Improved Model using Estimate Error for Daily Reservoir Inflow Forecasting

Yaowalak Ngamsanroaj[1] and **Kreangsak Tamee**[2]

ABSTRACT

Inflow forecasting is one of the important components for reservoir operation and resource management. To obtain enhanced accuracy for forecasting reservoir inflow, this paper proposes an improved model for forecasting the inflow of Bhumibol reservoir. The 3,169 records of daily inflow data from June 1, 2008, to February 1, 2017, had been collected to calculate the inflow into the reservoir by using Artificial Neural Networks (ANN) running the Back-Propagation Learning Algorithm for forecasting the inflow of the reservoir in the main model and error prediction model. The performance of the model is evaluated by four methods: the coefficient of determination (R^2), the Nash-Sutcliffe efficiency (NSE), the Root Mean Square Error (RMSE), and the Mean Absolute Error (MAE). The proposed main and error prediction models were combined to support the forecast of reservoir inflow. The performance of the proposed model can be determined using the following measured values: R^2 was 0.927, NSE was 0.925, RMSE was 6.805 and MAE was 3.611. This indicates that the improved model provides more accurate value than the model without estimate error.

Keywords: Artificial Neural Network, Model Forecasting, Reservoir Inflow, Error Prediction

1. INTRODUCTION

Nowadays, the problem of water crisises has intensified, such as water shortages in the dry season and floods in many areas during the rainy season. These problems occur due to the changing nature of the world's climate. Water management is more complicated with the changing circumstances. This is why effective water management is difficult. In order to conduct the proper management of water in the reservoir, the precise forecasted flow of water into the reservoir is required [1-3]. Forecasting the reservoir inflow is an important method for improving agricultural activities. The purpose of forecasting is to introduce a practical way for deciding how to use water and assist in decision making. The decisions made can significantly affect the main crop, with the associated economic consequences [4].

In the past, many researchers have developed accurate and easy-to-use models for predicting water flow in reservoirs. Conventional prediction techniques include time series models suggesting that Autoregressive Integrated Moving Average (ARIMA) was used as a statistical model with time series. This model provides good results for linear data, is not suitable for nonlinear data [5-7]. The hydrological daily time series processes can be modelled to generate synthetic inflow values, using different techniques such as conceptual and time series models. Conceptual models usually incorporate simplified, non-linear, time-invariant, and deterministic relationships, with parameters representing the watershed characteristics [8]. Some researchers have used the method of machine learning, especially artificial neural networks, which solves problems with both linear and nonlinear data. Therefore, this method provides better predictive efficiency than the ARIMA method [9-12]. However, in the above research, some errors have occurred in the simulation models of the neural networks in extreme weather conditions such in wet and dry years. During a wet year, the water flow in the reservoir is quite high. Some areas, such as Thailand, sometimes suffer from acute drought in a dry year, and then there is no water inflow into the reservoir. The prediction models cannot provide an efficient forecast for the amount of water entering the reservoir. At present, researchers try to take advantage of the error caused by the main model by integrating the forecasting error with the main model. This approach can yield better predictions. Xiaojing Zhang [13] manipulates the forecasting error with the main model for predicting reservoir water levels. The results show the forecasting performance of effective lead times can be enhanced by minimizing the difference between the forecasted and observed water levels. From the above-mentioned problem, the researchers have developed the idea of using the error information to create a model to estimate the error that will occur with the model. The main simulation purpose is to increase the efficiency of flow forecasting in the reservoir.

Manuscript received on June 27, 2019 ; revised on September 4, 2019.

Final manuscript received on November 1, 2019.

[1]The author is with Department of Computer Science and Information Technology, Faculty of Science Naresuan University Phitsanulok, 65000, Thailand., E-mail:ayaowalak@hotmail.com

[2]The author is with Department of Computer Science and Information Technology, Research Center for Academic Excellence in Nonlinear Analysis and Optimization, Faculty of Science Naresuan University Phitsanulok, 65000, Thailand, E-mail:kreangsakt@nu.ac.th

2. BACKGROUND

2.1 Artificial Neural Networks Modelling

An Artificial Neural Network (ANN) is a method inspired by the human brain and nervous system. ANNs consist of a set of processing elements (neurons) operating in parallel. As in the biological reality, the function of the ANN is determined basically by the connections between the neurons. ANNs have been used in various scientific fields to solve problems such as pattern recognition, particle identification, and classification. Furthermore, ANNs are a proved and efficient method to model complex input-output relationships. The networks learn the relationship directly from the data being modelled. Various fields of hydrology have been investigated with success using ANNs.

