Jurnal
Teknologi

Full Paper

EVENT-BASED RAINFALL-RUNOFF MODELING USING ADAPTIVE NETWORK-BASED FUZZY INFERENCE SYSTEM

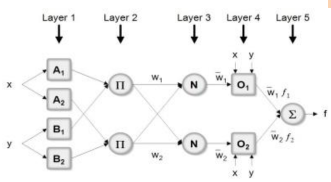
Nadeem Nawaz[a*], Sobri Harun[a], Amin Talei[b], Tak Kwin Chang[b]

[a]Department of Hydraulics and Hydrology, Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]Discipline of Civil Engineering, School of Engineering, Monash University Malaysia, Bandar Sunway, Selangor Darul Ehsan, Malaysia

Article history
Received
25 February 2016
Received in revised form
20 April 2016
Accepted
25 May 2016

*Corresponding author
nnadeem2@live.utm.my

Graphical abstract



Abstract

Population growth and transformation of agricultural or forest landscapes to built-up areas are the common phenomenon in the fast developing countries. Such changes have significant impact on hydrologic processes in the catchment which in turn may end up with an increase in both magnitude and frequency of floods in urban areas. Therefore, reliable rainfall-runoff models that are able to estimate discharge of a catchment accurately are in need. To date, several physically-based models are developed to capture the rainfall-runoff process; however, they require significant number of parameters which could be difficult to be measured or estimated. Beside these models, the artificial intelligence techniques have shown their ability to identify a direct mapping between inputs and outputs with less number of physical parameters. Adaptive network-based fuzzy inference system (ANFIS) is one of the well-practiced techniques in hydrological time series modeling. The aim of this study was to check the capability of ANFIS in event-based rainfall runoff modeling for a tropical catchment. A total of 70 rainfall-runoff events were extracted from twelve years hourly rainfall and runoff data of Semenyih River catchment where 50 of them were chosen for training and the remaining 20 for testing. An input selection method based on correlation analysis and mutual information was developed to identify the proper input combinations of rainfall and discharge antecedents. The results obtained by ANFIS model were then compared with an autoregressive model with exogenous inputs (ARX) as a bench mark. Results showed that ANFIS outperforms ARX model and has capabilities to be used as a reliable rainfall-runoff modeling tool.

Keywords: Rainfall-Runoff modeling; Event-based modeling; Neuro-fuzzy systems; ANFIS

Abstrak

Pertumbuhan penduduk dan transformasi landskap pertanian atau hutan ke kawasan binaan adalah fenomena biasa di negara-negara pesat membangun. Perubahan tersebut mempunyai kesan yang ketara ke atas proses hidrologi di kawasan tadahan yang seterusnya mungkin berakhir dengan peningkatan dalam kedua-dua magnitud dan kekerapan banjir di kawasan bandar. Setakat ini, beberapa model berasaskan fizikal dibangunkan untuk menjelaskan proses hujan-air larian; walaubagaimanapun, mereka memerlukan sejumlah besar parameter yang boleh menjadi sukar untuk diukur atau dianggarkan. Sistem pertimbangan logik fuzi berasaskan-adaptasi jaringan (ANFIS) adalah satu teknik yang diamalkan dalam hidrologi siri masa. Tujuan kajian ini adalah untuk memeriksa keupayaan ANFIS dalam pemodelan hujan air larian berasaskan peristiwa bagi kawasan tadahan tropika. Sebanyak 70 peristiwa hujan-air larian ini telah dipilih daripada data sela-jam dua belas tahun hujan dan air larian bagi tadahan Sungai Semenyih di mana 50 daripadanya dipilih untuk latihan dan baki 20 lagi untuk ujian. Satu kaedah pemilihan input berdasarkan analisis korelasi dan maklumat bersama telah dibangunkan untuk mengenal pasti kombinasi input yang betul. Keputusan yang diperolehi oleh model ANFIS kemudiannya dibandingkan dengan model regresi-auto model dengan input eksogenus (ARX) sebagai tanda aras. Hasil kajian menunjukkan bahawa ANFIS melebihi kebolehan model ARX dan mempunyai keupayaan untuk digunakan sebagai alat pemodelan hujan-air larian yang boleh dipercayai.

