Comparison between response surface models and artificial neural networks in hydrologic forecasting

Yu, Jianjun; Qin, Xiaosheng; Larsen, Ole; Chua, Lloyd Hock Chye

2014

Yu, J., Qin, X., Larsen, O., & Chua, L. H. C. (2014). Comparison between Response Surface Models and Artificial Neural Networks in Hydrologic Forecasting. Journal of Hydrologic Engineering, 19(3), 473-481.

https://hdl.handle.net/10356/79536

https://doi.org/10.1061/(ASCE)HE.1943-5584.0000827

Comparison between Response Surface Models and Artificial Neural Networks in Hydrologic Forecasting

Jianjun Yu[1], S.M.ASCE

Xiaosheng Qin [2*], A. M.ASCE

Ole Larsen[3]

Lloyd Hock Chye Chua[4], M.ASCE

ABSTRACT

Developing an efficient and accurate hydrologic forecasting model is crucial to managing water resources and flooding issues. In this study, the response surface (RS) models including multiple linear regression (MLR), quadratic response surface (QRS) and nonlinear response surface (NRS) were applied to daily runoff (e.g. discharge and water level) prediction. Two catchments, one in southeast of China and the other in western Canada were used to demonstrate the applicability of the proposed models. Their performances were compared with artificial neural network (ANN) models, trained with the learning algorithms of the gradient descent with adaptive learning rate (ANN-GDA) and Levenberg-Marquardt (ANN-LM).The performances of both RS and ANN in relation to the lags used in the input data, the length of the training samples, long term (monthly and yearly) predictions, and peak value predictions were also analyzed. The results indicate that the QRS and NRS were able to obtain equally good performance in runoff prediction, as compared with

[1] Ph.D. candidate, School of Civil and Environmental Engineering, Nanyang Technological Univ., 50 Nanyang Ave., Singapore 639798, Singapore. E-mail: yuji0011@ntu.edu.sg
DHI-NTU Water & Environment Research Centre and Education Hub, DHI Water & Environment (S) Pte. Ltd, 200 Pandan Loop Pantech 21, #08-03, Singapore 128388.
[*2]Assistant Professor, School of Civil and Environmental Engineering, Nanyang Technological Univ., 50 Nanyang Ave., Singapore 639798, Singapore. E-mail: xsqin@ntu.edu.sg
Principal Investigator, Earth Observatory of Singapore, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
[3] Director, DHI-NTU Water & Environment Research Centre and Education Hub, DHI Water & Environment (S) Pte. Ltd, 200 Pandan Loop Pantech 21, #08-03, Singapore 128388.
[4] Assistant Professor, School of Civil and Environmental Engineering, Nanyang Technological Univ., 50 Nanyang Ave., Singapore 639798, Singapore.

ANN-GDA and ANN-LM, but require lower computational efforts. The RS models bring practical benefits in their application to hydrologic forecasting, particularly in the cases of short-term flood forecasting (e.g. hourly) due to fast training capability, and could be considered as an alternative to ANN.

Keywords: hydrologic forecasting, regression, response surface model, artificial neural networks

INTRODUCTION

Hydrologic forecasting is substantial for effective operation of a water resources planning or a flood mitigation system. Data-driven models (e.g. regression models and artificial intelligent methods) have shown the characteristic of easy-to-develop compared with physical rainfall-runoff models, due to the fact that such models can extract the input-output relations directly based on the observed hydrological and meteorological data without precise knowledge of the hydrological process and sufficient data for contributing physical variables.

During the past decades, linear regression models, including autoregressive (AR) (Carlson et al., 1970), autoregressive moving average (ARMA) (Salas et al., 1985), and autoregressive moving average with exogenous inputs (ARMAX) (Haltiner and Salas, 1988), have been developed and offered reasonable results in a number of streamflow forecasting cases. Corley II (1980) and Haktanir and Sezen (1990) applied the probability density functions of the two-parameter gamma and the three-parameter beta distributions to fit the hydrograph, and suggested that they could also be used for the synthetic unit hydrograph. However, such models relied heavily on the presumption about the form of data distribution or the functional relations among the parameters concerned. They may suffer from insufficient accuracy in hydrologic modeling applications, where the exact mathematical representation of the data distribution for the rainfall-runoff process does not exist.

Recent studies have demonstrated the applicability of nonlinear techniques such as chaos-based approaches (Jayawardena and Lai, 1994; Jayawardena and Gurung, 2000; Elshorbagy et al., 2002), nonlinear prediction (NLP) (Porporato and Ridolfi, 1997, 2001;

Islam and Sivakumar, 2002), nonparametric techniques (Iorgulescu and Beven, 2004) and artificial intelligent methods (e.g. adaptive neural-based fuzzy inference system, genetic programming and support vector machine) (Wang et al. 2009) in hydrologic modeling. Particularly, artificial neural networks (ANN) have gained much interest due to their capabilities in characterizing complex nonlinear relationships with high prediction accuracy (Govindaraju, 2000a, b). Many studies have compared ANN with linear regression approaches (e.g. AR and ARMA) and proved that ANN can outperform statistical techniques (Raman and Sunilkumar, 1995; Abrahart and See, 2002; Kişi, 2004; Castellano-Méndeza et al., 2004; Adeloye, 2009). Jain and Indurthy (2003) aimed at a comparative analysis of deterministic unit hydrograph theory, statistical regression and ANN for the purpose of modeling an event-based rainfall-runoff process and proved that ANN consistently outperformed conventional models, particularly in prediction of peak discharge. He and Valeo (2009) demonstrated that ANN yielded higher accuracy in quantile estimation than the parametric methods in rainfall/flood frequency analysis. Londhe and Charhate (2010) compared ANN with genetic programming (GP) and model trees (MT) in forecasting river flow one-day ahead at Narmada catchment of India. They found that the ANN and MT performed almost equally well, but GP did better than both in terms of normal and extreme events.