The proposed model was developed with the ANN technique. The theory and mathematical basis of ANN have been described excellently by Shamseldin [14]. Essentially, the structure of ANN comprises an input layer, an output layer and one or more hidden layers as illustrated in Figure 1. The illustration in figure 1 has a single hidden layer which is generally enough to approximate any complex, non-linear function [15]. The layers contain nodes which are connected by weights. Determining optimal values for these weights and other parameters of the network is the purpose of the ANN training exercise.

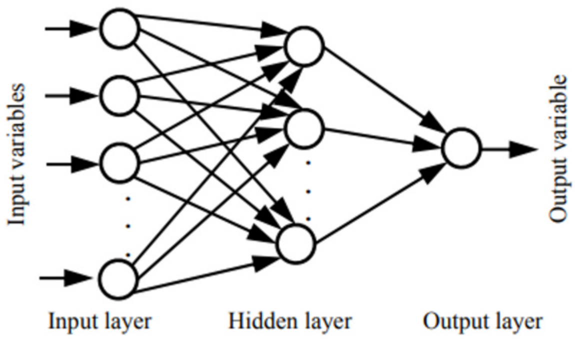


Fig.1: *Illustration of an artificial neural network.*

For a given problem, the number of nodes in the output layer is fixed by the problem. In the current work, it is the daily inflow forecast. The input nodes must be determined by the factors known to affect the output variable. The number of neurons in the hidden layer is much more difficult to arrive at, and is normally determined as part of the training by trial and error as described by Adeloye and De Munari [16].

Training is often improved through the use of an early-stop-rule (ESR) that helps to avoid over-fitting. In ESR, the available data is divided into three parts: (i) a training set, used to determine the network weights and biases, (ii) a validation set, used to estimate the network performance and decide when the training should be stopped, and (iii) a test set, used to verify the effectiveness of the stopping criterion and to estimate the expected performance in the future.

The importance of the ANN forecasting model is due to its capability to forecast future values of the inflows or any hydrological variable using only the historical record of this variable. The need for other variables is usually essential for forecasting inflow using other types of forecasting models. The accurate forecasting of future inflow to a reservoir can allow efficient operation of this reservoir. For example, if the persistence in the time series is well represented, the required actions for expected floods and draughts can be defined in advance, and preparations for these actions can be achieved.

2.2 Model Accuracy Indicators

The efficiency or performance of the model can be measured by using the following four methods.

- *The coefficient of determination (R^2)*

R^2 is an estimation of the distribution of the spread between the measurable real-valued dataset and the predicted dataset.

$$R^2 = \frac{\sum_{i=1}^{n}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=0}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(P_i - \bar{P})^2}}$$

O_i and P_i are the actual inflow and the predicted inflow, respectively. $\bar{O}$ and $\bar{P}$ are the mean of the actual inflows and predicted inflows to the reservoir respectively. n is the total number of data values. The range of R^2 is between 0 and 1, where the value of zero denotes no relation, while the value of 1 denotes the distribution of the predicted values are equal to the actual values.

- Nash-Sutcliffe efficiency (NSE)

NSE is the most commonly used index for model accuracy used to measure the efficiency and effectiveness of simulated models in predicting the desired values.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$$

NSE can range from $-\infty$ and 1.0. A perfect match of modeled data and the observed data is shown with 1.0. An efficiency of 0.0 indicates that the model predictions are as accurate as the mean of the observed data. NSE less than zero indicates that the observed mean is a better predictor than calculated value from the model.

- *Root mean square error (RMSE)*

RMSE, the square root of the average squared value, is a method of measuring the error tolerance of the predicted value from the model to the actual measured value.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2}$$

RMSE values are always greater than zero. If the value is close to zero, the model can be estimated as the actual measured value.

