different models based on characteristics of a watershed. In this study, rainfall is used for input node and discharge is for output/target node. The Scilab is an open source software for the purposes of numerical calculation and provides a powerful computing environment for engineering and scientific applications [13], can be obtained from [14]. It provides the Neural Network Toolbox for purposes of design, implementation and

visualization of the simulation models using neural network algorithm with the support of multilayered feed forward back propagation algorithm. This tool is very helpful for researchers who study the ANN technology because the features provided quite adequate for the purposes of research.

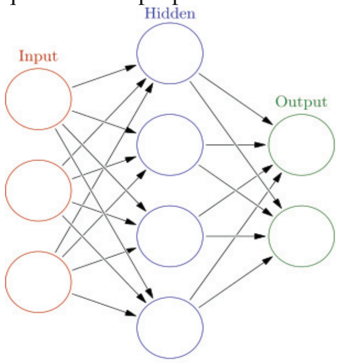


Figure 1. Structure of an Artificial Neural Network (ANN)[15]

3. Research Method

3.1 Location

This research was conducted in the watersheds located in the district of Bolon Simalungun, North Sumatra. The area of watershed is ± 790 km^2 and the length of the Bah Bolon river is 118 km. It crosses two districts i.e Simalungun and Asahan Districts. The land usage is generally for crops, industries, housings and tourisms [16].

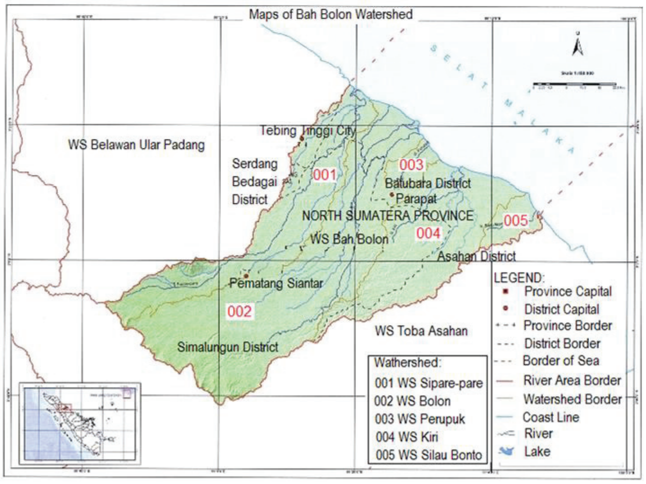


Figure 2. Map of Bah Bolon Watershed

3.2 Data and Research Tools

This study uses secondary data i.e rainfall and discharge as the main data and maps of North Sumatera as supporting data. The research data needed are as followings: Map of the watershed as well as the location of rainfall stations in the watershed Bah Bolon, rainfall data from the last 12 years (2001 to 2012) of each station (3 stations: Marihat, Bah Jambi and Sidamanik rainfall stations) in the watershed Bah Bolon, and discharge data in the recent 12 years of watershed Bah Bolon as well. Tools such as software that is used in this study are: AutoCad 2007 for the watershed map processing program, Microsoft Excel 2007 for rainfall data processing, and Scilab 5.4.1 is used to design the ANN models.

3.3 Research Stages

Stage1: Literature review and preliminary study; Stage 2: Data collection of rainfall and discharge; Stage 3: development of data model for ANN's input; Stage 4: Development of ANN model; Stage 5: Test of model reliability; Stage 6: Validation of simulation result; Stage 7: Calculation of dependable discharge and Stage 8: Conclusion and writing the research report.

3.4 Test of Model Reliability

Correlation is one of the analytical techniques used to measure how strong the relationship between the two variables observed. Correlation value of the two variables tested range between zeros to 1. How to calculate the correlation can be seen in formula 1 with a description of the gradation value of the correlation.

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\{n(\sum x^2)(\sum x)^2)\} - \{n(\sum y^2)(\sum y)^2)\}}} \quad (1)$$

with:
r : coeff of correlation, n : number of data, x = simulated discharge (m^3/s) and y = measured discharge (m^3/s).

