

# A Comparative Study of SWAT, RFNN and RFNN-GA for Predicting River Runoff

Hieu N. Duong<sup>1</sup>, Hien T. Nguyen<sup>2\*</sup>, Snasel Vaclav<sup>3</sup> and Lee Sanghyuk<sup>4</sup>

<sup>1</sup>Faculty of Computer Science and Engineering, HCM University of Technology, Vietnam; dnhiu@cse.hcmut.edu.vn

<sup>2</sup>Faculty of Information Technology, Ton Duc Thang University, Vietnam; hien@tdt.edu.vn

<sup>3</sup>Department of Computer Science, VSB-Technical University of Ostrava, Czech Republic; vaclav.snasel@vsb.cz

<sup>4</sup>Xi'an Jiaotong-Liverpool University, Suzhou, China; Sanghyuk.Lee@xjtlu.edu.cn

## Abstract

**Background/Objectives:** Data-driven models such as Recurrent Fuzzy Neural Network (RFNN) have been proven to be great methods for modeling, characterizing and predicting various kinds of nonlinear hydrologic time series data such as rainfall, water quality and river runoff. In modeling and predicting river runoff, the most important advantage of data-driven models is that they do not need as much data as do physical models such as the Soil and Water Assessment Tool (SWAT). In Vietnam, most of data which are required by SWAT are not available, thus data-driven models seem to be more suitable for predicting river runoff than SWAT. The objective of this study is to investigate the performance of SWAT, RFNN and an improvement of RFNN (RFNN-GA), which is a hybrid of RFNN and Genetic Algorithm (GA) in predicting the runoff of Srepok River in Central Highland of Vietnam. **Methods/Statistical Analysis:** Coefficient of correlation ( $R^2$ ) and mean absolute relative error (MARE) are used to analysis and compare the performance of SWAT, RFNN and RFNN-GA. **Findings:** The experimental results demonstrate that RFNN and RFNN-GA give the performance better than that of SWAT and they are able to be applied to real applications. Among these methods, RFNN-GA is the most superior. **Application/Improvements:** In the terms of MARE and  $R^2$ , RFNN-GA improves RFNN 0.9% and 2.2%, respectively; and improves SWAT 27.4% and 12.5%, respectively. RFNN-GA was deployed to predict the runoff of Srepok River in Central Highland of Vietnam.

**Keywords:** Srepok, Runoff, Prediction, Recurrent Fuzzy Neural Network, Genetic Algorithm

## 1. Introduction

River runoff prediction is very important for water resource planning and management. River runoff prediction has been studied by scientists for recent decades<sup>1-8</sup>. Generally, river runoff prediction models are classified into physical and data-driven based methodology. The first approach has complex structure and it needs rather deep mathematical knowledge. In the actual applications, researchers often apply the said models in water resource modelling, especially runoff of rivers. In<sup>3,10,11</sup>, authors proposed the method that use Soil and Water Assessment Tool (SWAT) and GIS techniques to make modeling to estimate and predict water

resources. The main disadvantage of the method is that SWAT requires diverse kinds of data, ranging from climate, water resource to soil map data. As a result, using SWAT is costly and time-consuming. Moreover, in Vietnam, the environmental and natural data are not available or unreliable. Consequently, the use of SWAT for modeling river runoff is just barely acceptable<sup>10</sup>.

In contrast, data-driven models require minimal information, are easy to develop, and have been found to be accurate in hydrologic predicting applications<sup>5</sup>. Among the data-driven models, Artificial Neural Network (ANN) is capable of simulating the sophisticated relationships of many kinds of complex data, especially hydrologic data.

\*Author for correspondence

In 2007, Ibrahim Can *et al.* applied two kinds of ANN that are Feed Forward Back Propagation and Generalized Regression Neural Network to predict Karasu River runoff<sup>2</sup>. In 2009, a study by Lance E. Besaw *et al.*<sup>1</sup> demonstrated that the use of neural networks is effective in driver runoff prediction. Besaw *et al.* pointed out that predictions based on hourly data are more efficient than those that use daily data because hourly data is considered the important relationships between climate and the runoff. In another study, Ankit Chakravarti *et al.* used artificial neural networks to simulate the relationships of rainfall and runoff data that are generated by a rainfall simulator called Advanced Hydrologic Systems (AHS)<sup>8</sup>. Besides, Saman Razavi *et al.*<sup>6</sup> and Nur Athirah Ashaary *et al.*<sup>9</sup> showed that utilizing ANNs to solve the prediction problems involving to reservoirs was effective for their case studies. In order to improve the performance of ANN, several researchers have proposed hybrid methods which are composed by neural networks and theories such as Hidden Markov Model<sup>17</sup>, Particle Swarm Optimization (PSO) algorithm<sup>18</sup>, chaotic theory<sup>19, 20</sup> and Genetic Algorithm (GA)<sup>13-16</sup>.