- *Mean absolute error (MAE)*

MAE is a method to measure the difference between the actual measured value and the predicted value of the model by calculating the mean for all recorded absolute errors.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i|$$

MAE is always positive. If the MAE is low, the model can predict a value close to the actual collected value.

3. DATA AND STUDY AREA

This study collected the amount of water entering the reservoir on a daily used basis from Bhumibol Dam. The Bhumibol Dam, one of the multipurpose dams of the Electricity Generating Authority of Thailand (EGAT), is a concrete arch dam that walls the Ping River at Khaokaew area of Sam Ngao district, Tak province. The geographic coordinates are 17°14'33"N 98°58'20"E, as shown in Figure 2. The reservoir capacity is 13,462 million cubic meters (MCM), with about 300 square kilometers of water surface. The reservoir begins in Hod district in Chiang Mai province and has a length of 207 kilometers.

The researchers collected 3,169 records of data for 8 years, from 01/06/2008 to 02/02/2017. The duration curve of the daily data sequences set for the period is given in Figure 3. The data is characteristic time series data which is stored regularly in the same range, such as daily, monthly, annual data, etc. The utilized data in this study is daily data, which has seasonal variations. The seasonal variation reveals that the amount of water entering the reservoir is not much different in the same period of the year, and the maximum inflow normally occurs in October of each year, which is the rainy season in Thailand. Sometimes the volume of water inflow is zero because there is no inflow water in the basin during that day. This mostly occurs during the dry season. In addition, some ranges of data are subject to irregular variations, which is the period when the amount of inflow water in the basin is extremely high because

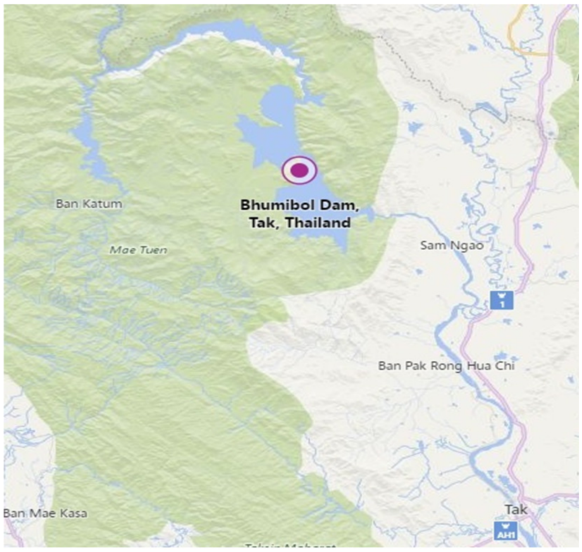


Fig.2: *The topography of Bhumibol reservoir, Tak Province, Thailand.*

of heavy rain in the upstream area and side flow occurring together with flooding during that period.

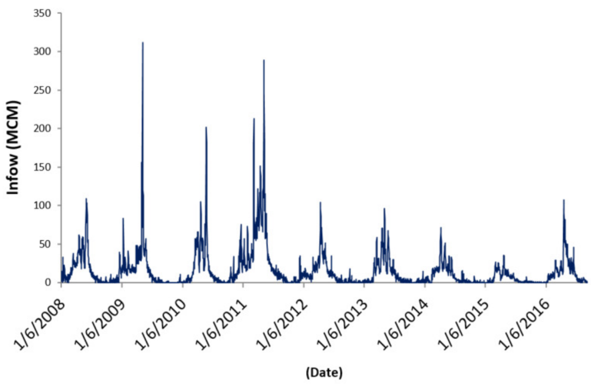


Fig.3: *Duration Curve of inflows.*

4. METHODOLOGY

The authors used a three-layer, feed-forward ANN and a Levenberg Marquardt training algorithm with backpropagation implemented in MATLAB Neural Network Toolbox. Training of the ANN was performed in a supervised manner using historical data. The output of the ANN provided daily inflow. The proposed model will now be presented in more detail.