The meaning value of r are: 0 : no correlation , 0 – 0,25 : very weak correlation, 0,25 – 0,5 : Fair correlation, 0,5 – 0,75 : strong correlation, 0,75 – 0,99: very strong correlation and 1 : perfect correlation.

3.5 Runs Test of Model

Runs test is conducted to ensure the validity of the simulation results. According to [18] in [17], as shown in Figure 3 based on theory of runs, m is positive run-length and n is negative run-length. Hence, the total run-length, r is $m + n$. Value of estimated r, $E(r)$, is expressed in probability q for estimated positive run-length, m, and p is estimated probability for negative run-length, n.

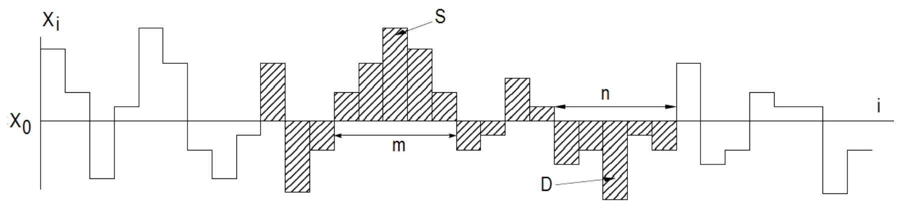


Figure 3. Positive run length (m), negative run length (n) of Theory of runs [18]

This relationship of them is stated by [18] as followings:

$$E(r_q) = \frac{1}{q(1-p)} = \frac{1}{pq}, \text{ with condition } 0<q<1 \quad (2)$$

$$p = \frac{m}{m+n}, q = \frac{n}{m+n} \quad (3)$$

$$\bar{r}_q = \frac{1}{k_r}\sum_{j=1}^{k_r} r_{q,j} \quad (4)$$

with: $\bar{r}_q$ is the total run length based on probability q, j=1,2,3,…, k_r and kr is number of total run-length.

The number of runs of the simulation result ($\bar{r}_q$) must meet the criterion as stated in Formula 5.

$$E(r_q) - \frac{t_{\alpha/2}}{pq}\left(\frac{p^3+q^3}{k_r}\right)^{1/2} \leq \bar{r}_q \leq E(r_q) + \frac{t_{\alpha/2}}{pq}\left(\frac{p^3+q^3}{k_r}\right)^{1/2} \quad (5)$$

with: α is significant value (5%) and t is the value of statistical t table according to the value of k_r .

4. Result And Discussion

4.1 Neuron Models

It is generally known that rainfall is dependent of time function. As the rainfall data is grouped into monthly rainfall and months are subset of year, here the input data are modeled into neurons of 12 month based and neuron of 4 year based (as the data have 12 year of period).

Based on the results of running the iteration parameters of 500 (optimum number of iterations) and tolerances (MSE) 0.0001 correlations in data validation as shown in Table 1 below. The best model of input neurons is 12 monthly and neurons in the hidden layer 1 and 2 are 2 and 3, which resulted in a correlation of 0.540071 with the iteration 500 times.

4.2 Results of Running Program

Based on Table 1 model, training data is conducted with running parameters of the ANN in accordance to number 3, but with the number of iterations is 1000 times. This process produced the value of Mean Squared Error (MSE) of the model= 0.0002061 (Figure 4). Figure 4 shows that the training process reach its maximum iteration i.e 1000 times instead of the value of MSE.

Table 1. Correlation value with a variety of approaches the number of neurons in the hidden layer and input neuron

No	Number of neurons		Correlation value	
	Hidden 1	Hidden 2	12 month based	4 year based
1	2	2	0.5283270	0.2197610
2	3	2	0.3447689	0.0539308
3	2	3	0.5450071	0.1466791
4	3	3	0.4759717	0.1478648

