

## A COMPARATIVE STUDY ON CALIBRATION METHODS OF NASH'S RAINFALL-RUNOFF MODEL TO AMMAMEH WATERSHED, IRAN

Vahid Nourani\*

*Faculty of Civil Engineering, University of Tabriz, Iran*

Received 7 December 2007; received in revised form 26 April 2008; accepted 8 June 2008

---

**Abstract:**

Increasing importance of watershed management during last decades highlighted the need for sufficient data and accurate estimation of rainfall and runoff within watersheds. Therefore, various conceptual models have been developed with parameters based on observed data. Since further investigations depend on these parameters, it is important to accurately estimate them. This study by utilizing various methods, tries to estimate Nash rainfall-runoff model parameters and then evaluate the reliability of parameter estimation methods; moment, least square error, maximum likelihood, maximum entropy and genetic algorithm. Results based on a case study on the data from Ammameh watershed in Central Iran, indicate that the genetic algorithm method, which has been developed based on artificial intelligence, more accurately estimates Nash's model parameters.

**Keywords:**

Rainfall-runoff modeling; parameter estimating; genetic algorithm; Nash model; Ammameh watershed

© 2008 *Journal of Urban and Environmental Engineering (JUEE)*. All rights reserved.

## INTRODUCTION

During last years, many efforts have been done to accurately estimate runoff within watersheds. Since 1930 various linear and nonlinear hydrologic models have been developed to simulate and forecast hydrologic processes and variables. Utilizing new methods and increasing knowledge about hydrologic processes helped these models to improve through the years. Though, because of the complexity in the rainfall-runoff process, mostly conceptual models have been used in simulations and analysis of this process (Nourani & Mano, 2007). Furthermore, linear reservoir model presented by Zoch, as the eldest and simplest and most applicable model in simulating rainfall-runoff and stream flow routing, is the base of most other conceptual models (Chow *et al.*, 1988). Indeed the linear reservoir model of Nash is the first conceptual model (Nash, 1957), in the base of linear reservoir concept, which utilized mathematical theory in developing instantaneous unit hydrograph (IUH) within a basin (Nourani *et al.*, 2007). This model has been addressed by many researchers such as Diskin *et al.* (1978), Ponce (1980), Croley (1980), Singh (1988) and Szilagyi (2003).

According to significant importance of accurately estimating the parameters of conceptual models such as Nash's model, various methods have been developed.

Some of the conventional parameter estimation methods are: moment method, least square error method, Maximum likelihood method, Maximum entropy method and linear and nonlinear programming method (see Singh, 1988).

Furthermore, recently artificial intelligent methods have been utilized to improve these conceptual models. During last years, many problems which didn't have a clear solution have been solved using artificial intelligence methods (Chau, 2007). Too many parameters affecting a physical process and completely nonlinear relations between these parameters could increase complications of assessing and analyzing this process. These intelligent methods, using potential knowledge therein available data, develop general relations between these data and adopt these relations within other conditions. Genetic algorithm is one of these intelligent methods (Goldberg, 1989) which is increasingly utilized in various optimization problems (especially since 1990 in the field of water resources engineering). In the parameter calibration process, the genetic algorithm can be considered as a robust approach to solving problems that are not yet fully characterized or too complex to allow full characterization, but for which some analytical evaluation is available (Santos *et al.*, 2003).

In hydrology, Wang (1991) used genetic algorithm to calibrate the Tank model. Then, Cieniawski *et al.* (1995) utilized genetic algorithm to optimize a ground water model. Franchini (1996) combined genetic

algorithm with a successive nonlinear programming method, to calibrate a conceptual model for rainfall-runoff. Aly & Peralta (1999) combined genetic algorithm with neural networks to optimize use of aquifers. Cheng *et al.* (2002) used genetic algorithm - fuzzy logic hybrid model in order to calibrate a multi-objective rainfall-runoff model. Jain & Srinivasulu (2004) used non-binary genetic algorithm with neural networks to improve rainfall-runoff modeling and presented a new class for these models. Also Jain *et al.* (2005) used this non-binary genetic algorithm to determine an optimal unit pulse response for rainfall-runoff model.

