Water Level Prediction of Lake Cascade Mahakam Using Adaptive Neural Network Backpropagation (ANNBP)

Mislan*[1], A F O Gaffar[2], Haviluddin[3], N Puspitasari[4]

[1]Department of Physics, Faculty of Mathematics and Natural Science, Mulawarman University, Indonesia
[2]Department of Information Technology, State of Polytechnic Samarinda, Indonesia
[3,4]Department of Computer Science, Faculty of Computer Science and Information Technology, Mulawarman University, Indonesia

*Corresponding author: airmasadepan@yahoo.co.id, haviluddin@gmail.com

Abstract. A natural hazard information and flood events are indispensable as a form of prevention and improvement. One of the causes is flooding in the areas around the lake. Therefore, forecasting the surface of Lake water level to anticipate flooding is required. The purpose of this paper is implemented computational intelligence method namely Adaptive Neural Network Backpropagation (ANNBP) to forecasting the Lake Cascade Mahakam. Based on experiment, performance of ANNBP indicated that Lake water level prediction have been accurate by using mean square error (MSE) and mean absolute percentage error (MAPE). In other words, computational intelligence method can produce good accuracy. A hybrid and optimization of computational intelligence are focus in the future work.

1. Introduction

A natural hazards and flood phenomenon are part of a natural phenomenon that always exists. For example, some floods are caused by dam damage, landslides including overflowing of lake water. In general, during dry period, a greater possibility of flooding and also river becomes very low. The rising sea level will make it more difficult for the rivers in the delta to drain into the sea. Rivers do not recognize national borders. The flood is a climatological phenomenon that is affected by geology, geomorphology, relief, soil, and vegetation. The meteorology and hydrology processes are actually predictable whether caused by natural factors, humans or both. Therefore, precautions can be evaluated and implemented as well as necessary[1]. Furthermore, all these processes are expected to help communities better understand the steps to reduce risk exposure [2, 3].

Therefore, in order to produce an accurate result of a cascade lake water level forecasting, several methods were developed. Among them, statistical models have been used to create waterfall levels of lakes. This model works with the decline of the data itself (SVE), decomposition, exponential smoothing method (ES) and autoregressive integrated moving average (ARIMA). Several studies have shown that lake water levels are cascaded because the data is not linear. However, in some cases, the falling waterfall method of cascade lakes is also capable of producing good and accurate predictions [4, 5].

Along with the development of computing technology, many researchers are trying to make predictions using the computational intelligence i.e., Fuzzy, neural network, Self-organizing Maps

(SOM) [6]etc. methods in the field of hydrology.Farajzadeh et.al have presented Feed-forward Neural Network and Autocorrelation Regressive Integrated Moving Average (ARIMA) methods in order to forecast the monthly rainfall in Urmia lake basin, Iran. The results showed that these methods could be predicted the rainfall for a 6-year period (2012-2017) [7]. Cramer et al have implemented and compared six machine learning algorithms such as Genetic Programming, Support VectorRegression, Radial Basis Neural Networks, M5 Rules, M5 Model trees, and k-Nearest Neighbors. This studied was used the rainfall time series across data sets for 42 cities (20 cities around Europe and 22 cities around USA). The results showed that machine learning based intelligent system have had accuracy prediction for rainfall [1]. This paper is purpose of prediction of Lake Cascade water level in order to describe flood pattern and its probabilities. In this study, the water level prediction of Lake Cascade Mahakam using Adaptive Neural Network Back-Propagation based on AR model have been performed.

2. Method

The result of observation of a particular event in the past can be expressed in terms of historical data, called time series. Where, this time series data model can be used to generate new data values for forecasting. Meanwhile, the accuracy of forecasting depends on how well the time series data model is built [2].