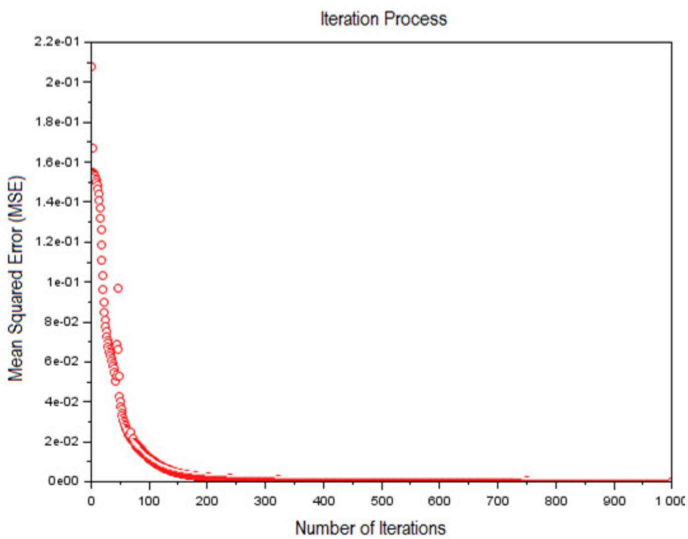
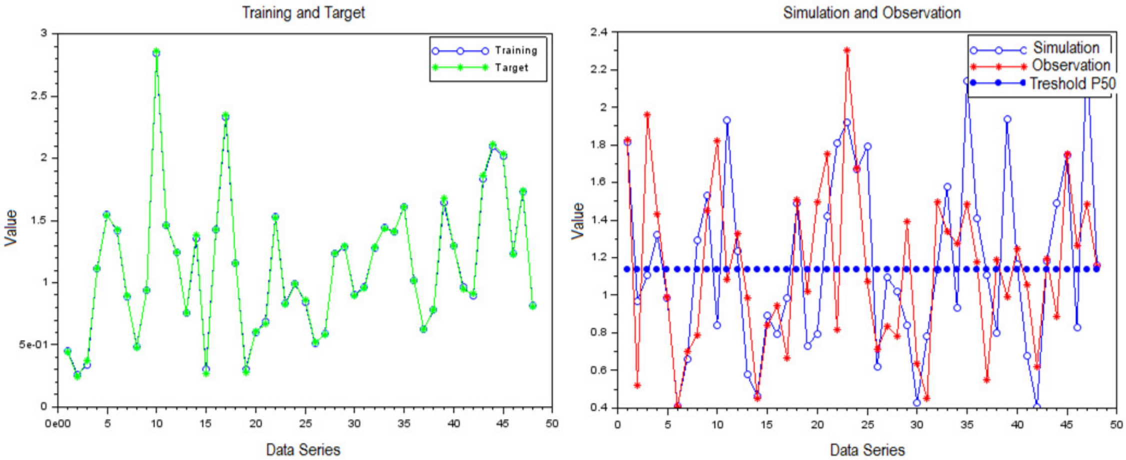


Figure 4. Iteration process of ANN

Figure 5a shows a graph between training and target outputs coincide with each other, which resulted in a correlation value: 0.9997761. For the simulation results and observations produces a correlation value= 0.5272530 (strong correlation based on the explanation of the formula 1), the median value (Q50)= 1.1373039, mean= 1.1706728 and standard deviation= 0.4861986 (Figure 5b).



(a) Plot of output and target training of ANN (b) Plot of simulation and validation of ANN

Figure 5. Result of training the ANN model for Bah Bolon Site

Figure 6 shows the discharge prediction for period year 2013 to 2016 with the median (Q_{50})= 1.2123416, mean= 1.2123416 and standard of deviation= 0.5193679.