The purpose of this study is to investigate the accuracy of three models when they are applied for predicting the Srepok River runoff in the Central Highland of Vietnam. In this study, SWAT, Recurrent Fuzzy Neural Network (RFNN), and the hybrid of RFNN and Genetic Algorithm

(RFNN-GA) are used to predict the Srepok River runoff at specific hydrologic stations. The results obtained are compared together to find out the most superior method.

The rest of this paper is organized as follows. Section 2 presents the Srepok River. In Section 2, we also introduce the challenge of predicting the Srepok runoff. Section 3 describes the three models SWAT, RFNN and RFNN-GA. In Section 4, we present the experimental results. The findings are reported in Section 4. Finally, we draw conclusion in Section 5.

## 2. Study Area

For this study, the Srepok River basin has been selected as a study region (Figure 1.). The Srepok River basin in the Central Highlands of Vietnam is about 18.200 km<sup>2</sup> area and about 406 km long. The Srepok River is located between latitudes 1° 53' to 13° 55' and longitudes 107° 30' to 108° 45'. The Srepok River has a very important role in life and production in the basin. The Srepok watershed has a plentiful lake system, districts evenly. Due to the sloped terrain, the water retention is not good and the small streams of the river run out of water in the dry season and the water level of several big lakes drops drastically. Together with the impact of climate change, the unusual change in the basin of the Srepok River that has been observed recently has posed a problem for the



Figure 1. Location of study area.

security of Vietnam’s water resources. After monitoring the Srepok runoff from 1990 to 2011, we realize that the Srepok runoff has been quite erratic and unpredictable. Naturally, the Srepok runoff varies seasonally; it is low in the dry season and high in the wet season. But in some years, the Srepok runoff has decreased suddenly in the dry season or increased suddenly in the wet season; these changes have a direct impact on the lives of residents of the Srepok basin. The challenge is how to model and predict the Srepok runoff for the benefit of water resource managers and the basin population.

### 3. Methodology

#### 3.1 SWAT

SWAT is a hydrologic quality model developed by United States Department of Agriculture’s Agricultural Research Service (USDA-ARS). The objective of the model is to predict the impact of management of water, sediment and agricultural chemical yields in a large basin. SWAT is a regression model that simulates the relationship between input and predicted parameters. For predicting river runoff, the input parameters are quite complex including several major hydrologic components such as weather, erosion map, sedimentation map, soil temperature, plant map and pesticide use<sup>12</sup>. SWAT model uses Nash Sutcliffe Index (NSI) (Equation 7) value and coefficient of correlation ( $R^2$ ) (Equation 8) to assess the quality of the simulating model.

The simulating quality of SWAT is assessed with four levels<sup>20</sup>:  $0.75 \leq NSI \leq 1$ : very good;  $0.65 \leq NSI \leq 0.75$ : good;  $0.5 \leq NSI \leq 0.65$ : satisfaction;  $NSI \leq 0.5$ : dissatisfaction, the factor of the model needs to consider clearly.

As mentioned before, the input data is very important when applying SWAT to simulate river runoff. For modeling the Srepok runoff, we must first gather the input data of the Srepok basin. Then, the input data of the levels from basins to sub-basins are arranged. The input data consist of:

1. Spatial data: Digital Elevation Model (DEM), soil map, land use map, the Srepok River map.
2. Climate data: average of temperature, maximum of temperature, minimum of temperature, average humidity, minimum of humid degree, rain quantity, evaporation per day and the number of daylight hours.

We process these input data by ArcGis software. The SWAT model is then applied to assess the impact of land use to the Srepok runoff. Finally, we determine the relevant SWAT model to simulate the Srepok runoff.

#### 3.2 RFNN

Fuzzy neural networks have been applied in numerous fields and RFNN is a well-known fuzzy neural network. We re-implemented the proposed RFNN in<sup>21</sup>. Figure 2 shows the structure of RFNN including four layers. Let  $u_i^{(k)}$  and  $O_i^{(k)}$  be the input and the output of the node  $i^{th}$  in the layer  $k$ , respectively. The process of RFNN is presented as follows.

**Layer 1:** This is the input layer that has  $N$  nodes, each of which corresponds with a parameter. In our data, input could be parameters of the climate data including average of temperature, rain quantity, evaporation per day, average humidity, and the number of daylight hours.

$$O_i^{(1)} = u_i^{(1)} = x_i(t), \text{ where } i = 1 \div N. \quad (1)$$

**Layer 2:** This is the membership layer. Nodes in this layers will be converted to the crisp data in fuzzy data by applying membership functions such as Gauss function. The number of neural nodes in this layer is  $N \times M$  where  $M$  is the number of fuzzy rules. Every node has three parameters:  $m_{ij}$ ,  $\sigma_{ij}$  and  $\theta_{ij}$ .

$$O_{ij}^{(2)} = \exp \left[ -\frac{(u_{ij}^{(2)} - m_{ij})^2}{(\sigma_{ij})} \right], \text{ where } i = 1 \div N \text{ and } j = 1 \div M. \quad (2)$$

In Equation 2,  $m_{ij}$  and  $\sigma_{ij}$  are the center and the variance of Gauss distribution function.