In this study, the water level of Mahakam Cascade Lake is calculated from sea level as reference point 0. Then, water level data can be categorized as time series data with seasonal and non-stationary pattern. The AR model is declared in Equation 1.

$$y(t) = -a_1 y(t-1) - \cdots - a_n y(t-n) + e(t) \tag{1}$$

Where,a_n is a constant, n is order number of system, and $e(t)$ is a white noise. Thus, Equation 1 can be decomposed in Equation 2.

$$
\begin{aligned}
y(t) &= (a_1 q^{-1} + \cdots + a_n q^{-n})y(t) + e(t) \\
&= -\left(\sum_{i=1}^{n} (a_i q^{-1} + \cdots + a_n q^{-n}) y(t) + e(t) \right)
\end{aligned}
\tag{2}
$$

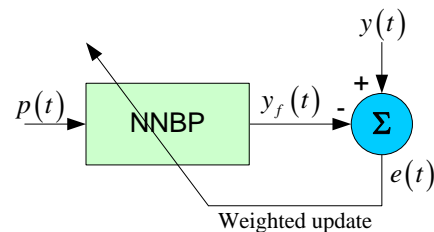Where,q^{-1} is the delay operator, and $A(q^{-1})$ is polynomial to be estimated.

The AR models are approximated by using Feed Forward Neural Network (FFNN) which declared in Equation 3.

$$(t) = N_{ff}(A(q^{-1})y(t)) + e(t) \tag{3}$$

By training $N_{ff}(\dots)$, such that$e(t) \to 0$ then $N_{ff}(\dots) \to y(t)$ is obtained. In its implementation, $e(t)$ is set as small as possible. The FFNN architecture based on AR (Equation 3) is shown in Figure 1.
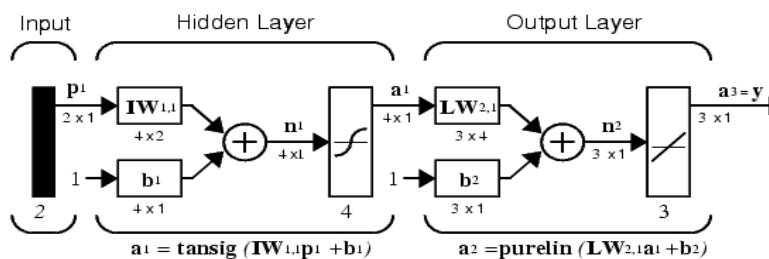
Figure 1. FFNN Architecture [2]   **Figure 2.** NNBP training

In principle, neural network backpropagation (NNBP) is doing back-propagation to fix the weight of each layer such that to achieve target error [3], as shown in Fig. 2. Under certain circumstances, for various reasons such as the amount of data, data patterns (i.e., non-linearity, non-stationary, and discontinuity), and error targets that are too small. Thus, NNBP maybe cannot reach the target error.

Furthermore, ANNBP is NNBP with adaptively weighted adjustment based on random selection both on input weights and on layer weights. The selected weights are updated by adding a very small random number which multiplied by current training error (MSE$_{train}$). The network weights are updated adaptively which conducted continuously to achieve the target error. The training model of ANNBP is shown in Fig. 3. Then, the algorithm as shown in Fig. 4.
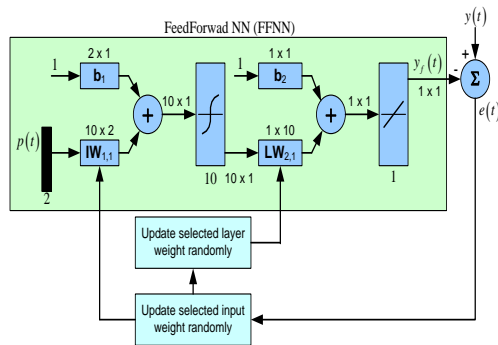
Figure 3. The training model of ANNBP
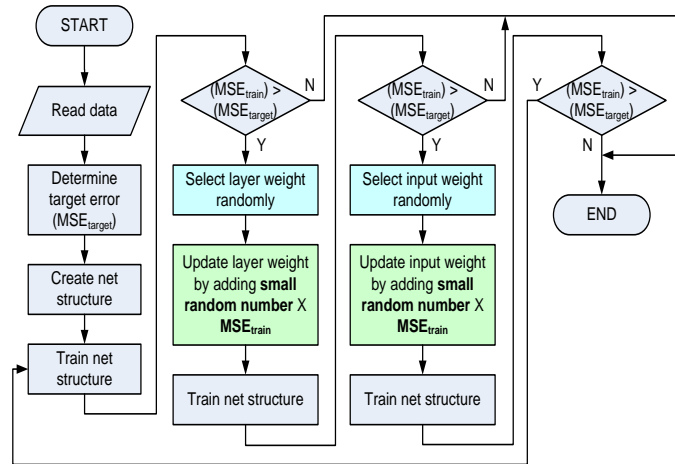
Figure 4. The algorithm of ANNBP

In this study, Lake Cascade Mahakam Water Data, East Kalimantan Province in 1989 – 2016 period or 1008 datasets has been used. The datasets consists of three types: maximum, minimum, and average water levels as shown in Tables 1, 2 and 3.

