Hydrology and Earth System Sciences Discussions

Real-time flood forecasting by employing artificial neural network based model with zoning matching approach

M. Sulaiman[1], A. El-Shafie[2], O. Karim[2], and H. Basri[2]

[1]Faculty of Civil Engineering and Natural Resources, University of Malaysia, Pahang, Malaysia
[2]Department of Civil Engineering, Faculty of Engineering,
University Kebangsaan Malaysia, Malaysia

Correspondence to: A. El-Shafie (elshafie@eng.ukm.my)

HESSD

8, 9357–9393, 2011

Real-time flood forecasting

M. Sulaiman et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

Flood forecasting models are a necessity, as they help in planning for flood events, and thus help prevent loss of lives and minimize damage. At present, artificial neural networks (ANN) have been successfully applied in river flow and water level forecasting
⁵ studies. ANN requires historical data to develop a forecasting model. However, long-term historical water level data, such as hourly data, poses two crucial problems in data training. First is that the high volume of data slows the computation process. Second is that data training reaches its optimal performance within a few cycles of data training, due to there being a high volume of normal water level data in the data training, while
¹⁰ the forecasting performance for high water level events is still poor. In this study, the zoning matching approach (ZMA) is used in ANN to accurately monitor flood events in real time by focusing the development of the forecasting model on high water level zones. ZMA is a trial and error approach, where several training datasets using high water level data are tested to find the best training dataset for forecasting high water
¹⁵ level events. The advantage of ZMA is that relevant knowledge of water level patterns in historical records is used. Importantly, the forecasting model developed based on ZMA successfully achieves high accuracy forecasting results at 1 to 3 h ahead and satisfactory performance results at 6 h. Seven performance measures are adopted in this study to describe the accuracy and reliability of the forecasting model developed.

²⁰ ## 1 Introduction

The nature of river flow is determined by many factors, such as sedimentation, dam operation, soil type, cross sectional area of the river, and rainfall. This leads to stochastic behavior that complicates the study of river flow. There are studies that use regression for forecasting river flow, but the computed results only provide a rough estimation
²⁵ of the flow. Extensive reviews on artificial neural network (ANN) applications in hydrological simulation and forecasting have been reported in ASCE (2000a, b), and ANN modeling can achieve good performance when used to predict river flow. ANN

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

is a parallel-computing model that mimics information processing in the human brain (El-Shafie et al., 2008; El-Shafie and Noureldin, 2011). ANN does not require data regarding the physical characteristics of the study area (Dawson et al., 2001), but does, however, require historical data on the subject being studied. This data is normally

5 separated into a training dataset and a validation dataset. ANN learns the hidden patterns in the historical data through the training dataset. Once the learning process is completed and the knowledge is saved, forecasting can be done using new data input. To verify the success of the data training, forecasting results using the validation dataset are evaluated. Historical records should also be as accurate as possible to

10 ensure reliable forecasting results.

In river flow studies, forecasting normally means either forecasting the water leve1 (Alvisi et al., 2006; El-Shafie et al., 2008; Campolo et al.,1999; Leahy et al., 2008) or the runoff (Thirumalaiah and Deo, 1998; Shamseldin et al., 2002; Shrestha et al., 2005; Toth et al., 2000; Dawson et al., 2001). The source for data training can be based on

15 the same data type as the forecasting or it can come from several data types. Data types that are used as the source for data training are river flow (El-Shafie et al., 2008; Thirumalaiah and Deo, 1998; Shrestha et al., 2005), rainfall (Toth et al., 2000; Dawson et al., 2001), water level and rainfall (Alvisi et al., 2006; Campolo et al., 1999; Lihua et al., 2004; Boucher et al., 2010), water level and sea level pressure (Leahy et al., 2008),

20 and flow, rainfall, temperature and snowmelt (Coulibaly et al., 2000).

The rainfall runoff method for river flow forecasting has the drawback that this approach cannot be easily implemented because the data from rainfall and runoff are not easily synchronized in real time. It also requires infrastructure to manage the different data inputs (Dawson and Wilby, 1998). Further, although river flow results can be con-

25 verted into water level using a rating curve, this does not necessarily provide the same results as forecasting based on water level. To date, no study has shown that river flow forecasting is comparable with water level forecasting.

The most common problems in flood forecasting are underestimating the flow or water level (Alvisi et al., 2006; Thirumalaiah and Deo, 1998; Shrestha et al., 2005)

either at peak or low level events. The lack of data at the high points is generally suggested (Toth et al., 2000) as the cause of underestimating the peak water level. This is true for the highest points reached above flood level. However, it is not true for events where the water level is within flood alert or dangerous levels. The problem of underestimating overestimating may be caused by the approach in developing the models. The selection of training data where high proportion of the data comes from the normal zone could contribute to underestimating the water level at its peak. This is true if the same data type is used for the data input and data output of an ANN-based forecasting model. However, if other data sources such as rainfall are included in the data input, the situation is different depending on whether the importance of rainfall as a data input is higher or lower than the flow or water level data. Confidence in ANN forecasting results has been an issue raised by many skeptics, so it is important to provide an indication of the accuracy of the model (Dawson et al., 2001). Kerh and Lee (2006) show accuracy and reliability of training results by using a chart showing the distribution in errors of discharge. Using more evaluation methods can help to describe the reliability and accuracy of the ANN forecasting model.

It is always better to have a long lead-time with high accuracy in a forecasting model, so that preventative measures and flood warnings can be carried out in sufficient time. In practicality, a longer lead-time decreases the forecasting accuracy. Typically, a shorter lead-time is used for operational flood warning and a longer lead-time is used as guidance to take precautionary measures in case of a flood event.

The objective of this study is to investigate an ANN-based forecasting model to improve predictions of water level above the flood alert level, and also to find the best lead-time in terms of providing both high accuracy and satisfactory performance in forecasting. The study area is Rantau Panjang, Johor, Malaysia and the lead-time tested in the development of the forecasting model is from 1 to 6 h. The reliability of forecasting results is also a major focus in this study, since this will build confidence in the forecasting model that is developed.

2 Study area

The study area concerned in the development of this forecasting model is Kota Tinggi, Johor, Malaysia. Kota Tinggi is located on the banks of Johor River and has been hit with more than 12 flood events since 1963. It is an administrative town with a large
5 population, making the development of the forecasting model relevant. However, due to the lack of historical water level data at Kota Tinggi station, Rantau Panjang station, which is also on the banks of Johor River upstream from Kota Tinggi, has been selected as the study area. The Johor River basin is shown in Fig. 1.

The distance between the two stations is about 40 km and the lag time for river
10 flow from Rantau Panjang to Kota Tinggi is about 24 h. The normal water levels at Rantau Panjang and Kota Tinggi are 4 m and 2 m respectively, while the flood levels at the two stations are 9 m and 2.1 m. Figure 2 shows the normal, alert and danger water levels at Rantau Panjang station. Both locations have been hit with more than 12 flood events since 1963. There is a correlation between the flooding events that
15 occurred at the two stations that will not be described in this paper, but successful forecasting at Rantau Panjang could later assist with studies of flood events at Kota Tinggi. All of the flooding events at Rantau Panjang occurred during the Northeast monsoon which happens between November and March, shown in Fig. 3. The Johor River is about 122.7 km in length and drains an area of 2636 km^2. Its main tributaries
20 are Sayong River and Linggiu River. The river originates from Mount Gemuruh (109 m) and discharges into the Straits of Johor (0 m). The average annual precipitation for the Johor River catchment is 2.47 m.

