Advances in River Sediment Research – Fukuoka et al. (eds)
© 2013 Taylor & Francis Group, London, ISBN 978-1-138-00062-9

Real-time sediment inflow prediction for sediment bypass operation at Miwa Dam in Japan

D. Nohara & T. Sumi
Disaster Prevention Research Institute, Kyoto University, Uji, Japan

S.A. Kantoush
Civil Engineering Department, German University in Cairo, New Cairo City, Egypt

ABSTRACT: Reservoir sedimentation has been an important issue at dams in the world. Evacuating sediments with sediment bypass tunnels, which allow sediments from upstream of reservoirs to pass through the reservoirs without being captured by the reservoirs, is expected to be an effective countermeasures against the issue. In this paper, real-time prediction models of sediment inflow are developed for efficient operation of a sediment bypass tunnel at Miwa Dam in the Mibu River basin in Japan. Two kinds of prediction models, namely a Multivariable Linear Regression (MLR) model and Artificial Neural Network (ANN) models, are respectively developed in order to predict sediment inflow for the coming three hours with hourly time resolution. Through application of developed prediction models, it was demonstrated that the ANN model showed better performance than MLR model, especially in terms of time series predictability.

1 INTRODUCTION

Reservoir sedimentation has been an important issue at dams in the world since it decreases the available capacity of the reservoirs and consequently brings decrease in the efficiency of the reservoirs. Effective sediment management is therefore needed for sustainable use of reservoirs.

As one of the countermeasures against reservoir sedimentation, sediment bypasses, which allow sediments from the upstream to pass through the downstream of reservoirs by diverting them with bypass tunnels during the flooding situation, have been installed at some reservoirs. This measure is considered as one of the effective measures to overcome the reservoir sedimentation issues, because it can pass through without letting them enter the reservoir and therefore has a potential to be a permanent solution for sediment management at reservoirs. It is, however, needed for reservoir managers to have foresight into the future sediment inflow in order to operate the sediment bypass tunnels effectively and adequately. Real-time prediction of the time series of sediment inflow to the reservoir is therefore important to decide when and how long the bypass tunnels to be operated.

Studies on estimation of sediment amount or concentration in the river water have been increasing in this decade. Nagy *et al*. (2002) developed Artificial Neural Network (ANN) models for estimation of sediment load concentration in rivers with physical parameters related to river flows or river channels. Cigizoglu & Alp (2006) developed two kinds of ANNs, namely, feed forward back propagation and generalized regression neural networks in order to model daily river sediment yield. Cobaner *et al*. (2009) developed estimating models of daily Suspended Sediment Concentration (SSC) by employing neural network techniques as well as conventional Sediment Rating Curves (SRC), and demonstrated advantages of neural network models for the estimation of SSC. Rajee *et al*. (2009) developed ANN and neuro-fuzzy models for simulation of daily SSC, and concluded that these models could be suitable substitutes for the conventional linear models such as Multi-Linear Regression (MLR) or SRC methods. These studies mainly focused on estimating sediment concentration or amount at daily basis.

On the other hand, reservoir operations are often conducted with shorter time step such as hourly basis during flooding situation especially at reservoirs which have small mountainous catchment areas where flood water runoffs quickly to the reservoirs. Sediment amount transported by flood waters of the rivers also changes quickly as flood situation changes. Thus prediction of sediment discharge from the upstream with shorter time resolution is needed for effective operation of sediment bypass tunnels so as to pass the sediments through the dams when more amount of sediment is discharged. A few studies have, however, been reported that focused on estimation of sediment yield with short time resolution (Rai & Mathur, 2008; Chutachindakate & Sumi, 2009) as it is more challenging to develop such models in terms of data availability and complexity in variability of sediment yield.

From the point of view described above, short-term prediction models of sediment inflow to a reservoir during flood events are developed in this study for effective decision making in the operation of a bypass tunnel installed into the reservoir. Two types of prediction models, namely, MLR and ANN models, are developed here in order to predict sediment inflow for the coming three hours with one hour resolution considering hydrological states such as rainfall or water inflow. The models are developed for Miwa Dam in the Mibu River, Central Japan, which has a precipitous catchment area and equips a sediment bypass tunnel to discharge a huge amount of sediment from the upstream during flood events.

