

การประยุกต์ใช้โครงข่ายประสาทเทียมเพื่อการทำนายระดับน้ำในลุ่มน้ำชี

Application of ANN to Prediction of Water Level in Chi River Basin

ธีรวงศ์ เหล่าสุวรรณ¹, พรเมษฐ์ จันทร์เพ็ง², ณัชชา เหล่าสุวรรณ³

Teerawong Laosuwan¹, Poramate Chunpang², Nutch Laosuwan³

บทคัดย่อ

วัตถุประสงค์ของการศึกษานี้เพื่อศึกษาการคาดการณ์ระดับน้ำในลุ่มน้ำชีโดยการใช้ข้อมูลปริมาณน้ำฝนที่ยังไม่สมบูรณ์ด้วยการใช้โครงข่ายประสาทเทียม (ANN) การศึกษาได้ดำเนินการในลุ่มน้ำชีในภาคตะวันออกเฉียงเหนือของประเทศไทย ด้วยการรับรู้ในหลายระดับ (Multilayer Perception: MLP) พร้อมด้วยคุณลักษณะของการแก้ไขกระบวนการซ้ำของโมเดล ANN ที่ได้รับการพัฒนาขึ้นมาใหม่ โดยข้อมูลปริมาณน้ำฝนที่ไม่สมบูรณ์นั้นจะสามารถที่จะมีความแม่นยำถึง 96.6 % การคาดการณ์ปริมาณน้ำฝนที่ไม่สมบูรณ์นั้นมีการนำมาใช้ในขั้นตอนปกติและการคาดการณ์ปริมาณระดับน้ำที่มีอยู่ ความแม่นยำในระดับ 84.9%ของการคาดการณ์ระดับน้ำจากวิธีการคาดการณ์แบบใหม่ ในขณะที่ความแม่นยำที่ระดับ 6.99 %ที่ได้จากกระบวนการเดียวกันโดยไม่มีการใช้ข้อมูลที่หายไป ผลการศึกษาครั้งนี้แสดงให้เห็นว่า ANN เป็นเครื่องมือที่มีประสิทธิภาพในการคาดการณ์ระดับน้ำไม่ว่าข้อมูลน้ำฝนนั้นจะสมบูรณ์หรือไม่ก็ตาม โดยโมเดลที่ได้รับการพัฒนาจากการศึกษานี้สามารถที่จะประยุกต์ใช้งานสำหรับนักอุทกศาสตร์ต่อไปในอนาคต

คำสำคัญ:

ว วิทย เทคโนโลยี มมส

Abstract

The objective of this study is to prediction of water level in Chi River basin under incomplete rainfall data by using Artificial Neural Network (ANN). The study has been carried out at Chi River Basin, located on Northeast, Thailand. By using two-layer Multilayer Perceptron (MLP) with modified backpropagation feature of newly developed ANN model, the missing rainfall data could be estimated with 96.6% accuracy. The estimated missing data was then used in the regular process and predicted water levels were obtained. An 84.9% accuracy of water level predictions was achieved from the process with estimated missing data while only 6.99% accuracy was obtained from the same process without missing data. It can be suggested from the obtained results that the ANN is still an effective tool in water level prediction no matter what the rainfall data is complete or not. The model developed in this study should be one of high potential methods used among worldwide hydrologists in future.

Keyword:

J Sci Technol MSU

¹ผู้ช่วยศาสตราจารย์, ²อาจารย์, ³เจ้าหน้าที่, ภาควิชาฟิสิกส์, คณะวิทยาศาสตร์ มหาวิทยาลัยมหาสารคาม อำเภอกันทรวิชัย จังหวัดมหาสารคาม 44150

^{1,2}Lecturer, ³Officer, Department of Physics, Faculty of Science, Mahasarakham University, Kantarawichai, Mahasarakham 44150

*Corresponding author: teerawong laosuwan, Department of Physics, Faculty of Science, Mahasarakham University, Kantarawichai,