Kata Kunci: Permodelan hujan-air larian, Sistem logic fuzi, ANFIS

1.0 INTRODUCTION

Modeling of the rainfall-runoff processes has an important role in hydrological sciences as it can be helpful in decreasing the damages of flooding and also managing the water resources during the drought periods [1]. Many types of mathematical models have been established that can transform the input data into corresponding output data which can be classified as physically based, conceptual or data driven. Physically based models support an understanding of the processes of the input output transformation but are often problematic to calibrate due to insufficient field data [2]. The implementation and calibration of conceptual models typically presents various difficulties, requiring sophisticated mathematical tools, significant amounts of calibration data, and some degree of expertise and experience with the model [3]. Empirical, stochastic and system theoretic models falls in the category of data driven models. System theoretic models are well reputed in rainfall-runoff modeling since few decades due to producing a direct mapping between input and output without considering the physical procedures of the catchment. Artificial neural networks, genetic algorithms, wavelet methods, support vector machines, and fuzzy logic are data driven approaches which has been used in several studies. Artificial neural network (ANN) techniques have several successful applications in rainfall-runoff modeling and can be found in literature [4-7]. Although ANNs have shown reasonably good performance in rainfall-runoff modeling but still suffering from several issues including long training time, non-transparent internal process, and requiring trial and error procedure to find an optimum structure. Fuzzy logic approach has also been applied in rainfall-runoff modeling in different studies [8-11]. The fuzzy rule based approach was first introduces by [12]. There are two basic fuzzy logic modeling techniques, the expert knowledge based proposed by [13] and the data based proposed by [14]. Both techniques have their own benefits, the Mamdani approach can be applied to the problems with verbal data. This approach is also suitable for situations with inadequate data but a set of problem related linguistic statements. In contrast, the Takagi-Sugeno approach needs only numerical data and cannot applied with verbal data. Therefore, the Takagi-Sugeno approach is regarded as a data driven method and provides superior results when numerical input output data are provided.

A new method of merging artificial neural networks and fuzzy logic known as neuro-fuzzy systems has gradually attracted scientists in different fields. The neuro-fuzzy system is a fuzzy system that uses the learning capability of neural networks to determine its fuzzy sets and rules by processing data trials [15]. Neuro fuzzy system has significant advantage of reduced training time in comparison with ANNs. Moreover, NFS is not completely a black-box model as it can give an insight about its internal process in terms of IF-THEN rules. The successive applications of neuro fuzzy systems in rainfall-runoff simulation can be found in literature [16-23]. Adaptive Network-based Fuzzy Inference System (ANFIS) is the most popular neuro fuzzy system [24]. ANFIS was established by [25] to integrate the conception of fuzzy logic into neural networks to streamline learning and adaptation. ANFIS uses Takagi-Sugeno approach and found to be an appropriate tool in non-linear mapping between input and output data such as rainfall-runoff modeling.

2.0 METHODOLOGY

2.1 ANFIS Model

ANFIS can be applied by flexible structure of neural network thus useful for application purposes. It uses a feed forward network to search for fuzzy decision rules that perform well on a given task. Using a given input–output data set it creates an FIS for which membership function parameters are adjusted using either a back propagation algorithm alone or a combination of a back propagation algorithm and a least-squares method. This allows the fuzzy systems to learn from the data being modelled. Figure 1 illustrates the basic structure of ANFIS model.

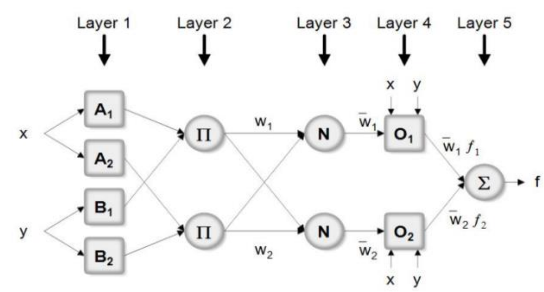


Figure 1 Adapted Network-based Fuzzy Inference System

Layer 1: Each node in this layer generates membership grades of an input variable. The node output OP_i^l is defined by:

$$OP_i^l = \mu_{Ai}(x) \quad \text{for} \quad i = 1, 2 \ \text{or} \tag{1}$$

$$OP_i^l = \mu_{Bi-2}(y) \quad \text{for} \quad i = 3, 4 \tag{2}$$

where x (or y) is the input to the node; A_i (or B_{i-2}) is a fuzzy set associated with this node, characterized by the shape of the membership functions in this node and can be any appropriate functions that are continuous and piecewise differentiable such as Gaussian, generalized bell shaped, trapezoidal shaped and triangular shaped functions. Assuming a generalized bell function as the membership function, Where (a_i, b_i, c_i) is the parameter set that changes the shapes of membership function with maximum equal to 1 and minimum equal to 0. The output can be computed as:

$$\mathrm{OP}_i^1 = \mu_{Ai}(x) = \frac{1}{1+(x-c_i/a_i)^{2b_i}} \tag{3}$$

Layer 2: Every node in this layer multiplies the incoming signals, denoted as π, and the output OP^2_i represents the firing strength of the rule and computed as:

$$\mathrm{OP}_i^2 = w_i = \mu_{Ai}(x)\,\mu_{Bi}(y), \quad i = 1, 2. \tag{4}$$

Layer 3: The i^{th} node of this layer, labelled as N, computes the normalized firing strengths as:

$$\mathrm{OP}_i^3 = w_i = \frac{w_i}{w_1 + w_2}, \quad i = 1,2. \tag{5}$$

Layer 4: Node i in this layer computes the contribution of the i^{th} rule towards the model output, with the following node function, where w is the output of layer 3 and (p_i, q_i, r_i) is the parameter set.

$$\mathrm{OP}_i^2 = w_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \tag{6}$$

Layer 5: The single node in this layer computes the overall output of the ANFIS as:

$$\mathrm{OP}_i^2 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{w_i} \tag{7}$$

2.2 ARX Model

ARX is a characteristic time-series estimating model, and consequently being used as a comparison model to assess the competency of any other model. Even though a direct model like ARX is probably not able to simulate the nonlinearity involved in rainfall-runoff process, but still it is a frequently used technique in exercise and it is suitable to set a bench mark for models comparison.

2.3 Study Area and Data Used

This study was performed on Semenyih River catchment which has four rainfall stations and one discharge station located at the outlet as can be seen in Figure 2.

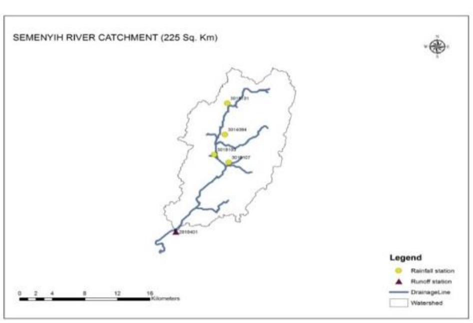


Figure 2 Map of Semenyih River catchment

Twelve years (2002-2013) of hourly rainfall and runoff data was arranged from the department of irrigation and drainage (DID), Malaysia.

2.4 Events Selections

The criteria adopted to select the events were based on the intensive rainfall causing a runoff higher than 20 m^3/s. Seventy of extreme events were extracted from the twelve years rainfall-runoff time series. Out of extracted events fifty events were used for training and the remaining twenty for the testing the model.

2.5 Input Selection

Antecedent rainfall and discharge are the potential inputs of a rainfall-runoff model. In order to find the proper inputs, coefficient of correlation (CC) can be calculated for rainfall with lead times ranging from R(t) to R(t-t_0) and the runoff Q(t) where t_0 is usually extended at least up to the time of concentration. In addition, mutual information (MI) was calculated using all possible combinations of two rainfall lead times from R(t) to R(t-t_0). Those combinations of two rainfall lead times which gave small MI values and showed large CC values were selected. In this study, the calculated CC and MI values were then normalized within the range of [0, 1] and a ranking procedure was applied to determine the inputs that gave the optimal combination of low MI and high CC values. The combinations with high ranking values are the potential candidates. This procedure can be carried out for different number of inputs including 2, 3, 4, etc. In each case, the k-highest ranked input combinations were chosen as candidates for input to the model. A verification process is needed to select the best combination and number of inputs from the selected candidates.