3 Methodology

This study aims to improve the current forecasting approach by developing an ANN-
25 based model for high water level events in real time with the zoning matching approach (ZMA). The initial target lead-time for testing the model is 3 h and the target forecasted

water level is above 8000 mm. Several performance measures such as Nash-Sutcliffe efficiency, the correlation coefficient, root mean square error, the scatter index and three offset errors are used to evaluate forecasting results.

3.1 Artificial Neural Network

An Artificial Neural Network (ANN) is a parallel-computing mathematical model for solving dynamic nonlinear time series problems. There are many types of ANN, the most common being the multilayer perceptron neural network (MLP-NN) (Zhang et al., 1998) that is used in this study. The architecture of the MLP-NN, shown in Fig. 4, contains three types of layer that are ordered in sequence. The first layer is an input layer, the last layer is an output layer and there can be one or more hidden layers in between. Each layer consists of one or more neurons. The function of the neurons in the input layer is to receive data input and pass this data to the neurons in the second layer. The function of neurons in the hidden and output layers is to receive the input and the weight of input from the neurons in the previous layer and compute the activation transfer function (ATF). There are many types of ATF, and again we use the most common of which is the sigmoid function (Zhang et al., 1998; Maier et al., 2000). The equations for computing the inputs are shown in below:

$$\text{Input} = I_{i=0}^{N}(W_{ij}.X_{ij}) \text{ where } x_0 = 1 \tag{1}$$

$$\text{Output} = \frac{1}{1 + e^{-k.\text{Input}}}, \tag{2}$$

where x is the output from the previous neuron, w is the weight of the output and k is the gradient of the sigmoid function. An extra neuron x_0 is added in the input layer and in each hidden layer as shown in Fig. 4 with a fixed output value of 1. This is called the bias and its function is to stabilize the computed output between 0 and 1. It does not have any links to previous neurons. In most studies, the numbers of input and hidden neurons are determined by trial and error (Coulibaly et al., 2000; Joorabchi

Real-time flood forecasting

M. Sulaiman et al.

et al., 2007; Solaimani and Darvari, 2008; Turan and Yurdusev, 2009). The number of outputs is normally one, which can be a forecasted week, day or hour, or a forecast at M-hour intervals. The neurons in the network architecture are interconnected between the layers. These interconnections represent the flow of computation in the ANN. The nonlinear equation for forecasting water level WL at time $t + 1$ with N data inputs is defined as below:

$$\text{WL}_{t+1} = f(\text{WL}_t, \text{WL}_{t-1}, \ldots, \text{WL}_{t-N}, \text{WL}_1, \ldots, w_K) \tag{3}$$

where w is a weight and K is the number of weights. The computation process starts from input neurons where data inputs are received, and then propagates to hidden neurons and further to the neurons in the output layer, which produce the model output. The computational process described above is called feed-forward computation. If the number of neurons and layers are established, the only unknown parameters in the computation are the weights, since K can be computed based on the network architecture.

The process of data training is used to determine the weights. Data training is the process of using sample historical data as the input and output of the network model so that it can simulate the sample data. The training process involves feed-forward and back-propagation computation cycles. Back-propagation computation adjusts the weights of the output and hidden neurons based on the gradient descent method. These weights are normally initialized with random values to speed up the training process. Among the performance measures to evaluate the simulation are mean squared error, root mean squared error and sum of squared error (Zhang et al., 1998). Once data training is successfully completed, data forecasting can be made with new data input. To evaluate forecasting performance, validation data are used for the input to the network where only feed-forward computation processes the data. Several performance measures are applied to the output of the model and the outputs are compared with observations from the validation dataset to determine the accuracy and reliability of the network model developed.

3.2 Datasets

Hourly historical water level data from the period 1963–2008 were collected from the Department of Irrigation and Drainage (DID) as shown in Fig. 5. The data was divided into training and validation datasets. The training dataset is 37 yr of hourly water level data from the period 1963–1999 while the validation dataset is hourly water level data from 2000–2008, that is, about 9 yr of data.

The dataset is organized into sets of inputs and output based on Eq. (3), where the number of data inputs depends on the network requirements. In this study, the number of inputs ranges from three to seven water level observations prior to the forecasted period. The reason a minimum of three data inputs are used to determine the best number of inputs for forecasting is that pre-analysis using less than three data points resulted in a poorer forecasting performance than using three or more data inputs. This could be because there are not enough patterns in the data when using fewer inputs. On the other hand, pre-analysis using more than seven inputs did not improve forecasting performance, possibly due to too many input patterns causing the loss of a distinct pattern within the training dataset. The results presented in this study use from three to seven data inputs, which is reasonable for developing a water level forecasting model.

To forecast the water level M hours ahead with N data inputs requires the data input to consist of hourly water level data at times t, $t-1$ to $t-N$, where the interval between each time step is M hours. The reason that the interval is the same as the lead-time is so that the rate of change in water level is measured on a consistent scale. Among the data inputs, the hourly water level at the forecasted time $t+1$ is also required. During data training the hourly water level at the forecasted time is needed for the learning process, while in data forecasting the hourly water level at the forecasted time is required for validation. A schematic of the model for an interval of 3 h is shown in Fig. 6, where the interval between each time step is 3 h and the lead-time is also 3 h. For a lead-time of M hours, the interval between each time step is M hours. The water

Real-time flood forecasting

M. Sulaiman et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◄◄ ►►

◄ ►

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

level data are normalized for the network model based on the equation:

$$N_i = \frac{(O_i - O_{min})}{O_{max} - O_{min}},\qquad(4)$$

where N is the normalized value, O is the observed water level, O_{max} is the highest observed water level and O_{min} is the lowest observed water level.

3.3 Network model

This study uses network models with the following features. One hidden layer is used, since this is adequate for approximating non-linear equations, based on the universal approximation theorem (Hornik et al., 1989; Maier and Dandy, 2000). The number of input neurons is the same as the number of the data inputs, and the number of neurons in the hidden layer is the same as the number of neurons in the input layer. Pre-analysis using more hidden neurons than input neurons did not produce any significant improvement in forecasting performance, but it was apparent that more hidden neurons made the data training process slower. The activation transfer function used in the hidden and output neurons is the sigmoid function. Most studies use 1 as the value of the steepness coefficient in the sigmoid function, but here we use a value of 0.075. This is based on the study by Sulaiman et al. (2011), which shows that a steepness coefficient between 0.17 and 0.025, with similar numbers of hidden and input neurons, successfully achieves optimal daily water level forecasting performance.

We stop the data training process in this study when there is no improvement to the data training performance. Many ANN studies have reported that this approach could cause over-fitting, that is, the performance of the data training increases while the validation performance deteriorates. In this study, a small number of hidden neurons are used to avoid over-fitting. The final data forecasting performance is compared to the data training performance to verify that over-fitting does not occur. Data training performance is evaluated using Nash-Sutcliffe efficiency.

