Eng. & Tech. Journal, Vol. 27, No. 13, 2009

Prediction of Tigris River Stage in Qurna, South of Iraq, Using Artificial Neural Networks

Ali H. Al Aboodi* Ammar S. Dawood* & Sarmad A. Abbas*

Received on:31/8 /2008
Accepted on:2/4/2009

Abstract

Artificial neural networks (ANNs) with back-propagation algorithm are performed for predicting the stage of Tigris River in Qurna city, Basrah, south of Iraq. This model was adopted to investigate the applicability of ANNs as an effective tool to simulate the river stage for short term. By using the neural network toolbox in Matlab R2007b, three models are constructed as the first experiment. Multilayer percpetron with one hidden layer is used in the architecture of network. The best model is selected according to the trial and error procedure based on three common statistic coefficients (coefficient of correlation, root mean square error, and coefficient of efficiency). The best model from first experiment is used to predict the stage river for one, two, and three days ahead as the second experiment. Results indicated the ANNs with back-propagation algorithm are a powerful technique to predict the short term stage of Tigris River.

Keywords: Prediction, Tigris, Stage, Neural, Networks

التنبؤ بمنسوب نهر دجلة في القرنة, جنوب العراق باستعمال الشبكات العصبيةِ الاصطناعية

الخلاصة

تمت تهيئة الشبكات العصبية الصناعية باستخدام طريقة انسياب الخطأ لأعداد نموذج التنبؤ بمنسوب نهر دجلة في قضاء القرنة, محافظة البصرة, جنوب العراق. استخدم النموذج المعد للتحري عن إمكانية الشبكات العصبية الصناعية كأداة فاعلة في محاكاة منسوب النهر للمدى القصير باستعمال صندوق عدة الشبكات العصبية في برنامج (Matlab R2007b). ثلاثة نماذج أنشأت كتجربة أولية حيث استخدمت الشبكة العصبية من النوع متعدد الطبقات لبناء معمارية الشبكة. أفضل نموذج اختير طبقا لعملية التجربة والخطأ مستندا على ثلاث معاملات إحصائية معروفة هي (معامل الارتباط, جذر معدل الخطأ ألتربيعي, ومعامل الكفاءة). أفضل نموذج انبثق كحصيلة لنتائج المرحلة الأولى, استعمل لتنبؤ بمنسوب النهر ليوم, يومين, وثلاثة أيام لاحقة. بينت النتائج, إن الشبكات العصبية الصناعية المعايرة باستعمال طريقة انسياب الخطأ تقنية كفؤءة للتنبؤ بمنسوب نهر دجلة للمدى القصير.

* Civil Engineering Department, University of Technology / Baghdad

https://doi.org/10.30684/etj.27.13.13
2412-0758/University of Technology-Iraq, Baghdad, Iraq

1-Introduction

The term stage refers to the height of a river (or any other body of water) above a locally defined elevation. This locally defined elevation is a reference level, often referred to as datum. For example, for Tigris River, reference level or datum is sea level (0.0 m).

When the discharge of a river increases, the channel may become completely full. Any discharge above this level will result in the river overflowing its banks and causing a flood. The stage at which river will overflow its bank is called bankfull stage or flood stage. Discharge is not linearly related to stage, because discharge depends on both depth and width of the stream channel, or more precisely, on the cross-sectional shape of the channel. Stage refers only to the height of the water above some reference level. The factors that cause flooding are heavy rainfall, sudden or heavy snow melt, and dam failure. All of these factors can suddenly increase discharge of water into stream, within streams, and out of streams. Furthermore, when the discharge causes the river to rise above flood stage, water runs onto the floodplain. In other words, if the river is heavily dammed (e.g. Tigris River) in order to provide water for irrigating the arid and semi-desert region boarding the river valley lead to a significant change in stage of the river. Therefore, constructing a model to predict the river stage is a good idea to represent the essential features of the river.

Artificial neural networks (ANN_s) are being used increasingly to predict and forecast water resources variables. Based on the consideration of the hydrological process, Dooge. [1] divides hydrological models into three categories:- physically based distributed models, lumped conceptual models and black box models. Neural network models which inherently involve mapping of input and output vectors, can be considered as a black box model. The excessive requirements of field data in the case of physically based distributed models and the large number of parameters and subsequent difficulty in calibration in the case of lumped conceptual models render such models less suitable in operational stage forecasting use. This is the reason why simple black box model or storage based model found to be used extensively as water resources models.[2]

In this study, prediction of Tigris River stage in Qurna city, Basrah, south of Iraq, for one, two,and three days ahead by using artificial neural network is the main objective. In addition to investigate applicability and performance of (ANN_s) as an effective tool for applying in water resources engineering.