2.6 Data Processing

All the rainfall and runoff data were normalized before analysis. Normalization concentrates the dispersed data into a defined interval. The normalization method used in this study follows [26] which can be given by:

$$x_n = F_{min} + \left[\frac{x_i - x_{min}}{x_{max} - x_{min}}\right] \times (F_{max} - F_{min}) \tag{8}$$

where F_{min} and F_{max} are the required minimum and maximum of the new domain (e.g. 0.1-0.9), x_n is the standardized data, x_{min} and x_{max} are the minimum and maximum observed data, respectively; and x_i is the observed data.

2.7 Model Performances

The following error statistics and goodness of fit measures were adopted in this study:

(a) Nash–Sutcliffe coefficient of efficiency or NCE

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q}_i)^2} \tag{9}$$

where $\bar{Q}$ is the average observed discharge and n is the total number of the observations. Q_i is observed flow rate and $\widehat{Q}_i$ is the simulated low rate. This measure has a range between 1 and -∞.

(b) Coefficient of determination (R^2)

$$R^2 = \left[\frac{\sum_{i=1}^{n}(Q_i-\bar{Q})(\widehat{Q}_i-\tilde{Q})}{\sqrt{\sum_{i=1}^{n}(Q_i-\bar{Q})^2} \times \sqrt{\sum_{i=1}^{n}(\widehat{Q}_i-\tilde{Q})^2}}\right]^2 \quad (10)$$

where Q is the average simulated discharge and R^2 shows the degree of co-linearity between the observed and simulated time series and has a range of 0.0–1.0, with higher values indicating a higher degree of co-linearity.

(c) Root mean square error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Q_i-\widehat{Q}_i)^2}{n}} \quad (11)$$

The RMSE accords extra importance on the outliers in the data set and is therefore biased towards errors in the simulation of high flow rates.

(d) Mean Absolute error (MAE)

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|Q_i-\widehat{Q}_i|}{n} \quad (12)$$

MAE computes all deviations from the original data regardless of sign and is not weighted towards high flow values.

(e) Relative Peak Error (RPE)

In addition to goodness of fit, the accurate prediction of the peak flow is also important, therefore RPE is also included.

$$\mathrm{RPE} = \frac{|(Q_p)-(\widehat{Q}_p)|}{(Q_p)} \quad (13)$$

where Q_p is the observed peak discharge and $\widehat{Q}_p$ is simulated peak discharge. Values of RPE closer to zero indicate better estimation of the peak flow.

3.0 RESULTS AND DISCUSSION

An input selection method based on correlation analysis and mutual information was developed to identify the proper input combinations of rainfall and discharge antecedents. The catchment has four rainfall stations therefore the developed method analyzed the rainfall input at different Lags for each rainfall station to the Desired output Q(t). Moreover addition of one discharge antecedent Q(t-1) enhanced the model output. The best combination obtained from this method was R4(t-2), R4(t-4) and Q(t-1). The same input combination was used to develop ANFIS model. Moreover 40 no of epochs and Gaussian membership function were found appropriate for developing ANFIS model. The conventional statistical model ARX model was developed using all the rainfall stations. The predicting abilities and generalization capabilities in capturing rainfall-runoff dynamics of ANFIS model and ARX model in testing phase were evaluated by different statistics such as CE, R^2, RMSE, MAE, and RPE. The obtained performances from ANFIS and ARX models can be seen in Table 1.

Table 1 Performances of ANFIS and ARX models for testing events.

