Malaysian Journal of Civil Engineering 25(2):239-253(2013)

IMPROVEMENT OF STREAMFLOW SIMULATION FOR GAUGED SITE OF HYDROLOGICAL MODEL

Ponselvi Jeevaragagam[1]* and Slobodan P. Simonovic[2]

[1] *Industrial System Research Group (ISRG), Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia.*

[2]*Department of Civil and Environmental Engineering, University of Western Ontario, London, Ontario N6A 5B9, Canada.*

*Corresponding Author: *ponselvi@utm.my*

Abstract: The paper presents an improvement procedure for streamflow simulation at gauged site of a semi-distributed river basin model. In addition to streamflow and precipitation, meteorological observations that are not employed in the HEC-HMS model calibration are used as inputs in the procedure. Some of the available meteorological variables may be of limited values in calibrating a large range of streamflow hydrographs for obtaining the optimum state variables and parameters of a river basin model. This study presents the integration of the Bayesian regularization neural network with the HEC-HMS model to provide most accurate streamflow simulations at gauged site, for a wide range of streamflow hydrographs pertinent to the hydrometeorological conditions. The artificial neural network is capable of generating a good generalization with given hydrometeorological patterns.

Keywords: *River basin, hydrometeorological, bayesian network, gauged site*

1.0 Introduction

The hydrometeorological observations and a rainfall-runoff model are the basic components of streamflow simulation study in a river basin. They provide accurate streamflow simulation for prediction, planning and management of the water resources. The hydrometeorological observations include streamflow, precipitation and various meteorological data. The rainfall-runoff model may use some of the hydrometeorological data in the simulation of hydrological processes to produce streamflow hydrograph as output. Rainfall-runoff modeling can be done using either an empirical, conceptual or physically based model. According to the spatial distribution of hydrologic parameters, those models can be lumped, semi-distributed or distributed (Beven 2001; Cunderlik 2003). In many studies, the deterministic physically based models use mathematical representations to describe the selected hydrological processes, as well as a number of state variables and also parameters of the model.

Model calibration using a large number of time series observation data provides values for model parameters. However, models may not be ideal for accurately simulating streamflow for a wide range of hydrometeorological conditions. Uncertainty in model output can be resulted by the model limitations: (i) model structure; (ii) time variation; (iii) insufficient data; and other factors.

The output error of gauged site in a deterministic physically based model can be reduced by implementing an updating procedure based on the artificial neural network (ANN) technique. Neural network is an adaptive system, capable of learning nonlinear hydrological functions between inputs and outputs without analyzing the internal structure of the hydrological processes. This technique provides improvement to the model performance with faster model development and shorter computation time.

The neural network method is widely used in hydrological applications over the last two decades. A few studies have been conducted to review the theory and applications of the ANN in hydrology (ASCE 2000; Govindaraju and Rao 2000). Previous studies have also shown that ANNs are appropriate for modeling nonlinear relationships of rainfall-runoff processes (Zealand *et al*. 1999; Rajurkar *et al*. 2004; Ahmad and Simonovic 2005; Cullmann *et al*. 2006; Akhtar *et al*. 2009); stream flow forecasting (Anctil *et al*., 2004; Moradkhani *et al*., 2004; Kisi, 2007); precipitation forecasting (Toth *et al*., 2000; Luk et. al., 2001); river stage forecasting (Thirumalaiah and Deo, 1998; Bhattacharya and Solomatine, 2000; Liong *et al*., 2000); and groundwater modeling (Rogers and Dowla, 1994; Coulibaly *et al*., 2001). Meteorological data (such as, air temperature, snowmelt or snow depth, relative humidity, sunshine hours, evapotranspiration, number of cloudy days, ENSO index, wind velocity, wind direction etc.) were included in the research to improve the ANN prediction (Poff *et al*., 1996; Dolling and Varas, 2002; Anctil and Rat, 2005; Jain and Srinivasulu, 2006; Wardah *et al*., 2008; Aytek *et al*., 2008). There are many different updating approaches available in the hydrologic literatures that are most appropriate for the study of rainfall-runoff processes in the gauged watersheds. (Xiong and O'Connor 2002; Anctil *et al*., 2003; Abebe and Price, 2004; Xiong *et al*., 2004; Goswami *et al*., 2005; Abrahart and See, 2007). It was found that various input-output combinations of observations and/or simulated results used in these procedures could minimize the associated uncertainty, and improve the overall efficiency of the hydrological model. The most common updating approaches include the use of optimization methods in the ANN weight updates, use of output error of a physically based model in the streamflow forecasting, emulation of hydrological knowledge in a numerical model, and the development of a hybrid system coupled by two (or more) linear and/or nonlinear models.

In this study, an updating approach is introduced to improve the errors in simulated discharge at gauging station over a river basin with the assistance of ANN model (the Levenberg-Marquardt algorithm with Bayesian regularization), using available additional hyrometeorological data.

The paper starts with the presentation of a model output updating methodology in the next section. The implementation of the methodology to Mitchell gauging station of the Upper Thames River basin follows. The paper ends with the presentation of results and conclusions obtained from the study.