However, ANN possesses a number of disadvantages, such as the tedious trial and error process to determine its structure, possibility of getting stuck in local minima, and long training period required. Efficient and robust methods are still being attempted to make the prediction closer to reality with minimized errors. The response surface (RS) model, as an alternative method to nonlinear prediction, has received success in application to chemical and biochemical processes (Andersson and Adlercreutz, 1999; Ranade et al., 2010). However, only limited studies have discussed the applicability of the RS models in streamflow forecasting or compared RS models with ANN (Singh and Deo, 2007).

Thus, the objective of this study is to investigate the applicability of RS models in hydrologic forecasting, and compare their performances with ANN. Based on first-order and second-order polynomial format, RS models involved in this study include multiple linear regression (MLR), quadratic response surface (QRS) and nonlinear response surface (NRS). ANN models, trained with the learning algorithms of the gradient descent with adaptive learning rate (GDA) and Levenberg-Marquardt (LM), will be setup for comparison. Two real-world study cases were selected for demonstration. The first is a watershed in the southeast China where the rainfall dominates the streamflow and the other is in western Canada where the rainfall-runoff relations are complex due to the snowmelt. The performances of the models under various modeling scenarios, such as different lags used in the input data, the length of the training samples, long term (monthly and yearly) predictions and peak value prediction, were also analyzed.

METHODOLOGY

Response Surface Models

RS models are multivariate polynomial models, which employ a group of mathematical equations that describe the relationship between the response and the independent variables (Myers et al, 2004). Considering the situation in which the response y depends on n variables: x_1, x_2, ... x_i ..., x_n, the general equation for the RS model is given as follows:

$$y = f(x_1,\ x_2,\ \ldots x_i\ \ldots,\ x_n) + \varepsilon \tag{1}$$

where f is the unknown response function; ε is the statistical error represents the sources of variability not accounted for in f, such as measurement errors, variations inherent in the rainfall-runoff process, and other effects. The first-order polynomial based RS model or MLR, is used to describe the linear relationship between response and predictors. It is represented by the following equation:

$$y_l = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \tag{2}$$

However, a plausible MLR requires that the input variables should have a high linear correlation with the output variables; this may lead to the lack of fitting for hydrologic forecasting. The second-order polynomial, written as:

$$y_q = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1, i<j}^{n} \beta_{ij} x_i x_j + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \varepsilon \tag{3}$$

is usually employed in the QRS to approximate the nonlinearity, where β_0 is the intercept, $\beta_i x_i$ are the linear terms, $\beta_{ij} x_i x_j$ are the quadratic interaction terms, and $\beta_{ii} x_i^2$ are the square terms. High-order polynomials are seldom used since so many parameters need to be estimated which may lead to inefficiency and over fitting.

The NRS is an alternative approach to model the nonlinearity and complexities of rainfall-runoff processes. In this study, an NRS model based on the form of quadratic polynomials, with additional exponential coefficients for the linear, interaction and quadratic terms, will be tested. The general equation of NRS could be written as:

$$y_n = \beta_0 + \sum_{i=1}^{n} \beta_i x_i^{\theta_i} + \sum_{i=1, i<=j}^{n} \beta_{ij} \left(x_i x_j\right)^{\gamma_{ij}} + \varepsilon \tag{4}$$

where θ_i is the exponential coefficient for linear terms; γ_{ij} is the exponential coefficient for interactive and quadratic terms. They will be estimated together with β during the fitting process. After tried various simplified forms of Eq. (4) through reducing the number of exponential coefficients, we found that the following equation gave better results:

$$y_n = \beta_0 + \sum_{i=1}^{n} \beta_i x_i^{\theta} + \sum_{i=1, i<=j}^{n} \beta_{ij} \left(x_i x_j \right)^{\gamma} + \varepsilon \tag{5}$$

Compared with Eq. (4), the equation with reduced number of exponential coefficients by using the same exponential coefficient for all linear terms and for all interactive and quadratic terms (as shown in Eq. (5)) brings the advantage of getting a quick response for computations carried out on a desktop computer. Different from MLR and QRS, whose coefficients are estimated through the method of least squares, NRS employs the LM algorithm (Seber and Wild, 1989). The following vectors and matrices are originated from the estimated coefficients $\hat{\beta}$:

$$\mathbf{B_1} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 2\hat{\beta}_{11} & \hat{\beta}_{12} & \cdots & \hat{\beta}_{1n} \\ \hat{\beta}_{12} & 2\hat{\beta}_{22} & \cdots & \hat{\beta}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\beta}_{1n} & \hat{\beta}_{2n} & \cdots & 2\hat{\beta}_{nn} \end{bmatrix}$$

