Rainfall-Runoff Simulation using MIKE11/NAM and ANFIS Models (A Case Study: Qleh Shahrokh Basin in Iran)

Elham Kakaei Lafdani a, Alireza Moghaddam Nia b*, Ahmad Pahlavanravi c, Azadeh Ahmadi d

a Department of Range and Watershed Management, University of Zabol, Zabol, Iran

b Associate Professor of Mountain and Arid Regions Reclamation Department, Faculty of Natural Resources, University of Tehran, Karaj, Iran

c Assistant Professor of Range and Watershed Management Department, University of Zabol, Zabol, Iran

d Assistant Professor of Civil Engineering Department, Isfahan University of Technology, Isfahan, Iran

* Corresponding author: a.moghaddamnia@ut.ac.ir

Abstract: Rainfall-Runoff modeling is considered as one of the major hydrologic processes and is essential for water resources management. In this study rainfall has been predicted using Adaptive Neural Fuzzy Inference System (ANFIS) and the best input combination has been identified using Gamma Test (GT) for the rainfall prediction. Then, runoff was simulated by a conceptual hydrological MIKE11/NAM model and the results were compared together. The study area is Qaleh Shahrokh basin located in Iran. The ability of ANFIS and MIKE11/NAM models were evaluated based on Root Mean Square Error (RMSE), correlation coefficient ($R^2$) and Efficiency Index (EI).

Keywords: Rainfall-Runoff, ANFIS, ANN, Gamma Test, MIKE11/NAM Model

1. Introduction

Several different models have been developed to simulate hydrological processes such as rainfall-runoff process, which can be divided into three “physical”, “conceptual” and “black box” models. Because of its non-linear and multi-dimensional nature, rainfall-runoff convert modeling is extremely complicated (Lipi Wattanakarn et al., 2004). The hydrological Model of NAM1 is an integrated and conceptual model of rainfall-runoff which is able to simulate surface flow, subsurface and base flow; this model has been developed by Danish Hydraulic Institute (DHI) in 1972 (DHI, 1999). In addition, artificial intelligence techniques such as Adaptive Neural Fuzzy Inference System (ANFIS) have been frequently used due to their flexibility in modeling the non-linear processes such as rainfall-runoff model. Moreover, one of the major phases in modeling using artificial intelligence techniques is identifying the best input combination of (Kakaei Lafdani et al., 2013). Various techniques are applied to do so including

1 Nedbore-Afstromings-Model (Rainfall-Runoff in Danish)
genetic algorithm, Gamma Test (GT) and etc. So many studies have been conducted using ANFIS and NAM models. Shamsdin and Hashim (2002) applied NAM model for predicting the runoff rate in Liang River located in northern part of Malaysia. The results showed that the predicted amounts by the NAM model were in accordance with the historical data appropriately and in general the results were satisfactory. Lipiwattanakarn et al. (2004) compared the performance of ANN and MIKE11-NAM models. The results showed that ANN model was more efficient in simulating the discharge peak while MIKE11/NAM was more capable in simulating the basic flow or discharge. Liu and Sun (2010) suggested a novel sensitive analysis scheme for MIKE11/NAM rainfall-runoff model which indicated the sensitivity analysis problem in a general multi-objective framework. Aldrian and Djamil (2008) used adaptive multivariate ANFIS system using relative humidity, temperature, pressure and rainfall data with a 15-minute interval for predicting daily rainfall prediction. The results showed that ANFIS model showed a higher capability in rainfall prediction using relative humidity values. Dastorani et al. (2010) used ANN and ANFIS models to predict rainfall in arid areas (Yazd Province) in Iran. Final results indicated that ANN and ANFIS were efficient tools for rainfall prediction in that area. Noori et al. (2011) investigated the role of input parameters pre-processing using Principal Component Analysis (PCA) techniques, GT and Forward Selection (FS) method to test the performance of the support vector machine (SVM) model for the discharge flow prediction.

In this study, first the best input combination of variables has been identified using GT and rainfall was predicted using ANFIS. Thereafter, the runoff was simulated through the hydrological model of MIKE11/NAM.