2 STUDY AREA

Miwa Dam is a multi-purpose dam for flood control, power generation and irrigation with 29,952,000 m^3 of total storage capacity, and is located in the Mibu River, a tributary of the Tenryu River in Central Japan. The catchment area of Miwa Dam lies in the mountainous area around Itoigawa-Shizuoka Tectonic Line and receives approximately 2,000 mm of annual precipitation, which causes intensive sediment yield to the river and subsequently to the dam reservoir. In order to mitigate sediment deposition in the reservoir, a sediment bypass tunnel has been installed to Miwa Dam since 2005 to divert inflow water containing a large quantity of sediment and to discharge them directly to the downstream of the dam without letting the sediment come into the reservoir during major flooding events. The bypass tunnel is designed to be operated when inflow water amount is expected to exceed 300 m^3/s so that it can divert flood water with a large amount of sediment effectively while it can store river water with less amount of sediment for the purposes other than flood control.

On the other hand, Miwa Dam has the small catchment which surface area is 311.1 km^2 with steep geography in the mountainous region. Water and sediment are therefore discharged quickly and cause rapid increase in water and sediment inflows at the reservoir within several hours after heavy rainfall has started. Due to these characteristics, short-term prediction of sediment inflow at the reservoir with short time resolution like one hour unit is crucially important for the effective operation of the sediment bypass tunnel at Miwa Dam during flood events.

3 METHODOLOGY

3.1 *Outline of methodology*

Prediction models of hourly sediment inflow for the coming three hours are developed in this study in order to support real-time decision makings for operation of the sediment bypass tunnel. Hourly SSC of the inflowing water to the reservoir is employed as the predicted variable for sediment inflow.

Candidates for predictors are selected from among variables observed at Miwa Dam including rainfall in the upstream, inflow and SSC at earlier hours according to results of correlation analysis respectively conducted between them and SSC at time to be predicted. Two prediction models, which are based on MLR and ANN techniques respectively, are then developed employing variables strongly correlated with SSC to be predicted as result of the correlation analysis.

Although Miwa dam has observed a number of flood events after the bypass tunnel for sediment passing was installed, hourly SSC had been continuously monitored through one entire flood event since before increase in inflow water only in several events, as inflow water normally increases rapidly after rainfall and therefore makes preparation for monitoring difficult. The correlation analysis and development of MLR and ANN models must be conducted paying attention to the situation that the number of data does not seem to be fully enough much for statistics and model development in this study.

3.2 *Correlation analysis*

Correlation analysis between SSC and hydrological observation at the reservoir basin is conducted to find potential candidates for predictors of SSC. Correlation with time lags more than three hours was analyzed here so as to develop a prediction model of SSC which can estimate SSC before three hours according to the equation described below:

$$\rho(k, l) = \frac{\sigma_{xy}(k, l)}{\sigma_x \cdot \sigma_y(k)} \tag{1}$$

where $\rho(k, l)$ = correlation between SSC and hydrological variable k with time lag of l; σ_x = standard deviation of SSC; σ_y = standard deviation of hydrological variable k; and $\sigma_{xy}(k, l)$ is covariance between SSC and hydrological variable k described as follows:

$$\sigma_{xy}(k, l) = \sum_i x(k, i - l) \cdot y(i) \tag{2}$$

where $x(k, i - l)$ = ith sample of hydrological variable k with consideration of time lag l and $y(i)$ = ith sample of SSC in inflow water.

Four hydrological variables, namely hourly water inflow at the current time t ($q(t)$), hourly rainfall ($r(t)$), and accumulated rainfall amounts over the past two and three hours ($r_2(t)$ and $r_3(t)$ respectively), which can be respectively considered to have a relationship with SSC in inflowing water at the reservoir, are considered in the correlation analysis with SSC in inflowing water. Hydrological variables strongly correlated with SSC of inflow water at the reservoir are considered as candidates of predictors when the prediction models for SSC are developed.