2.0 A Methodology for Output Updating a River Basin Model

The calibration of hydrologic model for gauged river basin using a large number of streamflow hydrographs is a very tedious and time-consuming process. The model calibration was performed manually by trial-and-error process of adjusting parameters via visual inspection to judge the goodness of fit of the model simulations to observations. The study presented in this paper introduces a new procedure based on the ANN technique for reducing the output error of a deterministic physically based model to gauging river basin. One of the advantages of this procedure is the inclusion of additional real-time data on the hydrometeorological environment. While the additional real-time data is not used in the model calibration, it is used by ANN model to provide more accurate flow values for a wide range of flow hydrographs. Other advantages of this procedure are that it can consume less computation time and provides faster and more accurate updates for the output errors of the physically based model for both recent and future flow hydrographs.

The overall output updating procedure for the physically based model, as illustrated in Figure 1, is presented next. First, the physically based model is run by using input variables to compute the streamflow values. The neural network model (a multilayer feed-forward network with the Levenberg-Marquardt algorithm and Bayesian regularization) is then applied by using available hydrometeorological observation data to improve the output error of the physically based model for the selected gauged streamflow site in a watershed. The steps in the proposed methodology based on the computational engine of the HEC-HMS (USACE, 2000) are summarized as follows:

1. Estimate the streamflow error at the gauged site of the HEC-HMS model, $e_{gHMS}(t-1)$ as:

   $$e_{gHMS}(t-1) = Q_o(t-1) - Q_{gHMS}(t-1) \qquad (1)$$

   where, $Q_o(t-1)$ is an average observed streamflow and $Q_{gHMS}(t-1)$ is the HEC-HMS computed flows at the gauged site, and t = 1 to N is the time step.

2. Determine the improved streamflow value(s) at the gauged site(s), $iQ_{gHMS}(t)$, as below:

$$iQ_{gHMS}(t) = Q_{gANN}(t) \quad (2)$$

$$Q_{gANN}(t) = \text{f}(e_{gHMS}(t\text{-}1);\ Q_o(t\text{-}1);\ \text{and other available meteorological data}) \quad (3)$$

The streamflow error ($e_{gHMS}(t-1)$) is improved with the assistance of the ANN approach. The simulated streamflow generated by the ANN model, $Q_{gANN}(t)$, as shown in Eq. 2, becomes the improved streamflow at the gauged site, $iQ_{gHMS}(t)$. The previous and/or recent streamflow error, observed streamflow, mean-areal rainfall and snowmelt, and additional meteorological variables are used by the ANN model in the output updating procedure.

The next section of the presented methodology is the Levenberg-Marquardt algorithm with Bayesian regularization, and the model evaluation criteria.

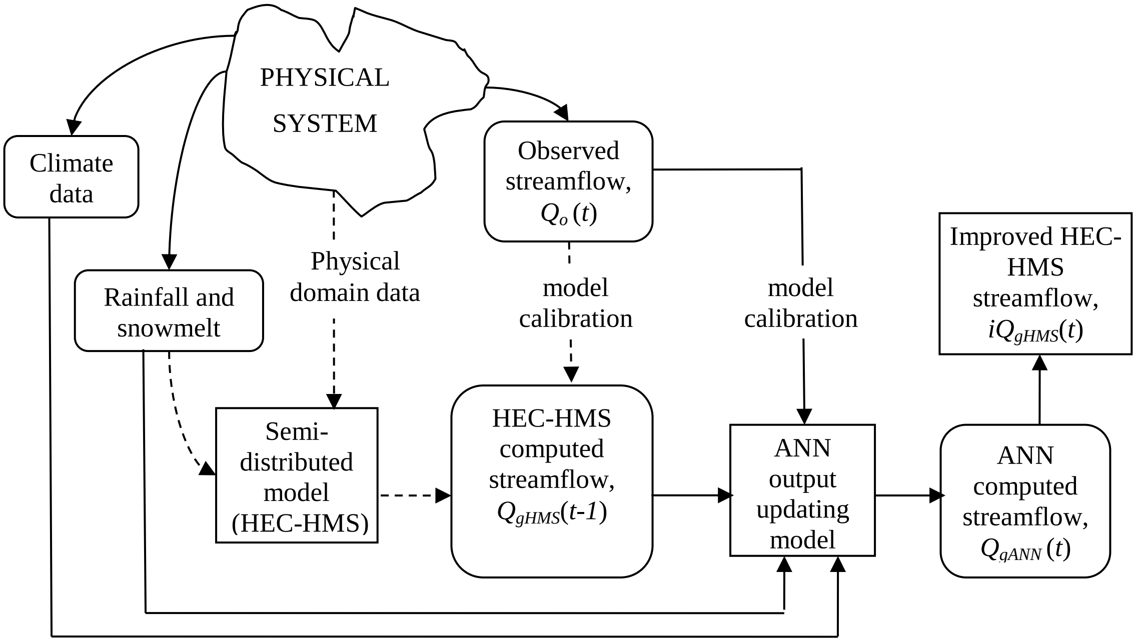

Figure 1: Schematic diagram of the output updating procedure

2.1 *The LM Algorithm with Bayesian Regularization*

The Bayesian regularization neural network can ensure accurate prediction of flow values, through automatically avoiding the overfitting and underfitting of the datasets. It also applies an early learning stopping procedure as soon as the overtraining signal starts to appear. In many applications, a multilayer feed-forward network associated with the Levenberg-Marquardt (LM) algorithm using Bayesian regularization had proven faster and more effective in finding optimal results (Foresee and Hagan, 1997; Anctil *et al*., 2003; Parent *et al*., 2008). The LM algorithm that belongs to the second-order nonlinear optimization techniques usually demonstrates the best performance (Hagan and Menhaj, 1994). In the Bayesian framework, a term that consists of mean of the sum of squares of the network weights and biases, F_w, is automatically added, to the mean sum of squares of the network errors, F_e, to improve generalization, as below (MacKay, 1992):

$$F = \beta F_e + \alpha F_w = \beta \sum_{i=1}^{N} e_i^2 + \alpha \sum_{i=1}^{M} W_i \tag{4}$$

where, F is the error function; e, is the network error, the difference between the desired flow, Q_o, and the network output, Q_{gANN}, for N number of training inputs; W is the network weights and biases for M total number of weights; α and β are the error function parameters.