Real-time flood forecasting

M. Sulaiman et al.

3.4 Zoning Matching Approach (ZMA)

Two approaches to the development of a water level forecasting model are compared in this study. The first is the standard approach (SA), which is the common way of developing forecasting model. In this approach, all data in the training dataset is used in the training process and all data in the validation dataset is used in data forecasting. The forecasting performance for this approach is more general, meaning that the performance represents forecasting of all water levels, whether low, normal or high. The approach introduced in this study is ZMA, which selects the training dataset based on the target water level for forecasting. The target water level in this study is 8000 mm, which is 1 m below flood danger level. The alert water level (7000 mm) is not used as the target water level, because water levels above 8000 mm are more critical. However, the final network model will be tested using the alert water level to look at the possibility of whether a single forecasting model can be used successfully to not only to forecast water levels above 8000 mm, but also above 7000 mm.

3.5 Stages in the analysis

In this study, there are four stages in the development of forecasting models for high water level events.

The first stage is to evaluate five network models that use SA to forecast water level above 0 mm and above 8000 mm with a lead-time of 3 h. We use 3 h initially as a guide since it is in between 1 h and 6 h, and we also intend to find the best lead-time in this study. The total number of records in the training dataset is about 290 000. Each record consists of a data set for inputs and output of the network model. The architecture of the five network models Net_3 to Net_7 is shown in Table 1. The first objective is to show the current weaknesses in SA when it is used to forecast high water level events. The second objective is to determine the best model to compare with and evaluate ZMA in forecasting high water level events.

The second stage is to find the best training dataset using ZMA, combining it with the network model found in the first stage to improve forecasting performance of high water level events with a lead-time of 3 h. Four training datasets, ZMA_7, ZMA_8, ZMA_9 and ZMA_10, consisting of water level data above 7000 mm, 8000 mm, 9000 mm and 10 000 mm respectively, are used for data training. In comparison to SA, which uses about 290 000 records, ZMA has fewer than 2900 records, or about 1 % of the total data training records.

The third stage is to determine the best lead-time for achieving high forecasting accuracy model, in addition, the best satisfaction lead-time forecasting results for water levels above 8000 mm. this is accomplished by evaluating the proposed lead-time range from 1 to 6 h utilizing the best dataset achieved in the second stage. It should be noted that, the lead-time 3 h is already evaluated, then, the model is examined for lead-time of 1, 2, 4, 5 and 6 h. The data inputs and output model for the lead-time testing is shown in the Fig. 6, where the interval between time steps is adjusted to equal the lead-time.

Finally, in the fourth stage we evaluate the performance of forecasting water levels above 7000 mm using the two forecasting models found in the third stage. The aim here is to extend the capability of the two models in terms of forecasting water levels above the alert level.

3.6 Performance measures

There are seven performance measures that are used to evaluate the results in this study. Four of these come from common statistical indices that evaluate the goodness of fit between two data series. These are the Nash-Sutcliffe efficiency coefficient (NSC), the correlation coefficient ($R2$), root mean square error (RMSE) and the scatter index (SI). The equations for these are shown below:

$$NSC = 1 - \frac{I_{i=1}^{N}(O_i - F_i)^2}{I_{i=1}^{N}(O_i - \overline{O})^2},$$

(5)

$$R2 = \frac{I_{i=1}^{N}(O_i - \overline{O})(F_i - \overline{F})}{\sqrt{I_{i=1}^{N}(O_i - \overline{O})^2 I_{i=1}^{N}(F_i - \overline{F})^2}}, \tag{6}$$

$$RMSE = \sqrt{\frac{I_{i=1}^{N}(F_i - O_i)^2}{N}}, \tag{7}$$

$$SI = \frac{RMSE}{\overline{O}}, \tag{8}$$

where i is the record number, O is the observed value, F is the forecast value and N
is the number of records evaluated. The performance ranking of the indices depends
on their value as follows. For NSC and $R2$, a value of one means a perfect fit and for
RMSE and SI, a value of zero means the best fit. The unit for RMSE is millimeters.
In this study, a value for NSC and $R2$ between 0.95 and 1 means a strong perfor-
mance, a value between 0.9 and 0.95 is satisfactory, and below 0.9 is unsatisfactory.
RMSE below 100 mm is strong accuracy, a value between 100 mm and 500 mm is sat-
isfactory and greater than 500 mm is unsatisfactory. An SI value between 0 and 1 is
a strong performance, a value between 1 and 2 is satisfactory, and greater than 2 is
unsatisfactory.

Additional performance measures that are included in this study are percentage of
offset error above 200 mm, percentage of offset error above 500 mm and highest offset
error. Offset error is the difference between the forecasted and observed water levels.
The equation for the offset error is:

$$Offset_i = (F_i - O_i), \tag{9}$$

where i is at the time step, F is the forecasted value and O is the observed value.
The allowable offset error is used as an evaluation tool because it describes the pre-
cision of the forecasting results in physical values, which can help water authorities to
understand and sense the accuracy of the forecasting model developed. In this study,

less than 5 % of offset errors above 200 mm means a very accurate forecasting result is achieved. This means that 95 % or more of the offset errors are below 200 mm. We use 200 mm as the criteria for high accuracy since 200 mm is very small in terms of water level. A break point of 500 mm is used as an acceptable offset error for a satisfactory forecasting model in this study. Less than 5 % of offset errors above 500 mm means a satisfactory forecasting performance is achieved. This means that more than 95 % of offset errors are below 500 mm. Lastly, the highest offset error describes the worst case of water level error expected from the forecasting model developed. Charts showing the distribution of offset errors are presented to show the accuracy and reliability of the forecasting models.

4 Results and discussion

As described earlier, the process of finding the two best forecasting models and the corresponding best lead-times for forecasting with high accuracy and satisfactory performance is divided into four stages.

The results for the first stage are shown in Table 2. The table shows the forecasting performance results for the five network models that are based on the SA method. The number of iterations in the data training is only 20, which is quite a small number. The reason for this is that the training performance for the entire SA had already achieved an NSC greater than 0.995 after the first 20 iterations. It takes about 15 min to compute the 20 iterations for each of the network models, which is time consuming. The reason for the slow computing speed is the number of records in the training dataset, which as described earlier is 62 000. The forecasting results for the five network models are divided into two sections. The first section evaluates the performance for forecasting water levels above 0 mm, which means using all the water level data in the validation datasets. The second section evaluates the performance for forecasting water levels above 8000 mm, which is the focus of this study. The results for the first section correspond to the SA forecasting model that uses all available data from the training

dataset and makes forecasts for all data in the validation dataset. In this section, the five networks model show strong forecasting performance according to NSC, $R2$, and RMSE. The five network models have NSC above 0.99, $R2$ above 0.99 and RMSE below 100 mm. This shows a strong forecasting performance by all five models. This shows that any of the five network models is suitable as a water level forecasting model at Rantau Panjang. However, further evaluation using SI, the number of offset errors above 200 mm, the number of offset errors above 500 mm and the highest offset error shows otherwise. The SI for each of the network models is above 2.0, much greater than 0, which is the best possible value. Of the 62 000 forecasted water levels, the number of offset errors above 200 mm is between 650 and 1245 records (1–2 %), and the number of offset errors above 500 mm is between 178 and 255 records (0 %), depending on the lead-time. Even though the percentage is very small the number of errors is high. The highest offset error in the forecasted water level is about 1800 mm or 1.8 m, which is also quite high. The last four parameters measuring the performance indicate that there are many errors in the forecasting results using the SA-based model. The second section is the evaluation of forecasting performance for water levels above 8000 mm. The results emphasize the poor forecasting performance of SA for high water level events. The NSC values for all the models are below 0.8. However, the $R2$ values show strong performances, with values above 0.9, and the RMSE values indicate acceptable performance, that is, within 500 mm. The SI for all SA network models is above 5, which shows poor performance. The number of offset errors above 200 mm is between 234 and 334 (31–44 %) and the number of offset errors above 500 mm is more than 130 (17 %). The results show that there is a high percentage offset errors in the forecasting results for high water events. Of the five network models, NET_5 produce the best forecasting performance for water levels above 8000 mm. However, Table 2 still shows that none of the SA forecasting models is suitable for forecasting water levels above 8000 mm with a 3-h lead-time.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