Mahasarakham 44150. E-mail: teerawong@msu.ac.th

1. Introduction

Thailand possesses twenty-five main basins and three of them located in Northeastern area. Chi River basin is interesting one. This basin is main part of Chi River; the longest river in Thailand (about 765 km long), originates from Phetchabun mountains and flows through many provinces such as Chaiyaphum, Khon Kaen, Maha Sarakham, Roi Et, Yasonthon until joins the Mun river at Sisaket provinces. The river has been estimated to carry approximately 9,300 cubic kilometers of water per year. It covers an area of about 49 square kilometers. The average rainfall in this area was about 1,200 mm per year (range from 1,000 to 1,600 mm/year). It was known that the amounts of rainfall were obviously influenced by the northeast and southwest monsoons which cause rain in this area during May to October every year. It was known that floods always occurred in this area during heavy rainfall periods which the significantly high rainfall amounts have been estimated as 2-3 years interval.

Two kinds of flood (overflow and stagnate) usually appeared at upstream and downstream, respectively. Cause of overflow flood is that amounts of influence water is very much higher than effluence and accumulated water is over than basin capacity while the cause of stagnate flood is that the effluence water is obstructed by high water level of river end (Mun river) which the effluence cannot be drained out adequately. The floods around the Chi River basin are annually occurred and cause many disasters both to residential and agricultural areas, in particular most of plain lands. The flood also reduces any agricultural production even in survive areas. In addition to property and life and farm damages; the floods also destroy the forest in form of diversity, population, and ecology.

Together with artificial deforestation, losing of forest also causes next flooding more and more sever. Moreover, there are many constructions, especially transportation structures, rise along the Chi River. These constructions are known to obstruct water drainage and reduce drainage efficiency of the river. From historical information, the most severe floods occurred in 1978, 1995, 2000, and 2001.

In order to protect or relieve the disaster, accurate water level forecasting may be vital solution. It will allow people more time to protect their properties, escape to safer place, and avoid as much damage as possible. Anyway, there was also forecasting procedure available so far. However, it worked with quite low accuracy due to incomplete water rainfall data existed. This because the existing forecasting system composts of eleven rainfall stations and only one water level station distributed over the basin. So, the complete water rainfall values were not possible to measure with reliable amounts. The objective of this study is to establish water level prediction under incomplete rainfall data using Artificial Neural Network (ANN). The location of the basin is depicted in Figure 1.



Figure 1. The study area

2. Methodology

Artificial neural networks are a computational tool that is able to capture and represent complex input/output Relationships^{1,2,3,4,5}. Thus, these properties of ANN are well suited to the problem of weather forecasting^{6,7,8,9}. In briefly, the algorithm in three steps was used to develop the ANN model for water level forecasting under incomplete rainfall data. Firstly, the rainfall and water levels were predicted using conventional ANN procedure. Secondly, new ANN model was developed. Finally, the water level was predicted again using such developed ANN model. All three steps were described below.

2.1 Rainfall and water level prediction using ANN model

The first step began with ANN model selecting. A two-layer Multilayer Perceptron (MLP) backpropagation network^{10,11,12} (see in Figure 2) was chosen in this study.

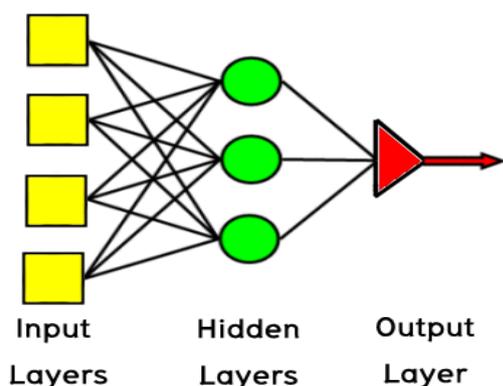


Figure 2. The Multilayer Perceptron

The network works by feeding input data to input nodes. After processed in hidden nodes which the data was multiply by a certain weight, bias would be excluded from data and results were sent to normalize using a non linear transfer function in order to adjust the interconnection weights using proper training algorithm. At this

point, different transfer functions, tangent sigmoid for backpropagation procedure with in hidden layer and pure linear transfer for output layer, were also used in hidden and output layers in order to try out for the best ANN structure which would be used over the study. The same operation was repeated again at the output node. This is a kind of popular backpropagation learning methods mentioned artificial intelligent areas and this simulation technique has been proven as one of outstanding forecasting performance in hydrology area and favor for most hydrologists^{13,14,15,16}. All calculations were carried out by MATLAB program^{17,18,19,20,21}. Eventually, all network performances were measured in form of correlation coefficient (R).

