

# Impact of climate change on runoff in Lake Urmia basin, Iran

Hadi Sanikhani<sup>1</sup> · Ozgur Kisi<sup>2</sup> · Babak Amirataee<sup>3</sup>

Received: 16 August 2016 / Accepted: 27 February 2017 / Published online: 24 March 2017  
© Springer-Verlag Wien 2017

**Abstract** Investigation of the impact of climate change on water resources is very necessary in dry and arid regions. In the first part of this paper, the climate model Long Ashton Research Station Weather Generator (LARS-WG) was used for downscaling climate data including rainfall, solar radiation, and minimum and maximum temperatures. Two different case studies including Aji-Chay and Mahabad-Chay River basins as sub-basins of Lake Urmia in the northwest part of Iran were considered. The results indicated that the LARS-WG successfully downscaled the climatic variables. By application of different emission scenarios (i.e., A1B, A2, and B1), an increasing trend in rainfall and a decreasing trend in temperature were predicted for both the basins over future time periods. In the second part of this paper, gene expression programming (GEP) was applied for simulating runoff of the basins in the future time periods including 2020, 2055, and 2090. The input combination including rainfall, solar radiation, and minimum and maximum temperatures in current and prior time was selected as the best input combination with highest predictive power for runoff prediction. The results showed that the peak discharge will decrease by 50 and 55.9% in 2090 comparing with the baseline period for the Aji-Chay and Mahabad-Chay basins, respectively. The results

indicated that the sustainable adaptation strategies are necessary for these basins for protection of water resources in future.

**Keywords** Climate change · LARS-WG · GEP · Aji-Chay · Mahabad-Chay

## 1 Introduction

Changing climate in different regions is due to increasing concentration of greenhouse gases and changing general circulation (Kaleris et al. 2001). Water resources have been facing severe challenges during the last decades all over the world. This condition is intensified by decreasing trend of precipitation and increasing trend of temperatures. It is expected that most of the regions in the world will experience a negative impact of climate change on water resources and freshwater ecosystems (Intergovernmental Panel on Climate Change, IPCC 2007). However, the characteristics of the impact may vary from one region to another. The prediction was done by the most general circulation models (GCMs) that indicated an increase in the number and occurrence of huge climatic variation (i.e., precipitation) in different regions over the world. Such climatic variations can affect the situation of water resources in the near future.

Various hydrological models have been applied for quantifying the impact of hydrologic climate change by using GCMs information. However, the spatial resolution between the output of GCMs and the required data for hydrological models is not sufficient, and it has always been a challenge. For addressing this type of problem, regional climate models (RCMs) and statistical downscaling methods (SDSMs) have been developed. The main problem of RCMs is that

---

✉ Hadi Sanikhani  
hsanikhani12@gmail.com

<sup>1</sup> Water Engineering Department, Faculty of Agriculture, University of Kurdistan, Sanandaj, Iran

<sup>2</sup> Center for Interdisciplinary Research, International Black Sea University, Tbilisi, Georgia

<sup>3</sup> Water Engineering Department, Urmia University, Urmia, Iran

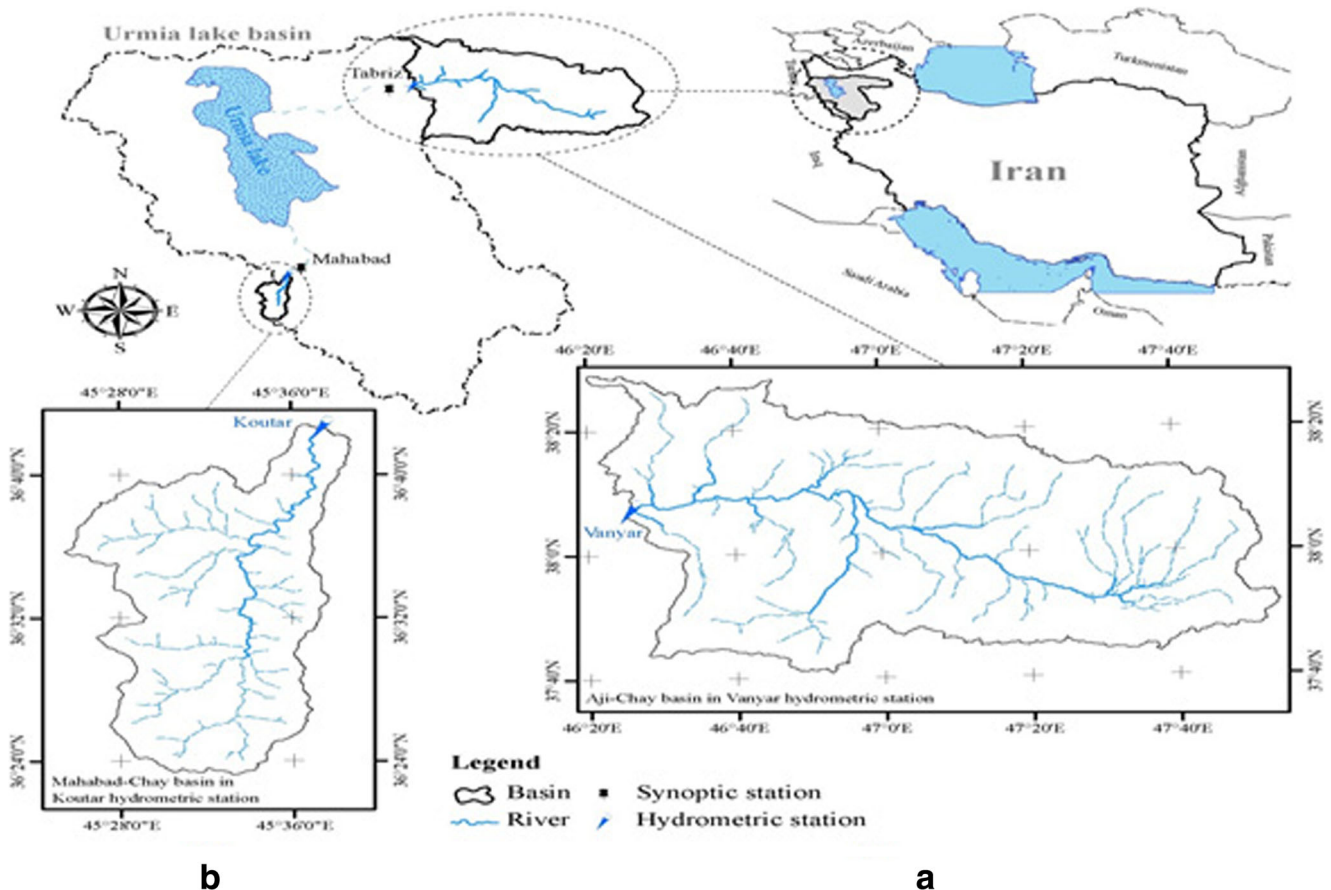
they are computationally expensive. So, they are applicable in some special regions. Furthermore, application of RCMs in a small watershed is problematic, and some scenarios defined at watershed scales are necessary.

The SDSM was developed to overcome the aforementioned problems. SDSM plays the role of an intermediate between GCMs or RCMs with other variables (represent the small scale such watersheds). SDSM is computationally cheap and simple runoff model (Srikanthan and McMahon 2001). Long Ashton Research Station Weather Generator (LARS-WG) is classified in the category of SDSM. Some successful applications of LARS-WG were reported by Semenov et al. (1998), Khan et al. (2006), Zarghami et al. (2011), Hassan et al. (2014), and Parajuli et al. (2016).