The results for stage two are shown in Table 3. For the four training datasets in the model with ZMA, this table shows the effect on water level forecasting performance for water levels above 8000 mm with a lead-time of 3 h. Recall that NET_5, which is the best model developed in the first stage, is used as the basic model in which ZMA is tested. There are dramatic improvements in performance with ZMA as compared to the forecasting performance of SA. The NSC values for ZMA are all above 0.99, whereas the best NSC for SA is only 0.796, representing an improvement from poor performance to high accuracy. There is not much difference between the $R2$ values for SA and ZMA. In Table 3, the $R2$ for ZMA is above 0.997, while $R2$ in Table 2 is 0.970. However, RMSE for ZMA again shows another dramatic improvement over SA, with RMSE values around 100 mm, indicating high accuracy in performance. The RMSE for SA is about 500 mm, so there is about an 80 % improvement with ZMA. The SI for all ZMA trials also show accurate forecasting, with all SI values less than 1. The number of offset errors greater than 200 mm for ZMA is between 31 and 59, while for SA it is above 234, and while there are more than 283 cases of offset errors greater than 500 mm using SA, for ZMA there are at most two. The highest offset error in the SA model is 1669 mm, while for ZMA it is 479 mm. The last three performance measures show a high accuracy of forecasting performance by ZMA. There is no clear best ZMA model for the 3-h time-step, since each of the models has a different advantage in terms of the performance criteria. In fact, all of these models produced good forecasting results. The authors select ZMA_9 as the best, since its RMSE is 90 mm, the SI is 0.986 and it has about 4 % offset errors that greater than 200 mm. The charts shown in Fig. 7 help to visualize the effects of ZMA in high water level forecasting. In the figure, all the ZMA models have forecasted values very close to the observed values except at the highest water level events, where, ZMA_9 has the closest fit. The number of epochs for each ZMA is shown in Table 3 along with the forecasting performance measures, although this study focuses on achieving the best data training performance and the computational time is not critical. In this study, ZMA_9 for the 3-h time-step took about an hour to compute.

The results for the third stage are shown in Table 4, which shows the performance of NET_5 and ZMA_9 for lead-times from 1 to 6 h. The accuracy of forecasting results for 1 to 3 h are similarly high, except that the performance at 3 h is slightly worse than for 1 and 2 h. However, all three of these lead-times show high accuracy, with NSC above 0.99, RMSE below 100 mm, and SI below 1. Further, there is only one offset error of more than 500 mm, and this is also the highest error at 509 mm. Thus, the best lead-time to ensure a high accuracy forecasting model is 3 h, since the longer lead-time allows for more advanced predictions of flooding events. For lead-times of 4 to 6 h, the forecasting performances are all satisfactory, with only 2 % offset errors above 500 mm and a highest error of 702 mm. The other measures show strong performance with the NSC above 0.97, RMSE below 200 mm, and the SI below 2. Hence the 6-h lead-time is selected as the satisfactory model since it has the greater lead-time time. The overall results show that as the lead-time increases, the forecasting performance decreases. This is to be expected, due to the fact that a shorter time interval provides better forecasting accuracy. Figures 8 and 9 show comparisons between the observed and forecasted values for lead-times of 3 h and 6 h respectively, representing the high accuracy model and the satisfactory model found in this study. To highlight the accuracy of the 3- and 6-h lead-times in monitoring flood events, two cases of flooding that occurred in December 2006 and January 2007 in Rantau Panjang are shown in Figs. 10 and 11. The general pattern to observe in the figures is that the offset error is quite high as the flooding starts, but the offset is small as the flood recedes. During the peak water level event in January 2007, the model under-estimated the peak water level, possibly due to the fact that not many historical flooding events above 12 000 mm have occurred. In contrast, for the December 2006 event the model is able to accurately forecast peak water levels.

The results for the fourth stage, forecasting water levels above 7000 mm, are also shown in Table 4. The performance for a lead-time of 3 h is similar to the result for water levels above 8000 mm. The percentages of offset errors above 200 mm are 4 % and 3 % for water levels above 7000 mm and 8000 mm respectively, and there are only

Real-time flood forecasting

M. Sulaiman et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

one or two cases of an offset error above 500 mm. Forecasting performance for a 6-h lead-time at water levels above 7000 mm and 8000 mm also show small differences in the number of offset errors and the statistical indices. Thus, both models can be used for forecasting water levels above 7000 mm and 8000 mm, with one model having high accuracy and the other being satisfactory.

The error distribution charts shown in Fig. 12 highlight the accuracy of the forecasting results for the best SA model with a 3-h lead-time, the best ZMA model with a 3-h lead-time, and the best ZMA model with a 6-h lead-time in forecasting water levels above 8000 mm. Figure 12a for the SA model clearly shows that most of the offset errors are non-zero, while a high percentage of the offset errors are greater than 500 mm. Meanwhile, in Fig. 12b the offset errors are within 500 mm and most of the errors are close to zero. Even in Fig. 12c, most of the offset errors are within 500 mm, with a few points that are above 500 mm, but still below 1000 mm. Another result in the study is that all data training performances using ZMA are slightly better than the forecasting performances shown in Tables 3 and 4. This means that over-fitting has not occurred in this study, since the small number of hidden neurons helps to avoid this.

5 Conclusions

In this paper, an ANN based model with ZMA is reviewed in the development of real time water level forecasting at Rantau Panjang station, specifically to monitor high water level events. The forecasting models developed using ZMA are able to forecast water levels above 8000 mm with high accuracy with a lead-time of up to 3 h, and with satisfactory performance at 6 h. Two cases of high flooding events that occurred in December 2006 and January 2007 are successfully tracked by the model that is developed. The strength of ZMA is that the selection of training data is based on the forecasting target. The trial and error method using several datasets is needed to obtain the best training dataset for forecasting the target water level. On the other hand, SA is not suitable for forecasting high water levels, because the high volume of normal

water level in the training dataset causes high water level data to be ignored during the training process. Since the properties of high water level events is not learned well by the model, it cannot make good predictions, and thus the forecasting results for high water level events are poor.

Another important finding of this study is that the common statistical performance measures such as NSC, $R2$ and RMSE cannot be assume to indicate specific target water level forecasting performance, especially when the weight of target data in the training dataset is minimal. Additional measures such as SI, an offset distribution chart, scatter plot, allowable offset error, and highest offset error are needed to provide a clearer view of the quality and reliability of forecasting results so that relevant authorities can have confidence in the model. In summary, the complexity of water level forecasting can be addressed by using ANN models. The forecasting model developed may assist appropriate water-monitoring authorities to take preventative measures before flooding occurs, and help managing flood operations. Further study of ZMA on low and normal water level data could enhance the real time forecasting capability at Rantau Panjang station.