The predicted value of the NRS may be obtained from the resulting fitted response surfaces, represented as matrix form:

$$y_n = \hat{\beta}_0 + \mathbf{x_1 B_1} + \frac{1}{2} \mathbf{x_q B_2 x_q^T} \tag{6}$$

where $x_1, x_2, \ldots x_i \ldots, x_n$ are the predictors; $\mathbf{x_q} = [x_1^{\hat{\gamma}} \; x_2^{\hat{\gamma}} \; \cdots \; x_n^{\hat{\gamma}}]$ and

$\mathbf{x}_1 = [x_1^{\hat{\theta}}\ x_2^{\hat{\theta}}\ \cdots\ x_n^{\hat{\theta}}]$ are the vectors of predictor x and estimated exponential coefficients $\hat{\theta}$ and $\hat{\gamma}$.

Artificial Neural Networks

ANN has been developed to describe the complex, nonlinear relationships in hydrologic modeling with the characteristic of self-learning ability from data. It is a network consisting of nodes, analogous to biological neurons, in multiple layers which are interconnected by weighted links. Each neuron employs a transfer function (e.g. the step, liner or sigmoid function) to transmit the single scalar input or a weighted sum of inputs through a connection that multiplies its strength by the scalar weight to generate the output (Zealand et al., 1999). One or more neurons can be combined in a single layer, and three or more layers are typically used to form the ANN. The outputs generated by neurons in one layer form the inputs of the next layer, and there is no restriction that the number of neurons in a particular layer must equal the number of inputs to that layer. The feed forward neural network (FFNN), which is characterized as not having interconnection between nodes in the same layer and with connections oriented from the input towards the output nodes only, is by far the most widely used network in water resources applications (Hornik et al., 1989; Kůrková, 1992). After determining the architecture of the network (e.g. nodes, layers, transfer functions and interconnection between nodes and layers), it can be trained with various learning algorithms (e.g. back-propagation, conjugate gradient, and LM) in order to simulate the prediction. Detailed discussions about the application, metrics and shortcomings of ANN are provided by Maier and Dandy (2000).

The training, similar as the parameter estimation in regression models, is a process to adjust and optimize the weights and biases in order to minimize the global error and find the global solution in response to the inputs (Kişi, 2007). In this study, FFNN trained with GDA and LM algorithms are adopted. The GDA algorithm adjusts the weights and biases in the direction where the performance function decreases most rapidly (as back-propagation (BP) algorithm does) using the first-order local optimization method, written as (Rumelhart et al., 1986):

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{a}_k \mathbf{g}_k \tag{7}$$

where $\mathbf{x}_k$ is a vector of current weights and biases, $\mathbf{g}_k$ is the current gradient, and $\mathbf{a}_k$ is the adaptive learning rate to keep the learning step size as large as possible while to the extent that the network can learn without large error increases. Thus, it is considered to be the improved BP algorithm. The LM algorithm is designed to approach the second-order training speed and overcome the problem to compute Hessian matrix in Newton's method, which is complex and expensive (Moré, 1978). It employs the Jacobian matrix to approximate Hessian matrix in the following Newton-like updating function (Moré, 1978):

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}\right]^{-1} \mathbf{J}^T \mathbf{e} \tag{8}$$

where $\mathbf{J}$ is the Jacobian matrix that contains the first derivatives of the network errors with respect to the weights and biases; $\mathbf{e}$ is a vector of network errors; $\mathbf{I}$ is the identity matrix and μ is a scalar value, which will decrease after each successful step and increase only when a tentative step would increase the value of the performance function. It thus can shift toward Newton's method as quickly as possible. During the training process, the weights and biases are updated iteratively until the global performance (e.g. mean square error in this study) reaches a pre-defined objective (e.g. less than 0.0001 in this study). Otherwise it will stop when the iteration exceeds the specified maximum times (e.g. 15000 for GDA and 6000 for LM in this study).

Performance Metrics

The performance of the prediction models can be evaluated by the following four evaluation criteria: coefficient of determination (R^2) (O'Connell et al., 1970), Nash-Sutcliffe efficiency coefficient (E) (Nash and Sutcliffe, 1970), root mean squared error ($RMSE$) (Karunanithi et al., 1994) and mean absolute percentage error ($MAPE$) (Hu et al., 2001). These metrics are calculated as:

$$R^2 = \left[\frac{\sum_{i=1}^{n} \left(Q_i^o - \overline{Q}^o\right)\left(Q_i^m - \overline{Q}^m\right)}{\sqrt{\sum_{i=1}^{n} \left(Q_i^o - \overline{Q}^o\right)^2} \sqrt{\sum_{i=1}^{n} \left(Q_i^m - \overline{Q}^m\right)^2}} \right]^2 \tag{9}$$

$$E = 1 - \frac{\sum_{i=1}^{n} \left(Q_i^o - Q_i^m\right)^2}{\sum_{i=1}^{n} \left(Q_i^o - \overline{Q}^o\right)^2} \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Q_i^o - Q_i^m\right)^2} \tag{11}$$

and

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Q_i^o - Q_i^m}{Q_i^o} \right| \times 100 \tag{12}$$

where Q_i^o is the ith observed runoff level (e.g. stage or discharge), Q_i^m is ith the predicted runoff level, $\overline{Q}^o$ is the mean of the observed runoff level and $\overline{Q}^m$ is the mean of the predicted runoff level. The R^2 indicates how well the correlation between the predicted and observed values is, and E is often used to assess the predictive capability of hydrological models. The closer the value of R^2 or E is to 1, the more accurate the model. The *RMSE* evaluates the residual between the predicted and observed values, and the *MAPE* is a weighted average of the absolute errors. The smaller the value of *RMSE* or *MAPE* is, the more accurate the model will be.