Considering the variety of calibration methods of rainfall-runoff models, evaluation and comparison of these methods can be an instructive study for the hydrologist in real world applications. This paper tries to calculate parameters of Nash's conceptual rainfall-runoff model which is the most popular rainfall-runoff model, utilizing some conventional methods as well as genetic algorithm by using sufficient statistical data of a real case study. Then, the accuracies of the methods have been evaluated and compared.

The rest of the paper has been organized as follows. Next section presents a brief description about Nash's model. The used calibration methods are introduced in the other section. Study area and results are presented in two separated sections, respectively and conclusions will be the last section of the paper.

## NASH'S MODEL

Nash (1957) developed his instantaneous hydrograph model through this assumption that the watershed is formed of a successive linear reservoirs cascade with rainfall input at the first reservoir. Applying this assumption within linear reservoir model, gamma distribution function based on  $n$  (number of reservoirs) and  $k$  (reservoir storage coefficient) will be as follow (Singh, 1988):

$$h(t) = \frac{1}{\Gamma(n)} \left( \frac{t}{k} \right)^{n-1} \frac{e^{-t/k}}{k} \quad (1)$$

where  $\Gamma$  is gamma function,  $t$  is time and  $h$  indicates instantaneous unit hydrograph model of Nash. Parameters of  $n$  and  $k$  can be obtained through various estimation methods.

Peak time,  $t_p$ , can be computed through derivation of **Eq. (1)** related to  $t$  and letting  $dh/dt=0$ :

$$t_p = k(n-1) \quad (2)$$

Total time lag,  $t_L$ , is:

$$t_L = nk \quad (3)$$

The lag time (lag of a watershed) is defined as time interval between area center of excess rainfall hyetograph and area center of direct hydrograph (Singh, 1988).

Sometimes it is taken as the time interval between onset of excess rainfall and center of surface through direct runoff. Therefore lag time of a reservoir  $k$  can be obtained using Eq. (3) when the number of the reservoirs  $n$  and the watershed total lag  $t_L$  are known.

**PARAMETER ESTIMATION METHODS**

**Moment method**

While  $t$  is a continuous variable and  $f(t)$  is its function,  $r$ th momentum,  $M_r$ , of  $f(t)$  about origin is defined as follows (Singh, 1988):

$$M_r = \int_0^{+\infty} (t)^r f(t) dt \tag{4}$$

and parameters of Nash model are computed through moment theorem as follows (Singh, 1988):

$$M_1(Q) - M_1(I) = nk \tag{5}$$

$$M_2(Q) - M_2(I) = nk^2(n+1) + 2nkM_1(I) \tag{6}$$

$I$  and  $Q$  are input and output hydrographs as functions of  $t$ .

**Least square method**

Assume a function as  $y=f(x,a_1,a_2,\dots,a_m)$ , in which parameters of  $a_i$  ( $i=1,2,3,\dots,m$ ) are to be determined. The least square method, evaluate the parameters by minimizing least square differences between observed and computed  $y$ ; this summation can be expressed by:

$$S = \sum_{i=1}^n a_i^2 = \sum_{i=1}^n [y_0(i) - y_c(i)]^2 = \sum_{i=1}^n [y_0(i) - f(x_i, a_1, a_2, \dots, a_m)]^2 \tag{7}$$

where  $y_0(i)$  is the  $i$ th observed  $y$ ,  $y_c(i)$  is  $i$ th computed  $y$  and  $n > m$  is the number of observations. By derivation of  $S$  related to each parameter and letting it equal to zero Eq. (7) will be minimized as follows:

$$\begin{aligned} \frac{\partial \sum_{i=1}^n [y_0(i) - f(x_i, a_1, a_2, \dots, a_m)]^2}{\partial a_1} &= 0 \\ \frac{\partial \sum_{i=1}^n [y_0(i) - f(x_i, a_1, a_2, \dots, a_m)]^2}{\partial a_2} &= 0 \\ &\vdots \\ \frac{\partial \sum_{i=1}^n [y_0(i) - f(x_i, a_1, a_2, \dots, a_m)]^2}{\partial a_m} &= 0 \end{aligned} \tag{8}$$

Then through these  $m$  equations, which are called as normal equations,  $m$  parameters of  $a_i$  can be obtained.