3.3 *Multi-variable Linear Regression model (MLR)*

A MLR model for prediction of SSC in inflowing water is described as follows.

$$y(t) = a_0 + \sum_k a_k \cdot x(k, t) + \varepsilon(t) \tag{3}$$

where a_0 = constant term; a_k = coefficient for variable k which are selected as predictors from result of correlation analysis with SSC of water inflow; and ε = error term for estimation of prediction model of SSC in the inflow water.

3.4 *Artificial Neural Networks (ANN)*

An ANN is a network model which is inspired by the functioning of the brain and biological nervous systems (Tokar & Markus, 2000). Because of its capability for modeling non-linear relationships, ANNs have been employed for modeling hydrological processes (e.g. French *et al.*, 1992; Dawson & Wilby, 1998; Nagy *et al.*, 2002; Nohara *et al.*, 2006). Although various types of ANNs have been proposed including layer typed networks and mutually connected ones, a standard multi-layer feedforward model (see Fig. 1) with backpropagation training is employed as an ANN model in this study.

In this multi-layer ANN model, input values to the first layer (called as the input layer hereafter) are transformed according to the weight value of the connection between units of the two adjacent layers and conversion function of units (called as response function of units) in each middle layer

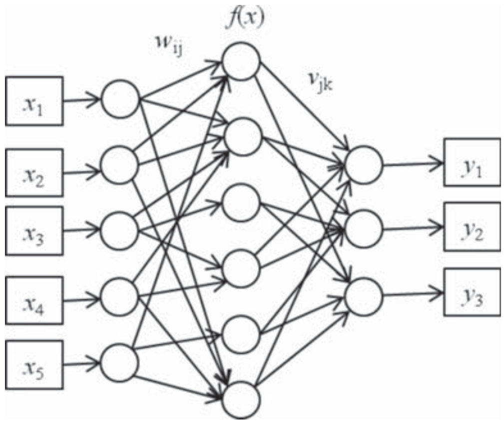

Figure 1. A feedforward ANN model with three layers.

(called as hidden layer hereafter), and the set of output values are obtained from the last layer (called as the output layer hereafter). Input and output values of the units in layer h are respectively calculated as following equations from the output values of the precedent layer (layer $h - 1$):

$$u_m^h = \sum_j w_{jm}^{h-1,h} o_j^{h-1} \tag{4}$$

$$o_m^h = f(u_m^h + \theta_m^h) \tag{5}$$

where u_m^h = input value to unit i in layer h; $w_{jm}^{h-1,h}$ = weight parameter of the connection between unit j in layer $h - 1$ and unit m in layer h; o_m^{h-1} and o_i^h are respectively output of unit j in layer $h - 1$ and one of unit m in layer h; θ_m^h = offset of unit m in layer h. $f(\cdot)$ is response function of the unit in the hidden layers, which is defined as sigmoid function described as below in this study:

$$f(x) = \frac{1}{1 + \exp(-2x/u_o)} \tag{6}$$

where u_o = coefficient. Adjustment of weight parameters of connection between units, which is called as training of the network, is iteratively conducted so as to minimize the value described as follows according to backpropagation algorithm:

$$E_\tau = \sum_p E_p = \sum_p \sum_n (T_{pn} - O_{pn})^2/2 \tag{7}$$

where O_{pn} = output of unit n in the output layer for training pattern p; T_{pn} = desired output signal which is also called as supervisory signal and normally derived from observed data to be estimated for unit n for training pattern p; E_p = integrated error of estimation for desired output values for training pattern p; and E_τ = total training error for all training patterns. The network is adjusted so as to be capable to output good estimation from data used for training through the processes described above.