In general, the results of applying neural networks to water level forecasting are promising. However, the proposed ANN models still lack an appropriate method for finding the optimum architecture. In addition, preprocessing of the data is an essential step for time series forecasting and more survey and analysis may lead to better accuracy in this application. The selection of the parameter set and components within an ANN model and the variable selection procedures (input pattern) were all attempted in this study. However, optimal selection of the parameters requires augmenting the ANN model with some other optimization model, such as the genetic algorithm or particle swarm optimization methods. On the other hand, variable selection (input pattern) in the ANN model is always a challenging task due to the complexity of the hydrologic process. Another advanced ANN model, namely the Dynamic Neural Network (DNN), considers the time-dependent interrelationships between the input and output patterns and may provide better modeling results. Furthermore, more robust input

pattern selection approaches (for example, systematic searching for optimal or near optimal variable combinations in DNN with the ensemble procedure) can be explored and may lead to important new methods for water level forecasting.

In addition to improved accuracy, forecasting is concerned with assessing uncertainty. Traditional error measures, such as the mean square error (MSE), do not provide a reliable basis for comparison of forecasting methods. The median absolute percentage error is more appropriate, because it is scale-invariant and not influenced by outliers. When comparing methods, especially when the number of series is small, we can control for the degree of difficulty by using the median relative absolute error, which compares the error for a given model against errors for the naïve "no change" forecast. The fit of a model to historical data is a poor way to estimate prediction intervals, as it typically results in confidence intervals that are too narrow. It is best to simulate the actual forecasting procedure as closely as possible, and use the distribution of the resulting ex ante forecasts to assess uncertainty.

References

Alvisi, S., Mascellani, G., Franchini, M., and Bárdossy, A.: Water level forecasting through fuzzy logic and artificial neural network approaches, Hydrol. Earth Syst. Sci., 10, 1–17, doi:10.5194/hess-10-1-2006, 2006.
ASCE Task Committee on the application of ANN in Hydrology: Artificial Neural Networks in Hydrology. I: Preliminary Concepts, J. Hydrol. Eng., 5, 115–123, 2000a.
ASCE Task Committee on the application of ANN in Hydrology: Artificial Neural Networks in Hydrology. II: Hydrological Applications, J. Hydrol. Eng., 5, 124–137, 2000b.

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Boucher, M.-A., Laliberté, J.-P., and Anctil, F.: An experiment on the evolution of an ensemble of neural networks for streamflow forecasting, Hydrol. Earth Syst. Sci., 14, 603–612, doi:10.5194/hess-14-603-2010, 2010.

Campolo, M., Andreussi, P., and Soldati, A.: River flood forecasting with a neural network model, Water Resour. Res., 35, 1191–1197, 1999.

Coulibaly, P., Anctil, F., and Bobee, B.: Daily reservoir inflow forecasting using artificial neural networks with stopped Training Approach, J. Hydrol., 230, 244–257, 2000.

Dawson, C. W. and Wilby, R. L.: An artificial neural network approach to rainfall-runoff modelling, Hydrol. Sci. J., 43, 47–66, 1998.

Dawson, C. W. and Wilby, R. L.: Hydrological modelling using artificial neural networks, Prog. Phys. Geogr., 25, 80–108, 2001.

El-Shafie, A. and Noureldin, A.: Generalized versus non-generalized neural network model for multi-lead inflow forecasting at Aswan High Dam, Hydrol. Earth Syst. Sci., 15, 841–858, doi:10.5194/hess-15-841-2011, 2011.

El-Shafie, A., Noureldin, A. E., Taha, M. R., and Basri, H.: Neural network model for Nile river inflow forecasting based on correlation analysis of historical inflow data, J. Appl. Sci., 8, 4487–4499, 2008.

Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, Neural Networks, 2, 359–366, 1989.

Joorabchi, A., Zhang, H., and Blumenstein, M.: Application of Artificial Neural Networks in Flow Discharge Prediction for Fitzroy River, Australia, J. Coastal Res., SI 50, 2007.

Kerh, T. and Lee, C. S.: Neural networks forecasting of flood discharge at an unmeasured station using river upstream information, Adv. Eng. Softw., 37, 533–543, 2006.

Leahy, P., Kiely, G., and Corcoran, Gearo 'id: Structural optimisation and input selection of an artificial neural network for river level prediction, J. Hydrol., 355, 192– 201, 2008.

Maier, H. R. and Dandy, G. C.: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, Environ. Modell. Softw., 15, 101–124, 2000.

Modarres, R.: Multi-criteria validation of artificial neural network rainfall-runoff modeling, Hydrol. Earth Syst. Sci., 13, 411–421, doi:10.5194/hess-13-411-2009, 2009.

Shamseldin, A. Y., Nasr, A. E., and OConnor, K. M.: Comparison of different forms of the Multilayer Feed-Forward Neural Network method used for river flow forecasting, Hydrol. Earth Syst. Sci., 6, 671–684, doi:10.5194/hess-6-671-2002, 2002.

Real-time flood forecasting

M. Sulaiman et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

| ◀◀ | ▶▶ |
| ◀ | ▶ |

| Back | Close |

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Shrestha, R. R., Theobald, S., and Nestmann, F.: Simulation of flood flow in a river system using artificial neural networks, Hydrol. Earth Syst. Sci., 9, 313–321, doi:10.5194/hess-9-313-2005, 2005.

Solaimani, K. and Darvari, Z.: Suitability of Artificial Neural Network in Daily Flow Forecasting, J. Appl. Sci., 8, 2949–2957, 2008.

Sulaiman, M., El-Shafie, A., Karim, O., and Basri, H.: Improved water level forecasting performance by using optimal steepness coefficients in an artificial neural network, Water Resour. Manage., 25, 2525–2541, 2011.

Thirumalaiah, K. and Deo, M. C.: Real time flood forecasting using neural networks, Comput.-Aided Civ. Inf., 13, 101–111, 1998.

Toth, E., Brath, A., and Montanari, A: Comparison of short-term rainfall prediction models for real time flood forecasting, J. Hydrol., 239, 132–147, 2000.

Turan, M. E. and Yurdusev, M. A.: River flow estimation from upstream flow records by artificial intelligence methods, J. Hydrol., 369, 71–77, 2009.

Xiong, Lihua, OConnor, Kieran M., and Guo, Shenglian: Comparison of three updating schemes using artificial neural network in flow forecasting, Hydrol. Earth Syst. Sci., 8, 247–255, doi:10.5194/hess-8-247-2004, 2004.

Zhang, G., Patuwo, B. E., and Hu, M. Y.: Forecasting with artificial neural networks: The state of the art, Int. J. Forecasting, 14, 35–62, 1998.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀◀ | ▶▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Network models in the study.

Network Model	Input neurons	Hidden neurons	Output neuron
NET_3	3	3	1
NET_4	4	4	1
NET_5	5	5	1
NET_6	6	6	1
NET_7	7	7	1

Table 2. Forecasting performances for the five SA network models.