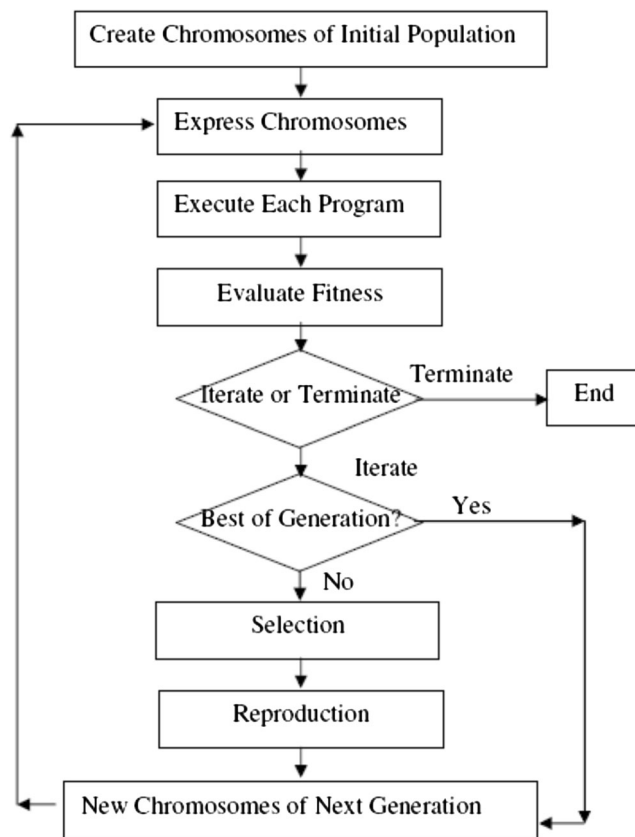
In recent years, many researchers focused on the impact of climate change on the runoff by using RCMs and SDSM in different scales. Chiew et al. (1995) investigated the impact of climate change on runoff and soil moisture of Australian basins. Matondo et al. (2004) evaluated the impact of climate change on hydrology and water resources in Swaziland. They used 11 GCMs for assessing the climate changes, and three of them were found to simulate the observed precipitation very well in the region. Jiang et al. (2007) compared the hydrological impact of climate change simulated by different hydrological models in

the Dongjiang basin, China. Zheng et al. (2007) utilized an RCM for modeling water resources changes. Fujihara et al. (2008) applied dynamically downscaled data for hydrologic modeling and assessing the impact of climate change on water resources of a basin in Turkey. Fatichi et al. (2011) used a weather generator for modeling of future climate scenarios. Chang et al. (2014) investigated the impact of climate change and human activities on runoff in Weihe River Basin in China. Nkomozepe and Chung (2014) studied the impacts and uncertainty associated with climate change on water resources in the Geumho River Basin, Republic of Korea. The results indicated that climate change will lead to lower water resources levels in future compared to current situation. Lu et al. (2016) applied an integrated statistical and data driven framework for flood risk analysis under climate change. The results revealed that the maximum monthly and annual streamflows would both increase at the middle and end of this century. In another study, Vallam and Qin (2016) presented a combined LARS-WG and multi-site rainfall simulator RainSim approach to investigate flow regimes under future conditions in the Kootenay Watershed, Canada.

In the recent years, artificial intelligent (AI) paradigms including artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS), support vector machine



**Fig. 1** Location of the study area in Iran. **a** Overall position. **b** In respect to Lake Urmia



**Fig. 2** The flowchart of GEP (see Aytok and Kisi 2008)

(SVM), and gene expression programming (GEP) methods have been used in water science and hydrology (Savic et al. 1999; Firat and Gungor 2007; Wang et al. 2009; Nourani et al. 2014; Mehdizadeh et al. 2017). GEP, as a branch of genetic

programming (GP), involves computer programs of different sizes and shapes encoded in linear chromosomes of fixed lengths. In the structure of GEP, the chromosomes are simple entities and the expression trees are exclusively the expression of their respective chromosomes (Ferreira 2006). Among AI, the GEP has the capability in presentation of relationship between dependent variable (target) and independent variables (inputs). This advantage is very important in modeling of various phenomena. Describing literature review related to GEP is beyond the scope of this study and only some studies are presented here.

Aytok and Alp (2008) utilized GEP for modeling rainfall-runoff process. Guven and Aytok (2009) introduced GEP for determining the stage-discharge relationship. Hashmi et al. (2011) applied GEP for downscaling of watershed precipitation. The results showed that GEP could be used in the process of precipitation downscaling. Kisi et al. (2013) compared the efficiency of ANNs, ANFIS, and GEP in rainfall-runoff modeling. Hashmi and Shamseldin (2014) used GEP in order to regionalize the flow duration curve. Zorn and Shamseldin (2015) estimated the peak of flood for Auckland region using GEP model. Zounemat-Kermani et al. (2017) applied GEP and decision tree methods for estimation of air demand in bottom outlet conduits of dams.

In this paper, the authors try to investigate the impact of climate change on runoff in two selected case studies from the northwest part of Iran. In the first part of the study, LARS-WG model was applied in order to downscale the climatic variables including temperature, rainfall, and solar radiation. In the second part of the study, the GEP model was used for prediction of runoff based on different input combinations extracted from the climate model.

## 2 Material and methods

The framework of this study includes two major steps. In the first step, the GCM data for two case studies under different emission scenarios downscaled. In the second step, the impacts of climate change on the runoff using GEP model were assessed. In the following sections, the study area, data, and methods are illustrated.

### 2.1 Case studies and used data

For investigation on the impact of climate change on runoff, two case studies including Aji-Chay and Mahabad-Chay River basins are selected, which are located in the northwest part of Iran. Figure 1 represents the geographical location of the Aji-Chay River and Mahabad-Chay River basins and the gauge stations.

Aji-Chay River basin, as one of the main sub-basin of Lake Urmia, is located between 37° 42' and 38° 30' north latitude

**Table 1** Statistical parameters of the daily data sets for the two case studies

Basin	Data set	Unit	$X_{\text{mean}}$	$X_{\text{min}}$	$X_{\text{max}}$	$S_X$	$C_V$	$C_{SX}$
Aji-Chay	T <sub>Min</sub>	C	7.5	-21.6	28.0	9.4	1.3	-0.1
	T <sub>Max</sub>	C	18.4	-9.2	40.6	11.5	0.6	-0.2
	R	mm	0.7	0.0	37.0	2.4	3.5	5.7
	SH	hour	7.8	0.0	13.5	3.9	0.5	-0.7
	Q	m <sup>3</sup> /s	5.2	0.005	97.1	2.9	1.4	2.7
Mahabad-Chay	T <sub>Min</sub>	C	7.0	-19.4	26.4	8.2	1.2	-0.3
	T <sub>Max</sub>	C	19.3	-8.2	42.0	11.1	0.6	-0.2
	R	mm	1.1	0.0	68.0	4.0	3.6	6.2
	SH	hour	8.0	0.0	13.6	3.8	0.5	-0.7
	Q	m <sup>3</sup> /s	6.8	0.09	206.9	13.0	1.9	4.1

T<sub>Min</sub> minimum temperature, T<sub>Max</sub> maximum temperature, R daily rainfall, SH sunshine hours, Q runoff

**Table 2** Different input combinations for runoff prediction by GEP

No.	Input combinations
1	$R(t)$
2	$R(t), T_{\text{mean}}(t)$
3	$R(t), T_{\text{max}}(t), T_{\text{min}}(t)$
4	$R(t), T_{\text{max}}(t), T_{\text{min}}(t), \text{SR}(t)$
5	$R(t), R(t - 1), T_{\text{max}}(t), T_{\text{max}}(t - 1), T_{\text{min}}(t), T_{\text{min}}(t - 1)$
6	$R(t), R(t - 1), T_{\text{max}}(t), T_{\text{max}}(t - 1), T_{\text{min}}(t), T_{\text{min}}(t - 1), \text{SR}(t), \text{SR}(t - 1)$