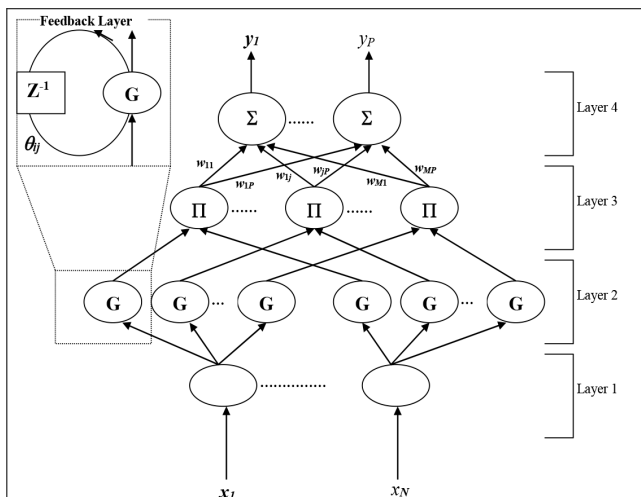


Figure 2. RFNN structure<sup>21</sup>.

$$u_{ij}^{(2)}(t) = O_i^{(1)} + \theta_{ij} O_{ij}^{(2)}(t-1), \text{ where } i = 1 \div N \text{ and } j = 1 \div M. \quad (3)$$

In Equation 3,  $\theta_{ij}$  denotes the weight of a recurrent node.

We see that the input of the nodes in this layer has the factor  $O_j^{(2)}(t-1)$ . This factor denotes the remaining information of the previous learning step. Therefore, after replacing  $u_{ij}^{(2)}$  in Equation 1 by Equation 2, we arrive at Equation 3.

$$O_{ij}^{(2)} = \exp \left[ - \frac{[O_i^{(1)} + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right] \\ = \exp \left[ - \frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right], \quad (4)$$

where  $i = 1 \div N$  and  $j = 1 \div M$ .

**Layer 3:** This is the layer of fuzzy rules. Each node in this layer conforms to a fuzzy rule. Connecting Layer 3 and Layer 4 presents a fuzzy conclusion. Each node in this layer corresponds with an AND expression. Each AND expression is defined as follows.

$$O_j^{(3)} = \prod_{i=1}^N O_{ij}^{(2)} \\ = \prod_{i=1}^N \exp \left[ - \frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right], \quad (5)$$

where  $j = 1 \div M$ .

**Layer 4:** This is the output layer including  $P$  nodes. In our model,  $P$  will be set to 1; this is the river runoff value. Nodes of this layer are responsible for converting fuzzy to crisp.

$$y_k = O_k^{(4)} = \sum_{j=1}^M u_{jk}^{(4)} w_{jk} = \sum_{j=1}^M O_j^{(3)} w_{jk} \\ = \sum_{j=1}^M w_{jk} \prod_{i=1}^N \exp \left[ - \frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right], \quad (6)$$

where  $k = 1 \div P$ .

After defining process of RFNN and the detailed operation of every layer, to train RFNN, we use a Back-Propagation (BP) algorithm that was first published by Werbos in 1974<sup>23</sup>. In our study, we improve BP by applying the momentum technique<sup>22</sup>; the pseudo-code of BP algorithm is as Algorithm 1.

### Algorithm 1: Pseudo-code of Back-Propagation algorithm

**Input:** coefficients of RFNN structure, training set  $D$ .

**Output:** RFNN satisfies one of terminating conditions

1. **While** terminating conditions are not satisfied **do**

2. **For each** training tuple  $X_i$  in training set  $D$  **do**

3. **For each** input layer unit  $i$  **do**

4.  $O_i^{(1)} = u_i^{(1)} = x_i(t)$

5. **For each** membership layer unit  $ij$  **do**

6.  $O_{ij}^{(2)} = \exp \left[ - \frac{(u_{ij}^{(2)} - m_{ij})^2}{(\sigma_{ij})^2} \right]$

7. **For each** layer of fuzzy rules unit  $j$  **do**

8.  $O_j^{(3)} = \prod_{i=1}^N O_{ij}^{(2)} = \prod_{i=1}^N \exp \left[ - \frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right]$