Testing	ANFIS					ARX				
Events	CE	R^2	RMSE	MAE	RPE	CE	R^2	RMSE	MAE	RPE
Event 1	0.81	0.84	4.62	2.60	0.04	0.40	0.45	8.77	6.90	0.04
Event 2	0.84	0.86	1.53	0.84	0.06	-0.97	0.09	6.09	5.41	0.21
Event 3	0.71	0.79	2.63	1.59	0.02	0.26	0.48	4.68	3.82	0.26
Event 4	0.64	0.68	2.88	1.92	0.01	0.23	0.42	3.92	3.33	0.06
Event 5	0.94	0.94	4.10	2.60	0.16	0.02	0.05	17.65	14.05	0.32
Event 6	0.95	0.96	3.77	2.10	0.12	0.66	0.69	10.60	8.93	0.17
Event 7	0.83	0.89	2.89	2.33	0.03	0.21	0.23	6.99	6.29	0.26
Event 8	0.94	0.94	4.10	2.60	0.04	0.23	0.27	21.84	18.51	0.42
Event 9	0.62	0.75	5.52	2.67	0.04	0.34	0.37	6.95	4.73	0.26
Event 10	0.71	0.76	4.22	2.33	0.03	0.40	0.45	6.68	5.36	0.20
Event 11	0.84	0.86	6.73	3.71	0.06	0.27	0.39	14.30	8.80	0.39
Event 12	0.90	0.93	1.84	1.31	0.02	0.21	0.21	5.35	3.43	0.33
Event 13	0.61	0.73	2.96	1.40	0.02	-0.58	0.03	6.37	5.36	0.16
Event 14	0.89	0.91	2.99	2.21	0.00	-0.24	0.01	11.60	9.91	0.26
Event 15	0.90	0.90	2.36	1.68	0.03	0.22	0.24	6.75	5.31	0.20
Event 16	0.93	0.93	1.58	0.87	0.03	0.22	0.22	4.54	2.93	0.16
Event 17	0.69	0.73	4.87	2.03	0.04	0.56	0.57	5.15	3.14	0.21
Event 18	0.89	0.91	2.99	2.21	0.01	0.24	0.35	20.52	16.07	0.27
Event 19	0.59	0.60	9.77	4.59	0.08	0.29	0.33	12.71	7.14	0.39
Event 20	0.83	0.84	2.35	1.65	0.03	0.26	0.28	4.74	3.30	0.13

As can be seen, ANFIS was able to simulate discharge successfully for all testing events. ANFIS model produced lower RMSE and MAE as well as higher CE and R^2. The less training time was taken by ANFIS and on the other hand model had shown good predicting ability of the high flows almost for all testing events.

Figure 3 illustrates the observed and simulated discharge by ANFIS model for the 20 testing events.

The performances obtained from the conventional ARX model were not satisfactory. The ANFIS model performed much better in term of all the statistics compared to ARX model.