Table 1. Average Water Level of Lake Cascade Mahakam, years 1989 – 2016.

Years	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1989	1,60	1,09	0,74	0,77	2,10	2,65	3,39	2,75	2,35	1,87	3,16	3,28
1990	4,48	2,43	1,93	2,76	2,48	0,94	-1,09	-0,98	-0,78	-0,69	0,11	-0,19
1991	0,43	0,47	1,71	2,12	4,25	4,84	3,37	0,67	0,03	-0,10	0,53	2,14
1992	2,18	0,95	1,54	1,98	2,14	2,68	1,39	1,78	1,22	1,91	2,63	4,64
1993	4,02	2,21	2,03	3,76	3,45	2,30	1,77	0,25	0,29	0,80	0,88	2,53
1994	3,42	3,31	2,74	3,92	4,89	4,88	3,80	1,70	1,16	0,93	0,97	2,75
1995	3,38	2,79	2,47	3,83	5,01	3,31	1,43	2,14	1,34	2,01	2,46	1,32
1996	2,66	4,65	4,19	3,68	4,33	2,38	1,14	-0,16	2,33	1,22	1,92	4,33
1997	3,84	2,91	1,22	-0,19	-1,32	-2,25	-1,63	-0,66	-1,62	-2,07	-2,17	-2,56
1998	-0,85	0,94	2,38	3,70	3,08	2,71	4,13	4,82	3,43	2,95	2,95	3,22
1999	3,99	2,77	3,02	4,42	4,82	4,11	2,80	1,30	0,06	0,67	1,01	0,09
2000	1,50	2,95	2,08	0,46	-0,86	-1,41	-0,28	0,53	-0,20	1,24	3,14	3,99
2001	2,87	4,22	5,43	3,85	4,68	2,19	-1,62	-1,02	-0,15	2,47	2,78	3,31
2002	4,26	4,79	4,93	5,35	5,20	4,30	2,15	-0,75	0,20	2,38	2,98	3,99
2003	2,98	1,61	0,83	0,75	0,57	1,56	1,04	0,62	0,18	0,46	-0,31	-0,21
2004	1,44	3,87	4,97	3,56	3,69	1,87	1,35	0,49	0,76	0,61	2,31	3,50
2005	2,38	2,59	2,58	5,24	5,54	4,59	2,87	1,54	1,96	2,29	3,92	5,65
2006	4,49	3,19	4,37	3,75	5,51	5,97	3,94	0,93	-0,97	-2,03	-1,60	-0,07
2007	3,19	4,09	2,50	2,96	7,67	6,38	2,57	0,88	1,50	1,08	1,91	2,57
2008	1,94	1,56	2,13	2,21	2,67	2,21	1,83	1,12	2,02	3,31	3,55	3,29
2009	3,96	1,48	2,73	4,22	2,79	-0,16	-0,48	0,62	0,66	0,32	1,95	3,50
2010	4,29	3,94	3,16	1,56	1,20	2,23	0,91	-0,03	1,08	3,36	5,69	4,96
2011	4,01	3,11	3,02	4,14	4,34	4,08	2,79	1,20	-0,04	0,77	1,01	0,12
2012	2,58	2,59	2,88	5,24	5,34	4,59	2,97	1,64	2,06	2,39	4,02	5,55
2013	2,94	2,56	2,13	2,21	2,67	2,21	1,91	1,12	2,02	3,37	3,68	3,35
2014	3,39	3,86	3,66	1,58	1,90	2,23	0,91	0,32	1,28	3,36	5,50	4,92
2015	3,92	1,49	0,51	0,60	0,57	1,38	0,93	0,49	0,35	0,27	-0,31	-0,21
2016	0,12	0,48	2,00	2,25	2,79	0,84	0,95	0,62	0,66	0,32	1,72	2,50

Table 2. Maximum Water Level of Lake Cascade Mahakam, years 1989 – 2016.