APPLICATION TO HYDROLOGIC FORECASTING

In order to compare the applicability of RS models (MLR, QRS and NRS) with ANN, two study cases with different rainfall-runoff characteristics are selected. The first case is water level prediction in Heshui catchment, China, where the effect of rainfall on runoff is predominant. The second case is river flow forecasting in Coquitlam catchment, Canada, where the hydrologic response to rainfall is highly nonlinear. The RS models (MLR, QRS and NRS) and ANN models trained with GDA (ANN-GDA) and LM algorithms (ANN-LM) were applied to the two catchments.

Water Level Prediction in Heshui Catchment

The Heshui catchment is located in Yongxin County, Jiangxi Province, China (Fig. 1a). It has an area of 2,228 km^2 and characterized by the subtropical climate, with higher temperatures and rainfall in the summer season. Within the catchment, the average rainfall per year is about 1,509 mm, and most of the rain falls between the months of April and June, with an average at 850.5 mm. The average evaporation per year is 1,066 mm (Huang et al., 2010).

Place Figure 1 here

The meteorological and hydrological data are measured and collected in Beimen station (26°50'N and 114°15'E). The dataset contains the information of daily rainfall and water level for a period of 13 years (1988-2000), of which 11 years (from 1988 to 1998, 4018 data points) were used for calibration and 2 years (from 1999 to 2000, 731 data points) were used for validation. Statistical information of Beimen station is provided in Table 1.

Place Table 1 here

One-day ahead forecasting based on the previous three days'(e.g. through trial and error) water levels and rainfall was carried out using the MLR, QRS, NRS, ANN-GDA and ANN-LM models. The ANN models can be illustrated by the following expression:

$$Q_t = ANN_{6,5,1}(Q_{t-1},\ Q_{t-2},\ Q_{t-3},\ R_{t-1},\ R_{t-2}, R_{t-3}) \qquad (13)$$

where Q_t is the mean daily water level [m] for current day under prediction; t is the time point representing the day of concern (i.e. current day); $ANN_{6,5,1}$ represents the ANN model with 6 inputs, 5 nodes in one hidden layer and 1 output node; Q_{t-i} and R_{t-i} (i = 1, 2, 3) are the mean daily water level and daily rainfall [mm/day], respectively, in the previous three days.

Table 2 shows the performances of the models, as well as the computational time for training. All the five models provide acceptable results in terms of R^2 (ranging from 0.90

to 0.93), E (ranging from 0.89 to 0.93), $RMSE$ (ranging from 0.094 to 0.115 m) and $MAPE$ (ranging from 6.2 to 7.4), out of which NRS and QRS outperform ANN and MLR. From Table 2, parameters of MLR and QRS were determined in 5 seconds; NRS took a slightly longer time (e.g. 20 seconds) than MLR and QRS; the ANN-GDA required 18.5 minutes and the ANN-LM required 7.5 minutes. Generally, the QRS and NRS model proposed in this study not only performed as good as ANN, but also were able to provide a faster training.

Place Table 2 here

River Flow Forecasting in Coquitlam Catchment

The Coquitlam catchment is located in the south of the Georgia Basin, Canada which is a tributary of the Lower Fraser River (Fig. 1b). The catchment has an area of 191 km^2 and average elevation of 1,154 m, and is characterized by a west-coast maritime climate with cool wet winters and warm dry summers. Runoff from the Coquitlam catchment is generated by both rainfall and snowmelt. Melt of the mild and upper elevation snow-packs continues over several months and usually generates the largest inflows in May or June. The second largest inflows occur in November and December, primarily as a result of heavy rain and associated melt of moderate-elevation snow-packs.

Daily discharge is available from Water Survey of Environment Canada and meteorological data is initially provided by Greater Vancouver Regional District (GVRD) and Environment Canada, including daily precipitation, snow water equivalent, and maximum and minimum daily temperature over the study catchment (Huang et al., 2007).

As showed in Fig. 1(b), the daily discharge is collected at the hydrometric station 08MH141 during year 1982 to 2002. The daily rainfall and mean temperature are collected at the climate station 1101890.The statistical characteristics of the data used in this study are summarized in Table 3. The dataset from year 1990 to 1999 (3652 data points) was used for calibration and the dataset from 2000 to 2001 (731 data points) was used for verification.