2.2 *Model Performance Criteria*

The performance predictions of the ANN model and physically based model at the gauged streamflow location are evaluated for training, testing and validation datasets. The overall performance of the ANN model is evaluated using the coefficient correlation of linear regression, R, in Eq. 5. A high number of R = 1.0 means perfect statistical correlation. The success measurement of sensitivity analysis for choosing the input variables is based on the root mean square error (*RMSE*), given by Eq. 6, which measures the level of fitness between the ANN model output and the observed data. The peak flow criterion (*PFC*) in Eq. 7, is used to identify the more accurate ANN model for flood flow simulation. The mean absolute error (*MAE*), given by Eq. 8, measures the global goodness of the fit of the forecasted error (the difference between the observed data and the model predicted output). The correlation between the predicted hydrograph and the observed hydrograph is evaluated using the Nash-Sutcliffe coefficient of efficiency (Nash and Sutcliffe, 1970), *EI*, given by Eq. 9, which ranges from negative infinity to 1.0. An *EI* value of 1.0 means a good agreement between the observed and predicted hydrographs. Finally, both observed and predicted streamflow hydrographs for the validation dataset are plotted for visual evaluation of the output for periods of low and high streamflows.

$$R = \frac{\sum_{t=1}^{N}(Q_t - Q_{ave})(\hat{Q}_t - \hat{Q}_{ave})}{\sqrt{\sum_{t=1}^{N}(Q_t - Q_{ave})^2(\hat{Q}_t - \hat{Q}_{ave})^2}} \tag{5}$$

$$RMSE = \left[\frac{1}{N}\sum_{t=1}^{N}(Q_t - \hat{Q}_t)^2\right]^{1/2} \tag{6}$$

$$PFC = \frac{\left(\sum_{t=1}^{NP}(Q_t - \hat{Q}_t)^2 Q_t^2\right)^{1/4}}{\left(\sum_{t=1}^{NP}Q_t^2\right)^{1/2}} \tag{7}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|Q_t - \hat{Q}_t\right| \tag{8}$$

$$EI = 1 - \frac{\sum_{t=1}^{N}(Q_t - \hat{Q}_t)^2}{\sum_{t=1}^{N}(Q_t - Q_{ave})^2} \tag{9}$$

where $\hat{Q}$ is the ANN predicted streamflow for the gauged location, Q is the observed streamflow at t, recent time step, Q_{ave} is the average streamflow; N is the number of observations, and NP is the number of peak flows greater than one-third of the mean peak flow.

3.0 Case Study

The proposed methodology for output updating on gauged sites of a physically based model is presented using the study of the Upper Thames River (UTR) watershed, located in the southwestern Ontario, Canada. The region is comprised of four counties such as Perth, Middlesex, Huron and Oxford. There are two main tributaries of the Thames River: the North branch (1,750 km^2) and the East branch (1,360 km^2). They converge at forks near the centre of the city of London. The Thames River then flows westwards and exits the outlet of watershed near Byron. The slope at the upper reaches of the Thames basin is close to 1.9 m/km and much flatter at lower reaches - less than 0.2 m/km (after Wilcox *et al*., 1998). The dates of more recent floods in this watershed include March 1977, September 1986, July 2000, April 2008, and December 2008

(UTRCA, 2009). Flooding most frequently occurs after the spring snow melts and summer storms (Prodanovic and Simonovic, 2006).

The rainfall-runoff model for the UTR watershed was originally developed using the Hydrologic Modeling System (HEC-HMS) version 2.2.2 (USACE, 2000), a product of the Hydrologic Engineering Center within the U.S. Army Corps of Engineers. The details of this work can be found in the report by Cunderlik and Simonovic (2004). The HEC-HMS hydrologic model for the Upper Thames River watershed, as illustrated in Figure 2, consists of thirty two sub-watersheds, twenty one river reaches, and three reservoirs. The hydrologic model is divided into a number of modules. Firstly, the snow module is used separately to provide the precipitation and temperature input adjustments and to simulate solid precipitation accumulation and melt. The rainfall and snowmelt is then used as the input data in the calibrated HEC-HMS model for the water losses estimation. The losses module accounts for the amount of water moisture movement through various conceptual reservoirs within a watershed, canopy, land surface, soils, and groundwater. The losses module output includes evapotranspiration, surface excess, baseflow, and ground water recharge. The surface excess is used by the transform module to generate surface runoff. This is done by performing a convolution of the unit hydrograph with the precipitation excess. The surface runoff is then combined with the baseflow to produce the direct runoff. Finally, the flood routing computation module uses the direct runoff as input to propagate the flood wave along a stream channel.