Years	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1989	2,74	2,51	0,98	1,14	2,82	3,32	4,01	3,71	2,81	2,45	4,01	4,32
1990	4,87	3,83	2,57	3,11	2,96	1,87	-0,11	-0,53	-0,12	0,00	0,49	0,56
1991	1,23	1,98	2,12	3,12	5,56	5,57	4,95	1,99	0,47	0,37	3,13	2,80
1992	4,49	4,79	5,43	5,35	7,67	6,38	4,13	4,82	3,43	3,37	5,69	5,65
1993	5,43	2,85	2,70	4,65	4,43	2,81	2,19	1,97	0,60	1,08	2,03	2,99
1994	3,76	3,72	3,21	4,51	5,18	5,13	4,61	2,88	1,75	1,50	1,57	3,51
1995	3,70	3,09	3,12	4,79	5,47	4,26	2,31	2,45	2,13	2,97	3,17	2,04
1996	3,99	5,27	5,22	4,35	5,09	2,89	1,92	1,50	2,94	2,00	3,13	5,05
1997	4,64	3,29	2,30	0,15	-0,76	-1,85	-0,99	-0,38	-1,12	-1,62	-1,59	-2,16
1998	0,15	1,53	3,13	3,90	3,69	3,37	4,63	5,09	4,20	3,23	3,37	3,76
1999	4,46	3,46	3,73	4,93	5,08	4,76	3,32	2,02	0,49	1,31	1,52	3,54
2000	3,62	3,27	2,86	1,18	-0,26	1,97	0,21	0,80	0,23	2,11	4,21	4,79
2001	3,43	5,03	5,98	4,72	5,50	4,23	-0,34	-0,18	0,78	3,01	3,45	3,82
2002	4,79	5,09	5,13	5,75	5,44	4,79	3,49	-0,18	1,27	2,85	4,34	4,79
2003	3,42	2,45	1,08	1,26	0,80	3,62	1,35	1,26	0,86	0,80	-0,15	1,50
2004	2,84	4,82	5,30	3,96	4,59	3,37	3,08	3,13	1,00	1,21	3,65	3,71
2005	2,69	2,82	2,74	6,42	6,05	5,26	3,51	2,22	2,35	2,45	5,93	5,99
2006	5,14	4,82	5,20	4,57	6,26	6,37	6,28	1,84	0,01	-1,68	-1,01	1,59
2007	4,08	4,40	3,27	5,66	8,68	7,96	4,32	1,67	2,22	1,33	3,02	3,19
2008	2,43	1,95	2,50	2,75	3,13	2,50	2,14	1,44	2,50	4,18	4,28	4,23
2009	4,50	2,66	81,94	4,96	4,93	0,38	0,17	1,15	1,13	1,00	2,90	5,04
2010	4,64	4,73	4,04	4,01	2,00	2,62	1,62	0,43	2,13	4,30	6,90	6,35
2011	2,38	2,59	2,58	5,24	5,54	4,59	2,87	1,54	1,96	2,29	3,92	5,65
2012	4,49	3,19	4,37	3,75	5,51	5,97	3,94	0,93	-0,97	-2,03	-1,60	-0,07
2013	3,19	4,09	2,50	2,96	7,67	6,38	2,57	0,88	1,50	1,08	1,91	2,57
2014	1,94	1,56	2,13	2,21	2,67	2,21	1,83	1,12	2,02	3,31	3,55	3,29
2015	3,96	1,48	2,73	4,22	2,79	-0,16	-0,48	0,62	0,66	0,32	1,95	3,50
2016	4,29	3,94	3,16	1,56	1,20	2,23	0,91	-0,03	1,08	3,36	5,69	4,96