2. Experimental procedures

2.1. Conceptual Hydrological MIKE11/NAM model

The hydrological NAM model is an integrated and conceptual model of runoff-rainfall. NAM model simulates the rainfall-runoff process using the linkage rule between the four reservoirs which are connected together and each is the representative of different physical specifications (Figure 1).
These four reservoirs are: snow storage, surface storage, groundwater storage, root zone storage. The required basic data for NAM model are: model parameters, initial conditions, meteorological data and data for hydrometric calibration and validation of the model. In addition, NAM model has the ability of simulating the changes made by human in hydrologic cycle (e.g. irrigation and wells pumping), meanwhile time series of irrigation and using rate of groundwater aquifers will be required.

2.2. Gamma Test (GT)

GT is one of the non-linear modelling tools whereby an appropriate combination from input parameters can be investigated for modelling the output data as well as establishing a smooth model. Suppose there is a set of data as the following:

\[(x_1, \ldots, x_m, y) = (X, y)\]  

(1)

where \(X = (x_1, \ldots, x_m)\) is the input vector in the output vector’s areas of \(y\) and \(C \in R^n\). Gamma Test includes a list of \(k(1 \leq k \leq p)\) the \(k^{th}\) neighbour for each vector \(X_i\) \((1 \leq i \leq M)\). Delta function calculates the mean squared distance of the \(k^{th}\) neighbour.

\[\delta_M(k) = \frac{1}{M} \sum_{i=1}^{M} |Y_{X_i[k]} - X_i|^2\]  

(2)

where \(\cdot\) indicates the Euclidean distance and its corresponding Gamma function:

\[\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^{M} (y_{Y_{X_i[k]}} - y_i)^2\]  

(3)

where \(y_{Y_{X_i[k]}}\) is the value of \(y\) corresponding to the \(k^{th}\) neighbour of \(X_i\) in the equation 14. In order to calculate \(\Gamma\), the linear regression is fitted from \(p\) spot to values of \(\delta_M(k)\) \& \(\gamma_M(k)\).

\[\gamma = A \delta + \Gamma\]  

(4)

2.3. Adaptive Neural Fuzzy Inference System (ANFIS)

ANFIS is the ability of combining the concepts of linguistic terms of fuzzy systems along with the numerical strength of neural networks. Generally, the ANFIS structure is composed of five layers.

Layer 1(input conjunctions): Every i node in this layer is a node with following function:

\[Q_i^1 = \mu_A(x)\]  

(5)
Layer 2: Every node (conjunction) in this layer is a circle node marked with II which multiples the input signals and removes the obtained multiplication. Every output node shows the weight rule. For example:

\[ w_i = \mu_A(x) \times \mu_B(y), \quad i = 1, 2 \]  

(6)

Layer 3: Every node in this layer is a circle node marked with N. \( i^{th} \) node is calculated relative to the fire power of \( i^{th} \) rule from the sum of all weight of the rules.

\[ w_i = \frac{w_j}{w_1 + w_2}, \quad i = 1, 2 \]  

(7)

Layer 4: Every \( i \) node in this layer is a square with following equation:

\[ Q_i^4 = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i) \]  

(8)

Where \( \overline{w}_i \) is the output of the third layer and \( \{p_i, q_i, r_i\} \) are the parameters set.

Layer 5: The single node in this layer is a circle node marked with \( \Sigma \) which calculates the total output as the sum of all input signals and so forth.

\[ Q_i^5 = \text{overall output} = \sum \overline{w}_i f_i = \sum \frac{w_i f_i}{\sum w_i} \]  

(9)

2.4. Study Area

The study area is the Qaleh Shahrokh basin, located in Iran as shown in Figure 2. Its most important river is Zayandehrood River. The data used in this study were daily river discharge, daily rainfall, daily evaporation data and daily temperature. The catchment area is 1525.67 km². In this study, statistical periods of 1999-2006 and 2006-2009 were selected as the training (calibration) period and testing (verification) period of, respectively.
3. Results and discussion

In order to predict rainfall the three day, five day, and seven day moving averages of rainfall were used (from one up to seven time lags). The best input combination for rainfall prediction was identified by the Gamma test. According to the principals of the GT, the combination with the minimum gamma value and the minimum RMSE would be the best combination for modeling. The results obtained from Gamma test are shown in Table 1.