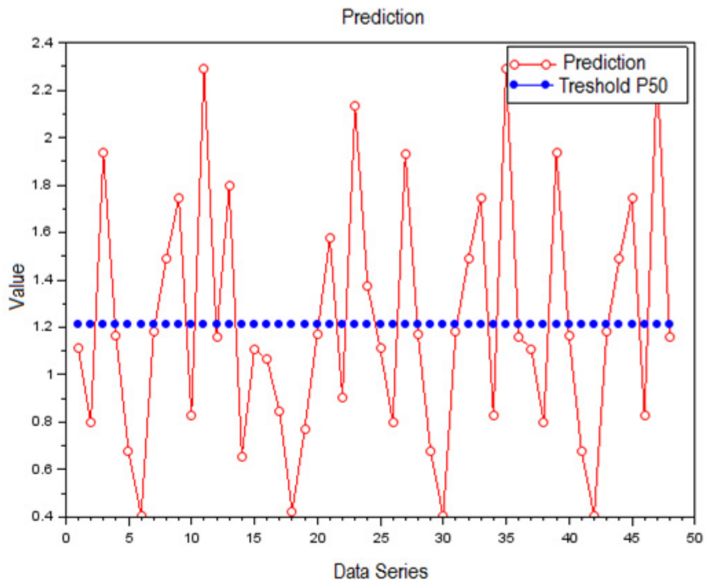


Figure 6. The discharge prediction for year 2013 to 2016 period

4.3 Validation of Simulation Results (Discharge)

Validation of data must be conducted to ensure the results of the simulation models can be justified scientifically correct. Validation of the model is done by performing runs test (according to Formula 2, 3, 4 and 5), with the aim to obtain homogeneity of simulation data with the original data (observations). The following values generated by the ANN model with 12 input neurons monthly, 2 and 3 neurons in the hidden layer 1 and 2. The occurring run is between the lower limit and upper limit to the value of the resulting discharge qualify for use.

Table 2. Validation of discharge model

Lower limit	Value of run	Upper limit
22.688489	25	27.311511

The Table 3 show the result of discharge predicted for period of year 2013 to 2016.

Table 3. Result of discharge predicted for period year 2013 to 2016(m^3/s)

year	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2013	1.112397	0.798605	1.933656	1.167062	0.678246	0.405596	1.183108	1.487564	1.746132	0.826979	2.290074	1.160095
2014	1.797911	0.655288	1.108866	1.065157	0.845499	0.42191	0.772274	1.173717	1.577355	0.905403	2.131359	1.371825
2015	1.110366	0.80085	1.933134	1.168599	0.678475	0.405649	1.182889	1.487383	1.745996	0.827076	2.28978	1.160181
2016	1.109957	0.797696	1.933971	1.166303	0.678135	0.405583	1.183239	1.487645	1.746201	0.826953	2.290207	1.160051

4.4 Dependable Discharge

Dependable discharge is the amount of discharge is available to meet the water needs of the risk of failure that has been taken into account. The goal is to determine the expected design discharge is always available in the river [19]. For MHP generally use a dependable discharge of 90% (Q90), which means a high risks to be faced because of the less discharge smaller than dependable discharge is 10% the number of observations. In calculation of dependable discharge with probability of 90 %, the *Weibull formula* is used:

$$p = \frac{m}{n+1} 100\% \tag{6}$$

with : $P =$ probability (%), $m =$ order and $n =$ number of data

Table 4. Calculation of dependable discharge for year 2013 to 2014(m^3/s)

2013				2014			
Month	Order	Discharge	Probability	Month	Order	Discharge	Probability
May	2	0.678246	84.62%	Feb	2	0.655288	84.62%

Table 5. Calculation of dependable discharge for year 2015 to 2016(m3/s)

2015				2016			
Month	Order	Discharge	Probability	Month	Order	Discharge	Probability
May	2	0.678475	84.62%	May	2	0.678135	84.62%

In determining the dependable discharge, the next process is used; reliability = 90% (R90), n = 12 (number of months in a year), so R90= 12/[100/(100-90)]+1, then R90= (12/10)+1 and as a result R90= 2,2 ≈2. It can be conclude that dependable discharge of year 2013 is at order of 2 which is in May of 0.678246 m^3/s. So that dependable discharge for next 3 year consecutively are 0.655288 m^3/s, 0.678475 m^3/s dan 0.678135 m^3/s, for year 2014 to 2016.

5. Conclusion

From this research, it can be concluded:

- GUI interface of ANN for rainfall-runoff simulation of Bah Bolon Site has been developed with neuron model (input model) is at best for 12 month base with 2 and 3 neurons for each hidden layer with correlation value of 0.5450071, and
- Dependable discharges predicted for year 2013 to 2016 consecutively are as follows: 0.678246 m^3/s, 0.655288 m^3/s, 0.678475 m^3/s and 0.678135 m^3/s at Bah Bolon Site.