Table 3. Minimum Water Level of Lake Cascade Mahakam, years 1989 – 2016

Years	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1989	0,57	0,48	0,44	0,41	1,21	1,85	2,58	1,83	1,64	1,01	2,53	2,18
1990	3,82	-0,37	1,25	2,15	1,65	-0,06	-1,52	-1,37	-1,34	-1,05	-0,39	-0,73
1991	-1,03	-1,13	1,30	1,74	2,22	4,17	2,24	-0,36	-0,50	-0,73	-0,27	1,28
1992	1,72	0,15	0,30	0,44	1,15	1,90	0,46	0,56	0,82	1,08	1,52	3,88
1993	2,07	1,42	1,36	2,78	2,52	1,83	1,31	-0,84	0,00	0,40	0,31	2,13
1994	3,05	2,80	2,32	3,27	4,53	4,67	2,95	0,85	0,71	0,56	0,56	0,94
1995	2,87	2,30	1,86	3,17	4,46	2,40	0,84	1,56	0,87	1,28	1,70	0,76
1996	0,80	3,80	3,09	3,10	3,20	1,94	-0,50	-0,52	0,79	0,76	1,04	3,23
1997	2,91	2,40	0,15	-0,70	-1,80	-2,47	-2,21	-1,04	-2,26	-2,45	-2,77	-2,80
1998	-1,39	0,27	1,61	3,25	2,35	2,15	3,48	4,30	2,60	2,68	2,53	2,53
1999	3,18	2,15	2,26	3,83	4,61	3,29	2,14	0,56	-0,29	0,16	0,32	-0,33
2000	0,71	2,34	1,30	-0,22	-1,58	-1,83	-0,93	0,18	-0,53	0,35	2,22	3,34
2001	2,35	3,51	4,15	3,38	3,83	-0,31	-2,51	-3,47	-0,75	1,03	2,14	2,86
2002	3,70	4,35	4,69	4,15	4,75	3,54	0,82	-1,65	-0,75	1,40	2,03	3,34
2003	2,50	0,71	0,52	0,32	0,31	0,80	0,21	-0,05	-0,56	-0,13	-0,53	-0,52
2004	0,26	2,86	4,04	2,76	2,60	0,12	0,16	-0,72	0,43	0,10	0,26	3,20
2005	2,03	2,39	2,33	2,68	5,30	3,55	2,29	1,21	1,33	2,08	2,48	5,17
2006	3,50	1,91	3,19	3,05	4,63	5,68	1,87	-0,12	-1,86	-2,16	-4,90	-1,34
2007	1,65	3,38	1,61	1,46	5,81	4,44	1,21	0,48	0,70	0,78	0,97	2,20
2008	-1,36	1,27	1,42	1,81	2,20	1,90	1,20	0,87	1,49	2,53	2,61	2,40
2009	2,73	0,74	0,88	2,99	0,44	-1,01	-1,04	0,19	0,24	0,00	1,04	3,04
2010	3,79	3,34	2,22	0,98	0,74	1,65	0,14	-0,44	0,11	2,20	4,32	5,12
2011	2,91	2,01	1,92	3,04	3,24	2,98	1,69	0,10	-1,15	-0,33	-0,09	-0,98
2012	1,48	1,49	1,78	4,14	4,24	3,49	1,87	0,54	0,96	1,29	2,92	4,45
2013	1,84	1,46	1,03	1,11	1,57	1,11	0,81	0,02	0,92	2,27	2,58	2,25
2014	2,29	2,76	2,56	0,48	0,80	1,13	-0,19	-0,78	0,18	2,26	4,40	3,82
2015	2,82	0,39	-0,59	-0,50	-0,53	0,28	-0,17	-0,61	-0,75	-0,83	-1,41	-1,31
2016	-0,98	-0,62	0,90	1,15	1,69	-0,27	-0,15	-0,48	-0,44	-0,78	0,62	1,40

In this study, all data is compiled into time series data. Then, the AR with 240 model has been applied using Equation 1, as stated.