4.1 Proposed Model

The proposed model for predicting water inflow will be divided into two parts: inflow forecasting or main model, and predicted error model. The main model is processed for forecasting the inflow of water into the reservoir by using daily data of inflow as input data and the ANN technique for modeling. The other model is using predictive error modeling using the difference data between the collected existing inflow data and predicted inflow result from the model with the same ANN technique. The predictions of both models are integrated to produce a model for predicting the inflow of water. This approach for the integration of the two main models can be depicted as shown in Figure 4.

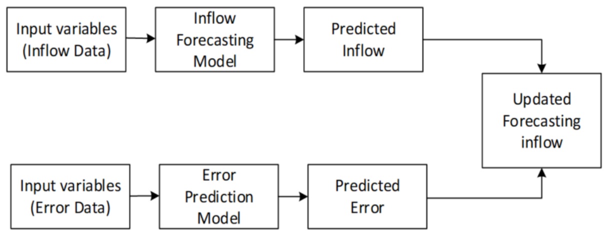


Fig.4: *Integration of Inflow Forecasting Model from Main and Error Prediction Models.*

4.2 Experimental

4.2.1 Data preparation and Data Cleaning

To make the training process of the ANN more effective, the available data has been prepared and cleaned. Anomalous data can negatively impact inflow forecasting by causing model parameters to be incorrectly estimated, and this is especially true for anomalies in time series data. Data cleaning is a process that consists of detecting and imputing anomalous data. In the hydrological domain, training accurate forecasting models requires data that correctly captures the underlying system. However, hydrological data often contain anomalies, which can be due to various causes such as human error (e.g., mistyping) or system error (e.g., erroneous measurement). During model training, anomalous hydrological data yields erroneous forecasting models. Applying an erroneous forecasting model on error data yields inaccurate forecasts. Therefore, in this study, the researcher had to edit the information correctly to make it ready to be used in the forecasting problem. For the data analysis, the actual inflow data is processed by a boxplot, which can eliminate the outliers of actual inflow time series and better catch the features of these series. This method removes the effects of outliers on the forecasting results and improves the accuracy and efficiency of inflow forecasting.

4.2.2 Model Building

The important task of ANN modeling for a time series is to choose an appropriate number of hidden nodes as well as to select the dimension of the input vectors (the lagged observations). However, it is difficult to determine the number of input nodes and hidden nodes in advance, as there are no theoretical developments that can guide the selection process. Hence, in practice, experiments are often conducted to select the appropriate values input nodes and hidden nodes.

At the beginning of this experiment, the researchers used the 3,169 records of water inflow to create a model with the ANN technique for predicting the inflow of water into the reservoir by using a MATLAB program. 70% of the data is used for the training set, 15% of the data is used for the testing set, and the remaining 15% of the data is used for the validating set. Therefore, the preparation of data for learning, or data for prediction, must be prepared in the matrix data table form to train the neural networks. Data is divided into two parts, Input and Output, which can be used in comparison with the results. These two parts are described as follows:

$$Input = \begin{bmatrix} V_1 & V_2 & \cdots & V_k \\ V_2 & V_3 & \cdots & V_{k+1} \\ \vdots & \vdots & \cdots & \vdots \\ V_{n-(k+1)} & V_{n-(k+2)} & \cdots & V_{n-1} \end{bmatrix}$$

$$Output = \begin{bmatrix} V_{k+1} \\ V_{k+2} \\ \vdots \\ V_n \end{bmatrix}$$

V_i is the observed water flow into the reservoir daily. k is the number of days that required to use in forecasting such as 10, 20 and 30 days. n is the total number of recorded days.

The researcher designed an algorithm to specify the most suitable parameters of the ANN: the input data of the training data set, and the number of neurons in the hidden layer. Therefore, the method used to determine the architecture of the ANN was to start with a small network (one hidden layer and four nodes), then to gradually increase the number of nodes and choose the network with the best performance. In such a way, the historical inflow was used as input data for training of the ANN. The study used 10, 20, and 30 days in the input layer and 4, 8, and 12 neurons in the hidden layer for each input layer. To evaluate the forecast performance with the selected indicators, competing forecasts made by ANNs with different parameters were compared to the actual observed data and assessed by indicators, which are described in section 2.2 and are presented in Table 1.