9. **For each** output layer unit  $k$  **do**

10.  $y_k = O_k^{(4)} = \sum_{j=1}^M u_{jk}^{(4)} w_{jk} = \sum_{j=1}^M O_j^{(3)} w_{jk}$

11. **For each** output layer unit  $k$  **do**

12.  $e(t)_k = y_k^{(d)}(t) - y_k(t)$

//  $y_k^{(d)}(t)$  is the real river runoff and  $y_k^{(t)} = O_k^{(4)}(t)$ . The target of the BP algorithm is how to minimize the sum

square error (SSE):  $E = \frac{1}{2} \sum_{k=1}^P (y_k^{(d)}(t) - y_k(t))^2 = \frac{1}{2} \sum_{k=1}^P (y_k^{(d)}(t) - O_k^{(4)}(t))^2$

```

// updating parameters of RFNN by gradient descent method.
//  $\eta$  is learning rate and  $\beta$  is momentum.
4   For each center of membership function  $m_{ij}$  do
15.      $m_{ij}(t+1) = m_{ij}(t) - \eta \frac{\partial E}{\partial m_{ij}} + \beta \times \Delta m_{ij}(t-1)$ 
16.   For each variance of membership function  $\sigma_{ij}$  do
17.      $\sigma_{ij}(t+1) = \sigma_{ij}(t) - \eta \frac{\partial E}{\partial \sigma_{ij}} + \beta \times \Delta \sigma_{ij}(t-1)$ 
18.   For each connection weight  $w_{jk}$  do
19.      $w_{jk}(t+1) = w_{jk}(t) - \eta \frac{\partial E}{\partial w_{jk}} + \beta \times \Delta w_{jk}(t-1)$ 
20.   For each recurrent weight  $\theta_{ij}$  do
21.      $\theta_{ij}(t+1) = \theta_{ij}(t) - \eta \frac{\partial E}{\partial \theta_{ij}} + \beta \times \Delta \theta_{ij}(t-1)$ 
22.   End While

```

### 3.3 Hybrid of RFNN and GA

The back-propagation algorithm has a big disadvantage that the training process usually falls into local minima. Although there is an improvement of momentum technique, it is trapped by local minima. One famous solution to this problem is to combine BP algorithm and an evolutionary algorithm such as Genetic Algorithm<sup>15</sup>. Moreover, ANNs are black boxes for end users and thus it is hard for them to find the most suitable combination of ANN coefficients. Typically, the end users must rely on their

experience and run the model several times with many different combinations of coefficients. While applying hybrids of BP algorithm and Genetic Algorithm, we can compensate for this disadvantage<sup>13-16</sup>. However, one of the drawbacks of evolutionary algorithms is their running time. Due to the searching strategy of the evolutionary algorithms which is based on stochastic exploration, the running time is very high. In this study, we try to predict monthly river runoff for the long-term, so the running time is not as important as if it is being compared with performance criterion.

#### Algorithm 2: Pseudo-code of training phase of RFNN-GA

```

Input: coefficients of RFNN individual structure, coefficients of back-propagation and genetic algorithm.
Output: the best RFNN individual satisfy one of criteria
1. Initialize the generation  $G_0$  containing NP RFNN individuals. Connection weights of every RFNN individual are random in range [0, 1]
2. While terminating conditions are not satisfied do
3.   For each RFNN individual  $i^{\text{th}}$  do
4.     Training every RFNN by back-propagation algorithm
5.     If terminating conditions are satisfied then
6.       Break out For loop
7.     End if
8.   End For
9.   If terminating conditions are not satisfied then
10.    //Create the next generation  $G_i$  from  $G_{i-1}$  by applying evolutionary operators
11.    Selection
12.    Crossover
13.    Mutation
13.   End If
14. End While

```

A typical Genetic Algorithm consists of three stages: 1) Initial population generation: Genetic Algorithm generates a set of chromosomes (individuals) called the first generation; 2) Computing the fitness of every individual and 3) construction of new generation in which Genetic Algorithm establishes the next generation by performing three evolutionary operators: selection, crossover, and mutation. Genetic Algorithm coefficients are population sizes, mating and mutation rates, and the numbers of generations.

In order to combine GA with RFNN, three questions must be answered: 1) How to encode a RFNN individual as a chromosome; 2) How to execute evolutionary operators such as selection, crossover, and mutation between two next generations; and 3) What fitness function is chosen. In our study, we employ a binary encoding algorithm called GENITOR to encode a RFNN individual as a chromosome<sup>24</sup>. This method is very popular because it is quite easy to understand and easy to answer the two first questions. For fitness function, we use the Sum Square Error (SSE). The training process of RFNN-GA is presented as Algorithm 2. In the Algorithm, we utilize the strength of BP that is able to improve the quality of each individual before proceeding the assessment and evolutionary operators on all individuals. *The idea of RFNN-GA is inspired by the nature of human society in which people should be trained (about education, physique, spirit, etc) to become better ones and to be able to produce better children.* Whereas the main task of GA in the hybrid method is to expand the search space and do not miss any potential areas of optima in the search space.