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) \ldots - a_{240} y(t-240) + e(t)$$

Then, by using FFNN is stated.

$$N_{ff} = \big(y(t-1), y(t-2), \ldots, y(t-240)\big) \rightarrow y(t)$$

Where, $y(t$-1$)$ is water level data in 1989 – 1994; $y(t$-2$)$ is water level data in 1990 – 1995; $y(t$-240$)$ is water level data in 2003 – 2008 as a training data; $y(t)$ is water level data in 2009 – 2014 as a testing data.   Before the ANNBP process, the normalized datasets into the interval [-1 ... 1] has been performed using Equation 4.

$$X_n(i) = \frac{2x(X(i) - X_{min})}{X_{max} - X_{min}} - 1 \qquad (4)$$

In this study, the ANNBP net structure by using five hidden neurons and error target $e(t) = 10^{-4}$ have been explored. Then, the performance of training result by using MSE (Mean Squared Error), Equation 5 and MAPE (Mean Absolute Error), Equation 6 have been implemented.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y(t_1) - y_{ff}(t_i))^2 \qquad (5)$$

$$APE_i = \frac{abs(y(t_i) - y_{ff}(t_i))}{y(t_i)} x100\%$$

$$\qquad (6)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} APE_i$$

Where, N is the number of training data; $y(t_i)$ is the ith training target; $y_{ff}(t_i)$ is the ith output of adaptive NNBP. In this experiment, MATLAB for ANNBP have been used.

3. Result and discussion

In this section, the water level prediction of Lake Cascade Mahakam by using Adaptive Neural Network Backpropagation (ANNBP) models is presented.

The time series is a dataset of observations ordered in time. A time series is an ordered sequence of observations and many ways are used to forecast the time series data. In this experiment, Lake Cascade Mahakam Water Datasets in 1989 – 2016 period or 1008 datasets consists of three types: maximum, minimum, and average water levels has been used.