Place Table 3 here

Models applied here to predict the daily discharge were developed in the same manner as for the Heshui catchment. Considering the complexity caused by snow melt, a new predictor (i.e. daily mean temperature) was added since temperature is supposed to be one of the main factors influencing snow melt. Thus, 9 predictor parameters were used including 3-day previous rainfall, daily discharge and mean temperature. The equations for the RS models were obtained by substituting these parameters into Eqs. (2), (3) and (5). ANN models could be represented by the following expression:

$$Q_t = ANN_{9,8,3,1}(Q_{t-1},\ Q_{t-2},\ Q_{t-3},\ R_{t-1},\ R_{t-2},\ R_{t-3},\ T_{t-1},\ T_{t-2},\ T_{t-3}) \tag{14}$$

where Q_t is the daily discharge [m^3/s] for current day under prediction; $ANN_{9,8,3,1}$ represents the ANN model with 9 inputs, 8 and 3 nodes in two hidden layers and 1 output node; Q_{t-i} , R_{t-i} and T_{t-i} (i = 1, 2, 3) are the previous 3-day daily discharge [m^3/s], rainfall [mm/day] and daily mean temperature [^{o} C], respectively.

Table 4 lists the performance and computation time of the models applied to daily discharge forecasting. The results show that all of the models could only obtain R^2 and E

values from 0.65 to 0.73, *RMSE* values from 3.6 to 4.1 m^3/s and *MAPE* value from 186 to 239. One of the reasons for the relatively poorer performance by the models is the complex rainfall-runoff process characterized by continuous snow melt and periodic precipitation. There is also a suspicion that the data collected at the climate station may not be truly representative of meteorological conditions in the catchment as the station is not located within the catchment (Fig. 1b). In addition, some of the missing temperature data had to be augmented using historical temperature data for the same period; this invariably introduces uncertainties with regards to the input data. Nevertheless, from the perspective of model comparison, the results appear to have the same characteristics as observed for the Heshui catchment. MLR is the worst performing model and the other four models show higher levels of accuracy (average 10.77% improvement in R^2 or E compared to MLR). The parameter estimation process of NRS has increased by 4 minutes due to the increase in the number of predictors. In total, 12 and 24 minutes were required for training the ANN-GDA and ANN-LM, respectively. It is still obvious that the QRS and NRS could be more cost-effective options to ANN.

Place Table 4 here

DISCUSSIONS

Performance Analysis for Different Modeling Scenarios

The performance of the RS and ANN are compared under the following four scenarios: (1) using inputs (e.g. discharge, rainfall) with different lags; (2) long term (monthly, yearly) predictions; (3) peak runoff values; (4) using training samples from different time periods.

Fig. 2 shows the results of the predictions conducted using the input variables lagged from 1 to 8 days. The results demonstrate that the models are not sensitive to the lag in the input data for the Heshui catchment which exhibits moderately nonlinear behaviour. The NRS obtained the best performance for each scenario. For Case 2 where the system is highly nonlinear, the model results are sensitive to the prediction using the observed data for the 3 or 4 previous days as inputs. With the exception of MLR, all models, deteriorated in performance when the inputs were lagged by more than 4 days; in particular, the performances of QRS and NRS decreased drastically. Generally, the prediction models using hydrological and meteorological data observed in the previous 3 or 4 days are suitable considering both the amount of data and prediction reliability.

Place Figure 2 here

The computation time for training ANN models (ANN-GDA and ANN-LM) and NRS is also compared in Fig. 2. MLR and QRS are not included since their coefficients estimation process is not significantly affected by the inputs varying with different lags. The same architecture and training dataset are utilized in ANN models to estimate the computation time required. The number of inputs is the only variable, associated with the different lags used. The results show that the computation time for the ANN models is not sensitive to the different lags used in the inputs. The architecture, learning algorithm employed and the size of the training dataset may be the main factors. However, it was evident that the computational time of the NRS model will increase with the increase of the number of parameters. A large number of coefficients in the NRS model may lead to intensive computational burden and cause trouble in practical applications, particularly in short term forecasting (e.g. hourly or minutely flood forecasting).

The monthly and yearly averages of the predictions were then calculated based on the average of the predicted daily values for appropriate time periods. Fig. 3 illustrates the comparison of the daily, monthly and yearly predictions by different models. As mentioned previously, QRS and NRS outperform ANN and MLR models in daily prediction. The differences between the models in monthly and yearly predictions can be ignored since all models performed well; however, the QRS and NRS models are marginally better.

Place Figure 3 here

Attention is often paid to the forecasts of the peak discharge or water level due to its importance in water resources and flood management applications. The performance of the peak values prediction was thus considered as another metric in model evaluation. The following formula defines that the runoff level (e.g. discharge or water level) is considered to be the peak value if and only if the runoff level of current day is no less than the value of previous day and the next day as well:

$$Q \in \{ Q \mid [Q_t \geq Q_{t+1}] \cap [Q_t \geq Q_{t-1}] \} \tag{15}$$

where Q_t, Q_{t-1} and Q_{t+1} denote the runoff of the current day, previous day, and next day respectively.

In total, 151 and 123 peak values were extracted from the verification set for Case 1 and Case 2 (731 data points for each case), and the performance of the peak value was re-calculated using Eqs. (9) and (10). Figure 4 shows a comparison of the results of the peak value and whole dataset of the models for the two catchments. The comparison shows that MLR, QRS, NRS and ANN-GDA appear to perform better for the peak value for the Heshui catchment which has moderate nonlinearity while the improvement in performance is not as obvious for the Coquitlam catchment, which is highly nonlinear. This can also be seen from the low Nash-Sutcliffe coefficients for the peak values obtained by the RS models; although the results for the whole dataset appear reasonable. This drawback of the RS models in forecasting the peak values should be noted in practical applications.