According to Table 1, the forecasting model with different parameters has been manipulated for various input nodes. The forecasted reservoir inflow from the model is correlated with the actual reservoir inflow, which is estimated from 4 indicators: $R^{2,}$ NSE, RMSE, and MAE. The R^2 and NSE should be closer to 1.0 while the RMSE and MAE should be closer to zero for indicating the performance of the model. The

Table 1: *Demographic and Converted Data.*

Input Node	**Single Hidden Layer**											
	4 nodes				**8 nodes**				**12 nodes**			
	$\mathbf{R^2}$	**NSE**	**RMSE**	**MAE**	$\mathbf{R^2}$	**NSE**	**RMSE**	**MAE**	$\mathbf{R^2}$	**NSE**	**RMSE**	**MAE**
10	0.907	0.905	7.614	3.779	0.909	0.908	7.503	4.050	0.895	0.894	8.033	4.207
20	0.907	0.906	7.580	3.557	0.903	0.878	8.631	5.818	0.907	0.907	7.561	3.624
30	0.904	0.903	7.699	3.565	0.906	0.886	8.376	5.468	0.916	0.915	7.242	3.652

most efficient parameters for the main model used to forecast the amount of water entering the reservoir at the best performance, estimated from the 4 indicators, is an artificial neural network model with a structure of 30-12-1 (30 nodes in the input layer and 12 nodes in the hidden layer).

When the inflow water flowing into the reservoir is estimated from the main model, the error can be found as in equation 1:

$$\varepsilon(t) = V(t) - P(t)$$

$P(t)$ represents the water inflow in the reservoir predicted from the main model during the day $t, \varepsilon(t)$ means the difference value between the actual value and the value obtained from the forecast of the day t. The difference is defined as the error value.

The preparation of error information for learning requires the data must be in the form of matrix data tables. To train neural networks, data is divided into two parts, Input and Output, for being used in comparison with the results. These are as follows.

$$Input = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_k \\ \varepsilon_2 & \varepsilon_3 & \cdots & \varepsilon_{k+1} \\ \vdots & \vdots & \cdots & \vdots \\ \varepsilon_{n-(k+1)} & \varepsilon_{n-(k+2)} & \cdots & \varepsilon_{n-1} \end{bmatrix}$$

$$Output = \begin{bmatrix} \varepsilon_{k+1} \\ \varepsilon_{k+2} \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

ε_i means the difference value between the actual value and the value obtained from the forecast of the day t. This is called the error value. k is the number of days that required to use in forecasting 30 days. n is the total number of recorded days.

In the error model, the authors used 30 nodes in the input layer because the study has to combine this error model with the main model. The selected main model is the model with the best structure by the 30 input nodes. To search for the best model, the error model uses different numbers of nodes in hidden layers (4, 8 and 12) and evaluates the performance with four indicators. The results of the evaluation of the error model are shown in Table 2.

According to Table 2, the error forecasting model that provides the best performance is the model that has 8 nodes in the hidden layer, resulting in an error prediction model with the structure of the neural network being 30-8-1 (30 nodes in the input layer and 8 nodes in the hidden layer).

The proposed model is the model that integrates the main model and the error prediction model and is calculated according to equation 2:

$$\hat{P}(t) = P(t) + \hat{\varepsilon}(t) \tag{1}$$

$\hat{\varepsilon}(t)$ represents the forecasted error predicted from the error model during the day t. $\hat{P}(t)$ means the water inflow in the reservoir from the proposed model during the day t.

5. RESULT AND DISCUSSION

In this section, the proposed model approach is evaluated using the historical data of Bhumibol Dam. To show the superiority of the proposed model, the forecasting performance is also compared with four different methods.

Defining the suitable neural network structure for nonlinear data series, the three sets of nodes, 10, 20 and 30 nodes should be considered. These sets represent the number of days (k), such as 10, 20, and 30 days. Due to the efficiency of the neural network, the number of nodes in the hidden layer should be differently assigned by dividing the number of nodes into 3 levels: 4, 8 and 12 nodes.

The neural network structure with the learning feature was evaluated by defining the suitable neural network structure with machine learning for adjusting the weight and bias values in the forecasting process. 70% of the data was used for the training set, 15% of the data was used for testing set, and the remaining 15% of data was used for validating set. The learning cycle had to reach 1,000 cycles as a learning condition according to the Levenberg Marquardt algorithm standard configuration [17].