$$R = \frac{\sum_{i=1}^{i=n} \left(\left[Q_m - \bar{Q}_m \right] \left[Q_s - \bar{Q}_s \right] \right)}{\sqrt{\sum_{i=1}^{i=n} \left[Q_m - \bar{Q}_m \right]^2 \sum_{i=1}^{i=n} \left[Q_s - \bar{Q}_s \right]^2}} \quad (1)$$

Where;

- Q_m = Actual value of rainfall/ water level
- Q_s = Simulated value of rainfall/ water level
- \bar{Q}_m = Average of actual value of rainfall/ water level
- \bar{Q}_s = Average of simulated value of rainfall/ water level

2.2 Prediction of rainfall missing

All rainfall data was recorded at the same gauging station located beside Chi River. Two techniques were used to predict for data missing, normal ration and ANN. Normal ratio method began with estimations of rainfall missing values from concurrent observations of the sites that both close and away from the station (see in equation 2).

$$\frac{P_x}{N_x} = \frac{1}{n} \left(\frac{P_1}{N_1} + \frac{P_2}{N_2} + \frac{P_3}{N_3} + \dots + \frac{P_n}{N_n} \right) \quad (2)$$

Where;

P_x	=	Missing rainfall for station X
P_1, P_2, P_3, P_n	=	Rainfall at neighboring station for the concurrent period
N_x	=	Normal long-term rainfall at station X
N_1, N_2, N_3, N_n	=	Normal long-term rainfall for neighboring station

The ratio of normal annually rainfalls were obtained from interpolation and index values (weighting factor). This technique gave more satisfy value than formally straight average technique. Artificial Neural Network (ANN) developed in this study was trained with different learning algorithms, learning rates, and number of neurons in hidden layer to optimize the condition and obtain more reliable results. There are three different backpropagation learning algorithms involved in the training procedure; Resilient Backpropagation (trainrp), Fletcher-Reeves Conjugate Gradient (traincgf), and Scale Conjugate Gradient (trainscg). The Resilient Backpropagation (trainrp) bases on the approach that the effects of small magnitude gradient will be eradicated from inputs as the magnitudes have very least effects on weighted data.

Only the weighted data was used to determine the data direction by using sigmoid transfer function. The trainrp algorithm looks faster and easier than the standard steepest descent algorithms. For Fletcher-Reeves Conjugate Gradient (traincgf), this algorithm converges in

less iteration than trainrp does, but each of iteration requires more complicate processing. In generally, the conjugate gradient algorithms usually take shorter time then variable learning rate backpropagation in calculation and they can be even faster than trainrp. Moreover, this algorithm needs only smaller storage space than other simple algorithms so that it is usually used in networks with high throughput weighting tasks. For the third algorithm used in hidden layer, Conjugate Gradient (trainscg), it has been designed to avoid line searching which is time consuming step. Thus, this point make it differs from other conjugate gradient algorithms which need line searching in their iteration procedures. However, the routine of trainscg usually requires more iteration to converge, but this may be compensated by significantly decrease in number of computation in each iteration.