Artificial Neural Networks (ANN_s)

ANN_s are based on the present understanding of the biological nervous system. An ANN is a massively parallel-distributed information processing system that has certain performance characteristics resembling biological neural networks of the human brain [2]. The network consists of layer of parallel processing elements, called neurons.

In most networks, the input layer receives the input variables for the problem in hand. This consists of all

* Civil Engineering Department, University of Technology / Baghdad

quantities that can influence the output. The output layer consists of values predicted by the network and thus represents model output. Between the input and output layer there may be one or more hidden layer. The neurons in each layer are connected to the neurons in a proceeding layer by a weight (W_{ji}), which can be adjusted during training. Fig.(1) illustrates a three layer neural network consisting of four neurons in input layer, four neurons in hidden layer and two neurons in output layer, with interconnection weights between layers of neurons.

ANN has been used in water resources engineering over the last decade. These include flood forecasting [3], rainfall-runoff modeling [4], stream flow prediction [5] and water level prediction [6].

The back-propagation network is the most popular among ANN paradigms [7]. The back-propagation algorithm gives a prescription for changing the weight (W_{ji}) in any feed forward network to learn a training vector of input-output pairs. It is a supervised learning method in which an output error is fed back through the network, altering connection weight so as to minimize the error between the network output and the target output. Therefore, the back-propagation is used here as an approach to ANN training.

The outputs of the hidden layer are gathered and processed by the last or output layer, which delivers the final output of the network. A neuron is a processing unit with n inputs ($x_1, x_2,\ldots,x_n$), and only one output (y), with

$$y = f(x_1, x_2,..., x_n) = A\left[\left(\sum_{i=1}^{n} w_i x_i\right) + b\right] \quad \ldots\ldots (1)$$

Where:

w_i: the weight of the neuron

b : the constant bias

A : the activation or transfer function

Study Area and Data Set

The Tigris River is one of the main rivers in the Middle East. Its total length is 1900 kilometers out of which 1415 kilometers run inside Iraq. It has a catchment area of about 235,000 square kilometers [8]. This river is heavily dammed in both Iraq and Turkey, in order to provide water for irrigating the arid and semi-desert region boarding the river valley. In accordance with the character of feeding and distribution of precipitation, one can distinguish three periods in the annual cycle of the Tigris River water regime. Flood period (February-June) connected with snow thawing in the mountains, summer low-water period (July-October) and a period of rain flooding (November-February) within the flood period the Tigris River conveys about 75 percent of annual flow, in the dry period about 10 percent and in the period of autumn-winter flood about 15 percent [8]. Tigris and Euphrates meet near Qurna, in Iraq. The combined flow, called Shatt Al Arab empties in Arabian Gulf.

The daily observed measurements of Tigris River stage for the period (5/10/2004-6/2/2006) are measured in Qurna. This city is a pleasant little place 74 kilometers North West of Basrah at the very tip of the point of Shatt Al Arab as shown in Fig.(2). The observed data of Tigris River are obtained from Qurna Irrigation OfficeMosul (Al Mawsil) shares the severe alternations of temperature experienced by upper Mesopotamia . The summer heat is extreme, and in winter frost is not unknown . Nevertheless the climate is considered healthy and agreeable; copious rains fall in general in winter.

Methodology

Methodology of this study is to use the available data (observed data) for the period (5/10/2004-6/2/2006). Hence 490 instances are available for predicting the stage of river by using the neural network toolbox in Matlab R2007b. These data are divided into three statistically parts: - 60% for training, 20% for validation, and 20% for testing. The training of the neural network is accomplished by adjusting the interconnecting weights till such time that the root mean square error (RMSE) between the observed and predicted set of values is minimized. The adjusting of inter-connecting weights is accomplished using the back-propagation algorithm. Multilayer perceptron with one hidden layer is used in the architecture of the network. A trial and error procedure based on root mean square error (Eq.2), coefficient of efficiency (Eq.3), and coefficient of correlation (Eq.4) are used to select the best network architecture and perform of ANN for predicting the stage of Tigris River.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{i=n}\left(Q_m - Q_s\right)^2} \quad \text{.... (2)}$$

$$CE = 1 - \frac{\sum_{i=1}^{i=n}\left(Q_m - Q_s\right)^2}{\sum_{i=1}^{i=n}\left(Q_m - \bar{Q}_m\right)^2} \quad \text{.... (3)}$$

$$R = \frac{\sum_{i=1}^{i=n}\left(Q_m - \bar{Q}_m\right)\left(Q_s - \bar{Q}_s\right)}{\sqrt{\sum_{i=1}^{i=n}\left(Q_m - \bar{Q}_m\right)^2 \sum_{i=1}^{i=n}\left(Q_s - \bar{Q}_s\right)^2}} \quad \text{.... (4)}$$