Network model	NSC	RMSE	Offset errors			Totalrecords	Epochs
	$R2$	SI	> 200 mm (%)	> 500 mm (%)	Highest		
Section 1: forecasting performance for water level above 0 mm							
NET_3	0.996	78	650 (1 %)	178 (0 %)	1908	62 753	20
	0.998	2.144					
NET_4	0.995	85	789 (1 %)	198 (0 %)	1890	62 708	20
	0.998	2.328					
NET_5	0.995	85	943 (2 %)	211 (0 %)	1891	62 663	20
	0.998	2.333					
NET_6	0.994	94	1192 (2 %)	255 (0 %)	1891	62 618	20
	0.998	2.588					
NET_7	0.994	94	1245 (2 %)	248 (0 %)	1876	62 573	20
	0.998	2.577					
Section 2: forecasting performance for water level above 8000 mm							
NET_3	0.765	488	250 (33 %)	135 (18 %)	1794	762	20
	0.976	5.446					
NET_4	0.756	498	255 (33 %)	139 (18 %)	1815	762	20
	0.973	5.552					
NET_5	0.796	454	234 (31 %)	133 (17 %)	1669	762	20
	0.973	5.009					
NET_6	0.701	551	334 (44 %)	165 (22 %)	1856	762	20
	0.970	6.202					
NET_7	0.748	505	289 (38 %)	157 (21 %)	1728	762	20
	0.968	5.654					

Table 3. Forecasting performances for the four ZMA models and NET_5 for water levels greater than 8000 mm.

Network model	NSC	RMSE	Offset errors			Total records	NSC in Data Training & Epoch
	$R2$	SI	> 200 mm (%)	> 500 mm (%)	Highest		
ZMA_7	0.985	122	59 (8 %)	0 (0 %)	495	762	0.99711601
	0.996	1.335					
ZMA_8	0.991	93	44 (6 %)	0 (0 %)	479	762	0.99812236
	0.997	1.021					
ZMA_9	0.992	90	31 (4 %)	1 (0 %)	509	762	0.99917507
	0.998	0.986					
ZMA_10	0.990	101	36 (5 %)	2 (0 %)	541	762	0.99820154
	0.998	1.111					

Table 4. Forecasting performances for ZMA_9 and NET_5 for 1 to 6 hour intervals.

| Network model | NSC | RMSE | Offset errors | | | Total records | NSC in Data Training & Epoch |
	$R2$	SI	> 200 mm (%)	> 500 mm (%)	Highest		
			Section 2: forecasting performance of water level above 8000 mm				
1	0.997	56	16 (2 %)	0 (0 %)	288	762	1.00015539
	0.999	0.608					
2	0.993	84	31 (4 %)	0 (0 %)	390	762	0.99912266
	0.997	0.922					
3	0.992	90	31 (4 %)	1 (0 %)	509	762	0.99812236
	0.998	0.986					
4	0.989	107	46 (6 %)	3 (0 %)	576	762	0.99713544
	0.996	1.168					
5	0.984	127	64 (8 %)	7 (1 %)	654	762	0.99411844
	0.994	1.385					
6	0.976	157	115 (15 %)	14 (2 %)	702	762	0.99112645
	0.99	1.717					
			Section 2: forecasting performance of water level above 7000 mm				
3	0.996	73	42 (3 %)	2 (0 %)	524	1513	0.99812236
	0.999	0.882					
6	0.985	138	156 (10 %)	26 (2 %)	737	1513	0.99112645
	0.994	1.666					

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
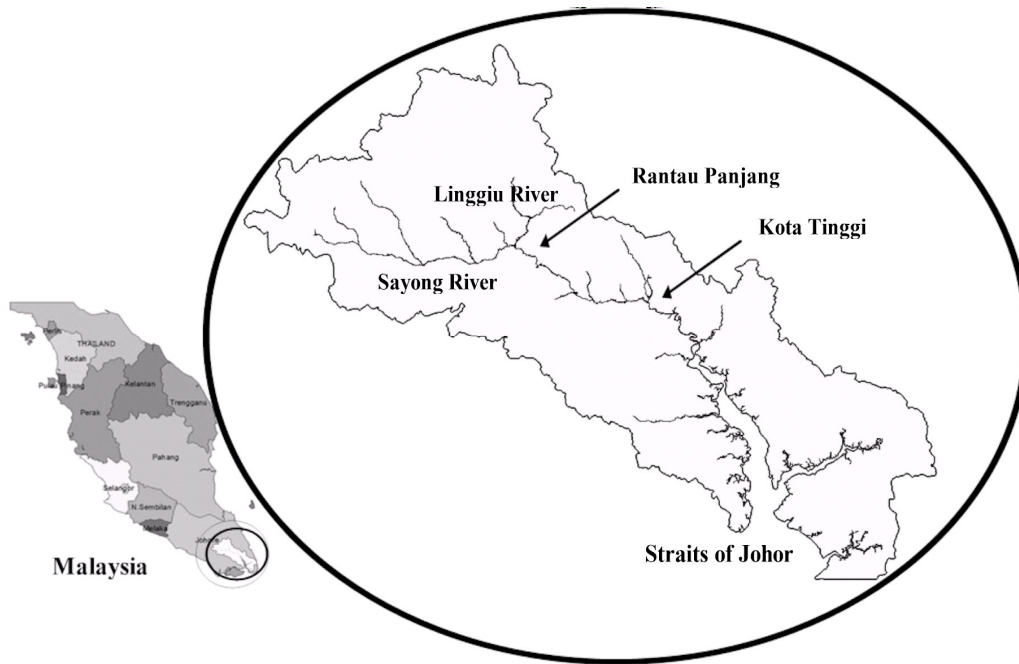
Fig. 1. Johor River basin, Malaysia.
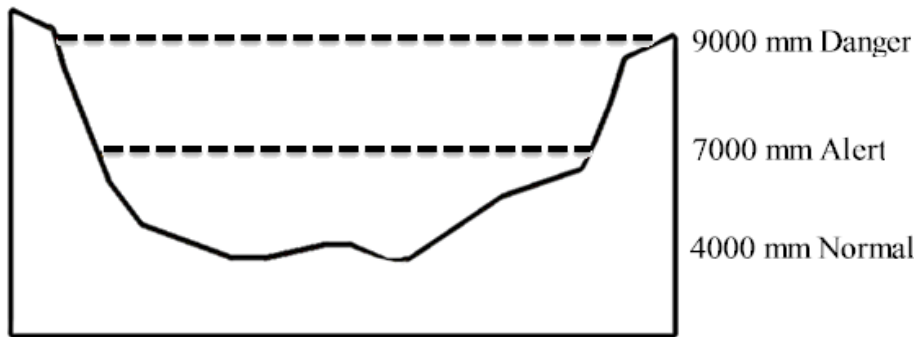
Fig. 2. Normal, alert and danger levels at Rantau Panjang station.

9000 mm Danger

7000 mm Alert

4000 mm Normal

Fig. 3. Flooding events recorded at Rantau Panjang station.

Input   Input Layer   Hidden Layer   Output Layer   Output

$i_0=1$

WL[t-1] → i_1 $j_0=1$

$j_0=f(\sum w_{ij}.x_{ij})$

WL[t-2] → i_1 $j_1=f(\sum w_{ij}.x_{ij})$ $k=f(\sum w_{jk}.x_{jk})$ → WL[t]=k_0

..   ..   ..