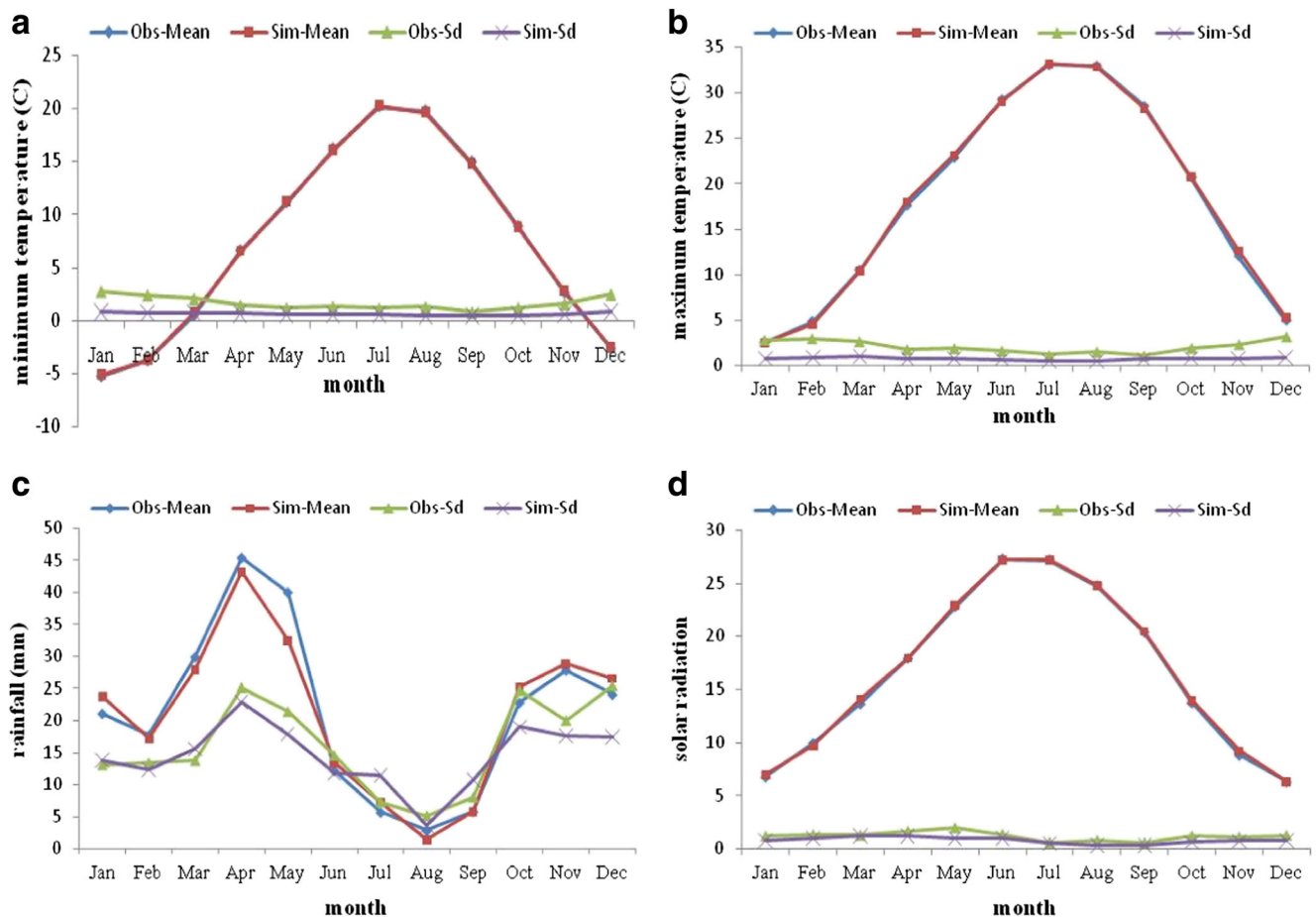
and 45° 40'–47° 53' east longitude and has a total area of 12,790 km<sup>2</sup>. The altitude varies from 1270 to 3400 m. The variables of the daily rainfall, minimum temperatures, maximum temperatures, and sunshine hours were used as input of LARS-WG model. The model automatically converts the sunshine hours to solar radiation.

In Aji-Chay River basin, the climatic variable data obtained from Tabriz synoptic station and the hydrometric data of the Aji-Chay River at the Vanyar station for the period 1982–2010 were used.

Another case study, Mahabad-Chay River basin, the fourth largest sub-basin of Lake Urmia, is located in the south of Lake

Urmia (between 36° 22' and 37° 10' N latitudes and between 44° 45' and 45° 56' E longitudes). The weather data related to Mahabad synoptic station and runoff data for Koutar hydrometric station was used in this case study. The historical data include a period of 23 years (1988–2010). For both case studies, the selected weather and runoff data have good quality.

Figure 1 shows the overall location of basins in Iran and the location of both stations respect to Lake Urmia. With considering the recent drought over Lake Urmia, the largest lake in the Middle East and the sixth largest salt-water lake on Earth due to climate change, dam



**Fig. 3** Monthly mean and standard deviation of observed and simulated weather data (Aji-Chay River basin). **a** Minimum temperature. **b** Maximum temperature. **c** Rainfall. **d** Solar radiation

**Table 3** Bias and MAE values for weather data in Aji-Chay River basin

Variable	Bias	MAE
Rainfall	-0.212	2.078
Minimum temperature	-0.003	0.129
Maximum temperature	0.045	0.242
Solar radiation	0.118	0.182

construction over the upstream, agricultural consumption, etc., the results of current research can be useful in this matter. The statistical parameters of daily data sets ( $X$ ) are presented in Table 1. In this table,  $X_{\text{mean}}$ ,  $X_{\text{min}}$ ,  $X_{\text{max}}$ ,  $S_X$ ,  $C_v$ , and  $C_{S_X}$  denote the mean, minimum, maximum, standard deviation, coefficient of variation, and skewness coefficient values of each variable, respectively. It is apparent from the table that the daily rainfall and runoff data have highly skewed distributions and high variations for the both basins.

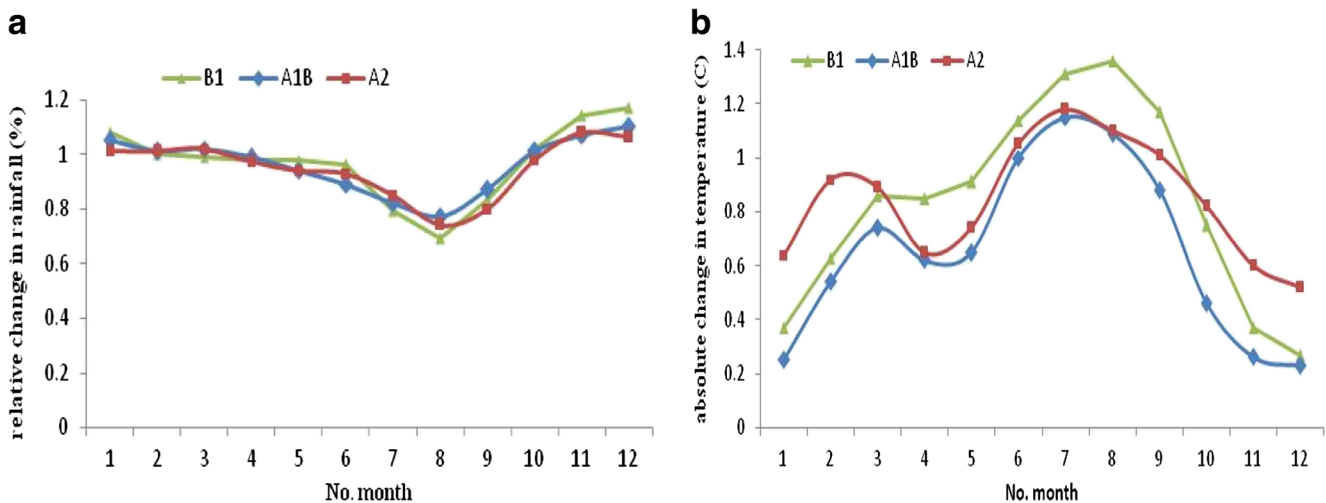
## 2.2 Description of LARS-WG

The weather generator models simulate the climatic variables in two steps. In the first step, rainfall event and its intensity were simulated and then, other remaining variables such as minimum and maximum temperature, solar radiation, humidity, and wind speed were simulated (Johnson et al. 1996). LARS-WG as a downscaling model was introduced by Racsko et al. (1991). Daily time series of maximum and minimum temperatures, solar radiation, and precipitation can be generated by LARS-WG. Semenov et al. (2010) stated that the LARS-WG can be applied for observed daily climatic data for a special site in order to compute the parameters of fitted distribution to climatic data.