In addition to three algorithms mentioned above, the process inside the hidden layer also composted of learning rate applications which were ranged from 0.20 to 0.80. This setting would affect overall network during the simulation took place. The number of neurons in the hidden layer was adjusted after trial and error step in order to find out an adequate number of neurons in supporting or withhold complex input elements of hidden layer. The proper numbers were 20, 40, 60, 80, and 100. Next, daily rainfall data from year 2007 to 2011 was trained and test. Finally, the missing rainfall data (of one station interpolated with other 5 stations nearby at specified date) was obtained from the output. The model equation of this study was shown in equation 3.

$$P_x(t) = f \left[P_1(t), P_2(t), P_3(t), P_4(t), P_5(t) \right] \quad (3)$$

Where;

$$\begin{aligned}
 P_x &= \text{Rainfall at station X (missing)} \\
 P_1, P_2, P_3 &= \text{Rainfall at 5 neighboring} \\
 P_4, P_5 &= \text{stations}
 \end{aligned}$$

2.3 Water level prediction

A model used to modify ANN was described elsewhere. The input node consists of antecedent water level, antecedent rainfall, and current rainfall processors (see in equation 4).

$$W(t) = f \left[\begin{array}{l} P(t-4), P(t-3), P(t-2), P(t-1), \\ P(t), W(t-4), W(t-3), W(t-2), \\ W(t-1), W(t) \end{array} \right] \quad (4)$$

Where;

t = Time (days)

P = Rainfall

W = Water level

The water levels of current dates were the expected outputs. The rainfall missing values previously estimated were used to predict the water level of Chi River. Three different set of input data were developed at this step and designated as Set A, B, and C. The estimated missing data of Set A were predicted by using the normal ratio method and incorporated into rainfall data which was used to estimate the water levels in developed ANN model. In Set B, the developed ANN model was used to estimate both rainfall missing and water levels from whole data set. For Set C, all missing data was excluded and only remained data was used in prediction.

Predicted missing rainfall values earlier are used to predict water level of Chi River Basin. A total of 3 different sets of input data were developed, namely Set A, B and C. For set A, missing rainfall data are predicted using Normal

ratio method. Data for set B utilized ANN to predict its missing rainfall, where as in set C, predictions were not made to missing data. Parameters of the recommended network from Bessaih¹² are tabulated in Table 1.

Table 1 Recommended ANN model for daily water level prediction from

No. of nodes in hidden layer	20
Antecedent time	5 days
Learning Algorithm	Trainscg
Learning Rate	0.80

3. Result and Discussion

3.1 Prediction of rainfall Missing

The trainrp was proven in this study as the best algorithm in prediction of rainfall missing under given simulation (see in Table 2).

Table 2 Comparison of different learning algorithm Used

Learning Algorithm	R (Training)	R (Testing)
Trainrp	1	0.966
Traincgf	1	0.903
Trainscg	1	0.951

The prediction result gave high correlation value during simulation process. No significant effect on simulation process was found from various learning rate introduction into the network system. All given values could be able to converge in both training and testing phases. Nevertheless, the simulation using larger learning rates caused the convergence process slow down and was regarded as unsatisfactory running. Thus, the smallest learning rate suitable to the network is 0.20. The simulation of rainfall missing estimation using ANN was dealt with 20 to 100

neurons within hidden layer and the result indicated that ANN fed with 60 neurons gave the highest correlation value (see in Table 3 and Table 4). It was found that the backpropagation feature of ANN model developed in this study worked very well in the simulation as the rainfall missing prediction gave up to 96.6% accuracy.

Table 3 Optimum parameters for missing rainfall prediction

No. of nodes in hidden layer	60
Learning Algorithm	Trainrp (Resilient Backpropagation)
Learning Rate	0.20

Table 4 Influence of number of neurons in hidden Layer

Number of Neurons	R (Training)	R (Testing)
20	0.996	0.915
40	0.998	0.896
60	1	0.966
80	1	0.898
100	1	0.911

3.2 Prediction of water levels

Both estimation of rainfall missing and prediction of water levels using developed ANN model (Set B) gave highest correlation value (0.849) followed by Set A (0.831). The lowest correlation value was found in Set C (0.699) (see in Table 5). The developed ANN model worked well at least for Chi River model in studied time period. Furthermore, the Set B model also gave highest correlation values in water level prediction. It was suggest that the incomplete dataset modification by incorporation of estimated missing data improve whole dataset reliability and usability

as higher correlation value obtained. It was clearly seen that the ANN model could better predict both missing values and water levels than normal ratio method. Eventually, the simulation under estimated missing values excluding gave inferior result than the simulation with such values existing.