The experiments were done to find the number of nodes to use in the input layer and the number of nodes to use in the hidden layer for the main model in predicting the amount of inflow water entering the reservoir. The results measured the efficiency of the model by R^2, NSE, RMSE, and MAE. From all manipulated data sets, the efficiency of the main model predicting the inflow of water into the reservoir at the best performance comes from a neural network model structure of 30-12-1 with input nodes of 30 days and the number of hidden nodes set to 12 nodes.

Table 2: *Performance of the error forecasting model with different numbers in the hidden layer .*

Input Node	Single Hidden Layer											
	4 nodes				8 nodes				12 nodes			
	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE	MAE
30	0.104	0.102	6.868	3.681	0.132	0.118	0.86.805	3.611	0.034	0.032	7.129	3.808

The best efficiency of inflow forecasting by using the testing data set is displayed in Figure 5. There is an error remaining between the actual or observed data and the predicted data of the inflow water. Therefore, to develop a model for predicting inflow water which is more accurate, it is necessary to use the error model (See Eq.2) to predict the error associated with the main model.

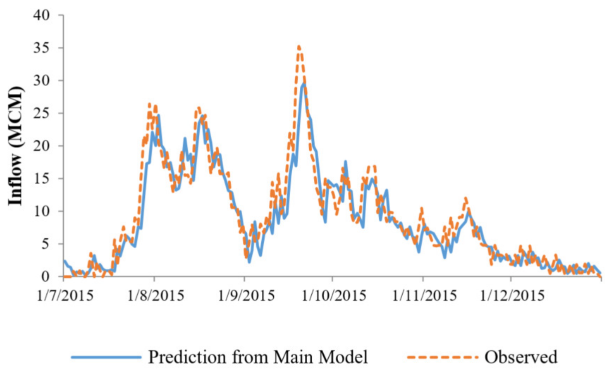


Fig.5: *Comparison of predicted inflow from the main model with observed inflow.*

In the development of the model for predicting the error, the researchers applied the ANN technique as well as developing the main model by setting the number of nodes in the input layer to 30 nodes, the same number selected for the main model. Therefore, in preparation of error information for learning, the data must be put in the form of matrix data tables. To train neural networks, data is divided into two parts: Input and Output. These are used in comparison with the results. The error model which provided the best performance with the neural network was a model of 30-8-1.

Improvement of the model for predicting inflow water volume is formulated from equation 2. The water inflow in the reservoir is predicted from the main model, while the predicted error values is manipulated from the predictive error model. These models were integrated by Equation 2 and measured by R^2, NSE, RMSE, and MAE. The performance of the main and modified models is depicted in Table 3.

Table 3: *The Performance of Main Model and Proposed Model.*

Model	Model Accuracy Indicators			
	R^2	NSE	RMSE	MAE
Main Model	0.916	0.915	7.242	3.652
Proposed Model	0.927	0.925	6.805	3.611

The result in Table 3 presents the related model accuracy indicators measured for the main model and proposed model. The coefficient of determination (R^2) value in the main model and proposed model are 0.916 and 0.927 respectively. This is determined that the proposed model R^2 is closer to 1.0 than the main model and implies that the distribution of the predicted values is very close to the actual values. The Nash-Sutcliffe efficiency (NSE) value in the main model and proposed model are 0.915 and 0.925 respectively. The proposed model NSE is closer to 1.0 than the main model and it is indicates that the proposed model predictions are more accurate. The root mean square error (RMSE) value in the main model and proposed model are 7.242 and 6.805 respectively. It is observed that the proposed model RMSE value is lower and it is closer to zero than the main model. The approximated output from the proposed model is very closed to the actual measured value. The mean absolute error (MAE) value in the main model and proposed model are 3.652 and 3.611 respectively. The proposed model MAE is nearer to zero than the main model. This shows that the prediction of the proposed model is quite close to the actual collected value. In conclusion, all four indicators provide satisfactory results and are in good agreement. This proved that the proposed model for reservoir inflow forecasting is valid. The improved model has proven to be effective and reliable for reservoir inflow forecasting. The integration of the proposed model and optimization technique can provide more efficient solutions for multi-purpose operation of the reservoir system and thereby improve the economy of hydropower production. The developed model can be used for planning purposes with different seasonal periods even in flood or drought years. In the monsoon period, the inflow forecast is used to release more water for increasing capacity of the reservoir to be ready when the huge amount of inflow occurs according to the prediction. In the case of a drought year, the inflow forecast can be used as a guideline for warning the downstream farmers not to do multi-crop cultivation due to the decreased inflow to the reservoir. Using the developed reservoir inflow model could have a significant impact on effective reservoir operation.