<table>
<thead>
<tr>
<th>Model Input.</th>
<th>Type of moving average</th>
<th>$\Gamma$</th>
<th>$SE$</th>
<th>Best input combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{R}_i$</td>
<td>3 days</td>
<td>0.00167</td>
<td>0.0039</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5 days</td>
<td>0.00056</td>
<td>0.00158</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>7 days</td>
<td>0.00030</td>
<td>0.000091</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{R}<em>i, \bar{R}</em>{i-1}$</td>
<td>3 days</td>
<td>0.00079</td>
<td>0.00058</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>5 days</td>
<td>0.000078</td>
<td>0.00089</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>7 days</td>
<td>0.00031</td>
<td>0.00010</td>
<td>01</td>
</tr>
<tr>
<td>$\bar{R}<em>i, \bar{R}</em>{i-1}, \bar{R}_{i-2}$</td>
<td>3 days</td>
<td>0.00087</td>
<td>0.00058</td>
<td>011</td>
</tr>
<tr>
<td></td>
<td>5 days</td>
<td>0.00019</td>
<td>0.00023</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>7 days</td>
<td>0.00024</td>
<td>0.000055</td>
<td>101</td>
</tr>
<tr>
<td>$\bar{R}<em>i, \bar{R}</em>{i-1}, \bar{R}<em>{i-2}, \bar{R}</em>{i-3}$</td>
<td>3 days</td>
<td>0.00069</td>
<td>0.00018</td>
<td>0101</td>
</tr>
<tr>
<td></td>
<td>5 days</td>
<td>0.00033</td>
<td>0.000039</td>
<td>1011</td>
</tr>
<tr>
<td></td>
<td>7 days</td>
<td>0.00024</td>
<td>0.000034</td>
<td>1001</td>
</tr>
<tr>
<td>$\bar{R}<em>i, \bar{R}</em>{i-1}, \bar{R}<em>{i-2}, \bar{R}</em>{i-3}, \bar{R}_{i-4}$</td>
<td>3 days</td>
<td>0.00059</td>
<td>0.000126</td>
<td>10101</td>
</tr>
</tbody>
</table>
According to the Gamma Test, using seven-day moving average rainfall with seven time lags with combination of (1101011) is selected as the best input combination. Small amounts of gamma show that the data with the provided combination has the possibility to achieve a better result in modeling. Following determination of the best combination from the input variables, the rainfall has been predicted using ANFIS model. In order to prediction by the use of ANFIS model, the performance of ANFIS model was investigated using different membership functions (including Gaussian, Gaussian combination, bell-shaped, two-sigmoidal, trapezoidal and triangular). In order to review the performance of ANFIS and NAM models, Root Mean Square Error (RMSE), NRMSE, Correlation coefficient ($R^2$) and EI (Efficiency Index) statistical criteria were used. These coefficients are calculated according to the following equations:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2}
\]  \hspace{1cm} (10)

\[
NRMSE = \frac{1}{\sigma} \left( \sum_{i=1}^{n} (O_i - P_i)^2 \right)^{\frac{1}{2}}
\]  \hspace{1cm} (11)

\[
R^2 = \frac{\sum_{i=1}^{n} (P_i - \overline{P})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n} (P_i - \overline{P})^2 (O_i - \overline{O})^2}}
\]  \hspace{1cm} (12)

\[
EI = \frac{\sum_{i=1}^{n} (O_i - \overline{O})^2 - \sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2}
\]  \hspace{1cm} (13)

In these equations, $O_i$ and $P_i$ are observed and simulated values, respectively, $\overline{O}$ and $\overline{P}$ are mean values for observed and simulated values and $n$ is the number of samples. Also $\sigma$ is standard deviation
The results showed that Trapezoidal membership function with lowest RMSE equal to 0.72 (mm) and highest EI equal to 0.85 and $R^2$ equal to 0.89 had a better performance than the other membership functions. Figure 3 indicates the predicted runoff using ANFIS model during the testing period.