4 CASE STUDY

4.1 *Correlation analysis and estimation of candidates for predictors*

Correlation analysis between SSC of inflow water to Miwa Reservoir and hydrological variables was conducted. Three flood events from 2006 to 2007shown in Table 1 in which SSC was successfully monitored were considered in correlation analysis. As mentioned in Chapter 3, four hydrological variables including hourly water inflow at the current time t ($q(t)$), hourly rainfall ($r(t)$), and accumulated rainfall amounts over the past two and three hours ($r_2(t)$ and $r_3(t)$ respectively) were considered as mentioned inChapter 3, and three time lags from three to five hours were also considered in the correlation analysis. Results of the correlation analysis are summarized in Table 2. As we can see in Table 2, water inflow before three hours ($q(t - 3)$) showed the strongest correlation (0.80) with the current SSC $SSC(t)$ in all variables related to water inflow. On the other hand, accumulated rainfall for three hours from before three hours ($r_3(t-3)$) showed the strongest correlation (0.61) with $SSC(t)$ in all variables related to averaged rainfall in the catchment area of the reservoir. These results are broadly consistent with the physical characteristics of the studied catchment area where rainfall water empirically runoffs to the reservoir within several hours after the soils are saturated with continuous rainfall.

Taking it into consideration that we are developing prediction models of SSC with not much data, more simple combination of variables is considered to be better for stability of the prediction for new data. Moreover, variables which strongly correlate each other should not be employed at once for a statistical prediction model like a regression model because it will cause multicollinearity and degrade predictability for new data. From these points of view, $q(t - 3)$, which showed the strongest correlation with $SSC(t)$, as well as $r_3(t - 3)$ which also showed the strongest correlation with $SSC(t)$ in the variables related to catchment rainfall and had weak correlation of 0.28 with $q(t - 3)$ were considered for a candidate of the predictors of SSC in development of MLR and ANN models.

Table 1. Three flood events which observational data was used for correlation analysis.

Event number	Date	Peak inflow (m^3/s)	Peak SSC (mg/l)
1	July 19th, 2006	366.2	16,900
2	July 15th, 2007	165.1	3610
3	Sep. 6th, 2007	565.3	25,000

Table 2. Results of correlation analysis.

Hydrological variable	Correlation with $SSC(t)$
$q(t-3)$	0.80*
$q(t-4)$	0.75
$q(t-5)$	0.72
$r(t-3)$	0.49
$r(t-4)$	0.55
$r(t-5)$	0.54
$r_2(t-3)$	0.57
$r_2(t-4)$	0.60
$r_2(t-5)$	0.55
$r_3(t-3)$	0.61*
$r_3(t-4)$	0.60
$r_3(t-5)$	0.56

* Employed for predictors in the prediction models.

4.2 *Development of MLR model*

A MLR model for prediction of SSC after three hours was developed with the predictors of inflow water at the current hour t and catchment rainfall accumulated for three hours from before three hours up to the current hour t based on the results of correlation analysis conducted in the previous section. Data from three flood events shown in Table 1 was also used for the development of the MLR model. The estimated MLR model for the prediction of $SSC(t + 3)$ (g/l) is described as the following equation with two predictors $q(t)$ (m^3/s) and $r_3(t)$ (mm).

$$SSC(t+3) = 11.30\, q(t) + 114.0\, r_3(t) \tag{8}$$

The developed regression model was applied to predict SSC for three hours ahead for a flood event observed in July in 2010 (see also Event 7 in Table 4). The time series of prediction is shown in Figure 2. It can be seen in Figure 2 that two peaks of SSC were predicted comparatively well while the MLR model overestimated SSC in the other parts of the time series. This is considered because the MLR model was developed using data mainly around peaks of SSC in the inflowing water of the three flood events used for the development of the model, and the model was not good at predicting comparatively low SSC condition.

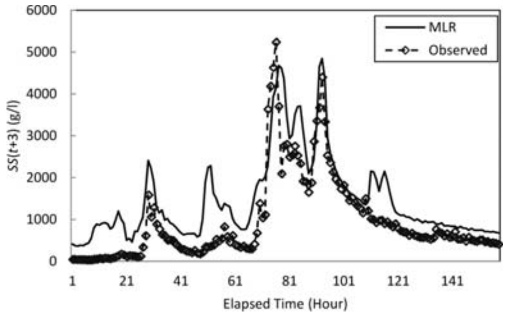

Figure 2. Time series of the prediction result for the flood event in July in 2010 by the developed MLR model.