**Maximum likelihood method**

Assume probability density function (PDF) of  $x$  as  $y=f(x,a_1,a_2,\dots,a_m)$ , in which parameters of  $a_i$  ( $i=1,2,3,\dots,m$ ) had to be determined.  $x_1,x_2,\dots,x_n$  are random samples of this density, combined probability density function would be as  $y=f(x_1,x_2,x_3, a_1,a_2,\dots,a_m)$ . The samples are random with independent assumption, so combined probability density function can be expressed as the PDFs multiply:

$$f(x_1,x_2,\dots,x_n,a_1,a_2,\dots,a_m) = \prod_{i=1}^n f(x_i,a_1,a_2,\dots,a_m) = L \tag{9}$$

Parameters can be determined by maximizing  $L$ . Therefore by derivation of  $L$  related to each parameter and letting it equal to zero,  $m$  equations will attain as follows:

$$\begin{aligned} \frac{\partial L(a_1, a_2, \dots, a_m)}{\partial a_1} &= 0 \\ \frac{\partial L(a_1, a_2, \dots, a_m)}{\partial a_2} &= 0 \\ &\vdots \\ \frac{\partial L(a_1, a_2, \dots, a_m)}{\partial a_m} &= 0 \end{aligned} \tag{10}$$

Solving these  $m$  equations,  $m$  unknown parameters can be determined.

Nash's model parameters through maximum likelihood method can be computed as follows (Rezai Pajand & Bozorgnia, 1995):

$$\ln A = \ln n - \ln\left(\frac{1}{k}\right) \tag{11}$$

$$\ln n - \Psi(n) = \ln A - \ln G \tag{12}$$

where  $A$  is arithmetic mean,  $G$  is geometric mean and  $\Psi$  is psi function.

**Maximum entropy method**

Entropy is a criterion to explain indeterminacy of variable  $x$  in  $f(x)$  and can be defined as follows (Singh, 1988):

$$\begin{aligned} I[f] = I[x] &= -k \int_0^{\infty} f(x) \ln[f(x)] dx \\ 0 \leq x &< \infty \end{aligned} \tag{13}$$

$k > 0$  is an arbitrary constant or a scale coefficient, that depends on type of measurement.

Nash's model parameters through entropy method can be computed as follows (Singh, 1988):

$$nk = E[t] \tag{14}$$

$$\ln(n) - \Psi(n) = \ln(E[t]) - E[\ln t] \tag{15}$$

where  $E[t]$  is expected function.

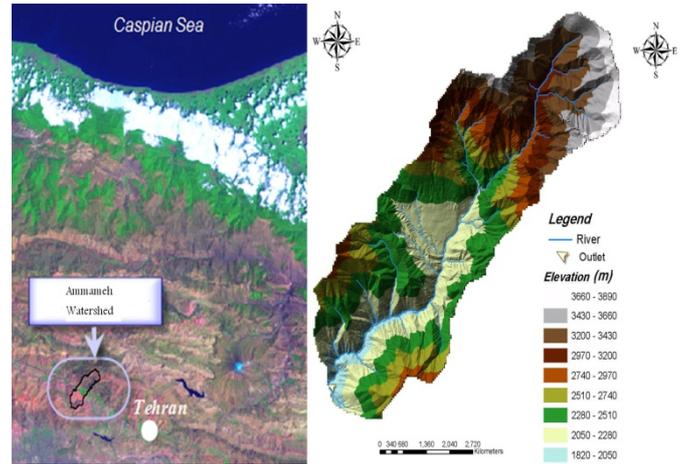
**Genetic algorithms**

This method has been developed based on nature and the role of inheritance on evolution of it and tries to optimize mathematical systems. Generally, genetic algorithm can be explained as follows. First, a population of chromosomes is selected through possible solutions and then target function value is calculated for each member of this population. In the next phase, new population is produced through existing population by a certain probability distribution function or another random operator and then target function value is calculated for each members of this new population. Comparing this new population (offspring) and parent population, new members can be selected for the next population. Though, selection procedures vary between different genetic algorithms.