The amount of water entering the reservoir from the predicted main model and proposed forecasting model from testing data set is depicted with the observed exiting data in Figure 6. It was found that the comparisons gave satisfactory results while the amount of water inflow from the proposed forecasting model was in quite good agreement to the actual

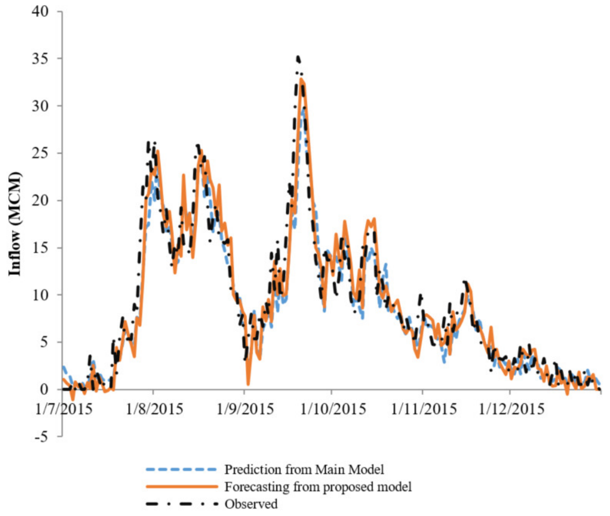


Fig.6: *Comparison of proposed model with main model and observed inflow.*

amount of water entering the reservoir. This indicates that the improved model provides more accurate value than the model without estimate error.

6. CONCLUSION

The results demonstrated that the best predictive efficiency for the main model in predicting water flow in the reservoir was the model with the neural network architecture of 30-12-1. The measured performance of the model as follows: the coefficient of determination was 0.916, the Nash-Sutcliffe efficiency was 0.915, the Root Mean Square Error was 7.242, and the Mean Absolute Error was 3.652. The error forecasting model with the best performance was the model with the architecture of the artificial neural network 30-8-1. When both models were used to improve the inflow model, the predicted inflow into the reservoir provided more accurate forecasting. The performance of the improved model was evaluated by using following measured values: the coefficient of determination was 0.927, the Nash-Sutcliffe efficiency was 0.925, the Root Mean Square Error was 6.805, and the Mean Absolute Error was 3.611. All values provided a significant is better predictive efficiency than the first model. The results showed that the ANN forecasts produced superior reservoir performance. The worst performing inflow situation was when there was a complete lack of knowledge about the inflow and release decisions were based on the starting storage alone. The improved model represents an objective demonstration of how good inflow forecast knowledge could allow more effective reservoir operation.

This research provides a useful approach for decision making in reservoir operation using more accurate predicted inflow. Further work should emphasize the application of machine learning techniques for determining the optimal operating policy of reservoirs. However, to improve the accuracy and reliability of inflow predictions from the models, natural and human conditions influencing the reservoir inflow must be considered.

ACKNOWLEDGEMENT

The authors wish to thank Bhumibol Dam, one of multi-purpose dams of the Electricity Generating Authority of Thailand (EGAT), for supplying hydrology and reservoir operational data. The authors would also like to express thanks for the financial support from Lampang Rajabhat University, Thailand.

References

[1] L. M. Lima, E. Popova and P. Damien, "Modeling and forecasting of Brazilian reservoir inflows via dynamic linear models," *International Journal of Forecasting*, vol.30, no. 3, pp.464-476, 2014

[2] C.T. Cheng, Z.K. Feng, W.J. Niu and S.L. Liao, "Heuristic methods for reservoir monthly inflow forecasting: A case study of Xinfengjiang reservoir in Pearl river, China," *Water*, vol.7, no. 8, pp.4477-4495, 2015.