Place Figure 4 here

Finally, the performances of the models trained using the data from different years are illustrated in Fig. 5. The results illustrate that MLR is insensitive to the size of the training samples, but at least 3 years' training data is necessary for ensuring the other models obtain acceptable prediction accuracies. The QRS and NRS show the same characteristics, where the performance would improve as the length of data used for training increases up to 3 years for both cases. After 3 years, R^2 and E would reach a plateau for Heshui catchment simulations, but those for the Coquitlam catchment would exhibit a slightly decreasing trend. The ANN models, on the other hand, exhibit a higher degree of fluctuations in the results, particularly in Case 2. The results of the RS models exhibit a more distinct trend when the length of calibration dataset increases, making it more straightforward to identify a suitable sample size for training. However, the ANN models are more sensitive to the length of the calibration dataset.

Place Figure 5 here

Model Comparison

In this study, RS models (MLR, QRS and NRS) are compared with ANN trained with GDA (ANN-GDA) and LM algorithms (ANN-LM) in two hydrologic forecasting applications with different rainfall-runoff characteristics. The results show that MLR may be only suitable for rainfall dominated hydrologic forecasting. The ANN can obtain results with high accuracy, which is consistent with the conclusions reported by most literatures, but the long training time may be an issue in some practical applications (e.g. hourly or minutely flood forecasting), where the dataset may be large and the time for training is limited. The QRS and NRS are able to achieve comparable performance as ANN, in spite of the shorter time required for calibration. This conclusion is different to the discovery of Singh and Deo (2007) who reported that the QRS model was highly unsatisfactory compared to ANN in their study. This suggests that the comparative performance of the QRS and ANN may be highly problem specific.

Having an explicit expression is one of the advantages of RS models. This brings a reduced amount of effort required for model manipulations (e.g. programming, parameter estimation, computation etc.). Particularly, MLR and QRS can obtain quick response for parameter estimation since the linear equations of the first-order and second-order polynomials are solved directly. NRS requires relative longer calibration time since the coefficients had to be estimated via a nonlinear fitting process (e.g. LM algorithm), although the existence of the exponential coefficients improves its capability for nonlinear modeling. Moreover, the correlation implied in RS models can be easily revealed and the explicit expression can facilitate post-optimization analysis (e.g. uncertainty and sensitivity analysis).

The ANN provides a flexible framework for establishing the network architecture and adjusting the learning algorithm and transfer function. It facilitates the modeler to develop the ANN model easily for specific application through trial and error until the calibration and validation are acceptable. For RS models, it is relatively cumbersome to explore the guidelines of how to adjust the equation expressions to obtain better performance. A relative longer time for training is expected for ANN due to the fact that the transfer function is applied in each neural and the weighted sum is calculated through the entire network to generate the output. As a result, the long training time (in the order of minutes to hours) may render the ANN models to be inadmissible to certain applications (e.g. hourly flood forecasting), where real time gauge data assimilation is important. RS models such as QRS and NRS, on the other hand, may fulfill the real time data assimilation tasks expeditiously.

CONCLUSIONS

The response surface (RS) models including multiple linear regression (MLR), quadratic response surface (QRS) and nonlinear response surface (NRS) were introduced and applied to predictions of daily runoff. Two catchments, one in southeast of China and the other in western Canada were used to demonstrate the applicability of the RS models. A comparison study was conducted on the RS models (MLR, QRS and NRS) and artificial neural network (ANN), trained with gradient descent with adaptive learning rate (ANN-GDA) and Levenberg-Marquardt learning algorithm (ANN-LM), respectively. The results indicate that QRS and NRS are able to obtain comparable performance within shorter training times, thus rendering these models as viable alternatives to the ANN models.

Other issues such as the performance of the models using inputs with different time lags, the ability of the models to make long term and peak value predictions, and the sensitivity to the length of training samples were also analyzed. It was concluded that the performance of the ANN was sensitive to the inputs with different numbers of the previous days and the length of the training data. The RS models' results showed distinct variation with the time lags in the input data. The computational time of NRS was highly sensitive to the number of inputs. QRS and NRS performed better in daily prediction comparing to ANN, whereas the advantage in monthly and yearly predictions was not obvious. The performance of the ANN-GDA model in predicting the peak value was better than the ANN-LM and RS models.

This paper was an attempt to apply RS models in hydrologic forecasting and to compare RS with ANN under various predictive scenarios. RS models can provide an alternative for researchers and engineers engaged in hydrologic forecasting. Future studies are still needed to investigate the performance of the RS models applied in various scales of watersheds and multistep ahead forecasting.

ACKNOWLEDGEMENT

This research was supported by Singapore MOE Academic Research Fund (AcRF) Tier 1, Earth Observatory of Singapore (M4080707.B50), and DHI Water & Environment (S) Pte. Ltd.. The authors deeply appreciate the reviewers' insightful comments, which have contributed much to improving the manuscript.