**STUDY AREA**

Ammameh watershed with area of 37.2 km<sup>2</sup> and elevations from 1990 up to 3868 m, is one of the sub-basins of Jajrood watershed, upstream Latian dam that located in south of central Alborz, Tehran and here is considered as the case study watershed. Ammameh river, as the drain of this watershed, directed from north east to south west and at Kamarkhani reaches Jajrood river and has a length about 13 km. About 200 hectares of the vegetation coverage include gardens and grass, and the remainder has vegetation coverage of bushes.

The geology formation of the watershed is hard volcanic and surface layer is constructed 15 cm approximately thick, dark brown in color with a varying texture of sandy to silt and clay. The topography is steep with average slope 11%. Hence the soils are susceptible to erosion to some extended. The prevailing climate of the study area is snowy and sub-humid having four well defined seasons viz. spring, summer, autumn and winter. During the wet season, the area is under the influence of middle-latitude westerlies, and most of the rain that occurs over the region during this period is caused by depressions moving over the area, after forming in the Mediterranean Sea on a branch of the polar jet stream in the upper troposphere. The mean daily temperatures vary from -22°C in January up to



**Fig. 1** Ammameh watershed map.

40°C in July with a yearly average of 9°C. The dominant winds over the area blow from the northeast and the southwest.

Some storm events within this basin have been recorded at time intervals of 30 minutes. **Figure 1** shows the watershed position and map.

**RESULTS AND DISCUSSION**

In order to determine parameters, data from 8 rainfall-runoff storm events, which have been recorded at station of Kamarkhani (located at output of Ammameh watershed), are used. Rainfall-runoff data of events have been recorded at time intervals of 30 minutes until the end of each event. **Table 1** presents information about these events. For the reassessment, the events of 1 to 6 are used for calibration and events 7 and 8 for verification of the accuracy of the methods. Columns of **Table 1** presents number of events, date of events, rainfall height, equivalent direct runoff height, rainfall rate and rainfall to runoff ratio, respectively.

Observed unit hydrograph of each event is determined as follows. First, using the constant slope method base flow of each event is calculated and then the observed direct hydrograph is determined. Then using unity principle and constant penetration method, the penetration value of each event is calculated and is subtracted from the observed hyetograph to determine effective hyetograph for each event. Finally, using the de-convolution method, observed unit hydrograph for each event can be determined (Chow *et al.*, 1988).

**Table 1.** Rainfall-runoff data

No. Event	Date	hrunoff (mm)	hrainfall (mm)	hrainfall / hrunoff	Ψ(mm/h)
1	8/10/1993	0.2	8.05	0.025	6.4
2	7/13/1996	1.85	10.50	0.176	1.2
3	7/04/1997	0.14	4.20	0.033	0.02
4	7/04/1992	2.63	6.35	0.415	1.51
5	04/17/1997	2.82	6.25	0.451	5.17
6	9/05/1992	0.2	11.5	0.017	8.3
7	03/27/1992	0.28	4.65	0.06	0.07
8	7/18/2004	2.74	4.95	0.553	1.11

In order to evaluate and compare the models performances, Nash-Sutcliffe criterion ( $E$ ) (Nash & Sutcliffe, 1970), Eq. (16), correlation coefficient between the observed and computed data ( $R$ ), Eq. (17), and absolute error ratio of peak flow ( $RAE_p$  (%)), Eq. (18), have been utilized in this paper, where  $Q_{i,obs}$  is the observed discharge at time  $t = i$ ;  $Q_{i,sim}$  is simulated discharge at time  $i$ ,  $Q_P$  denotes the peak flow and  $N$  is the number of observations.