Where:

Q_m : Measured value

Q_s : Simulated value

$\bar{Q}_m$: Average of measured values

$\bar{Q}_s$: Average of simulated values

n : Is the number of observations

Results and Discussion

Three models are adopted as the first experiment. Three common statistics are used for evaluating the models performance (root mean square error, coefficient of efficiency, and coefficient of correlation). Table (1) resumes results for modes M_1, M_2, and M_3 computed over the test set, with marked values corresponding with the best performance according to the criteria in each column.

These models are described as follows:-

M_1, $Q_t = f(Q_{t-1})$

M_2, $Q_t = f(Q_{t-1}, Q_{t-2})$
M_3, $Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$

Where:
Q_t : is the stage at a specified time
Q_{t-1}, Q_{t-2}, Q_{t-3} : are the stages at t-1, t-2, and t-3 respectively.

As shown in table (1), there is some decrease in the statistical parameter RMSE in Models (1-3). The same table shows that there is insignificant increase with reference to both the coefficient of efficiency (CE) and the coefficient of correlation (R^2). The latter witnesses relatively increase. Then, the Model (M_3) has selected for starting the second experiment (pr. Model (M_3) is selected to predict the river stage for one, two, and three days ahead. Table (2) shows the performance statistics of the second experiment and figures (3 to 6) show the observed versus calculated stages. According to table (2), one may note the size of a hidden layer is one of the most important considerations when solving actual problems using multilayer feed forward networks.

Too many neurons (case. 6) may not only require a large computational time for accurate training, but they may also result in overtraining [9]. A neural network is said to be overtraining when the network focuses on the characteristics of individual data point rather than just capturing the general patterns present in the entire training set.

Conclusions

A neural network model with back-propagation algorithm is developed for predicating the stage of Tigris River in Qurna, south of Iraq. The results shown that the artificial neural network with back-propagation algorithm is a powerful tool to simulate the river stage for short term. A large number of hidden neurons will ensure correct learning, and the network is able to correctly predict the data it has been trained on, but may also result in overtraining. With too few hidden neurons, the network may be unable to learn the relationships amongst the data and the error will fail to fall below an acceptable level. Thus, selection of the number of hidden neurons is a crucial decision.

References

1. J.C.I. Dooge .1974. Problems and Methods of Rainfall-Runoff Modeling, Mathematical Models for Surface Water Hydrology, John Wiley and Sons, 71-108.
2. S. Haykin (1994). Neural networks: a comprehensive foundation. Mac-Millan. New York.
3. R. Garcia-Bartual, “Short term river flood forecasting with neural networks”, Proceedings of iEMSs 2002, Switzerland, 2002.
4. H. Sobri, M.K. Amir Hashim, and A.N. Nor Irwan, “Rainfall-runoff modeling using artificial neural network”, Proceedings of 2nd World Engineering Congress. Kuching, 2002.
5. O.R. Dolling, and E.A. Varas, “Artificial neural networks for stream flow prediction”, Journal of Hydraulic Research. 40(5), 2002, pp. 547-554.

6. A.R. Patrick, W.G. Collins, P.E. Tissot, A. Drikitis, J. Stearns, P.R Michaud, and D.T. Cox, "Use of the NCEP MesoEta data in a water level predicting artificial neural network", in the Proceedings of 19th Conf. on weather Analysis and Forecasting/15th Conf. on Numerical. Weather Prediction, 2002.
7. R P Lippman. 'An Introduction to Computing with Neural Networks.' IEEE ASSP Magazine, 1987, pp 4-22.
8. Iraqi Ministries of Environment, water resources, Municipalities and public works (2006) " overview of present conditions and current use of the water in the water resources marshland area" book1, Italy-Iraq, 146p.
9. Brion, Gail Montgomery and Lingireddy, Srinivasa (1999). "A Neural Network Approach to Identifying Non-Point Sources of Microbial Contamination." Water Resources. Vol. 33 No. 14 pp. 3099-3106. 1999.