4.3 *Development of ANN models*

ANN models were also developed based on the results of correlation analysis between SSC and hydrological variables of the reservoir. Three layer feedforward ANN model was employed here, and hourly SSC for the coming three hours (i.e. $SSC(t + 1)$, $SSC(t + 2)$ and $SSC(t + 3)$) was employed as output variables (i.e. predictands) by the ANN models. Three ANN models which respectively have different combinations of input variables (i.e. predictors) were developed here with consideration of the results of the correlation analysis also employing SS(t) for consistence in time series of the prediction. The combinations of the input variables for ANNs are shown in Table 3.

The ANN models were respectively trained with data from six flood events from Event 1 through 6 shown in Table 4. Although data from only three flood events was used for development of MLR model in the previous section, six flood events including those of comparatively small flood events were considered for development of the ANN models, as ANN models have more parameters to be estimated and therefore seem to need more data for the estimation. The number of units in the hidden layer was set to five for each ANN model so as to minimize the number of units without decreasing in capability to model nonlinear input/output relationships with trial and error. The numbers of training of the networks were respectively set to 3,000. Figure 3 illustrates time series estimated by an ANN model for training data produced from the six flood events' data after the network was trained. The result for ANN model 3 is shown in Figure 3 as an example. It can be seen that estimated values by the ANN model generally agree with observed values.

After the training of the ANN models, prediction was also conducted for flood event data in July in 2010 (see also Event 7 in Table 4) by each developed ANN model. The results of predictions by three ANN models are summarized in Tables 5 through 7.

Root Means Square Errors (RMSEs) of predictions by the ANN models are shown in Table 5. ANN model 1 (described as ANN1 hereafter), which employed hourly rainfalls as input variables, predicted SSC for the coming hour with less RMSE value compared with the other two models, while ANN model 2 (described as ANN2 hereafter) and ANN model 3 (also described as ANN3 hereafter), both of which employed accumulated

Table 3. Input variables (predictors) of the developed ANN models.

ANN model	Number of input variables	Input variables (predictors)
1	5	$r(t-2)$, $r(t-1)$, $r(t)$, $q(t)$, $SSC(t)$
2	3	$r_3(t-1)$, $q(t)$, $SSC(t)$
3	3	$r_3(t)$, q(t), $SSC(t)$

Table 4. Six flood events which observational data was used for development of ANN models (Events 1 through 6) and one flood event for prediction as validation (Event 7).

Event number	Date	Peak inflow (m^3/s)	Peak SSC (mg/l)
1*	July 19th, 2006	366.2	16,900
2*	July 15th, 2007	165.1	3610
3*	Sep. 6th, 2007	565.3	25,000
4	Oct. 8th, 2004	132.8	1990
5	July 5th, 2005	78.1	2610
6	May 25th, 2007	102.8	3990
7	July 11th–16th, 2010	228.2	5160

* These events were also used for development of the MLR model.

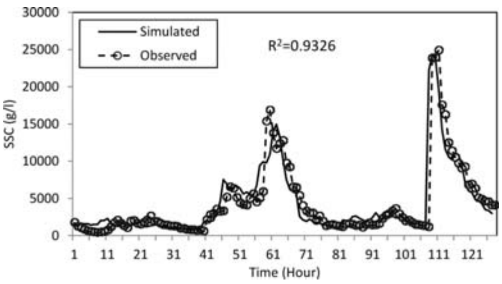

Figure 3. Estimation results of time series produced from data of six flood events for training by ANN model 3 after the model was trained.

rainfall as input variables, performed less RMSE values compared with ANN1 in the predictions of SSCs two and three hours ahead.

On the other hand, correlations between observations and prediction by the ANN models for the flood event are shown in Table 6. It can also be seen that ANN1 showed better correlation in the prediction of SSC for one hour and two hours ahead, although ANN2 and ANN3 showed stronger correlation in the prediction for SSC three hours ahead.