In the LARS-WG, rainfall modeling and probability of rainfall occurrence are done using semi-empirical method and Markov chain. Furthermore, the modeling of solar radiation and temperature is done based on semi-empirical method and Fourier series, respectively (Semenov et al. 2002; Zarghami et al. 2011).

In order to generate data using LARS-WG, firstly, the characteristics of each station (i.e., the name, geographical information, altitude, and the daily climatic data) should be defined as an inputs of model. Then, the model analyzed the data, and the outputs can be summarized in a text file, which including the statistical characteristics of historical data in the form of mean monthly and seasonal data. The model with the respect to the trend of observed data can be regenerated for a given period, and finally, by comparing the simulated and observed data, the performance of the model can be evaluated. After evaluation of model efficiency, for each station, the data set of future horizons based on climate change scenarios with respect to the output of GCMs can be generated (Semenov et al. 2002).

In this study, the latest version (5.5) of LARS-WG (<http://www.rothamsted.ac.uk/mas-models/larswg/download>) was applied for downscaling weather data. In this version, different information related to GCMs is provided. In the current study, the Hadley Centre coupled model (HADCM3) is utilized through the downscaling process. One attractive feature of the HADCM3 model is that, unlike the other models, HADCM3 requires no flux adjustments. Such adjustments could potentially impact aspects of climate variability, and so are undesirable (Dong and Sutton 2005). LARS-WG applies different statistical tests for comparing the data generated by the weather generator with historical data during the baseline period. The tests were implemented for the direct generation of synthetic observed data to test how well the LARS-WG is performing in



**Fig. 4** Relative change in rainfall and absolute change in temperature during 2011–2030 in respect to the baseline in Aji-Chay River basin using different scenarios (B1,A1B and A2) a) rainfall b) temperature

reproducing the characteristics of the observed data. The  $t$  test was used for evaluating the monthly means of weather data. For each month,  $F$  tests were utilized on the variances of all the daily values for the month across all the year and on the variances of monthly mean values for the different years. The chi-squared test was utilized for comparing distribution of the wet and dry periods (Semenov et al. 1998). The value of “significance level” was chosen as 5%.

Bias and mean absolute error (MAE) were utilized for comparing the simulated data by LARS-WG and observed data (in the baseline period). Bias and MAE relationships are as follows

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (S_i - O_i) \tag{1}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |S_i - O_i| \tag{2}$$

where  $i$  is the number of month, and  $S_i$  and  $O_i$  are the simulated and observed values, respectively. More details about LARS-WG can be found in Semenov and Barrow (1997) and Semenov et al. (1998).

### 2.3 Impact simulation by gene expression programming

GEP was first introduced by Ferreira in 1999 (Ferreira 2001). GEP, as an extension of GP, is a search paradigm that involves computer programs such as mathematical expressions and decision trees. GEP is a genetic algorithm (GA) that applies populations of individuals and chooses them proportionally to fitness and introduces genetic variation by application of some genetic operators. In GEP, the individuals are encoded as linear strings with fixed length (the genome or chromosomes) which

are then presented as nonlinear entities of different sizes and shapes (i.e., pars trees) (Ferreira 2001; Kisi and Sanikhani 2015). Figure 2 schematically presents the fundamental steps included in GEP.

In this study, a powerful soft computing model, GeneXpro Tools 4.0 was used to predict runoff using GEP. Runoff prediction by application of GEP includes some major steps (i.e., select appropriate fitness function, select terminal and function sets, choose chromosomal structure, and select the rate of generic operators). In the first step, fitness function was selected. The fitness ( $m_i$ ) of an individual program ( $i$ ) is described by Eq. 3.

$$m_i = \sum_{j=1}^N (S - |A_{(ij)} - B_j|) \tag{3}$$

where  $S$  is the selection range,  $A_{(ij)}$  is the value predicted by the individual program ( $i$ ) for fitness case ( $j$ ) (from between fitness cases), and  $B_j$  is the target value for fitness case ( $j$ ). In this study, root relative squared error (RRSE) and relative absolute error (RAE) were chosen as suitable fitness functions for Aji-Chay and Mahabad-Chay River basins, respectively. In the second step, terminal set (T) and function set (F) are selected. In this paper, the terminal set is including climatic variables (i.e., rainfall, maximum temperature, minimum temperature, solar radiation) in current and prior time steps. The selection of functions set is depending on the nature and complexity of the problem. In this study, different combinations of functions were used including basic arithmetic operators (+, −, ×, ÷) and some mathematical functions were as follows:  $(\ln x, e^x, x^2, x^3, \sqrt{x}, \sqrt[3]{x}, \sin x, \cos x, \arctan(x))$ .

Selecting the chromosomal architecture is the third step. In the next step, the type of linking function (i.e., addition,

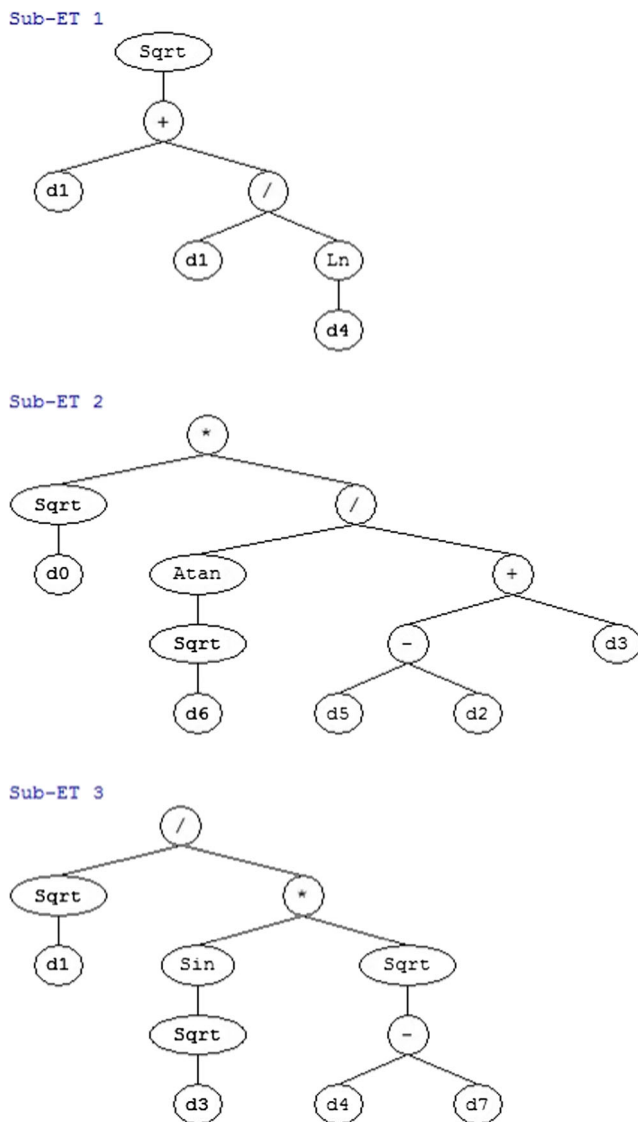
**Table 4** Validation results of different GEP basic and linking functions for the parse tree (Aji-Chay River basin)