WL[t-n] → i_n $j_n=f(\sum w_{ij}.x_{ij})$
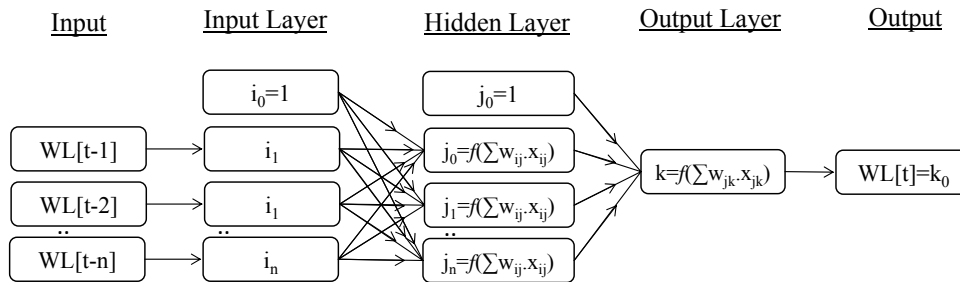
Fig. 4. Network model architecture.
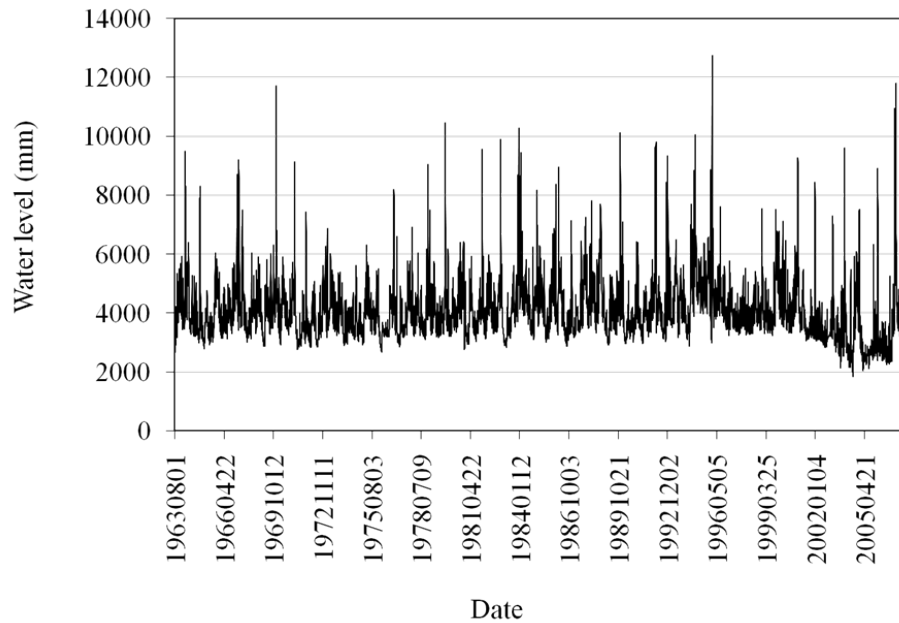
Fig. 5. Daily water level at Rantau Panjang station for period August-1963 to June-2008.

Time step	t-N	..	t-2			t-1			t			t+1
						Current			**forecast**			
Time	10-M*N	..	4	5	6	7	8	9	10	11	12	13

Fig. 6. Data inputs and output model for a lead-time of 3 h. M is the interval of the time-step and *N* is the number of data inputs.
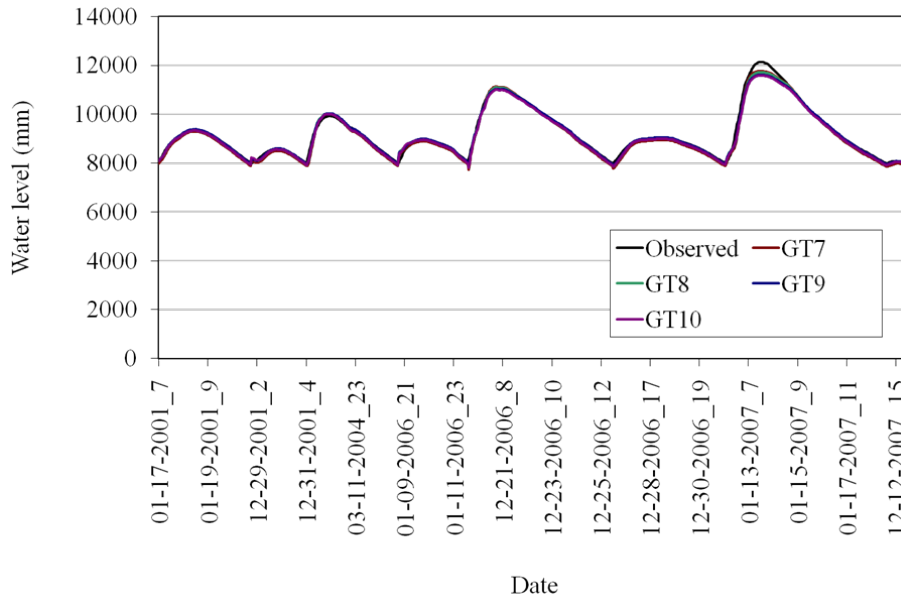
Fig. 7. ZMA_7 to ZMA_10 with 5five inputs and a lead-time of 3 h; forecasted values and observed values.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