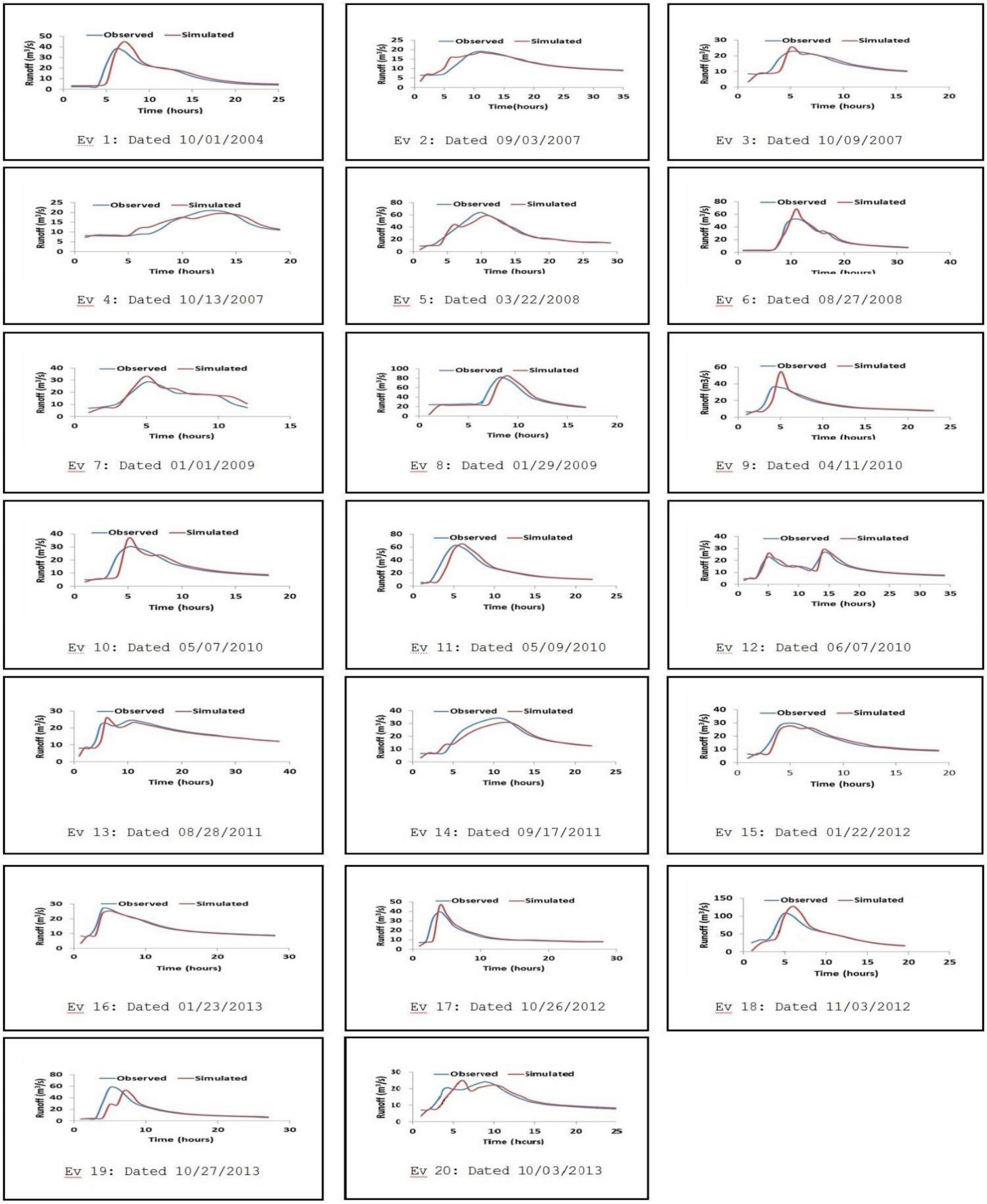


Figure 3 Observed and simulated hydrograph for the testing events

4.0 CONCLUSION

This study presented a successful application of the ANFIS model for event-based rainfall-runoff modeling in a tropical catchment. 70 major events were extracted from the hourly rainfall runoff data. Out of them 50 were chosen for training the ANFIS model and 20 were kept for tasting phase. ANFIS was able to simulate runoff well in terms of all statistics. Moreover, ANFIS was found to be superior to ARX model. In general, ANFIS was found to be a good potential to be used as a reliable computing tool in hydrologic modeling specifically for rainfall-runoff modeling. The model also took minimum time for the training. It is also concluded that more investigation on ANFIS model for simulation of rainfall-runoff processes is needed to attain the attraction of hydrologists.

Acknowledgement

The authors are grateful to ministry of education (MOHE) and Universiti Teknologi Malaysia for the financial support and Department of Irrigation and Drainage, Malaysia for provision of hydrological data for Semenyih River catchment.

References

[1] Aqil, M. Kita, I. Yano, A. and Nishiyama, S. 2007. A Comparative Study Of Artificial Neural Networks And Neuro-Fuzzy In Continuous Modeling Of The Daily And Hourly Behaviour Of Runoff. *Journal of Hydrology*. 337(1–2): 22-34.

[2] Jayawardena, A.W. Perera, E.D.P. Zhu, B. Amarasekara, J.D. and Vereivalu, V. 2014. A Comparative Study Of Fuzzy Logic Systems Approach For River Discharge Prediction. *Journal of Hydrology*. 514(0): 85-101.

[3] Babovic, V. and Keijzer, M. 2005. Rainfall-Runoff Modeling Based on Genetic Programming. Encyclopedia of Hydrological Sciences.

[4] Halff, A.H. Halff, H.M. and Azmoodeh, M. 1993. Predicting Runoff From Rainfall Using Neural Networks. *Engineering Hydrology*. 760-765.

[5] Tokar, A.S. and Johnson, P.A. 1999. Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering*. 4(3): 232-239.

[6] Nor, N. Harun, S. and Kassim, A. 2007. Radial Basis Function Modeling of Hourly Streamflow Hydrograph. *Journal of Hydrologic Engineering*. 12(1): 113-123.

[7] Kuok, K.K. Harun, S. and Shamsuddin, S.M. 2010. Particle Swarm Optimization Feedforward Neural Network For Modeling Runoff. *International Journal of Environmental Science and Technology*. 7(1): 67-78.