	Definition	RMSE (m <sup>3</sup> /s)
F1	{+, −, ×, ÷}	1.44
F2	{+, −, ×, ÷, ln, e <sup>x</sup> }	1.30
F3	{+, −, ×, ÷, √, ∛, x <sup>2</sup> , x <sup>3</sup> }	1.36
F4	{+, −, ×, ÷, ln, e <sup>x</sup> , √, ∛, x <sup>2</sup> , x <sup>3</sup> }	1.23
F5	{+, −, ×, ÷, ln, e <sup>x</sup> , √, ∛, x <sup>2</sup> , x <sup>3</sup> , sinx, cosx, arctgx}	1.16
Linking functions		
Addition		1.16
Multiplication		1.22
Subtraction		1.26
Division		1.31

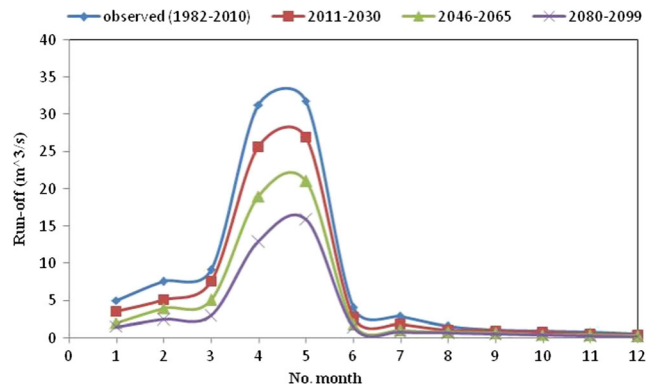
**Table 5** The performance criteria of GEP models using different input combinations in runoff simulation (Aji-Chay River basin)

Pattern no.	Training		Validation	
	$R^2$	RMSE (m <sup>3</sup> /s)	$R^2$	RMSE (m <sup>3</sup> /s)
1	0.54	2.34	0.45	2.43
2	0.64	2.01	0.60	2.23
3	0.76	1.76	0.71	1.89
4	0.84	1.62	0.81	1.74
5	0.88	1.30	0.84	1.34
6	0.91	1.11	0.89	1.16

subtraction, multiplication, and division) is selected. The final step is choosing the generic operators and their values (Ferreira 2006).



**Fig. 5** Expression tree obtained by GEP model (Aji-Chay River basin)



**Fig. 6** The comparison of observed and simulated runoff hydrographs in different future horizons in Aji-Chay River basin

The values allocated to each parameter are summarized as follows: number of chromosomes, 30; head size, 8; number of genes, 3; mutation rate, 0.044; inversion rate, 0.1; one point recombination rate, 0.3; two point recombination rate, 0.3; gene recombination rate, 0.1; gene transposition rate, 0.1; IS transposition rate, 0.1; and RIS transposition, 0.1 (Kisi et al. 2015; Kisi and Sanikhani 2015). It is notable that these values are the default values provided by GeneXpro Tools.

**2.4 Input combinations and performance criteria**

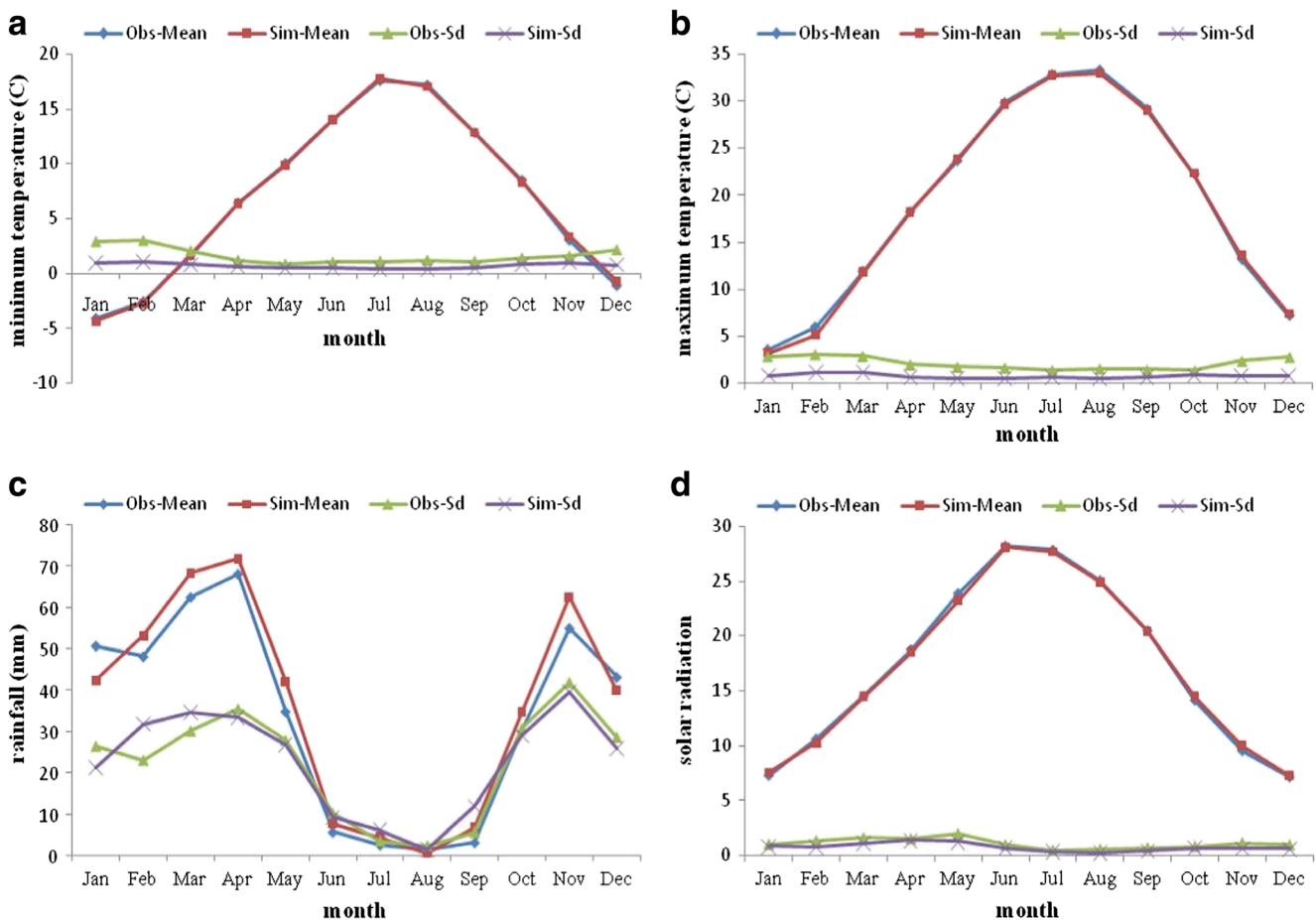
In this study, different input patterns were considered for the investigation of the impact of climate change on runoff using GEP model. Various input combinations including rainfall ( $R$ ), minimum temperature ( $T_{min}$ ), maximum temperature ( $T_{max}$ ), mean temperature ( $T_{mean}$ ), and solar radiation (SR) in current ( $t$ ) and prior time ( $t - 1$ ) were applied in GEP model. The input combinations are presented in Table 2.

For runoff prediction in different future horizons (i.e., 2020, 2050, and 2090) over the basins, historical data were used for training (calibration) of the models. Among the available data, 70% of data was used for training and the remaining 30% of data was used for validation of the models. Finally, the runoff prediction in future periods was made by application of the best input combinations during training and validation periods.