Acknowledgement

The author would like to acknowledge Sebelas Maret University Surakarta for supporting this research and also big thankyou to my colleague Dr. Rintis Hadiani for her full support during this research.

References

[1] Štěpán Kuchař et al, 2004, Simulation Of Uncertainty In Rainfall-Runoff Models And Their Statistical Evaluation In The Floreon+ System, available at http://floreon.vsb.cz/base/Download/1_ML.pdf
[2] Turban, E et al, 2005. Decision Support System and Intelligent System (Sistem Pendukung Keputusan dan Sistem Cerdas). Penerbit Andi. Yogyakarta.
[3] RR. Rintis Hadiani. 2009. Analisis Kekeringan Hidrologi (Studi Kasusdi Sub Das Kali Asem Lumajang). Malang: Universitas Brawijaya Malang.
[4] Imam Suprayogi, et al, 2013, Analisis Hujan Debit pada DAS Indragiri Menggunakan Pendekatan Model IHACRES, Konferensi Nasional Teknik Sipil 7 (KoNTekS 7), Universitas Sebelas Maret (UNS) - Surakarta, 24-26 Oktober 2013.
[5] R. Modarres, 2009, Multi-criteria validation of artificial neural network rainfall-runoff modeling, Hydrol. Earth Syst. Sci., 13, 411–421, 2009, Published by Copernicus Publications on behalf of the European Geosciences Union.
[6] Am Kalteh, 2008, Rainfall-Runoff Modelling Using Artificial Neural Networks (Anns): Modelling And Understanding, Caspian J. Env. Sci. 2008, Vol. 6 No.1 pp. 53~58
[7] Riad, S et al, 2004, Rainfall-Runoff Model Using an Artificial Neural Network Approach, Mathematical and Computer Modelling 40 (2004) 839-846, Elsevier.com
[8] Vos, N.J. de, 2003, Rainfall-Runoff Modelling Using Artificial Neural Networks, M. Sc. Thesis Report, Delft, Netherlands.
[9] Rajurkar, MP et al, 2002, Artificial neural networks for daily rainfall-runoff modeling, Hydrologkal Sciences-Jo umai-des Sciences Hydrologiques, 47(6) December 2002
[10] Hurkmans, R. T. W. L. et al, 2007, Water Balance Versus Land Surface Model In The Simulation Of Rhine River Discharges, Water Resources Research, Vol. 44, W01418, doi:10.1029/2007WR006168, 2008.
[11] Sosrodarsono, S., dan K. Takeda, 2003. Hidrologi untuk pengairan. Pradnya Paramita, Jakarta.
[12] Asdak, C., 1995. Hidrologi dan Pengelolaan Daerah Aliran Sungai.Gadjah Mada University Press, Yogyakarta.
[13] Baudin, Michael, 2011, Introduction to Scilab, Scilab Consortium, http://www.scilab.org.
[14] http://www.scilab.org/download
[15] Artificial neural network, http://en.wikipedia.org/wiki/Artificial_neural_network, accessed at 23 July 2015.
[16] Randi Gunawan, C. D. Soemarto, 2006, Analisis Sumberdaya Air Daerah Aliran Sungai Bah Bolon Sebagai Sarana Pengembangan Wilayah di Kabupaten Simalungun dan Asahan, Wahana Hijau Jurnal Perencanaan & Pengembangan Wilayah Vol2 No 1.
[17] Heri Eko Prasetyo et al, 2014, Analisis Data Runtun Waktu Debit Menggunakan Jaringan Syaraf Tiruan Di Das Wuryantoro Pada Awlr Kecamatan Wuryantoro, Matriks Teknik Sipil, Vol 2, No 2, Juni 2014
[18] Yevjevich, V (1972) Stochastic Processes in Hydrology, Colorado, USA: Water Resources Publications, Fort Collins.
[19] C. D. Soemarto, 1999, Hidrologi Teknik, Penerbit Erlangga, Jakarta.