Table (1) results of modes M_1, M_2, and M_3 computed over the test set

Testing Period				
Model	RMSE	CE	R^2	No. of hidden neurons
M_1	2.159	0.860	0.863	20
M_2	2.155	0.861	0.871	20
M_3	**2.085**	**0.870**	**0.878**	**20**

Table (2) results of the second experiment

Testing Period					
Case	**Model**	**RMSE**	**CE**	**R^2**	**No. of hidden neurons**
1	Q_{t+1}	2.231	0.8577	0.867	20
2	Q_{t+2}	3.274	0.7360	0.736	25
3	Q_{t+3}	4.558	0.5544	0.568	30
4	Q_{t+3}	4.496	0.5665	0.619	40
5	Q_{t+3}	4.329	0.5981	0.649	45
6	Q_{t+3}	4.637	0.5388	0.616	50

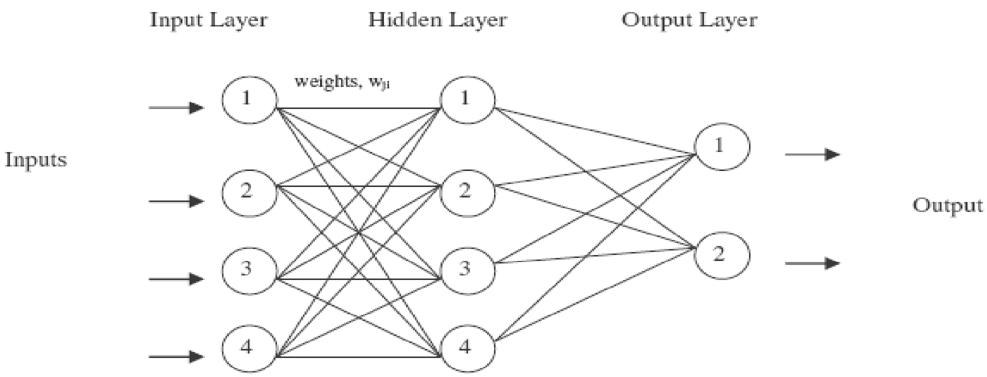

Figure (1) Configuration of three-layer neural network

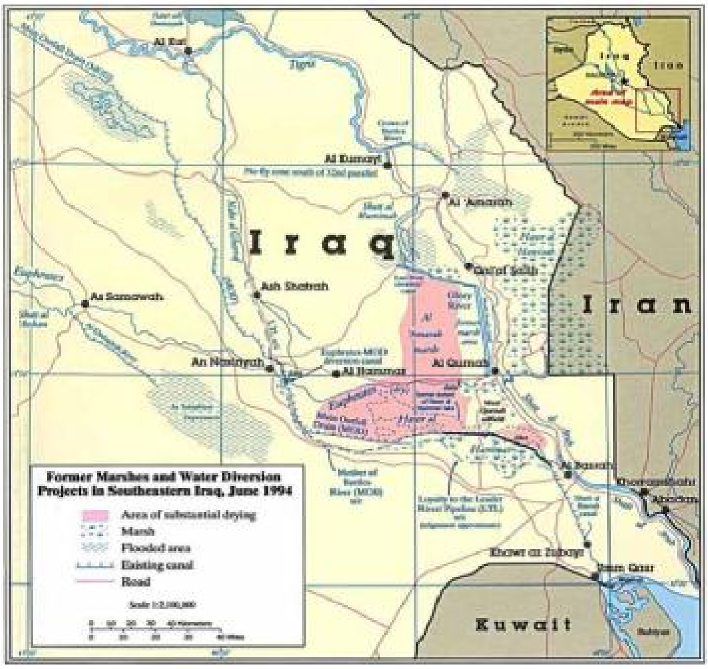

Figure (2) Water conveyance system in southern Iraq.