[8] Hundecha, Y. Bardossy, A., Werner, H.W. 2001. Development of a fuzzy logic-based rainfall-runoff model. *Hydrological Sciences Journal*. 46(3): 363-376.

[9] Nayak, P. Sudheer, K. and Ramasastri, K. 2005. Fuzzy Computing Based Rainfall–Runoff Model For Real Time Flood Forecasting. *Hydrological Processes*. 19(4): 955-968.

[10] Özelkan, E.C. and Duckstein, L. 2001. Fuzzy Conceptual Rainfall–Runoff Models. *Journal of Hydrology*. 253(1–4): 41-68.

[11] Maskey, S. Guinot, V. and Price, R.K. 2004. Treatment Of Precipitation Uncertainty In Rainfall-Runoff Modelling: A Fuzzy Set Approach. *Advances in Water Resources*. 27(9): 889-898.

[12] Zadeh, L.A. 1965. Fuzzy sets: Information And Control. 8(3): 338-353.

[13] Mamdani, E.H. 1974. Application Of Fuzzy Algorithms For Control Of Simple Dynamic Plant. *Proceedings of the Institution of Electrical Engineers*. 121(12): 1585-1588.

[14] Takagi, T. and Sugeno, M. 1985. Fuzzy Identification Of Systems And Its Applications To Modeling And Control: Systems. *Man and Cybernetics, IEEE Transactions on, SMC*. 15(1): 116-132.

[15] Lin, C.T. and Lee, C. 1996. Neural fuzzy systems: a neuro-fuzzy synergism to intelligent systems. Prentice-Hall, Inc.

[16] Talei, A., Chua, L.H.C. and Wong, T.S.W. 2010. Evaluation Of Rainfall And Discharge Inputs Used By Adaptive Network-Based Fuzzy Inference Systems (ANFIS) in rainfall–runoff modeling. *Journal of Hydrology*. 391(3–4): 248-262.

[17] Zounemat-Kermani, M. and Teshnehlab, M. 2008. Using adaptive neuro-fuzzy inference system for hydrological time series prediction. *Applied Soft Computing*. 8(2): 928-936.

[18] Nayak, P.C. Sudheer, K.P. Rangan, D.M. and Ramasastri, K.S. 2004. A Neuro-Fuzzy Computing Technique For Modeling Hydrological Time Series. *Journal of Hydrology*. 291(1–2): 52-66.

[19] Nawaz, N. Harun, S. and Talei, A. 2015. Application Of Adaptive Network-Based Fuzzy Inference System (ANFIS) For River Stage Prediction In A Tropical Catchment. *Applied Mechanics and Materials*. 735(0): 195-199.

[20] Talei, A. Chua, L.H.C. and Quek, C. 2010. A Novel Application Of A Neuro-Fuzzy Computational Technique In Event-Based Rainfall–Runoff Modeling. *Expert Systems with Applications*. 37(12): 7456-7468.

[21] Nourani, V. and Komasi, M. 2013. A Geomorphology-Based ANFIS Model For Multi-Station Modeling Of Rainfall–Runoff Process. Journal of Hydrology. 490(0): 41-55.

[22] Nasr, A. and Bruen, M. 2008. Development Of Neuro-Fuzzy Models To Account For Temporal And Spatial Variations In A Lumped Rainfall–Runoff Model. Journal of Hydrology. 349(3–4): 277-290.

[23] Vernieuwe, H. Georgieva, O, De Baets, B. Pauwels, V. R. Verhoest, N. E. and De Troch, F. P. 2005. Comparison Of Data-Driven Takagi–Sugeno Models Of Rainfall–Discharge Dynamics. Journal of Hydrology. 302(1–4): 173-186.

[24] Hong, Y.S.T. and White, P.A. 2009. Hydrological Modeling Using A Dynamic Neuro-Fuzzy System With On-Line And Local Learning Algorithm. Advances in Water Resources. 32(1): 110-119.

[25] Jang, J.S.R. 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. Systems. Man and Cybernetics: IEEE Transactions on. 23(3): 665-685.

[26] Van Ooyen, A. and Nienhuis, B. 1992. Improving The Convergence Of The Back-Propagation Algorithm. Neural Networks. 5(3): 465-471.