### 4. Experimental Results

In the Srepok River basin, there are several hydrologic stations that operate in the same way. In our study, we use data from the specific station called BUON DON. We gather 22 years of data (1990-2011) of the Srepok

River including daily climate and runoff data. We stored the data collected each day in a record, each of which consists of nine fields capturing information of that day such as average of temperature, maximum of temperature, minimum of temperature, average humidity, minimum of humid degree, rain quantity, evaporation per day, the number of daylight hours, and runoff. In total, we collected 8030 records of climate and runoff data. The data are used to make the experimental results of three models.

In our study, the performance of three methods is assessed by using three standard statistical performance evaluation criteria. The statistical measures considered are coefficient of correlation ( $R^2$ ), mean absolute relative error (MARE) and Nash Sutcliffe Index (NSI) as follows.

$$NSI = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}, \tag{7}$$

$$R^2 = \left[ \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right]^2, \tag{8}$$

$$MARE = \frac{1}{n} \sum_{i=1}^n \frac{|O_i - P_i|}{O_i}. \tag{9}$$

In Equations 7, 8 and 9,  $O_i$  is the observed runoff at time  $i$ ;  $\bar{O}$  is the average of observed runoff;  $P_i$  is the predicted runoff at time  $i$ ;  $\bar{P}$  is the average of predicted runoff and  $n$  is the number of observed data.

According to SWAT features, we have to gather some kind of extra-data such as soil data and map data, and preprocess these data by ArcGis software before simulating the Srepok runoff at the BUON DON station. Then we calibrate and validate the SWAT model. Four parameters

**Table 1.** SWAT sensitive parameters and calibrated values

Parameters	Description of parameters	Calibrated values		
		Fitted value	Min value	Max value
CN2	Initial SCS curve number II value	-0.17	- 0.20	0.20
ALPHA_BF	Base Flow Alpha factor	0.17	0.00	1.00
GW_DELAY	Groundwater delay	160.20	30.00	450.00
GWQMN	Threshold water depth in the shallow aquifer for flow	1.26	0.00	2.00

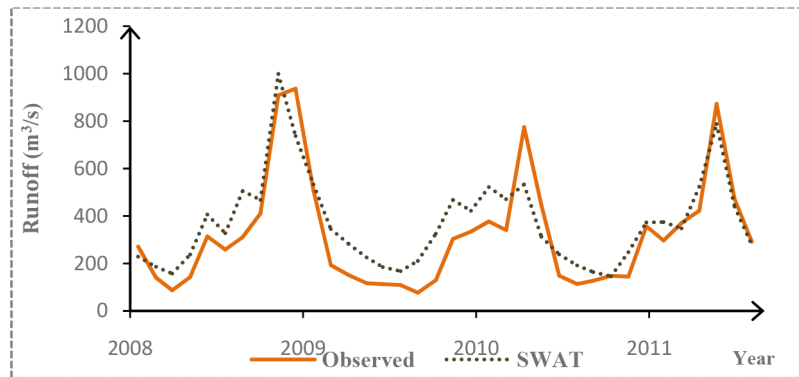
are chosen to calibrate the model: Curve Number (CN2), Base flow Alpha factor (ALPHA-BF), Groundwater Delay (GW-DELAY) and Threshold water depth in the shallow aquifer for flow (GWQMN). The result of calibration is shown in Table 1. After that, we use the calibrated result to run the SWAT model again. Consequently, we obtain higher values of NSI. Table 2 shows the calibration of SWAT model in 2004-2008. The fit of the simulated and observed runoff is acceptable because NSI is 0.68 and  $R^2$  is 0.75. Finally, we use the parameters obtained from the calibration to validate the model. In the result, the NSI

value reach 0.77 and  $R^2$  is 0.827. Figure 3 compares the simulated and observed runoff whereas Figure 4 presents the degree of correlation between the simulated and observed runoff in the validation phase. If we use the mean absolute relative error (MARE) to assess the model, the MARE of SWAT is quite large, approximately 0.401. Therefore, the results show that the SWAT model is just barely acceptable to simulate the Srepok runoff.