⊩ | ⊫

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
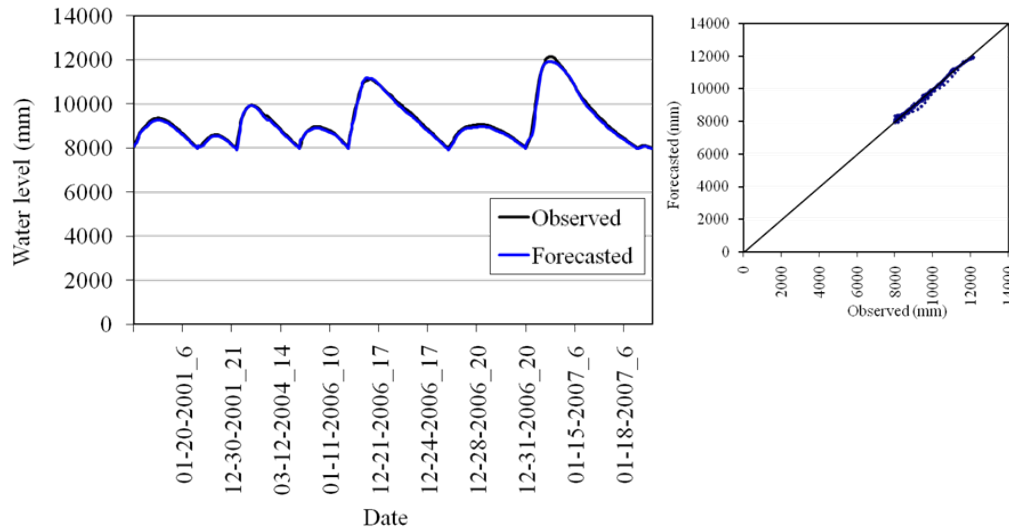
Fig. 8. Forecasting results at 3-hour intervals for water levels above 8 m using ZMA_9.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
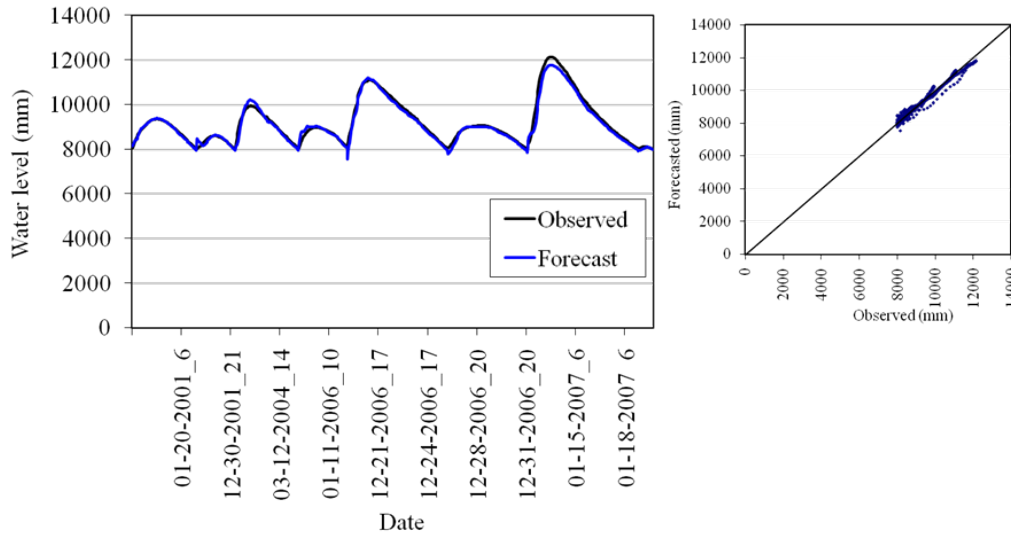
Fig. 9. Forecasting results at 6-h intervals for water levels above 8 m using ZMA_9.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
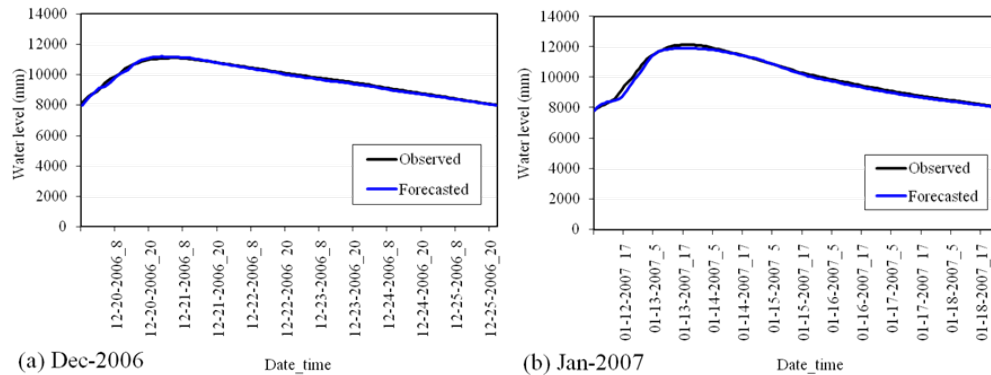
Fig. 10. Comparisons between observations and model outputs for 3-h intervals during 2006–2007 flood events.
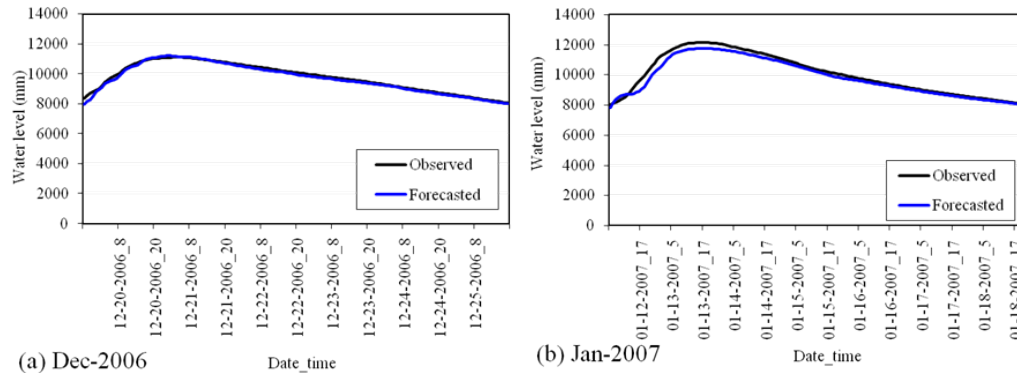
Fig. 11. Comparisons between observations and model outputs for 6-h intervals during 2006–2007 flood events.
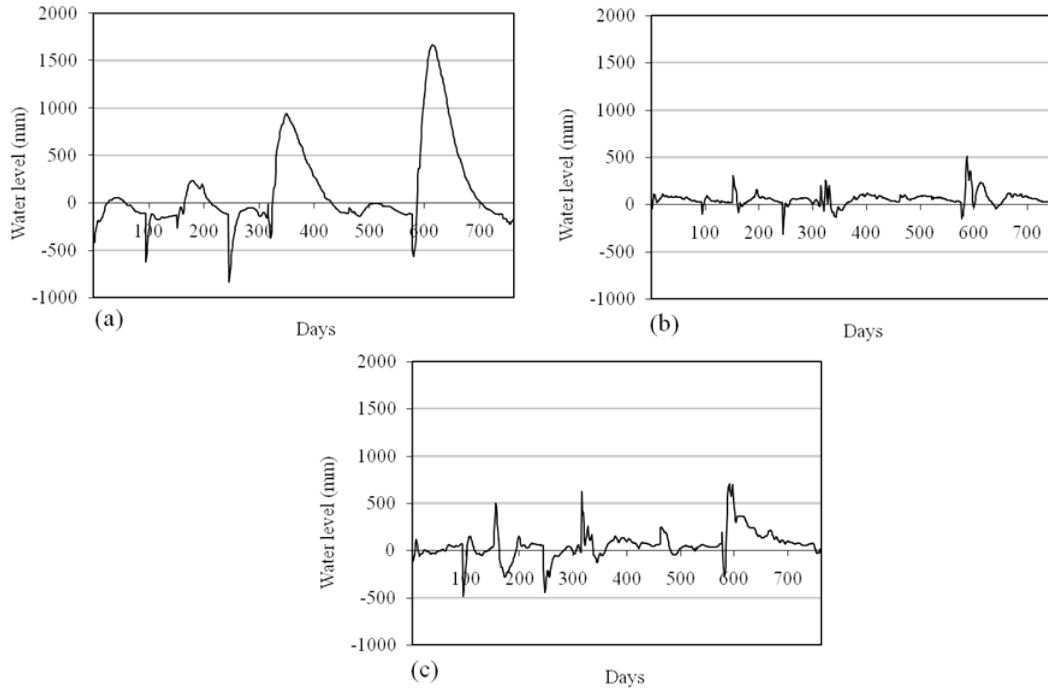
Fig. 12. Distribution of errors for forecasting results **(a)** SA at 3-h intervals **(b)** ZMA_9 at 3-h intervals **(c)** ZMA_9 at 6-h intervals.