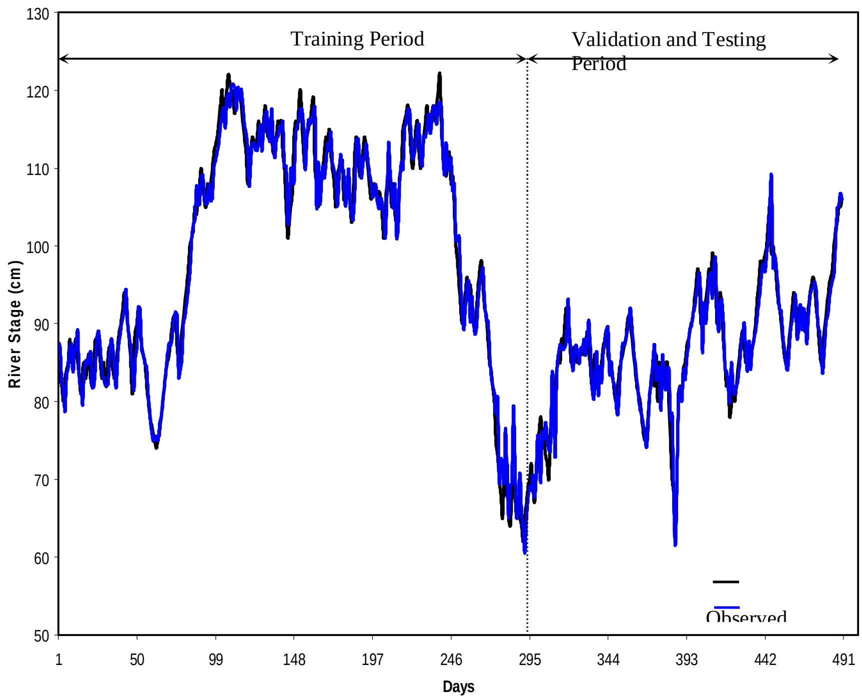

Figure (3) Observed and calculated river stages for one day ahead prediction

correct learning, and the network is able to correctly predict the data it has been trained on, but may also result in overtraining. With too few hidden neurons, the network may be unable to learn the relationships amongst the data and the error will fail to fall below an acceptable level. Thus, selection of the number of hidden neurons is a crucial decision.