Based on experiment, average performance training data (2009-2014) by using MSE of 8.78×10^{-4} and MAPE of 1.17% and performance testing data (2015-2016) MSE of 2.16 and MAPE of 3.73% of average water levels have been performed, Fig. 5, 6. Then, maximum performance training data (2009-2014) by using MSE of 1.14×10^{-1} and MAPE of 0.53% and performance testing data (2015-2016) MSE of 3.25 and MAPE of 18.8% of average water levels have been settled, Fig. 7, 8. Finally, minimum performance training data (2009-2014) by using MSE of 2.7×10^{-3} and MAPE of 0.77% and performance testing data (2015-2016) MSE of 2.25 and MAPE of 6.59% of average water levels have been settled, Fig. 9, 10. The next process, each ANNBP net structure 240-5-1 for each type of data that has been trained, then to predict the water level of 2017-2018 as shown in Fig. 11.
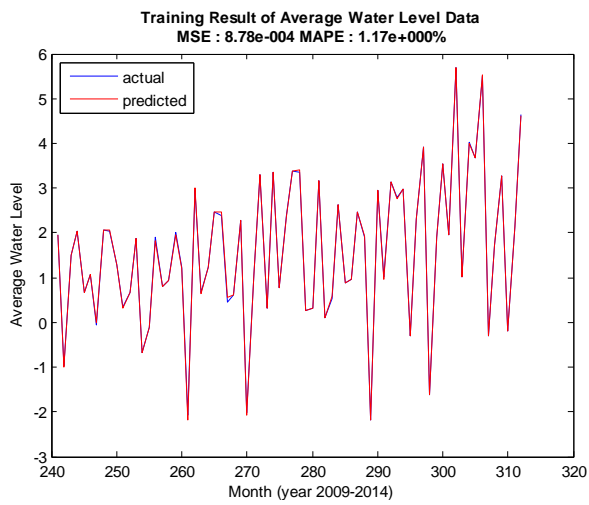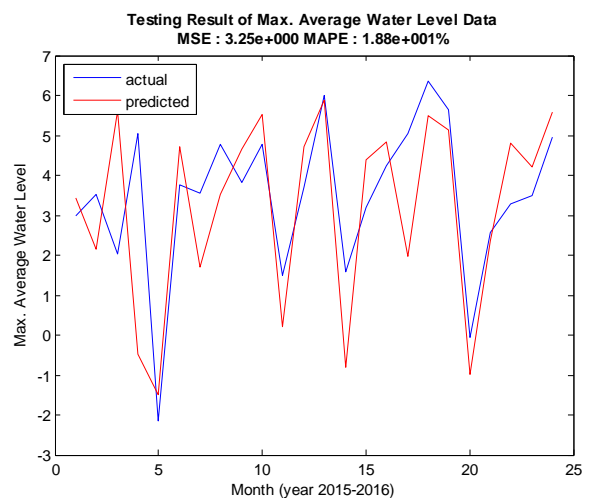
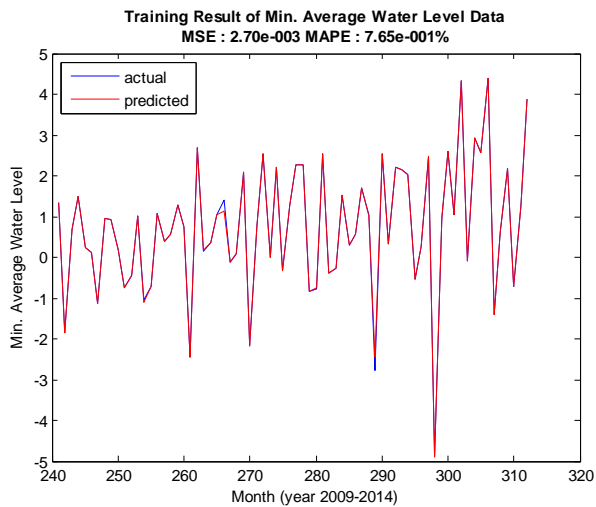**Figure 5.**Training Result of Average Water Level Data (year 2009-2014)
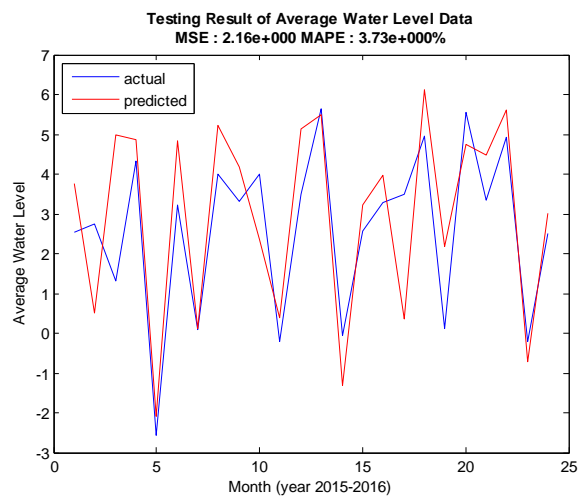
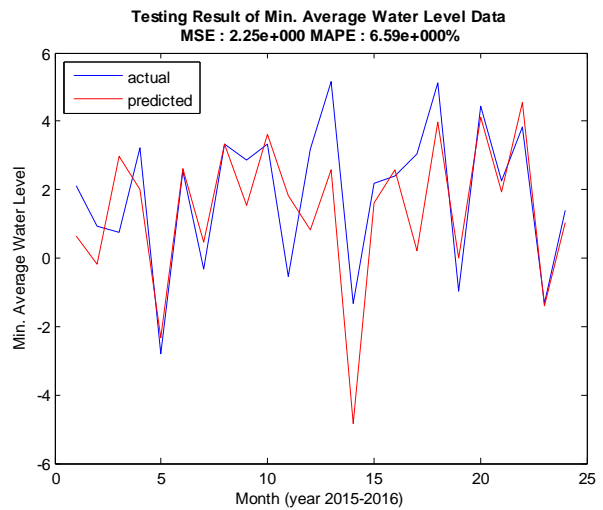**Figure 6.** Testing Result of Average Water level Data (year 2015-2016)
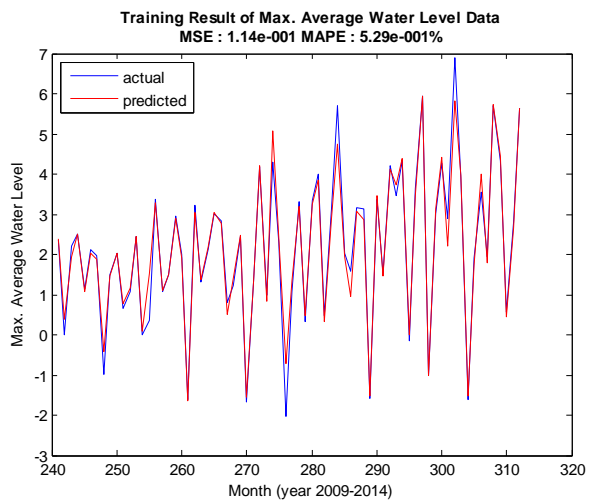
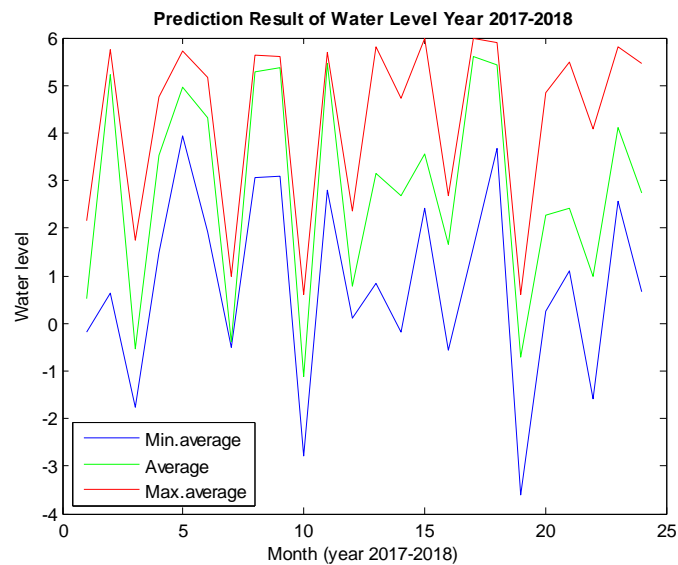**Figure 7.** Training Result of Max.Water Level Data(year 2009-2014)