3.1 The Input Data for Output Updating

The daily meteorological data from the nearest monitoring sites, such as Stratford (solar radiation), Wildwood Dam (evaporation) and London (air temperature, wind speed, wind direction, air station pressure, visibility, and humidity) are used in the procedure. These daily historic datasets are obtained from Environment Canada (EC) and the Upper Thames River Conservation Authority (UTRCA). A gauged streamflow of Mitchell river basin of the UTR watershed is selected to illustrate the output updating of the HEC-HMS continuous hydrologic model. The Mitchell SG in Figure 2 represents the Mitchell gauged site of the HEC-HMS model. The Mitchell river basin receives 921-1144 mm of annual precipitation (from year 2001-2005) and 4.9 m^3/s of estimated annual discharge. This river basin with 173 km^2 (5% of the UTR watershed) covers 93% agriculture, 5% forest and 2 % urban. The watercourse (total length 194 km) contributes 36% of the flow to the North Thames at Fanshawe Dam and 11% of the flow to the Thames downstream of London (UTR Watershed Report Cards, 2007).

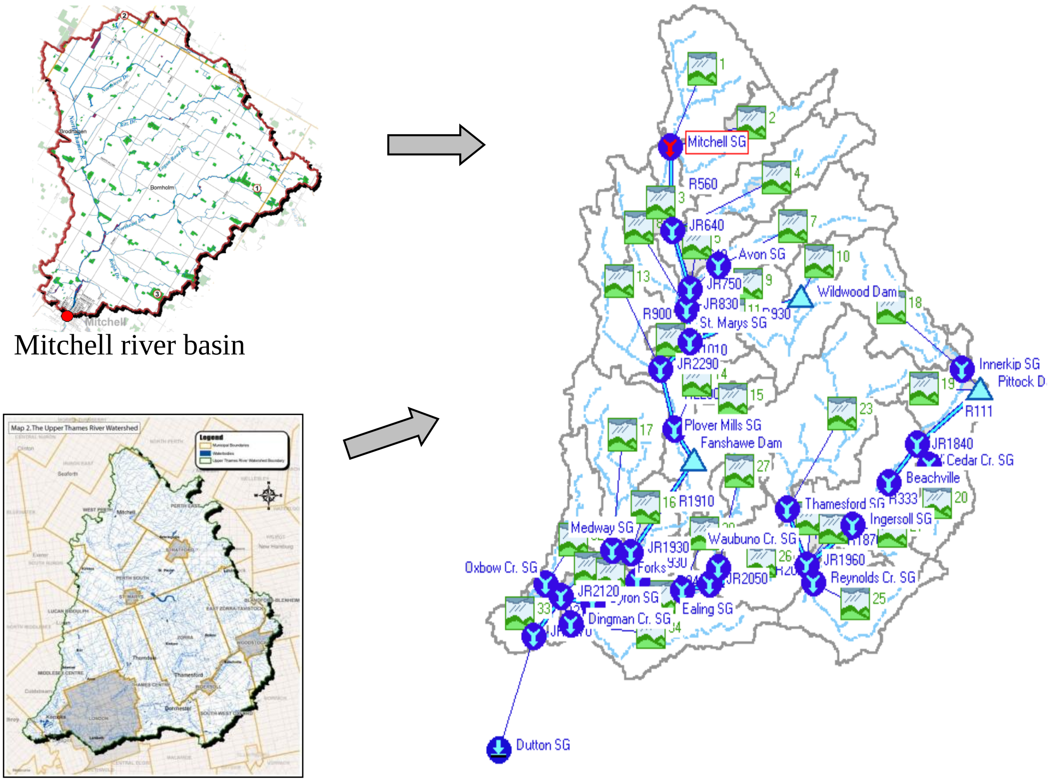

Figure 2: The HEC-HMS continuous hydrologic model of the UTR watershed

3.2 *The Neural Network Model*

The ANN input data includes the discharge error, average streamflow, average rainfall and snowmelt and meteorological variables (such as average visibility, average relative humidity, average wind speed, maximum solar radiation, minimum air station pressure, average evaporation, average air temperature, and average wind direction) at current and previous time steps. The autocorrelation and cross-correlation analyses are adopted to find all possible input variables for the ANN model. Table 1 shows the best correlation coefficient results (with lag-times up to 4 days) computed with 95% confidence interval for the input variables. For example, the Mitchell streamflow with lag-times of 1- to 3-day (Q_{t-1}, Q_{t-2} and Q_{t-3}); the average adjusted precipitation with lag-times of 0- to 2-day (NP_t, NP_{t-1}, and NP_{t-2}); and most of other meteorological variables with lag-time 1-day (e.g. H_{t-1}) are considered as potential input variables correlated with the current day Mitchell streamflow forecast (Q_t).