The performance of models was assessed with respect to determination coefficient ( $R^2$ ) and root mean squared error (RMSE) statistics. RMSE can be defined as follows

**Table 6** The percent of relative decrease of runoff in different future horizons comparing to baseline period (Aji-Chay River basin)

Variable	2011–2030	2046–2065	2080–2099
Mean discharge	19.8	41.5	58.7
Peak discharge	15.4	34.0	50.0



**Fig. 7** Monthly mean and standard deviation of observed and simulated weather data (Mahabad-Chay River basin). **a** Minimum temperature. **b** Maximum temperature. **c** Rainfall. **d** Solar radiation

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_o - Q_e)^2} \quad (4)$$

where  $N$  is the number of data sets, and  $Q_o$  and  $Q_e$  are the observed and predicted runoff values, respectively.

### 3 Application and results

#### 3.1 Aji-Chay River basin

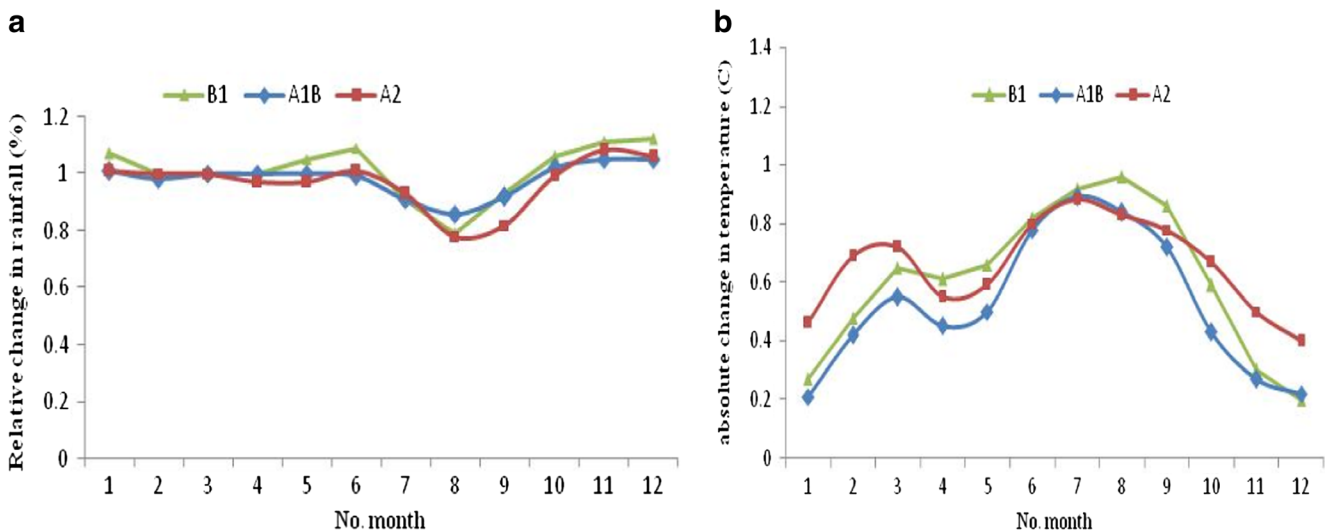
The monthly mean and standard deviation of observed and simulated weather data of Aji-Chay River basin for the period of 1982–2010 are compared in Fig. 3. The MAE and bias values for each data are also provided in Table 3. It is clear from the figure and table that the weather data are well simulated by LARS-WG, especially the minimum temperature, maximum temperature, and solar radiation. For rainfall, there are some deviations and significant differences in some

months (e.g., May, July, October, and December). It is also clear from the Table 3 that the rainfall has the highest MAE (2.078) and bias (−0.212) values. This may be due to the fact that the strength of the correlation between successive values varies considerably for the rainfall data, while the auto-correlation of the temperature (or solar radiation) data is very strong. Also, highly skewed distribution of rainfall data may be another reason of this. From Fig. 3c, it is clear that the maximum difference between the observed and simulated rainfall is 7.5 mm in May.

**Table 7** Bias and MAE values for weather data in Mahabad-Chay River basin

Variable	Bias	MAE
Rainfall	2.425	4.473
Minimum temperature	−0.027	0.172
Maximum temperature	−0.113	0.249
Solar radiation	−0.062	0.238





**Fig. 8** Relative change in rainfall and absolute change in temperature during 2010–2030 in respect to the baseline in Mahabad-Chay River basin using different scenarios (B1, A1B and A2) a) rainfall b) temperature

The weather data were predicted by using three different scenarios of LARS-WG climate model such as B1, A1B, and A2 for the period of 2011–2030. The relative change in rainfall and absolute change in temperature (comparing to historical period) are shown in Fig. 4 for the Aji-Chay River basin. Solar radiation is not provided in this figure, because the variation of solar radiation is approximately linear and there is not have any increasing or decreasing trend. The relative change in solar radiation by application of different scenarios (i.e., A1B, A2 and B1) is about 0.97–1.01. From Fig. 4a, it is clear that the rainfall generally shows decreasing trend especially in the summer months in which more water is required by the region. This lacking in rainfall will cause water stress and reduce the per capita water access. Temperature is another important parameter which significantly affects droughts. From Fig. 4b, it is clear that the temperature shows significantly increasing trend (0.8 °C) in summer months. Generally, according to Fig. 4, the Aji-Chay River basin will have less rain and hotter summers in the period of 2011–2030. However, there is an increasing trend in rainfall and decreasing trend in temperature for the month of September, October, and December.

Different GEP functions were tried in the study. The results of validation of used GEP functions are provided in Table 4. It is apparent from the table that the optimal GEP model was obtained by functions set of F5, which is default set in GeneXpro program and addition linking function. RRSE is found as the best fitness function in all GEP applications. Kisi et al. (2013) used GEP models in modeling rainfall-runoff process and they also found the RRSE as the best fitness function. The accuracy of the GEP models for each input combination (Table 2) is given in Table 5 for the both training and validation

periods. According to Table 5, it is clear that the GEP model comprising inputs of  $R(t)$ ,  $R(t-1)$ ,  $T_{max}(t)$ ,  $T_{max}(t-1)$ ,  $T_{min}(t)$ ,  $T_{min}(t-1)$ ,  $SR(t)$ , and  $SR(t-1)$  shows the best accuracy. Using only rainfall or together with temperature data gives poor accuracy in runoff prediction with respect to  $R^2$  and RMSE. It is apparent from Table 5 that adding previous values of rainfall, temperature, and solar radiation data as inputs to the GEP model considerably improves the prediction accuracy. The relative RMSE and  $R^2$  differences between the best (6th inputs) and the worst (1st input) input combinations are 52 and 98% in the validation period, respectively.

The equation of the optimal GEP model including sixth input combination for the Aji-Chay River basin is

$$Q(t) = A + \frac{\sqrt{R(t-1)}}{\sin \sqrt{SR(t-1)} \sqrt{T_{max}(t-1) - T_{min}(t)}} \quad (5)$$

$$A = B + \frac{\sqrt{R(t)} \text{Atan} \sqrt{T_{max}(t)}}{T_{min}(t-1) - SR(t-1) + SR(t)} \quad (5a)$$

$$B = \sqrt{R(t-1) + \frac{R(t-1)}{\text{Ln}(T_{max}(t-1))}} \quad (5b)$$

The expression tree of the optimal GEP model for Aji-Chay River basin is shown in Fig. 5. Figure 6 compares the observed and simulated runoff hydrographs in different future horizons for the Aji-Chay River basin. The hydrographs were predicted with respect to scenario A2 which assumes that the current socioeconomic situation will continue. Scenario A2 is better and more conservative than the A1B and B1. It is notable that the scenario B1 describes a convergent world with global population that peaks will occur in midcentury and decline

**Table 8** Validation results of different GEP basic and linking functions for the parse tree (Mahabad-Chay River basin)

	Definition	RMSE (m <sup>3</sup> /s)
F1	{+, -, ×, ÷}	1.83
F2	{+, -, ×, ÷, ln, e <sup>x</sup> }	1.73
F3	{+, -, ×, ÷, √, √, x <sup>2</sup> , x <sup>3</sup> }	1.52
F4	{+, -, ×, ÷, ln, e <sup>x</sup> , √, √, x <sup>2</sup> , x <sup>3</sup> }	1.63
F5	{+, -, ×, ÷, ln, e <sup>x</sup> , √, √, x <sup>2</sup> , x <sup>3</sup> , sinx, cosx, arctgx}	1.59
Linking functions		
Addition		1.52
Multiplication		1.57
Subtraction		1.62
Division		1.71

thereafter. Also, A1B, as a sub-set of scenario A1, emphasizes on a balance on all energy sources (Intergovernmental Panel on Climate Change, IPCC 2007). It is clear from the Fig. 6 that the runoff mean and hydrograph peak significantly decrease by increasing time horizons. Table 6 gives the percent decrement in mean runoff and hydrograph peak for the future horizons of 2011–2030, 2046–2065, and 2080–2099. It can be seen that the decrements in mean and peak discharges will be as high as 58.7 and 50% at the end of this century. This result is parallel with the increasing trend in temperature and decreasing trend in rainfall shown in Fig. 4. These results imply that new and sustainable adaptation strategies are needed to rescue the water supply in this region in the future.