$$E = 1 - \frac{\sum_{i=1}^N (Q_{i,obs} - Q_{i,sim})^2}{\sum_{i=1}^N (Q_{i,obs} - \bar{Q}_{obs})^2} \quad (16)$$

$$R = \frac{\sum_{i=1}^N ((Q_{i,obs} - \bar{Q}_{obs})(Q_{i,sim} - \bar{Q}_{sim}))}{\sqrt{\sum_{i=1}^N (Q_{i,obs} - \bar{Q}_{obs})^2 \sum_{i=1}^N (Q_{i,sim} - \bar{Q}_{sim})^2}} \quad (17)$$

$$RAE_p (\%) = \frac{|Q_{Pobs} - Q_{Psim}|}{Q_{Pobs}} \times 100 \quad (18)$$

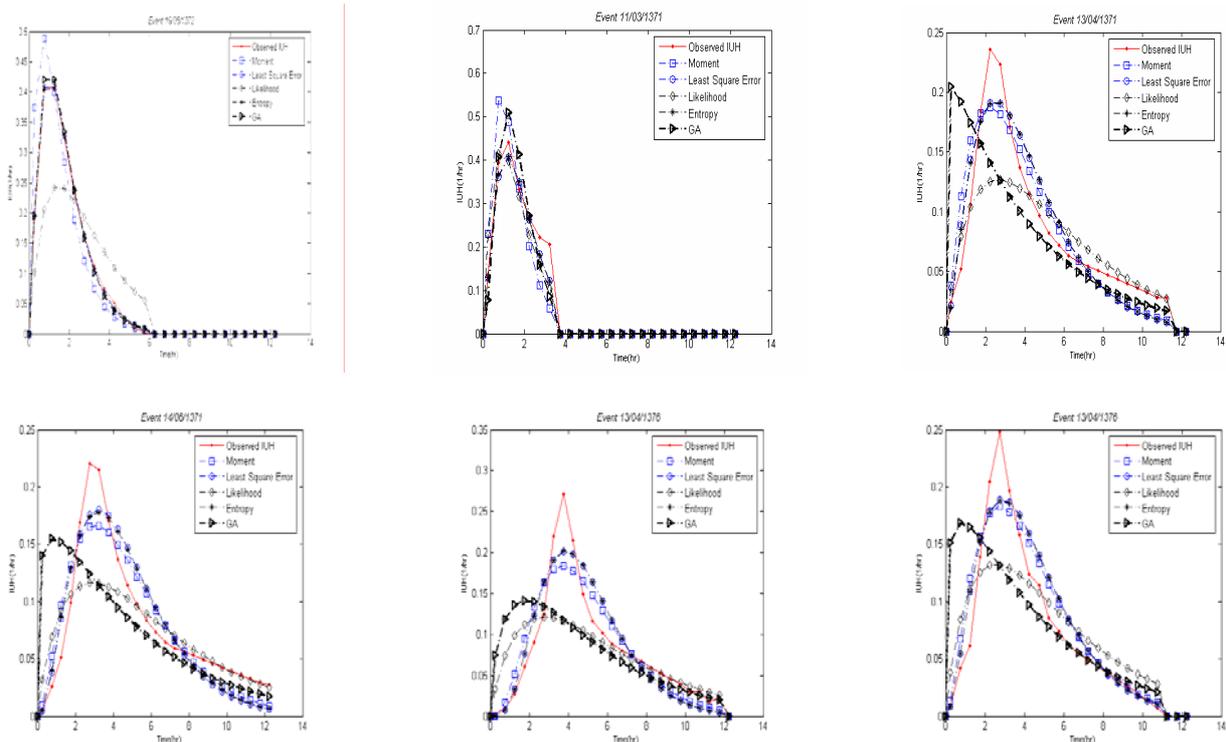
Parameters of Nash model are estimated for each event through various methods and finally the mean value of parameters for each method are listed in Table 2 for calibration and verification stages. For each event, the computed IUH against the observed IUH are illustrated in Figs 2 and 3.