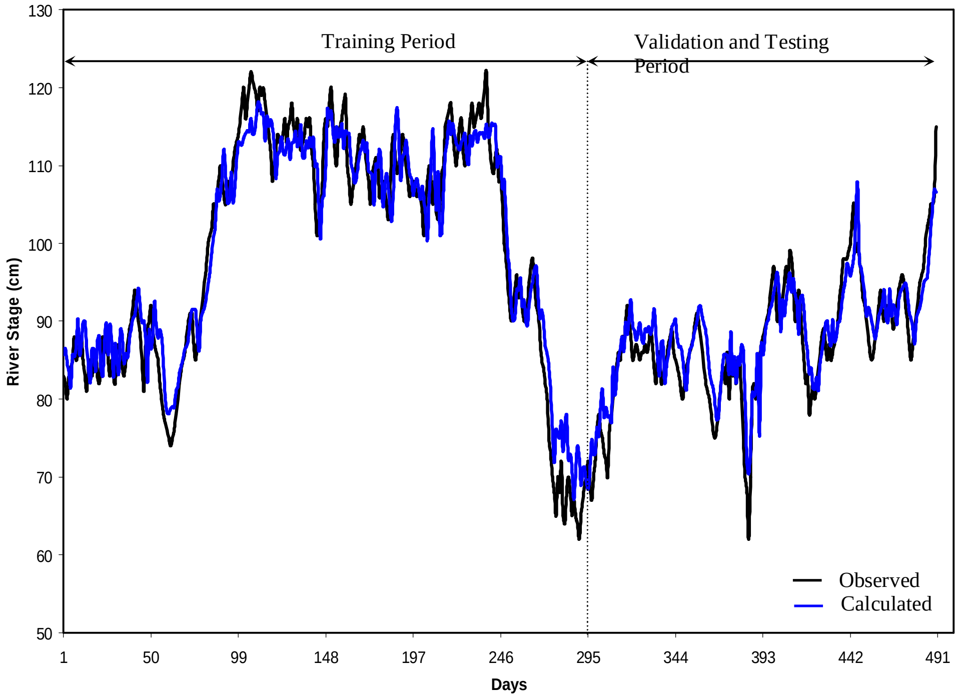

Figure (4) Observed and calculated river stages for two days ahead prediction

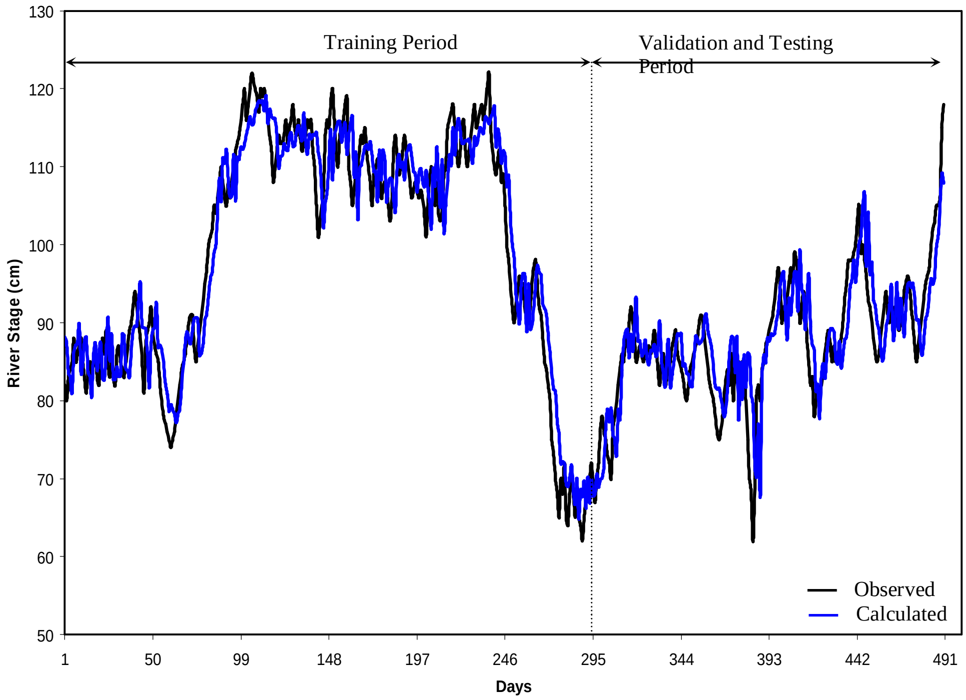

Figure (5) Observed and calculated river stages for three days ahead prediction (case. 3)

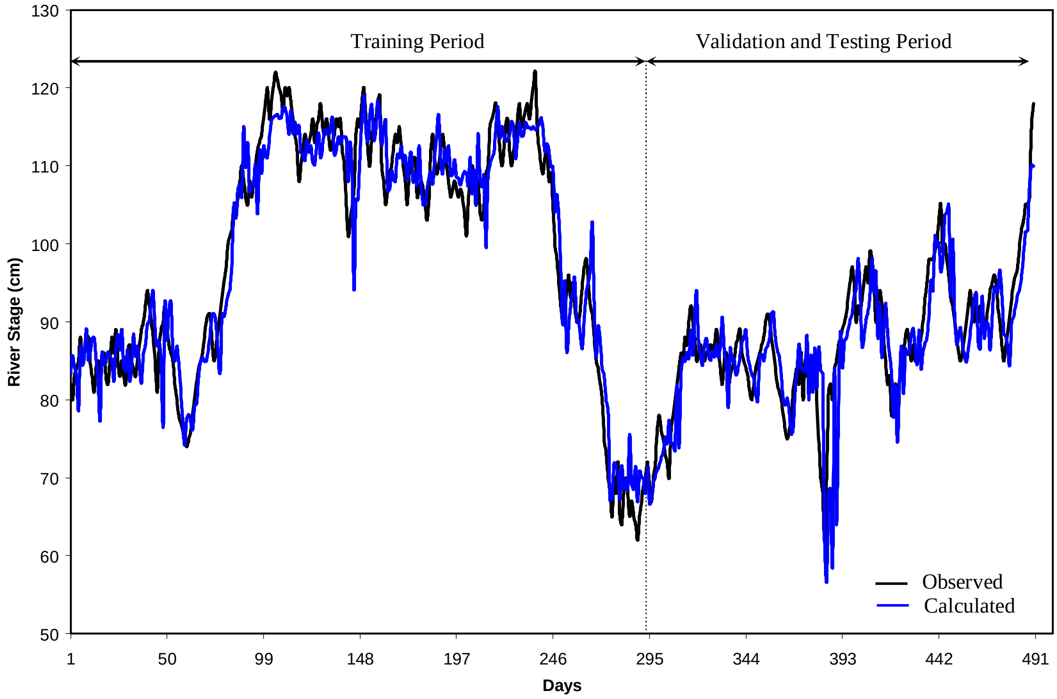

Figure (6) Observed and calculated river stages for three days ahead prediction (case.5)