Table 1: The autocorrelation (A) and cross-correlation (C) results for Mitchell streamflow

Lags	A(*Q*)	C(*NP*)	C(*Sr*)	C(*H*)	C(*SP*)	C(*T*)	C(*V*)	C(*Wd*)	C(*Ws*)	C(*E*)
0	1.0	**0.1**	-0.2	0.1	-0.2	**-0.1**	-0.1	**0.1**	0.2	-0.2
1	**0.7**	**0.4**	**-0.2**	**0.2**	**-0.2**	-0.1	**-0.2**	0.0	**0.2**	**-0.2**
2	**0.5**	**0.3**	-0.2	0.2	-0.1	-0.1	-0.2	-0.1	0.1	-0.2
3	**0.3**	0.1	-0.1	0.1	0.0	-0.1	-0.1	-0.1	0.1	-0.2
4	0.2	0.1	-0.1	0.1	0.0	-0.1	-0.1	0.0	0.1	-0.2

where e is the output/discharge error of HEC-HMS (m^3/s); Q is the average flow (m^3/s); NP is the total rainfall and snowmelt (mm); V is the daily average visibility (km); H is the daily average relative humidity (%); Ws is the daily average wind speed (km/h); Sr is the daily maximum solar radiation (MJ m^{-2} day^{-1}); Sp is the daily minimum air station pressure (kPa); E is the daily average evaporation (mm day^{-1}); T is the daily average air temperature (^oC); Wd is the daily average wind direction (10's Deg); and t is the recent time and delayed daily three times t-1, t-2 and t-3.

Table 2: ANN input variables for Mitchell station

M1: e_{t-1}
M2: e_{t-1} and Q_{t-1}
M3: e_{t-1} , Q_{t-1} and Q_{t-2}
M4: e_{t-1} , Q_{t-1}, Q_{t-2} and Q_{t-3}
M5: e_{t-1}, Q_{t-1}, NP_t, and NP_{t-1}
M6: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and NP_{t-2}
M7: e_{t-1}, Q_{t-1}, NP_{t-1} and NP_{t-2}
M8: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and T_{t-1}
M9: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and E_{t-1}
M10: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and Sr_{t-1}
M11: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and Ws_{t-1}
M12: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and Sp_{t-1}
M13: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and V_{t-1}
M14: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and H_{t-1}
M15: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1} and Wd_{t-1}
M16: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1} and V_{t-1}
M17: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1}, V_{t-1} and H_{t-1}
M18: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1}, V_{t-1}, H_{t-1} and Ws_{t-1}
M19: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1}, V_{t-1}, H_{t-1}, Ws_{t-1} and Sr_{t-1}
M20: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1}, V_{t-1}, H_{t-1}, Ws_{t-1}, Sr_{t-1} and E_{t-1}
M21: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1}, V_{t-1}, H_{t-1}, Ws_{t-1}, Sr_{t-1}, E_{t-1} and T_{t-1}
M22: e_{t-1}, Q_{t-1}, NP_t, NP_{t-1}, Sp_{t-1}, V_{t-1}, H_{t-1}, Ws_{t-1}, Sr_{t-1}, E_{t-1}, T_{t-1} and Wd_{t-1}

Based on the result from correlation analysis, different daily neural network input variables are developed that may highly correlated with the recent observed Mitchell streamflow, as summarized in Table 2. Sensitivity analyses are experimented on these input variables to find the best input variables and number of hidden nodes, for the ANN model with the Levenberg-Marquardt optimum algorithm and Bayesian regularization. From Tables 2 and 3, the input configuration M12 with station pressure variable has the minimum *RMSE* value of 3.51 m^3/s when compared with the model configurations M8 to M15 using other meteorological variables. In the case of M16 to M22, the analysis considers humidity, wind speed, solar radiation, air station pressure, evaporation, air temperature and lastly winds direction. These variables are not used in the output updating process. Multilayer feed-forward networks with a range of 5 to 20 hidden nodes are successively trained, and the best performance with the validation dataset is obtained within a pool of 25 repetitions. This implies that the selected configuration is among the top 14% of the distribution of all possible configurations, with 95% confidence, according to Iyer and Rhinehart (1999). The datasets used for training, testing and validating the ANN model, as in Table 4, are selected by implementing data cross-validation, extreme data partition, and trial and error methods. This result suggests improvement in the *RMSE* value of the trained network when additional meteorological data is used. For example, the optimal input configuration M12 with 5 input variables and 12 hidden nodes for a given training dataset, offers the minimum *RMSE* value of 3.51 m^3/s. The best ANN input configuration is presented as below for Mitchell station:

$$Q(t) = \mathrm{f}(e(t\text{-}1), Q(t\text{-}1), NP(t), NP(t\text{-}1), Sp(t\text{-}1)). \tag{10}$$

The trained network is then validated with unknown dataset. The comparison result of performance predictions of the neural network model and the HEC-HMS model for Mitchell station is presented in Table 4. The ANN model results for all datasets show that the Bayesian regularization network offers more accurate streamflow values compared to the HEC-HMS model. To provide for further comparison, the ANN simulated flow hydrograph is plotted in Figure 3.

4.0 Results and Discussions

The ANN training and test datasets in Table 4 presents the most accurate streamflow values with *MAE* about 1.600 m^3/s; slightly higher value of *EI* up to 0.890; smaller value of *PFC* than 0.255 and perfect fit the observed streamflow with *R* greater than 0.940. The overall performance measures for the HEC-HMS model with the training and test datasets are not as good: larger *MAE* up to 4.504 m^3/s; very low value of *EI* = 0.097; higher value of *PFC* up to 0.569; and lower value of *R* than 0.472 when compared to the ANN model. The negative value of *EI* = 0.341 on test dataset indicates that the output of the HEC-HMS model gives the more reliable mean of observed streamflow values. The ANN model performance for the validation dataset is also superior. This can be seen from the comparison with the HEC-HMS model: lower value of *MAE* equal to1.513 m^3/s opposed to a 4.031 m^3/s; satisfactory ANN model with higher *EI* value of 0.887 against 0.072; and slightly smaller value of *PFC* of 0.355 against 0.593. Furthermore, from Figure 3, the simulated streamflow hydrograph obtained using the ANN model is matched by the observed streamflow hydrograph with the correlation coefficient value of 0.942 that is much higher than 0.396 of the HEC-HMS model for the test dataset. These performance measures clearly indicate that the ANN model performs better than the HEC-HMS model.