Results given in Table 2 indicate that genetic algorithm and the least square error methods with highest values of correlation coefficient and Nash-Sutcliffe, lead the most accurate estimations. Furthermore, the minimum values of absolute error ratio of peak flow for these two methods indicate that peak value of calculated direct runoff very slightly varies with peak value of the observed direct runoff. As this difference decreases and even reaches to zero, more accurate regressions and consequently more accurate estimations of parameters can be obtained. Maximum entropy method and mostly maximum likelihood method because of their excessive approximations in numerical solutions attain unreliable estimations.

**Table 2.** Nash’s model parameters and efficiencies (Calibration and Verification)

Method	$n^*$	$k^*$ (h)	$E$		$R$		$RAE_p$		( $\%$ )	
			Cal.	Ver	Cal.	Ver	Cal.	Ver		
Moment	2.698	1.276	0.807	0.718	0.909	0.88	22.19	21.6		
Least Square Error	3.103	1.156	0.876	0.777	0.949	0.904	13.79	21.51		
Maximum Likelihood	1.913	2.502	0.646	0.522	0.847	0.844	38.73	47.06		
Maximum Entropy	1.802	2.752	0.38	0.48	0.678	0.834	22.54	49.27		
Genetic Algorithm	3.103	1.151	0.88	0.778	0.95	0.906	13.64	21.11		

\*The mean value of calibration step for 5 storms data



**Fig. 2** Observed and calculated unit hydrographs of calibration step.

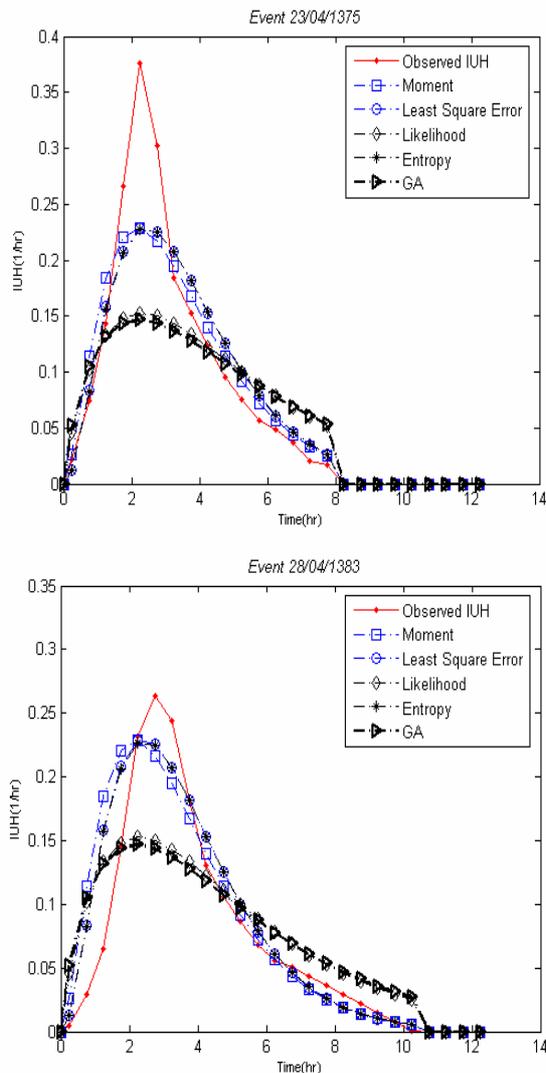


Fig. 3 Observed and calculated unit hydrographs of verification step.

Lower values of correlation coefficient and Nash-Sutcliffe criteria and higher value of absolute error ratio of peak flow for these two methods indicate unreliability of their estimations.

The mathematical framework of the maximum likelihood and maximum entropy is the same for the parameter estimation of the Nash's model and the results of the parameters estimation for these methods are approximately identical as can be seen in Table 2 (a few differences are related to the used numerical solutions errors) but when the estimated parameters are used for the runoff simulation, this low discrepancy may lead to different performances as can be seen in the obtained  $E$  and/or  $RAE_p$  for these two methods. Anyway it can be recommended to use just one of them in the real applications.

Although both least square and genetic algorithm methods have good efficiency, GA has some advantages over the other method such that it does not use function gradient and therefore, the discontinuities of the answer domain can be handled without trapping in the local minimum.