Figure 8. Testing Result of Max. Water level Data (year 2015-2016)

Figure 9. Training Result of Min.Water Level Data(year 2009-2014)

Figure 10. Testing Result of Min. Water level Data (year 2015-2016)

Figure 11. Prediction Result of Water level (year 2017-2018)

4. Conclusion

In this paper, ANNBP algorithm has been implemented and explored to model and predict water level of Lake Cascade Mahakam in East Kalimantan -Indonesia. After testing, the best MSE and MAPE values obtained were 9.7% with 240-5-1 NNBP architecture. In other words, this study have demonstrated that ANNBP models can be used as a predictive algorithm that provides a good predictive accuracy of water level Lake Cascade Mahakam. Future work is suggested to hybrid and optimization of a few computational intelligence methods in order to obtain more accurate prediction results.

References

[1] Cramer S, Kampouridis M, Freitas A A, and Alexandridis A K, 2017 An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives *Expert Systems with Applications* **85**, pp 169-81,.

[2] Mislan, Haviluddin, S Hardwinarto, Sumaryono, and M Aipassa, 2015 Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggarong Station, East Kalimantan - Indonesia," in Int Conf. on Computer Science and Computational Intelligence (ICCSCI 2015), Jakarta, Indonesia pp. 142-51.

[3] R Islam, R Kamaruddin, S A Ahmad, S J Jan, and A R Anuar, 2016 A Review on Mechanism of Flood Disaster Management in Asia International Review of Management and Marketing,.

[4] J. Wu, J. Long, and M. Liu, 2015 Evolving RBF neural networks for rainfall prediction using hybrid particle swarm optimization and genetic algorithm Neurocomputing **148** pp 136–42,.

[5] Haviluddin and A Jawahir 2015 Comparing of ARIMA and RBFNN for short-term forecasting," International Journal of Advances on Intelligent Informatics (IJAIN) **1**(1) pp. 15-22,

[6] Haviluddin, A. Yunianta, A. H. Kridalaksana, Z. Arifin, B. Kresnapati, F. Rahman, A. F. O. Gaffar, H. Y. Irawan, M. Mulyo, and A. Pranolo, "Modelling of Network Traffic Usage Using Self-Organizing Maps Techniques," in 2016 2nd International Conference on Science in Information Technology (ICSITech) pp 334-8.

[7] J Farajzadeh, A F Fard, and S Lotfi 2014 Modeling of monthly rainfall and runoff of Urmia lake basin using "feed-forward neural network" and "time series analysis" model Water Resources and Industry **7-8** pp. 38–48