### 3.2 Mahabad-Chay River basin

Figure 7 illustrates the monthly mean and standard deviation of observed and simulated weather data of Mahabad-Chay River basin for the period of 1988–2010. The MAE and bias values of each data are given in Table 7. Similar to the Aji-Chay River basin, here also, the weather data are well simulated by LARS-WG, especially the minimum temperature, maximum temperature, and solar radiation. Here also, some deviations and significant differences are clearly seen for rainfall in some months (e.g., January, February, March, and November). It is clear from Fig. 7 that the maximum difference between the observed and simulated rainfall is 8.3 mm in January. As seen from Table 7, the rainfall has the highest MAE (4.473) and bias (2.425) values.

The predictions of the rainfall and temperature data by three different scenarios of LARS-WG climate model such as B1, A1B, and A2 for the period of 2011–2030 are shown in Fig. 8. The figure indicates the relative

change in rainfall and absolute change in temperature (comparing to historical period) for the Mahabad-Chay River basin. Solar radiation is not provided in this figure because the relative change in solar radiation by application of different scenarios (i.e., A1B, A2, and B1) is about 0.98–1.01. As can be seen from Fig. 8a, rainfall generally shows decreasing trend especially in the summer months alike to previous basin (Aji-Chay). From Fig. 8b, it is clear that the temperature shows significantly increasing trend in summer months. It is clearly seen from the Fig. 8 that the Mahabad-Chay River basin will have less rain and hotter summers in the period of 2011–2030 but not as much as Aji-Chay River basin. The validation results of different GEP functions are given in Table 8. From the table, it is clear that the optimal GEP model was obtained by function set of F3 and addition linking function. In Mahabad-Chay River basin, relative absolute error (RAE) is found as the best fitness function in all GEP applications. Table 9 gives the accuracy of the optimal GEP models for each input combination in the training and validation period. It is clear from the table that the

**Table 9** The performance criteria of GEP model using different input combination in runoff simulation (Mahabad-Chay River basin)

Pattern no.	Train		Validation	
	R <sup>2</sup>	RMSE (m <sup>3</sup> /s)	R <sup>2</sup>	RMSE (m <sup>3</sup> /s)
1	0.50	2.79	0.53	2.84
2	0.62	2.54	0.59	2.65
3	0.73	2.07	0.72	2.12
4	0.86	1.86	0.83	1.95
5	0.88	1.69	0.85	1.74
6	0.92	1.50	0.90	1.52

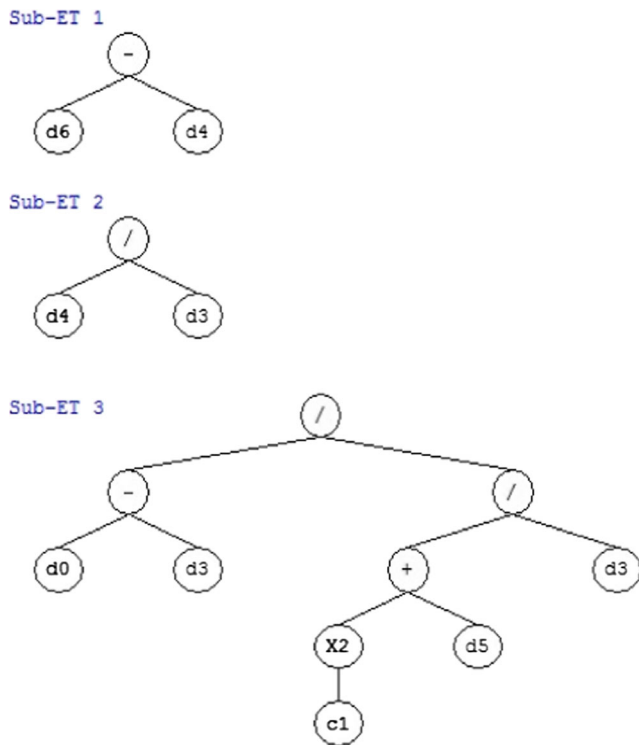


Fig. 9 Expression tree obtained by GEP model (Mahabad-Chay River basin)

GEP model comprising sixth input combination shows the best accuracy similar to the Aji-Chay River basin. Here also, using only rainfall and temperature data provides poor estimates for runoff prediction and adding previous values of rainfall, temperature, and solar radiation data as inputs to the GEP model considerably improves the prediction accuracy. The relative RMSE and  $R^2$  differences between the best (6th inputs) and the worst (1st input) input combinations are 46 and 70% in the validation period, respectively.

The equation of the optimal GEP model in predicting runoff of Mahabad-Chay River basin is

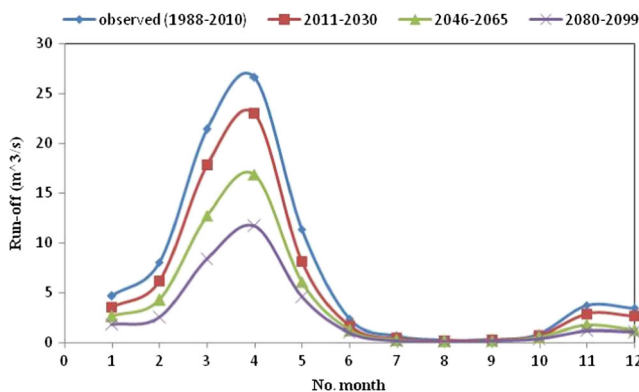


Fig. 10 The comparison of observed and simulated runoff hydrographs in different future horizons in Mahabad-Chay River basin

$$Q(t) = (T_{\max}(t) - T_{\max}(t-1)) + \frac{T_{\max}(t-1)}{SR(t)} + \frac{SR(t)[R(t) - SR(t-1)]}{26.5 + T_{\min}(t-1)} \tag{6}$$

The expression tree of the optimal GEP model for Mahabad-Chay River basin is shown in Fig. 9. The observed and simulated runoff hydrographs in different future horizons are compared in Fig. 10. Alike to the Aji-Chay River basin, here also, the mean runoff and hydrograph peak significantly decrease by increasing time horizons. The percent decrement in mean runoff and hydrograph peak is given in Table 10 for the future horizons of 2011–2030, 2046–2065, and 2080–2099. It is apparent from the table that the decrements in mean and peak discharges will be as high as 60.5 and 55.9% at the end of this century. This result confirms the increasing trend in temperature and decreasing trend in rainfall as shown in Fig. 8.