From the point of view that prediction accuracy in the peak concentration of suspended sediment in the inflow water is also important for the operation of the sediment bypass tunnel, percentages of errors of the predictions by ANN models to observations for the peak values of SSC are also summarized in Table 7. It can be seen in Table 7 that the absolute value of prediction error of ANN1 was smaller than those of ANN2 and ANN3 in the prediction of SSC for the next hour, while ANN3 respectively showed the smallest absolute errors in the three models in the prediction of SSC for two and three hours ahead.

These results suggest that accumulated rainfall for three hours in the near past can be good predictor for the prediction of future SSC like those two or three hours ahead as it was seen in the results of the correlation analysis, while recent hourly rainfall is encouraged to be considered in the prediction model for SSC in the coming hour at Miwa Dam. From the viewpoint that predictability of the peak of SSC from the early stage is considered important for successful operation of bypass tunnels, ANN3 with input of rainfall accumulated for the past three hours is considered effective as the prediction model in the three models studied in this paper.

Time series of predicted SSC by ANN3 are shown in Figures 4 through 6. Large errors were illustrated in the low SSC periods with any time lag. This can be considered because the prediction models were arranged so as to minimize sum of square error described as Equation (7) which considered important to give small errors during high SSC period in which the square error can easily become greater, and period when small values were observed were not considered well in the model fitting. On the other hand, large peaks of SSC were predicted with comparatively good accuracy while the accuracy of the prediction degraded as the lead time became longer.

Table 5. RMSEs of predictions for Flood Event 7 by ANN models.

ANN model	Root mean square error (g/l)		
	$SSC\,(t+1)$	$SSC\,(t+2)$	$SSC\,(t+3)$
1	629.4	805.5	960.4
2	663.5	740.0	860.7
3	690.8	721.1	743.8

Table 6. Correlations between observations and predictions by ANN models for Flood Event 7.

ANN model	Correlation		
	$SSC\,(t+1)$	$SSC\,(t+2)$	$SSC\,(t+3)$
1	0.947	0.902	0.851
2	0.918	0.898	0.892
3	0.928	0.882	0.888

Table 7. Percentage of errors of predictions by ANN models to observations for the peak values of SSC for Flood Event 7.

ANN model	Percentage of errors for the peak values (%)		
	$SSC\,(t+1)$	$SSC\,(t+2)$	$SSC\,(t+3)$
1	5.5	4.8	−12.8
2	10.2	16.2	11.1
3	−16.4	1.0	1.7

4.4 *Comparison of MLR and ANN models*

Comparing results of the prediction of SSC for three hours ahead by the MLR and ANN3 which are respectively illustrated in Figure 2 and Figure 6, it can be seen that ANN3 showed better performance in the low SSC period. ANN3 was especially less sensitive against small peak of SSC which can be seen in the earlier period of the flood event (Flood Event 7), and closer to the observed values. Similar characteristics can be seen after the second large peak of SSC, where ANN3 predicted a small peak while MLR predicted the larger peak. On the

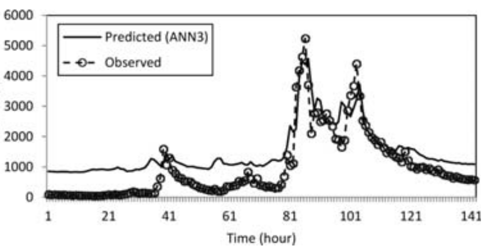

Figure 4. Time series of predicted SSC for the coming hour ($SSC(t+1)$) by ANN3 for Flood Event 7.

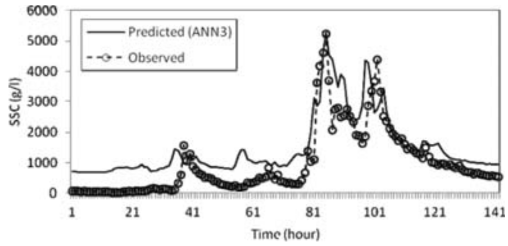

Figure 5. Time series of predicted SSC two hours ahead ($SSC(t+2)$) by ANN3 for Flood Event 7.