![Observed and predicted rainfall curve during testing period using ANFIS model](image)

After predicting rainfall by ANFIS model, runoff discharge was simulated by NAM model. In the verification period of the model, runoff values were simulated once based on the observed rainfall (we call it NAM) and once based on the predicted rainfall using ANFIS model (we call it ANFIS-NAM).

During calibration period $R^2$, EI and NRMSE of NAM model were equal to 0.71, 0.7 and 0.54 $(m^3/s)$ respectively. Also, Statistical analysis of the obtained results from verification of NAM model based on observed rainfall and predicted rainfall is shown in Table 2. Given the results of Table 2, both of two models showed an appropriate performance during verification period. According to the obtained results from the runoff simulation using NAM and ANFIS-NAM in Qaleh Shahrok basin showed that NAM with lowest NRMSE (0.59 $(m^3/s)$) and highest $R^2$ (0.69) and EI (0.88) had a higher ability to simulate runoff than ANFIS-NAM model with RMSE (0.62 $(m^3/s)$), $R^2$ and EI equal to 0.65 and 0.85 respectively. The EI and $R^2$ statistics of both models for the verification period were close together.

<table>
<thead>
<tr>
<th>Period</th>
<th>NAM $R^2$</th>
<th>NRMSE $(m^3/s)$</th>
<th>EI</th>
<th>ANFIS-NAM $R^2$</th>
<th>NRMSE $(m^3/s)$</th>
<th>EI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006-2007</td>
<td>0.69</td>
<td>0.58</td>
<td>0.66</td>
<td>0.79</td>
<td>0.55</td>
<td>0.69</td>
</tr>
<tr>
<td>2007-2008</td>
<td>0.73</td>
<td>0.61</td>
<td>0.63</td>
<td>0.56</td>
<td>0.79</td>
<td>0.37</td>
</tr>
<tr>
<td>2008-2009</td>
<td>0.66</td>
<td>0.67</td>
<td>0.56</td>
<td>0.54</td>
<td>0.7</td>
<td>0.51</td>
</tr>
<tr>
<td>Verification period</td>
<td>0.69</td>
<td>0.59</td>
<td>0.88</td>
<td>0.69</td>
<td>0.62</td>
<td>0.85</td>
</tr>
</tbody>
</table>
As displayed in table (4), the model’s range of change for the $R^2$ in simulation during the calibration period varies from 0.66 up to 0.73. Based on the observed rainfall, the EI in the verification period varies between 0.56 and 0.66. This value varies between 0.37 and 0.69 based on simulation by the predicted rainfall (ANFIS-NAM model). Comparing results based on the EI during the verification period shows that the NAM model will have a much better performance during the verification period rather than the ANFIS-NAM model. The difference between the $R^2$ and EI can be negligible at times and considerable at others. It should be noted about the EI that both during the verification period the value of this coefficient is less than the value of the $R^2$ which in terms indicates a systematic error in the simulation process. The 2007-2008 simulation can be referred to as an example of such error in which even though the $R^2$ was determined to be 0.73, the EI was found to be no more than 0.63 by NAM model.

Although Qaleh Shahrokh basin is in a good condition in terms of number of meteorological stations, Number of rainfall stations in this basin are small comparing to its basin size which can affect the performance of the models. Certainly, improvement of data quality and quantity in these stations can provide more appropriate outcomes. In addition, by improving the daily data of Qaleh Shahrokh hydrometric station, simulation with different time scales can be done much more accurately.

4. Conclusions

In this study, rainfall was predicted using ANFIS and runoff simulation by using the conceptual hydrological MIKE11/NAM model in Qaleh Shahrokh basin in Iran. Rainfall prediction was done using the three-day, five-day, and seven-day moving averages and the GT was used to determine the best input combination. The results showed that both models (NAM and ANFIS) had good capabilities in simulating discharge during calibration and verification periods. Using the predicted rainfall instead of the observed rainfall caused lower efficiency in the NAM model and runoff simulation. These results consequently indicated that the efficiency of the NAM model was much more dependent on the quantity and quality of the data input model.

References