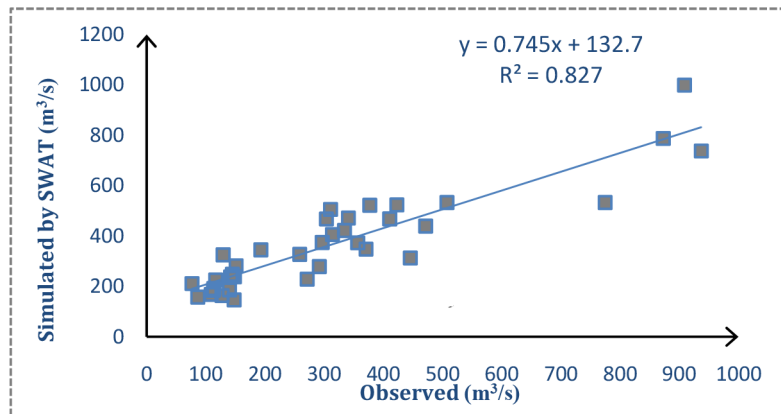
While making experimental results of RFNN and RFNN-GA, we also use the data of BUON DON Station in 1990- 2008 for training and in 2009-2011 for testing. In 4, we did not highlight temporal features so the result was not good. In this study, we analyze temporal features and consequently the performance of the model improved remarkably. In addition, we also prune some redundant attributes such as maximum of temperature, minimum of temperature, minimum of humid degree. Table 3, 4 show us the structure and performance of RFNN and RFNN-GA. Figure 5 and Figure 7 present the predicted runoff and the observed runoff of RFNN and RFNN-GA, respectively, in the testing phase. Figure 6 and Figure 8

**Table 2.** Model performance for the simulation of the Srepok runoff

Periods	Time steps	Values	
		$R^2$	NSI
Before calibration	Monthly	0.70	0.41
Calibration (2004-2008)	Monthly	0.75	0.68
Validation (2009-2011)	Monthly	0.82	0.77



**Figure 3.** Observed runoff and simulated runoff after validation by SWAT.



**Figure 4.** The degree of correlation between observed runoff and simulated runoff by SWAT.

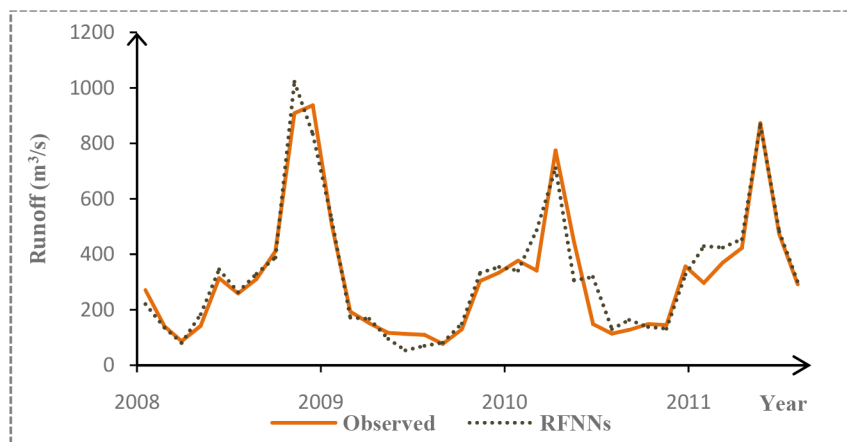
**Table 3.** Structure and performance of RFNN during training and testing phases

Fuzzy Rules	Epochs	MARE of training phases	Testing phases	
			MARE	R <sup>2</sup>
5	100.000	0.1173	0.1326	0.9267
10	100.000	0.1245	0.1774	0.9126
15	100.000	0.1070	0.1195	0.9392
20	100.000	0.1240	0.1689	0.9159
25	100.000	0.1230	0.1296	0.9320
30	100.000	0.1274	0.1345	0.9296
40	100.000	0.1312	0.1392	0.9291
Average		0.1245	0.1359	0.9305

present the degree of correlation between the observed runoff and predicted runoff of RFNN and RFNN-GA, respectively, in the testing phase. In testing phase, MARE of RFNN is about 0.1359 and MARE of RFNN-GA is about 0.1262 whereas MARE of SWAT is about 0.401. Moreover, R<sup>2</sup> of SWAT, RFNN and RFNN-GA are 0.82, 0.9305 and 0.9528, respectively. Therefore we can conclude that RFNN and RFNN-GA are superior to SWAT and RFNN-GA outperforms RFNN. However, SWAT is a physical model based on climate, soil, land use and water resource data, so if we have enough and exact data, SWAT is able to simulate and predict well river runoff. In Vietnam, because we lack present and future data for the SWAT model, it is hard to predict the Srepok runoff with SWAT as expected, but with RFNN or RFNN-GA, it is simpler.