Table 5 Performance of ANN in water level estimation of Sets A, B and C

Set	R (Training)	R (Testing)
A	1	0.831
B	1	0.849
C	1	0.699

4. Conclusions

It can be concluded from the results that ANN model developed in this study was proven as reliable and effective tool in forecasting. It not only works efficiently in rainfall-missing data estimation but also gives high accuracy in water level prediction at least for Chi River model. It can be said that this new developed ANN model is an important step for advances of the water resource engineering.

Reference

1. Rosmina Bustami, Nabil Bessaih, Charles Bong, Suhaila Suhaili, "Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River" IAENG International Journal of Computer Science, 34:2, 2007
2. Tokar, A.S., and Johnson, P.A., Rainfall-runoff modeling using Artificial neural networks, Journal of Hydrologic Engineering, 4(3), 1999, pp. 223-239.

3. ASCE Task Committee on Application of Artificial neural networks in Hydrology, "Artificial neural networks in hydrology (I & II)", *Journal of Hydrologic Engineering*, 5(2), April 2000, pp. 115-136.
4. Demuth, H., and Beale, M., *Neural Network Toolbox For Use with MATLAB*. Natick, MA: The Mathworks Inc., 2001.
5. Sobri, H., Amir Hashim, M.K., and Nor Irwan, A.N., "Rainfall-runoff modeling using artificial neural network", *Proceedings of 2nd World Engineering Congress*. Kuching, 2002.
6. Thirumalaiah, K. and Deo, M.C., Hydrological forecasting using artificial neural network, *Journal of Hydrologic Engineering*, 5(2), 2000, pp. 180-189.
7. Dastorani, M.T., and Wright, N.G., Artificial neural network based real-time river flow prediction, in the *Proceedings of Hydroinformatics 2002*.
8. Arvind Sharma and Manish Manoria, "A Weather Forecasting System using concept of Soft Computing," pp.12-20, 2006
9. MohsenHayati, and Zahra Mohebi, "Application of Artificial Neural Networks for Temperature Forecasting," *World Academy of Science, Engineering and Technology* 28, 2007.
10. Raúl Rojas, "The back propagation algorithm of Neural Networks-A Systematic Introduction," chapter 7, ISBN 978-3540605058.
11. Neaupane, K.M. and Achet, S.H. "Use of backpropagation neural network for landslide monitoring," *Engineering Geology* 74, 2004, pp. 213-226.
12. Gill J. et al "Training back propagation neural networks with genetic algorithm for weather forecasting" *IEEE 8th international symposium on intelligent systems and informatics serbia*, 2010.
13. Bessaih, N., Bustami, R., and Saad, M., Water level estimation for Sarawak River", *Proceeding of Rivers' 04*, USM, Penang, September 2004.
14. Wright, N.G. and Dastorani, M.T. "Effects of river basin classification on artificial neural networks based ungauged catchment flood prediction", in the *Proceedings of the International Symposium on Environmental Hydraulics*. Phoenix, 2001.
15. Wright, N.G., Dastorani, M.T., Goodwin, P., and Slaughter, C.W., A combination of artificial neural networks and hydrodynamic models for river flow prediction, in the *Proceedings of Hydroinformatics 2002*.
16. Dolling, O.R. and Varas, E.A. Artificial neural networks for streamflow prediction, *Journal of Hydraulic Research*. 40(5), 2002, pp. 547-554.
17. Garcia-Bartual, R., Short term river flood forecasting with neural networks, *Proceedings of iEMSs 2002*, Switzerland, 2002.
18. Mc.Cuen, R.H., *Hydrologic Analysis and Design* 2nd Ed. Upper Saddle River, N.J: Prentice Hall, 1997.
19. Gupta, R.S., *Hydrology and Hydraulic Systems*. New Jersey: Prentice Hall, 1989.
20. Haykin, S., *Neural Networks. A comprehensive Foundation*. Upper Saddle River, N.J.: Prentice Hall, 1994.
21. Mathur S. et al "A feature based neural network model for weather forecasting" *World academy of science, engineering and technology* 34, 2007.