## CONCLUSIONS

Given results through this investigation have led to the following conclusions:

Approximations and numerical solutions could lead to unreliable estimations. Among the methods presented in this study, genetic algorithm gives most accurate estimations of Nash hydrologic model parameters for the case study of Ammameh.

The most unreliable estimations were attributed to the maximum entropy method. Among the classic methods the most accurate estimations are obtained by the least square method.

## REFERENCES

- Aly, A.H. & Peralta, R.C. (1999) Optimal design of aquifer clean up systems under uncertainty using a neural network and genetic algorithm. *Water Resour. Res.* **35**(8), 2523–2532.
- Cheng, C.T., Ou, C.P. & Chau, K.W. (2002) Combining a fuzzy optimal model with a genetic algorithm to solve multi-objective rainfall-runoff model calibration. *J. Hydrol.*, **268**, 72–86.
- Chow, V.T., Maidment, D.R., Mays, L.W. (1988) *Applied Hydrology*. McGraw-Hill, New York, USA.
- Cieniawski, S.E., Eheart, J.W., Ranjithan, S. (1995) Using genetic algorithms to solve a multiobjective groundwater monitoring problem. *Water Resour. Res.* **31**(2), 399–409.
- Croley, T.E. (1980) Gamma synthetic hydrographs. *J. Hydrol.* **47**, 41–52.
- Chau, K.W. (2007) Reliability and performance-based design by artificial neural network. *Adv. Eng. Soft.* **38**(3), 145–149.
- Diskin, M.H., Ince, S. & Oben-Nyarko, K. (1978) Parallel cascades model for urban watersheds. *J. Hydraul. Div. ASCE*, **104** (HY2), 261–275.
- Franchini, M. (1996) Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models. *Hydrol. Sci. J.* **41**(1), 21–40.
- Goldberg, D.E. (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley Publishing Company, New York.
- Jain, A. & Srinivasulu, S. (2004) Development of effective and efficient rainfall-runoff models using integration of deterministic real-coded genetic algorithms and artificial neural network techniques. *Water Resour. Res.* **40**, W04302.
- Jain, A., Srinivasulu, S., Bhattacharjya, R.K. (2005) Determination of an optimal unit pulse response function using real-coded genetic algorithm. *J. Hydrol.* **303**, 199–214.
- Nash, J.E. (1957) The form of the instantaneous unit hydrograph. *IASH Publication* **45**(3), 114–121.
- Nash, J.E. & Sutcliffe, J.V. (1970) River flow forecasting through conceptual models I: a discussion of principles. *J. Hydrol.* **10**, 282–290.
- Nourani, V., Monadjemi, P. & Singh, V.P. (2007) Liquid analog model for laboratory simulation of rainfall-runoff process. *J. Hydrol. Enging, ASCE* **12**(3), 246–255.
- Nourani, V. & Mano, A. (2007) Semi-distributed flood runoff model at sub continental scale for south western Iran. *Hydrol. Processes* **21**(23), 3173–3180.
- Ponce, V. M. (1980) Linear reservoirs and numerical diffusion. *J. Hydraul. Div. ASCE* **106** (HY5), 691–699.
- Rezai Pajand, M. & Bozorgnia, O. (1995) *The gamma family and derived distributions applied in hydrology*. Ferdowsi University of Mashhad Press, Mashhad, Iran.

- Santos, C.A.G., Srinivasan, V.S., Suzuki, K. & Watanabe, M. (2003) Application of an optimization technique to a physically based erosion model. *Hydrol. Processes* **17**, 989–1003. doi: [10.1002/hyp.1176](https://doi.org/10.1002/hyp.1176).
- Singh, V.P. (1988) *Hydrologic Systems*. Vol. I. Rainfall-Runoff Modeling. Prentice-Hall, New Jersey.
- Szilagyi, J. (2003) State-Space discretization of the Kalinin-Milyukov-Nash-cascade in sample-data system framework for stream flow forecasting. *J. Hydrol. Engng. ASCE*, **8**(6), 339–347.
- Wang, Q.J. (1991) The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resour. Res.* **27**(9), 2467–2471.