Table 3: The RMSE performance results of the ANN network configuration sensitivity analysis for Mitchell station (unit: m^3/s)

	Number of hidden nodes															
Models	5	6	7	8	9	10	11	**12**	13	14	15	16	17	18	19	20
M1	7.19	7.23	7.30	7.14	7.24	7.23	7.15	7.15	7.08	7.13	7.06	7.07	7.03	6.90	7.04	6.91
M2	6.41	6.37	6.38	6.40	6.31	6.27	6.18	6.15	6.19	6.10	6.06	6.14	6.09	6.06	6.07	6.04
M3	6.23	6.20	6.14	6.17	6.13	6.16	6.15	6.23	6.15	6.19	6.18	6.10	6.09	6.12	6.11	6.15
M4	6.23	6.16	6.14	6.11	6.10	6.14	6.11	6.13	6.14	6.16	6.14	6.18	6.10	6.10	6.08	6.13
M5	4.28	4.24	4.04	4.28	3.88	3.91	4.22	4.05	3.93	4.03	3.93	4.04	3.77	3.93	3.83	3.92
M6	4.18	4.06	4.08	4.00	4.07	4.05	4.00	3.93	3.90	4.05	4.00	4.02	4.01	4.37	3.93	3.92
M7	4.36	4.31	4.26	4.32	4.23	4.31	4.18	4.34	4.25	4.15	4.17	4.19	4.24	4.26	4.28	4.16
M8	4.23	4.13	4.22	4.32	4.10	4.27	4.24	4.18	4.13	4.20	4.01	4.17	4.30	3.98	4.09	4.03
M9	4.35	4.14	4.01	4.09	4.10	4.13	4.09	4.21	4.03	4.32	4.17	4.13	4.07	4.08	4.09	4.22
M10	4.15	4.14	4.02	4.14	4.05	4.17	4.12	4.18	4.03	4.15	4.18	4.17	4.17	4.17	4.18	4.18
M11	3.91	3.77	3.81	3.93	3.80	3.70	3.87	3.77	3.95	3.76	3.87	3.75	3.53	3.93	3.82	4.01
M12	4.11	3.74	3.89	4.05	4.01	3.82	3.72	**3.51**	3.81	3.73	3.69	3.76	3.72	3.65	3.67	3.72
M13	4.16	4.27	3.98	3.97	4.22	4.19	4.12	3.93	4.06	3.83	3.90	4.09	4.00	4.08	3.99	4.13
M14	4.28	4.13	4.05	3.97	4.10	4.02	3.92	3.86	3.93	4.17	4.04	3.99	3.78	4.08	3.94	3.97
M15	3.94	3.95	4.09	4.07	3.85	4.19	3.91	4.00	3.77	3.93	3.73	3.82	3.84	3.96	3.88	3.90
M16	4.03	3.79	3.65	3.91	3.83	3.80	3.79	3.72	3.85	3.90	3.75	3.80	3.71	3.77	3.77	3.82
M17	4.13	3.91	3.98	3.88	3.88	3.87	3.89	3.90	3.80	3.69	3.67	3.60	3.89	3.72	3.82	3.92
M18	3.68	3.77	3.63	3.79	3.65	3.70	3.66	3.75	3.74	3.97	3.70	3.77	3.93	3.82	4.25	3.79
M19	4.00	3.72	3.79	3.75	3.89	3.75	3.85	4.02	3.88	3.83	3.99	3.99	3.78	3.98	3.93	3.95
M20	3.98	4.12	4.07	4.04	4.05	3.84	4.06	3.95	4.18	4.00	4.03	4.06	4.16	3.92	4.03	4.03
M21	4.06	3.83	4.08	4.15	3.90	4.04	4.09	3.95	4.11	4.16	4.03	4.41	4.22	4.30	4.17	4.18
M22	4.06	4.02	3.64	3.86	4.08	3.88	3.83	3.89	3.92	4.12	3.94	4.01	3.96	4.19	4.18	4.12

Table 4: The performance prediction between the ANN model and the HEC-HMS model

Dataset and period	Mean flow (m^3/s)	Simulated model	*RMSE* (m^3/s)	*MAE* (m^3/s)	*EI*	*PFC*	*R*
Training	4.504	HMS	9.778	3.786	0.097	0.395	0.472
(2002 to 2005)		ANN	3.405	1.594	0.890	0.217	0.945
Test	4.975	HMS	11.927	4.504	-0.341	0.569	0.456
(2000)		ANN	3.635	1.600	0.875	0.255	0.940
Validation	5.294	HMS	10.657	4.031	0.072	0.593	0.396
(2001)		ANN	3.727	1.513	0.887	0.355	0.942

Where, HMS denotes the HEC-HMS model, and ANN represents the artificial neural network model.