[3] Y. Bai, Z. Chen, J. Xie and C. Li, "Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models," *Journal of hydrology*, vol.532, pp.193-206, 2016.

[4] C. Srisawat and J. Payakpate, "Comparison of MCDM methods for intercrop selection in rubber plantations," *Journal of ICT*, vol.15, no. 1, pp.165-182, 2016.

[5] M. Valipour, M. E. Banihabib and S. M. R. Behbahani, "Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir," *Journal of hydrology*, vol.476, pp.433-441, 2013.

[6] M. E. Banihabib, A. Ahmadian and F. S. Jamali, "Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data," *GeoResJ*, vol.13, pp.9-16, 2017.

[7] H. Moeeni and H. Bonakdari. "Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model," *Stochastic environmental research and risk assessment*, vol.31, no. 8, pp.1997-2010, 2017.

[8] K. l. Hsu, H. V. Gupta and S. Sorooshian, "Artificial neural network modeling of the rainfall-runoff process," *Water resources research*, vol.31, no. 10, pp. 2517-2530, 1995.

[9] F. Othman and M. Naseri, "Reservoir inflow forecasting using artificial neural network," *International journal of physical sciences*, vol.6, no. 3, pp.434-440, 2011.

[10] R. H. Al-Suhili and R. A. Karim, "Daily inflow forecasting for Dukan reservoir in Iraq using ar-

tificial neural networks," *International Journal of Water*, vol.9, no. 2, pp.194-208, 2015

[11] C. Chiamsathit, A. J. Adeloye and S. Bankaru-Swamy. "Inflow forecasting using Artificial Neural Networks for reservoir operation," *Proceedings of the International Association of Hydrological Sciences*, vol.373, pp.209-214, 2016.

[12] A. Sauhats, R. Petrichenko, Z. Broka, K. Baltputnis and D. Sobolevskis, "ANN-based forecasting of hydropower reservoir inflow," in *2016 57th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)*,Riga, pp. 1-6, 2016

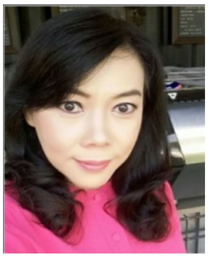
[13] X. Zhang, P. Liu, Y. Zhao, C. Deng, Z. Li, and M. Xiong, "Error correction-based forecasting of reservoir water levels: Improving accuracy over multiple lead times," *Environmental Modelling & Software*, vol.104, pp.27-39, 2018.

[14] A. Y. Shamseldin, "Application of a neural network technique to rainfall-runoff modelling," *Journal of hydrology*, vol.199, no. 3-4, pp.272-294, 1997.

[15] I. E. Mulia, T. Asano, and P. Tkalich, "Retrieval of missing values in water temperature series using a data-driven model," *Earth Science Informatics*, vol.8, no. 4 , pp.787-798, 2015.

[16] A. J. Adeloye and A. De Munari, "Artificial neural network based generalized storageyield-reliability models using the LevenbergMarquardt algorithm," *Journal of Hydrology*, vol.326, no. 1-4, pp.215-230, 2006.

[17] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics, vol.11, no. 2, pp 431-441, 1963.*



Yaowalak Ngamsanroaj received the B.A. degree in Business Computer from Yonok College, Lampang, Thailand, in 1993, the M.Sc. degree in Information Technology and Management from Chiangmai University(CMU), Chiangmai, Thailand, in 2003, respectively. She is PhD candidate in Information Technology at the Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Pitsanulok, Thailand. Her research interests include machine learning, AI and data science.

Kreangsak Tamee received the B.Sc. degree (Physics) from Chiangmai University (CMU), Chiangmai, Thailand, in 1997, the M.Eng. and D.Eng. degrees in electrical engineering from King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in 2001, and 2011, respectively. He has been a Lecturer at the Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Pitsanulok, Thailand. His research interests are in the areas of optical sensor, optical signal processing, machine learning, data analytics, optimization.