References

Abrahart, R. J. and See, L. (2000). "Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments." *Hydrological Processes*, 14(11-12), 2157-2175.

Adeloye, A.J., (2009). "Multiple Linear Regression and Artificial Neural Networks Models for Generalized Reservoir Storage–Yield–Reliability Function for Reservoir Planning.", *Journal of Hydrologic Engineering*, 14(7), 731-738.

Andersson, M. and Adlercreutz, P. (1999). "Evaluation of simple enzyme kinetics by response surface modelling." *Biotechnology Techniques*, 13, 903-907.

Carlson, R., MacCormick, A., and Watts, D. (1970). "Application of Linear Random Models to Four Annual Streamflow Series." *Water Resources Research*, 6(4), 1070-1078.

Castellano-Méndez, M., González-Manteiga, W., Febrero-Bande, M., Prada-Sánchez, J.M., and Lozano-Calderón, R. (2004). "Modelling of the monthly and daily behaviour of the runoff of the Xallas River using Box-Jenkins and neural networks methods." *Journal of Hydrology*, 296(1-4), 38-58.

Croley II, T.E., (1980). "Gamma synthetic hydrographs." *Journal of Hydrology*, 47, 41-52.

Elshorbagy, A., Simonovicb, S.P., and Panu, U.S. (2002). "Estimation of missing streamflow data using principles of chaos theory." *Journal of Hydrology*, 255(1-4), 123-133.

Govindaraju, R. S. (2000a). "Artificial neural networks in hydrology. I: Preliminary concepts." *Journal of Hydrologic Engineering*, 5(2), 115-123.

Govindaraju, R. S. (2000b). "Artificial neural networks in hydrology. II: Hydrologic applications." *Journal of Hydrologic Engineering*, 5(2), 124-137.

Haktanir, T. and Sezen, N. (1990). "Suitability of two-parameter gamma and three-parameter beta distributions as synthetic unit hydrographs in Anatolia." *Hydrological Sciences Journal*, 35(2), 167-184.

Haltiner, J.P. and Salas, J.D. (1988). "Short-term forecasting of snowmelt discharge using ARMAX models." *Journal of American Water Resources Association*, 24(5), 1083-1089.

He. J.X. and Valeo, C. (2009). "Comparative Study of ANNs versus Parametric Methods in Rainfall Frequency Analysis." *Journal of Hydrologic Engineering*, 14(2), 172-184.

Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer feedforward networks are universal approximators." *Neural Networks*, 2(5), 359-366.

Hu, T.S., Lam, K.C., and Ng, S.T. (2001). "River flow time series prediction with a range dependent neural network." *Hydrological Sciences Journal*, 46 (5), 729-745.

Huang, G.H., Sun, W., Nie, X.H., Qin, X.S., and Zhang, X.D. (2010). "Development of a decision support system for rural eco-environmental management in Yongxin County, Jiangxi Province, China." *Environmental Modelling & Software*, 25(1), 25-42.

Huang, G.H., Yin, Y.Y., Luo, B., Qin, X.S., Nie, X.H., He, L., Cai, Y.P., Li, H.L., Liu, Z.F., Lin, Q.G., and Zhang, H. (2007). "An optimization-simulation approach for watershed management under changing climate in the Georgia Basin." Prepared for CCIAD, Earth Sciences Sector, Natural Resources Canada.

Iorgulescu, I. and Beven, K.J. (2004). "Nonparametric direct mapping of rainfall-runoff relationships: An alternative approach to data analysis and modeling?." *Water Resources Research*, 40, W08403.

Islam, M.N. and Sivakumar, B. (2002). "Characterization and prediction of runoff dynamics: a nonlinear dynamical view." *Advances in Water Resources*, 25(2), 179-190.

Jain, A. and Prasad Indurthy, S.K.V. (2004). "Comparative Analysis of Event-based Rainfall-runoff Modeling Techniques—Deterministic, Statistical, and Artificial Neural Networks." *Journal of Hydrologic Engineering*, 8(2), 93-98.

Jayawardena, A.W. and Gurung, A.B. (2000). "Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach." *Journal of Hydrology*, 228(3-4), 242-264.

Jayawardena, A.W. and Lai, F. (1994). "Analysis and prediction of chaos in rainfall and stream flow time series." *Journal of Hydrology*, 153(1-4), 23-52.

Karunanithi, N., Grenney, W. J., Whitley, D., and Bovee, K. (1994). "Neural networks for river flow prediction." *Journal of Computational Civil Engineering*, 8(2), 201–220.

Kişi, Ö. (2004). "River Flow Modeling Using Artificial Neural Networks." *Journal of Hydrologic Engineering*, 9(1), 60-63.

Kişi, Ö. (2007). "Streamflow Forecasting Using Different Artificial Neural Network

Algorithms." *Journal of Hydrologic Engineering*, 12(5), 532-539.

Kůrková, V. (1992). "Kolmogorov's theorem and multilayer neural networks." *Neural Networks*, 5(3), 501-506.

O'Connell, P.E., Nash, J.E. and Farrell, J.P., 1970. River flow forecasting through conceptual models part II — The Brosna catchment at Ferbane. *Journal of Hydrology*, 10(4): 317-329.