**Table 4.** Structure and performance of RFNN-GA during training and testing phases

GA coefficients				BP coefficients		Results		
Populations	Generations	Crossover Probability	Mutation Probability	Epochs	Fuzzy rules	MARE of training phases	MARE of testing phases	R <sup>2</sup> of testing phases
100	50	0.4	0.1	10.000	15	0.1195	0.1254	0.9491
100	50	0.5	0.1	10.000	15	0.1234	0.1293	0.9560
100	50	0.6	0.1	10.000	15	0.1162	0.1244	0.9499
100	50	0.4	0.2	10.000	15	0.1196	0.1301	0.9577
100	50	0.5	0.2	10.000	15	0.1121	0.1219	0.9498
Average						<b>0.1158</b>	<b>0.1262</b>	0.9528



**Figure 5.** Observed runoff and average values of predicted runoff by RFNN in testing phases.



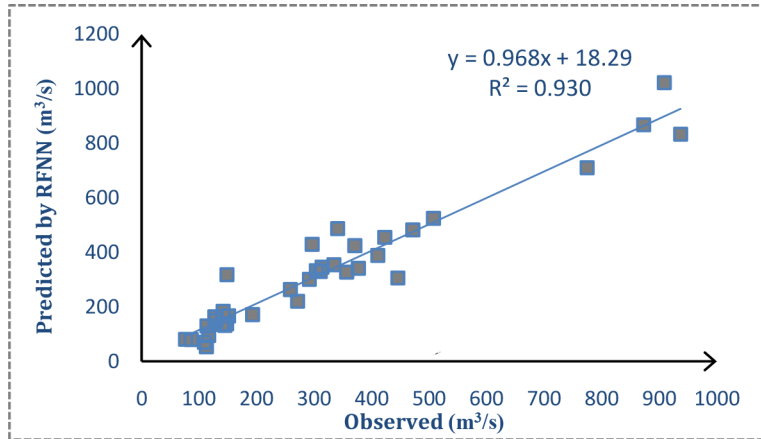


Figure 6. The average correlation between observed runoff and predicted runoff by RFNN in testing phases.

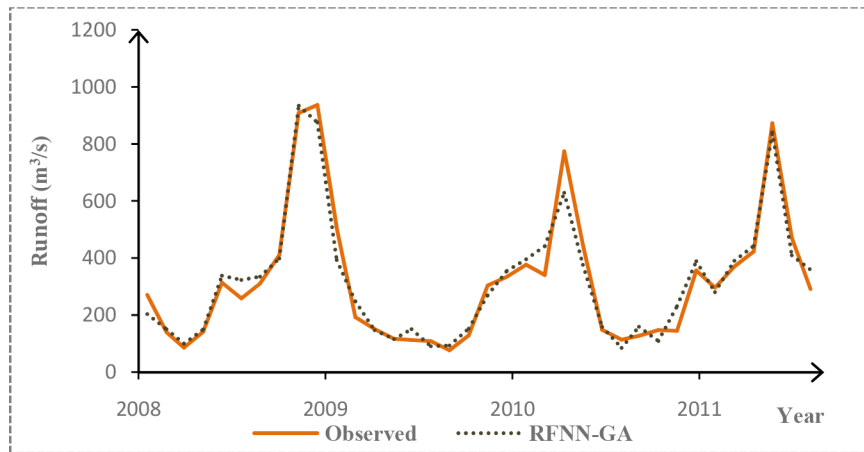


Figure 7. Observed runoff and average values of predicted runoff by RFNN-GA in testing phases.

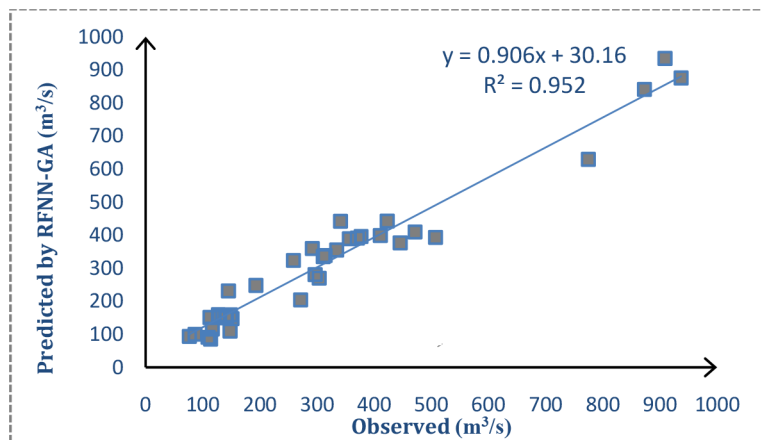


Figure 8. The average correlation between observed runoff and predicted runoff by RFNN-GA in testing phases.

## 5. Conclusion

In Vietnam, the Srepok River holds a central role in people's life and in production around the basin area. Therefore, if there are some methods of precisely predicting the runoff of the Srepok River it would be tremendously helpful for resource managers and for the public. In this paper, we compare RFNN, the hybrid of RFNN and GA (RFNN-GA) to SWAT. Generally, data-driven models are more suitable than physical-based models for dealing with problems of runoff prediction which are lacking in calibration data. The experimental results definitely point out that RFNN and RFNN-GA can predict exactly and outperform SWAT. While comparing performance of RFNN to RFNN-GA, we conclude that RFNN-GA is superior to RFNN.