### 4 Conclusion

In the current paper, the impact of climate change on runoff was investigated for two basins in northwest part of Iran. LARS-WG was utilized for downscaling the climatic variables including rainfall, solar radiation, and maximum and minimum temperatures. The results showed that LARS-WG could be successfully applied for downscaling of climatic variables. For both basins, the generated climatic variables in future horizon (2020) were compared with baseline period. The results indicated a decreasing rainfall trend, especially in the summer months in which more water was needed in the region. Furthermore, an increasing trend was found for temperature over the basins. In another part of the study, the GEP model was applied for prediction of runoff in the future horizons (2020, 2055, and 2090) by using A2 scenario. The GEP model whose inputs comprise the rainfall, solar radiation, and minimum and maximum temperatures in current and prior time was found to be the best one for runoff prediction. For Aji-Chay River basin, the mean and peak discharge were decreased by 58.7 and 50% in the 2090 horizon, respectively. For the Mahabad-Chay River basin, the decrements in mean and peak discharge will be

Table 10 The percent of relative decrease of runoff in different future horizon comparing to baseline period (Mahabad-Chay River basin)

Variable	2011–2030	2046–2065	2080–2099
Mean discharge	19.6	42.5	60.5
Peak discharge	13.8	36.7	55.9

as high as 60.5 and 55.9% at the end of this century, respectively. With respect to dramatic decrease of runoff in the future, it is necessary to find a solution to prevent or minimize the negative impact of climate change on runoff and the role of managers and decision makers in this matter is very essential in the region.

## References

- Aytek A, Alp M (2008) An application of artificial intelligence for rainfall-runoff modeling. *J Earth Syst Sci* 117(2):145–155
- Aytek A, Kisi Ö (2008) A genetic programming approach to suspended sediment modelling. *J Hydrol* 351(3):288–298
- Chang J, Wang Y, Istanbuluoglu E, Bai T, Huang Q, Yang D, Huang S (2014) Impact of climate change and human activities on runoff in the Weihe River basin. *Quaternary International*, China
- Chiew FHS, Whetton PH, McMahon TA, Pittock AB (1995) Simulation of the impacts of climate change on runoff and soil moisture in Australian catchments. *J Hydrol* 167(1):121–147
- Dong B, Sutton T (2005) Mechanism of Interdecadal thermohaline circulation variability in a coupled ocean–atmosphere GCM. *J Clim* 18:1117–1135
- Fatichi S, Ivanov VY, Caporali E (2011) Simulation of future climate scenarios with a weather generator. *Adv Water Resour* 34(4):448–467
- Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst* 13(2):87–129
- Ferreira C (2006) Gene expression programming: mathematical modeling by an artificial intelligence, 2nd edn. Springer, Germany
- Firat M, Gungor M (2007) River flow estimation using adaptive neuro fuzzy inference system. *Math Comput Simul* 75(3):87–96
- Fujihara Y, Tanaka K, Watanabe T, Nagano T, Kojiri T (2008) Assessing the impacts of climate change on the water resources of the Seyhan River basin in Turkey: use of dynamically downscaled data for hydrologic simulations. *J Hydrol* 353(1):33–48
- Güven A, & Aytek A (2009) New approach for stage–discharge relationship: gene-expression programming. *J Hydrol Eng* 14(8): 812–820
- Hashmi MZ, Shamseldin AY (2014) Use of gene expression programming in regionalization of flow duration curve. *Adv Water Resour* 68:1–12
- Hashmi MZ, Shamseldin AY, Melville BW (2011) Statistical downscaling of watershed precipitation using gene expression programming (GEP). *Environ Model Softw* 26(12):1639–1646
- Hassan Z, Shamsudin S, Harun S (2014) Application of SDSM and LARS-WG for simulating and downscaling of rainfall and temperature. *Theor Appl Climatol* 116(1–2):243–257
- IPCC, Intergovernmental Panel on Climate Change, (2007). Fourth Assessment Report, Climate Change.
- Jiang T, Chen YD, Xu CY, Chen X, Chen X, Singh VP (2007) Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China. *J Hydrol* 336(3):316–333
- Johnson GL, Hanson CL, Hardegree SP, Ballard EB (1996) Stochastic weather simulation: overview and analysis of two commonly used models. *J Appl Meteorol* 35(10):1878–1896
- Kaleris V, Papanastopoulos D, Lagas G (2001) Case study on impact of atmospheric circulation changes on river basin hydrology: uncertainty aspects. *J Hydrol* 245(1):137–152
- Khan MS, Coulibaly P, Dibike Y (2006) Uncertainty analysis of statistical downscaling methods. *J Hydrol* 319(1):357–382
- Kisi O, Sanikhani H (2015) Modelling long-term monthly temperatures by several data-driven methods using geographical inputs. *Int J Climatol* 35(13):3834–3846
- Kisi O, Shiri J, Tombul M (2013) Modeling rainfall-runoff process using soft computing techniques. *Comput Geosci* 51:108–117
- Kisi O, Sanikhani H, Zounemat-Kermani M, Niazi F (2015) Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. *Comput Electron Agric* 115:66–77
- Lu Y, Qin XS, Xie YJ (2016) An integrated statistical and data-driven framework for supporting flood risk analysis under climate change. *J Hydrol* 533:28–39
- Matondo JI, Peter G, Msibi KM (2004) Evaluation of the impact of climate change on hydrology and water resources in Swaziland: part II. *Phys Chem Earth Parts A/B/C* 29(15):1193–1202
- Mehdizadeh S, Behmanesh J, Khalili K (2017) Application of gene expression programming to predict daily dew point temperature. *Appl Therm Eng* 112:1097–1107
- Nkomozepi T, Chung SO (2014) The effects of climate change on the water resources of the Geumho River basin, Republic of Korea. *J Hydro Environ Res* 8(4):358–366
- Nourani V, Baghanam AH, Adamowski J, Kisi O (2014) Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. *J Hydrol* 514:358–377
- Parajuli PB, Jayakody P, Sassenrath GF, Ouyang Y (2016) Assessing the impacts of climate change and tillage practices on stream flow, crop and sediment yields from the Mississippi River basin. *Agric Water Manag* 168:112–124
- Racsko P, Szeidl L, Semenov M (1991) A serial approach to local stochastic weather models. *Ecol Model* 57(1):27–41
- Savic DA, Walters GA, Davidson JW (1999) A genetic programming approach to rainfall-runoff modelling. *Water Resour Manag* 13(3): 219–231
- Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. *Clim Chang* 35(4): 397–414
- Semenov MA, Brooks RJ, Barrow EM, Richardson CW (1998) Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Clim Res* 10(2):95–107
- Semenov MA, Barrow EM, Lars-Wg A (2002) A stochastic weather generator for use in climate impact studies. *User Manual*, Hertfordshire
- Semenov MA, Donatelli M, Stratonovitch P, Chatzidaki E, Baruth B (2010) ELPIS: a dataset of local-scale daily climate scenarios for Europe. *Climate Res (Open Access for articles 4 years old and older)* 44(1):3
- Srikanthan R, McMahon TA (2001) Stochastic generation of annual, monthly and daily climate data: a review. *Hydrol Earth Syst Sci* 5(4):653–670
- Vallam P, Qin XS (2016) Climate change impact assessment on flow regime by incorporating spatial correlation and scenario uncertainty. *Theor Appl Climatol*. doi:10.1007/s00704-016-1802-1
- Wang WC, Chau KW, Cheng CT, Qiu L (2009) A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J Hydrol* 374(3):294–306
- Zarghami M, Abdi A, Babaeian I, Hassanzadeh Y, Kanani R (2011) Impacts of climate change on runoffs in East Azerbaijan, Iran. *Glob Planet Chang* 78(3):137–146
- Zheng YQ, Qian ZC, He HR, Liu HP, Zeng XM, Yu G (2007) Simulations of water resource environmental changes in China during the last 20000 years by a regional climate model. *Glob Planet Chang* 55(4):284–300
- Zorn CR, Shamseldin AY (2015) Peak flood estimation using gene expression programming. *J Hydrol* 531:1122–1128
- Zounemat-Kermani M, Rajaei T, Ramezani-Charmahineh A, Adamowski JF (2017) Estimating the aeration coefficient and air demand in bottom outlet conduits of dams using GEP and decision tree methods. *Flow Meas Instrum* 54:9–19