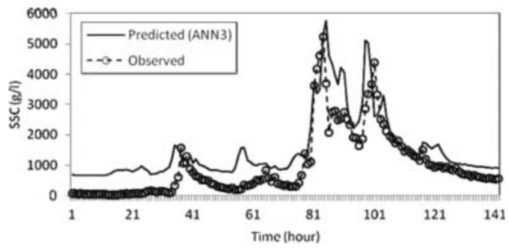

Figure 6. Time series of predicted SSC three hours ahead ($SSC(t+3)$) by ANN3 for Flood Event 7.

other hand, there was no significant difference in the predictability for the large peak of SSC in the two models. Synthesizing the results mentioned above, ANN3 model can be considered to show the better performance in terms of predictability in time series tendency, while there are still problems to be tackled in the performance of the prediction models especially in low SSC periods and for the small peak of SSC. Further monitoring of SSC or turbidity in the inflowing water will improve the models' performance by adding data for development of the models.

5 CONCLUSION

Real-time prediction models of sediment inflow were developed for efficient operation of the sediment bypass tunnel at Miwa Dam the Mibu River basin in Japan. Two kinds of prediction models, namely a MLR model and three ANN models, were developed in order to predict sediment inflow for the coming three hours with hourly time resolution. As a result of correlation analysis between SSC in inflow water and hydrological variables observed in the reservoir basin, amount of inflow water to the reservoir at the current time and catchment rainfall accumulated for the past three hours were respectively selected for the candidates of the predictors of SSC in inflow water three hours ahead. These results of the correlation analysis were effectively considered in the development of the prediction models to determine variables for predictors. Through application of developed prediction models, it was demonstrated that accumulated rainfall for the recent several hours can be considered as an important predictor for the prediction of SSC in the near future. It was also illustrated that ANN3 model, which employed catchment rainfall accumulated for the past three hours as a predictor, can be considered to show the better performance in terms of predictability in time series tendency, while more study with more observational data is considered necessary for more reliable conclusion.

REFERENCES

Cabaner, M., Unal, B. & Kisi, O. 2009. Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data, *J. Hydrol*. 367: 52–61.

Cigizoglu, H.K. & Alp, M. 2006. Generalized regression neural network in modeling river sediment yield, *Advances in Engineering Software*. 37: 63–68.

Chutachindakate, C. & Sumi, T. 2009. Real time sediment concentration forecast by rainfall information: Case study on Managawa River Basin, Japan, *Proc. of 33rd IAHR Congress: Water Engineering for a Sustainable Environment*: 430–437.

Dawson, C.W. & Wilby, R. 1998. An artificial network approach to rainfall runoff modeling, *Hydrological Sciences Journal*. 43(1): 47–66.

French, M.N., Krajewski, W.F. & Cuykendall, R.R. 1992. Rainfall forecasting in space and time using a neural network, *J. Hydrol*. 137: 1–31.

Nagy, H.M., Watanabe, K. & Hirano, M. 2002. Prediction of sediment load concentration n rivers using artificial neural network model, *J. Hydraul. Eng*. 128(6): 588–595.

Nohara, D., Kojiri, T. & Smith, P.J. 2006. Heuristic DSS for reservoir operation considering global meteorological information. In Van Alphen, van Beek & Taal (eds), *Floods, from Defence to Management—Symposium Proceedings, Proceedings of 3rd International Symposium on Flood Defence*: 715–724, Taylor & Francis Group, London.

Rai, R.K. & Mathur, B.S. 2008. Event-based sediment yield modeling using artificial neural network, *Water Resour Manage.* 22: 423–441.

Rajee, T., Mirbagheri, S.A., Zounemat-Kermani, M. & Nourani, V. 2009. Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models, *Science of the Total Environment.* 407: 4916–4927.

Tokar, A.S. & Markus, M. 2000. Precipitation-runoff modeling using artificial neural networks and conceptual models, *J. Hydrol. Eng.* 5: 156–161.