## 6. References

- Besaw LE, Rizzo DM, Bierman PR, Hackett W. Advances in ungauged stream flow prediction using artificial neural networks. *Journal of Hydrology*. 2010; 386:27-37.
- Can I, Yerdelen C, Kahya E. Stochastic modeling of Karasu River (Turkey) using the methods of artificial neural networks. *Hydrology Days*. 2007; p. 138-44.
- Demirela MC, Venancio A, Kahya E. Flow forecast by SWAT model and ANN in Pracana Basin Portugal. *Journal of Advances in Engineering Software*. 2009; 40:467-73.
- Hieu DN, Quyen NTN, Long BT, Hien NT, Vaclav S. Applying recurrent fuzzy neural network to predict the runoff of Srepok River. Vietnam: Proceedings of 13th International Conference on Computer Information System and Industrial Management Applications. 2014; p. 55-66.
- He Z, Wen X, Liu H, Du J. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal of Hydrology*. 2014; 509; 379-86.
- Razavi S, Araghinejad S. Reservoir inflow modeling using temporal neural networks with forgetting factor approach. *International Journal of Water Resource Management*. 2009; 23:39-55.
- Kar S, Dasb S, Ghoshb PK. Applications of neuro fuzzy systems: A brief review and future outline. *Applied Soft Computing*. 2014; 15:243-59.
- Ankit C, Nitin J, Himanshu P. Rainfall Runoff Analysis Using Artificial Neural Network. *Indian Journal of Science and Technology*. 2015; 8(14).
- Nur AA, Wan HWI, Ku RK-M. Neural Network Application in the Change of Reservoir Water Level Stage Forecasting. *Indian Journal of Science and Technology*. 2015; 8(13).
- Quyen NTN, Liem ND, Loi NK. Effect of land use change on water discharge in Srepok watershed, Central Highland, Vietnam. *International Soil and Water Conservation Research*. 2014; 2:74-86.
- Panhalkar S. Hydrological modeling using SWAT model and geoinformatic techniques. *The Egyptian Journal of Remote Sensing and Space Science*. 2014; 17:197-207.
- Santhi C, Arnold JG, Williams JR, Dugas WA, Srinivasan R, Hauck LM: Validation of the SWAT model on the large river basin with point and nonpoint sources. *Journal of American Water Resources Association*. 2001; 37:1169-1188.
- Jia W, Zhao D, Shen T, Su C, Hu C, Zhao Y. A New Optimized GA-RBF Neural Network Algorithm. *Computational Intelligence and Neuroscience*. 2014; 6 pages.
- Sarangi PP, Sahu A, Panda M. A hybrid differential evolution and back-propagation algorithm for feed forward neural network training. *Journal of American Water Resources Association*. 2013; 84(14).
- Harpham C, Dawson CW, Brown MR. A review of genetic algorithms applied to training radial basis function networks. *Neural Computing and Applications*. 2004; 13:193-201.
- Chang WY. An RBF neural network combined with OLS algorithm and genetic algorithm for short-term wind power forecasting. *Journal of Applied Mathematics*. 2013; 9 pages.
- Bhardwaj S, *et al*. Chaotic time series prediction using combination of hidden Markov model and neural nets. Poland: Proceedings of the 9th International Conference on Computer Information Systems and Industrial Management Applications. 2010; p. 585-89.
- Du B, Xu W, Song B, Dinh Q, Chu SC. Prediction of chaotic time series of RBF neural network based on particle swarm optimization. *Intelligent Data Analysis and its Applications*. 2014; 298:489-97.
- Ma QL, Zheng AL, Peng H, Zhong TW, Qin JW. Multi-step-prediction of chaotic time series based on co-evolutionary recurrent neural network. *Chinese Physics B*. 2008; 17:1674-1056.
- Todorov Y, Terziyska M. Modeling of chaotic time series by interval type-2 neo-fuzzy neural network. *Artificial Neural Networks and Machine Learning*. 2014; 8681:643-50.
- Lee CH, Teng CC. Identification and control of dynamic systems using recurrent fuzzy neural networks. *IEEE Transactions on Fuzzy Systems*. 2000; 8(4):349-66.
- Polyak BT. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*. 1964; 4(5):1-17.

23. Werbos P. Cambridge, MA: Harvard University: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD. Thesis. 1974 Aug.
24. Whitley D. The GENITOR Algorithm and Selection Pressure: Why Rank Based Allocation of Reproductive Trials is Best. Proceedings of the 3<sup>rd</sup> International Conference on Genetic Algorithms. 1989; 116-21.