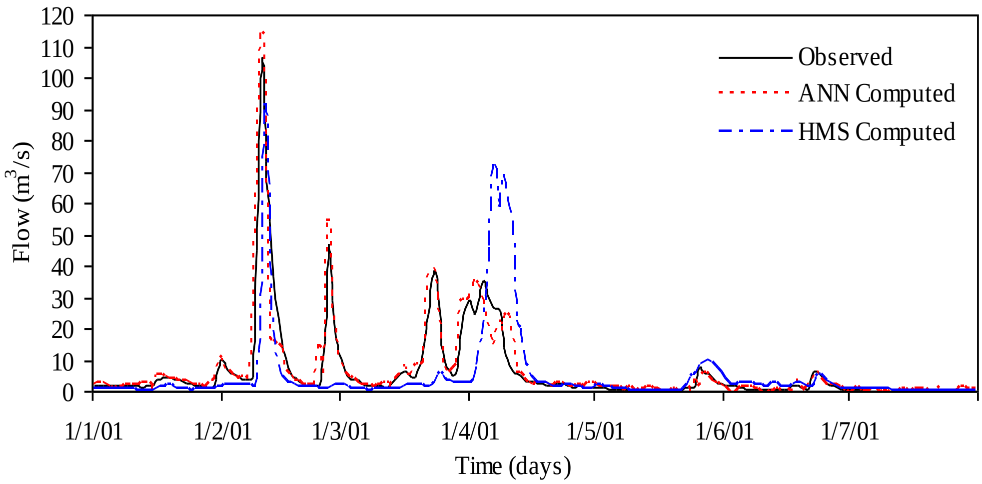

Figure 3: Improved streamflow hydrograph for validation dataset (starting January 2001)

5.0 Conclusions

This study presents an output updating procedure based on the ANN approach for gauged sites of river basins. The overall performance measures (such as *RMSE*, *MAE*, *EI*, *PFC* and *R*) of Bayesian regularization ANN model are superior to the HEC-HMS hydrologic model for Mitchell station. The ANN model shows the most accurate streamflow values with satisfactory model *EI* value higher than a 0.875, low values of *RMSE* and *MAE* ranging from 1.513 to 3.727 m^3/s, more accurate prediction of peak flow with lower value of *PFC* = 0.217 and a value of *R* above 0.940. It has also been shown that the use of additional meteorological data in the network training can considerably improve the trained network with a lower *RMSE* value. Furthermore, the results of analyses show that the implementation of LM algorithm and Bayesian regularization in the network training can provide a good approximation of input-output datasets. Therefore, the ANN model can successfully be applied to reduce discharge errors of gauged sites in a physically based model.

6.0 Acknowledgements

The authors wish to express their gratitude to Universiti Teknologi Malaysia (UTM Skudai), the Ministry of Higher Education Malaysia (MOHE) and the University of Western Ontario, London, Ontario, Canada for financial support of this study.

References

Abebe, A. J., and Price, R. K. (2004). "Information Theory and Neural Networks for Managing Uncertainty in Flood Routing." *Journal of Computing in Civil Engineering*, 18(4), 373-380.

Abrahart, R. J., and See, L. M. (2007). "Neural network modelling of non-linear hydrological relationships." *Journal of Earth System Science*, 11(5), 1563-1579.

Ahmad, S., and Simonovic, S. P. (2005). "An artificial neural network model for generating hydrograph from hydro-meteorological parameters." *Journal of Hydrology*, 315(1-4), 236-251.

Akhtar, M. K., Corzo, G. A., Van Andel, S. J., and Jonoski, A. (2009). "River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin." *Journal of Hydrology and Earth System Science*, 13, 1607-1618.

Anctil, F., Perrin, C., and Andreassian, V. (2003). "ANN output updating of lumped physically rainfall/runoff forecasting models." *Water Resources Bulletin*, 39(5), 1269-1279.

Anctil, F., Claude, M., Charles, P., and Vazken, A. (2004). "A soil moisture index as an auxiliary ANN input for stream flow forecasting." *Journal of Hydrology*, 286, 155-167.

Anctil, F., and Rat, A. (2005). "Evaluation of neural network streamflow forecasting on 47 watersheds." *Journal of Hydrology*, 10(1), 85-88.

ASCE Task Committee on Artificial Neural Networks in Hydrology (2000). "Artificial neural networks in hydrology." *Journal of Hydrological Engineering*, 5(2), 124-137.

Aytek, A., Asce, M., and Alp, M. (2008). "An application of artificial intelligence for rainfall-runoff modeling." *Journal of Earth System Science*, 117(2), 145-155.

Beven, K. J. (2001). *Rainfall-Runoff Modelling*, The Primer, John Wiley and Sons, Chichester, UK.

Bhattacharya, B., and Solomatine, D.P. (2000). "Application of neural networks in stage-discharge relationships." *Proc., 4-th International Conference on Hydroinformatics*, Iowa City, USA.

Coulibaly, P., Anctil, F., Aravena, R., and Bobe'e, B. (2001). "Artificial neural network modeling of water table depth fluctuations." *Water Resources Research*, 37(4), 885-896.

Cullmann, J., Schmitz, G., and Gorner, W. (2006). *A new system for online flood forecasting – performance and implications*, IHAR red book 303, Institute of Hydrology and Meteorology, University of Technology, Dresden, Germany.

Cunderlik, J. M. (2003). *Hydrologic Model Selection for the CFCAS Project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Condition*, Water Resources Research Report no. 046, Department of Civil and Environmental Engineering, The University of Western Ontario, London, Ontario, Canada, 38 pages. ISBN: (print) 978-0-7714-2622-3; (online) 978-0-7714-2623-0.