Londhe, S. and Charhate, S. (2010). "Comparison of data-driven modelling techniques for river flow forecasting." *Hydrological Sciences Journal*, 55(7), 1163–1174.

Maier, H.R. and Dandy, G.C. (2000). "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications." *Environmental Modelling & Software*, 15(1), 101-124.

Moré, J.J. (1978). "The Levenberg-Marquardt algorithm: Implementation and theory." *Numerical Analysis*, 630, 105-116.

Myers, R.H., Montgomery, D.C., Vining, G.G., Borror, C.M., and Kowalski, S.M. (2004). "Response surface methodology: A retrospective and literature survey." *Journal of Quality Technology*, 36(1), 53-77.

Nash, J.E. and Sutcliffe, J.V. (1970). "River flow forecasting through conceptual models part I — A discussion of principles." *Journal of Hydrology*, 10(3), 282-290.

Porporato, A. and Ridolfi, L. (1997). "Nonlinear analysis of river flow time sequences." *Water Resources Research*, 33(6), 1353-1367.

Porporato, A. and Ridolfi, L. (2001). "Multivariate nonlinear prediction of river flows." *Journal of Hydrology*, 248(1-4), 109-122.

Raman, H. and Sunikumar, N. (1995). "Multivariate modelling of water resources time series using artificial neural networks." *Hydrological Sciences Journal*, 40(2), 145-163.

Ranade, A. K., Pandey, M. and Datta, D. (2010). "Stochastic response surface based simulation of ground water modeling. In proceedings of International Conference on Modeling, Optimization, and Computing, 28-30 Oct. 2010: West Bengal, India.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986). "Learning Internal Representation by Error Propagation." *Parallel Distributed Processing, Vol 1: Foundations*, MIT Press, Cambridge, MA, 318-362.

Salas, J.D., Tabios III, G.Q., and Bartolini, P. (1985). "Approaches to multivariate modeling of water resources time series." *Journal of American Water Resources Association*, 21(4), 683-708.

Seber, G.A.F. and Wild, C.J. (1989). Nonlinear Regression. John Wiley & Sons Inc.

Singh, P. and Deo, M.C. (2007). "Suitability of different neural networks in daily flow forecasting." *Applied Soft Computing*, 7(3), 968-978.

Wang, W.C., Chau, K.W., Cheng, C.T. and Qiu, L. (2009). "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series." *Journal of Hydrology*, 374(3-4), 294-306.

Zealand, C.M., Burn, D.H. and Simonovic, S.P. (1999). "Short term streamflow forecasting using artificial neural networks." *Journal of Hydrology*, 214(1-4), 32-48.

Table Caption List:

Figure Caption Lists:

Table 1. Statistical information of rainfall and water level in Heshui catchment

Statistical Parameters	*Rainfall (mm/day)*		*Water level (m)*	
	1988-1998	1999-2000	1988-1998	1999-2000
Minimum	0.0	0.0	109.01	109.07
Maximum	104.3	100.7	112.97	112.19
Average	4.206	4.274	109.57	109.56
Standard deviation	10.113	10.201	0.369	0.363

Table 2. Performance in predicting water level for Heshui catchment

Models	*Calibration*				*Verification*				*Computation time (s)*
	R^2	*E*	*RMSE*	*MAPE*	R^2	*E*	*RMSE*	*MAPE*	
MLR	0.898	0.898	0.1178	6.7185	0.900	0.890	0.1148	7.3727	5
QRS	0.923	0.923	0.1023	5.8478	0.926	0.926	0.0988	6.4761	5
NRS	0.925	0.925	0.1012	5.7429	0.932	0.932	0.0945	6.2977	20
ANN-GDA	0.914	0.914	0.1082	6.0689	0.920	0.920	0.1028	6.5330	1110
ANN-LM	0.920	0.920	0.1037	5.8393	0.922	0.921	0.1016	6.4467	450

Table 3. Statistical information of hydrological and meteorological data in Coquitlam catchment

Statistical Parameters	*Rainfall (mm/day)*		*Mean Temperature (°C)*		*Discharge (m^3/s)*	
	1990-1999	2000-2001	1990-1999	2000-2001	1990-1999	2000-2001
Minimum	0.000	0.000	-11.30	-3.00	0.297	0.54
Maximum	282.60	137.6	24.80	22.30	107.0	58.80
Average	10.550	9.122	9.331	8.901	6.841	5.943
Standard deviation	21.493	17.316	6.340	5.705	9.640	6.814

Table 4. Performance in predicting daily discharge for Coquitlam catchment

Models	*Calibration*				*Verification*				*Computation time (s)*
	R^2	*E*	*RMSE*	*MAPE*	R^2	*E*	*RMSE*	*MAPE*	
MLR	0.594	0.594	6.1419	298.231	0.650	0.650	4.0368	238.458	5
QRS	0.691	0.691	5.5382	236.668	0.722	0.720	3.6089	190.278	5
NRS	0.704	0.704	5.2450	233.841	0.721	0.719	3.6189	186.444	240
ANN-GDA	0.654	0.654	5.6795	267.500	0.721	0.720	3.6120	203.748	720
ANN-LM	0.642	0.640	5.7875	268.129	0.725	0.711	3.6685	216.597	1470