Cunderlik, J. M., and Simonovic, S. P. (2004). *Assessment of water resources risk and vulnerability to changing climatic conditions: Calibration and verification data for the HEC-HMS hydrologic model*, Report No. II, Department of Civil and Environmental Engineering, The University of Western Ontario, London, Ontario, Canada.

Dolling, O. R., and Varas, E. A. (2002). "Artificial neural networks for streamflow prediction." Journal of Hydraulic Research, 40 (5), 547-554.

Foresee, F. D., and Hagan, M. T. (1997). "Gauss-Newton approximation to Bayesian regularization." *Proc., International Joint Conference on Neural Networks*, 1930-1935.

Goswami, M., O'Connor, K. M., Bhattarai, K. P., and Shamseldin, A. Y. (2005). "Assessing the performance of eight real-time updating models and procedures for the Brosna River." *Journal of Earth System Science*, 9(4), 394-411.

Govindaraju, R. S., and Rao, A. R. (2000). *Artificial Neural Networks in Hydrology*, Kluwer Academic Publisher, Dordrecht, The Netherlands.

Hagan, M. T., and Menhaj, M. B. (1994). "Training feedforward networks with Marquardt algorithm." *IEEE Transactions on Neural Networks*, 5(6), 989-993.

Rogers, L. L., and Dowla, F. U. (1994). "Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling." *Water Resources Research*, 30 (2), 457-481.

Iyer, M. S., and Rhinehart, R. R. (1999). "A method to determine the required number of neural-network training repetitions." *IEEE Transactions on Neural Networks*, 10(2), 427-432.

Jain, A., and Srinivasulu, S. (2006). "Integrated approach to model decomposed flow hydrograph using artificial neural network and physically techniques." *Journal of Hydrology*, 317 (3-4), 291-306.

Kisi, O. (2007). "Streamflow forecasting using different artificial neural network algorithms." *Journal of Hydrologic Engineering*, 12(5), 532-539.

Liong, S. -Y., Lim, W. -H., and Paudyal, G. N. (2000). "River stage forecasting in Bangladesh: neural network approach." *Journal of Computing in Civil Engineering*, 14(1), 1-8.

Luk, K. C., Ball, J. E., and Sharma, A. (2001). "An application of artificial neural networks for rainfall forecasting." *Mathematical and Computer Modelling*, 33 (6-7), 683-693.

MacKay, D. J. C. (1992). "Bayesian interpolation." *Neural computation*, 4(3), 415-447.

Moradkhani, H., Hsu, K. -L, Gupta, H., and Sorooshian, S. (2004). "Improved streamflow forecasting using self-organizing radial basis function artificial neural networks." *Journal of Hydrology*, 295(1-4), 246-262.

Parent, A., Anctil, F., Cantin, V., and Boucher, M. (2008). "Neural Network Input Selection for Hydrological Forecasting Affected by Snowmelt." *Water Resources Bulletin*, 44(3), 679-688.

Poff, N. L., Tokar, S., and Johnson, P. (1996). "Stream hydrological and ecological responses to climate change assessed with an artificial neural network." *Limnology and Oceanography*, 41(5), 857-863.

Prodanovic, P., and Simonovic, S. P. (2006). *Assessment of risk and vulnerability to changing climatic conditions: Inverse Flood Risk Modelling of The Upper Thames River Basin*, Report No. VIII, Department of Civil and Environmental Engineering, The University of Western Ontario, London, Ontario, Canada.

Rajurkar, M. P., Kothyari, U. C., and Chaube, U. C. (2004). "Modeling of the daily rainfall runoff relationship with artificial neural network." *Journal of Hydrology*, 285, 96-113.

Thirumalaiah, K., and Deo, M.C. (1998). "River stage forecasting using artificial neural networks." *Journal of Hydrologic Engineering*, 3(1), 26-32.

Toth, E., Brath, A., and Montanari, A. (2000). "Comparison of short-term rainfall prediction models for real-time flood forecasting." *Journal of Hydrologic*, 239, 132-147.

USACE (2000). *Hydrologic Modelling System HEC–HMS: Technical reference manual*, United States Army Corps of Engineers, Hydrologic Engineering Center, Davis, California.

UTRCA (2009). "Inspiring a Healthy Environment: Flooding on the Thames River." Upper Thames River Conservation Authority, London, Ontario, Canada. Available online at http://www.thamesriver.on.ca /Water_Management/flood_history.htm - last accessed January 29, 2010.

UTRCA (2007). "Upper Thames River Watershed Report Cards." Available online at www.thamesriver.on.ca – last accessed March 25, 2012.

Wardah, T., Abu Bakar, S. H., Bardossy, A., and Maznorizan, M. (2008). "Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting." *Journal of Hydrology*, 356 (3-4), 283-298.

Wilcox, I., Quinlan, C., Rogers, C., Troughton, M., McCallum, I., Quenneville, A., Heagy, E., and Dool, D. (1998). *The Thames River Basin: A Background Study for Nomination under the Canadian Heritage Rivers System*, Upper Thames River Conservation Authority, London, Ontario, Canada.

Xiong, L., and O'Connor, K. M. (2002). "Comparison of four updating models for real-time river flow forecasting." *Hydrological Sciences Journal*, 47(4), 621-639.

Xiong, L., O'Connor, K. M., and Guo, S. (2004). "Comparison of three updating schemes using artificial neural network in flow forecasting." *Journal of Earth System Science*, 8(2), 247-255.

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.