

# Evaluating climate change effects on runoff by statistical downscaling and hydrological model GR2M

Umut Okkan · Okan Fistikoglu

Received: 12 May 2013 / Accepted: 27 August 2013 / Published online: 19 September 2013  
© Springer-Verlag Wien 2013

**Abstract** The main purpose of this study is to evaluate the impacts of climate change on Izmir-Tahtali freshwater basin, which is located in the Aegean Region of Turkey. For this purpose, a developed strategy involving statistical downscaling and hydrological modeling is illustrated through its application to the basin. Prior to statistical downscaling of precipitation and temperature, the explanatory variables are obtained from National Centers for Environmental Prediction/National Center for Atmospheric Research reanalysis data set. All possible regression approach is used to establish the most parsimonious relationship between precipitation, temperature, and climatic variables. Selected predictors have been used in training of artificial neural networks-based downscaling models and the trained models with the obtained relationships have been operated to produce scenario precipitation and temperature from the simulations of third Generation Coupled Climate Model. Biases from downscaled outputs have been reduced after downscaling process. Finally, the corrected downscaled outputs have been transformed to runoff by means of a monthly parametric hydrological model GR2M to assess the probable impacts of temperature and precipitation changes on runoff. According to the A1B climate scenario results, statistically significant trends are foreseen for precipitation, temperature, and runoff in the study basin.

## 1 Introduction

Hydrometric and meteorological observations have shown that water resources and hydrologic process are being affected by climate changes, particularly greenhouse concentration and temperature increases. According to Fourth Assessment Report (AR4) of Intergovernmental Panel on Climate Change (IPCC), significant changes on temperature and precipitation which are the major climatic inputs to hydrologic system have been foreseen for many regions in the world due to climate change (IPCC 2007).

Changes in precipitation and temperature cause relative variation in runoff as they affect the rainfall–runoff processes in a basin. With regard to water systems, because of the close linkage between climate change effects and hydrological processes, there have been several studies that assess the impacts of climate change on water resources (Holt and Jones 1996; Lettenmaier et al. 1999; Landman et al. 2001; Arnell et al. 2001; Prudhomme et al. 2002; Phillips et al. 2003; Struglia et al. 2004; Milly et al. 2005; Gedney et al. 2006; Graham et al. 2007; Bates et al. 2008)

As known, the probable variations on major hydrological variables such as precipitation, temperature, evaporation, and runoff can be considered by using the coarse results of the general circulation models (GCMs) in which land, atmosphere, and ocean systems are numerically coupled. However, the GCMs may not represent the variations of local climate scales since the GCMs have horizontal resolutions of hundreds of kilometers. Hence, there is a need for high-resolution results to interpret the impact of the large-scale atmospheric patterns at the local scale (Wilby et al. 2002).

Two main downscaling approaches, namely dynamic downscaling and statistical downscaling, are used for the studies about climate scenario assessments at higher resolutions (Wilby et al. 2002; Anandhi et al. 2008). The dynamic downscaling method is associated with the physically based

---

U. Okkan (✉)  
Department of Civil Engineering, Faculty of Engineering–  
Architecture, Balikesir University, Cagis Campus,  
10145 Balikesir, Turkey  
e-mail: umutokkan@balikesir.edu.tr

O. Fistikoglu  
Department of Civil Engineering, Faculty of Engineering, Dokuz  
Eylul University, Tinaztepe Campus, 35160 Izmir, Turkey  
e-mail: okan.fistikoglu@deu.edu.tr

Regional Climate Models (RCMs) that are other numerical models in which GCM results constitute the boundary conditions for the local climate domain (Crane and Hewitson 1998). The RCMs are able to parameterize the physical atmospheric process and simulate the regional climate features (Frei et al. 2006; Leung et al. 2003). However, the drawback of RCMs is their complicated design, uncertainty, and high computational cost. Additionally, the applications of RCMs are not flexible to adapt to another region unlike in the applications of statistical downscaling methods (Fistikoglu and Okkan 2011).

The statistical downscaling methods involve deriving statistical relationships that transform large-scale atmospheric variables of GCMs to surface-level variables. There are three main types of statistical downscaling methods, namely weather classification methods, weather generators, and transfer functions (Khan et al. 2006). The most popular statistical downscaling approaches are the transfer functions, which statistically model the relationships between large-scale atmospheric variables and local surface variables (Tatli et al. 2004; Schoof et al. 2007; Fistikoglu and Okkan 2011). Applications of these transfer functions vary from linear and nonlinear regression types, artificial neural networks (ANNs), support vector machines, principal component analysis, canonical correlation to redundancy analysis (Crane and Hewitson 1998; Wilby et al. 2003; Maheras et al. 2004; Bardossy et al. 2005; Anandhi et al. 2008; Fistikoglu and Okkan 2011).

The presented study was designed to focus on evaluating the climate change effects on runoff, which result from forecasted changes in precipitation and temperature. The methodological steps covered the following: (a) generation of climate change scenarios to forecast changes on precipitation and temperature for a study region in Turkey, (b) application of the probable changes to the study area through a statistical downscaling approach and a hydrological model to estimate changes on runoff, and (c) interpreting the differences between future and past period for precipitation, temperature, and runoff. The application of methodology is carried on Tahtali watershed to understand possible effects of climate change on a critical region with Mediterranean climate characteristics. It is worth mentioning that to the best of our knowledge, the presented paper is one of the initial studies about assessing climate change effects on hydrometeorological variables in Turkey. In the study content, the results of third-generation Canadian General Circulation Model (CGCM3) cited in the AR4 of the IPCC have been evaluated in terms of a future climate scenario (A1B) and a scenario representing the climate of the twentieth century (20C3M). In order to get high-resolution results of these scenarios, a statistical downscaling strategy has been improved by using Levenberg–Marquardt algorithm-based feed forward neural networks (LM-FFNN) with stopped training approach. The following processes have been carried out for this study. First, explanatory climatic variables (predictors), which represent

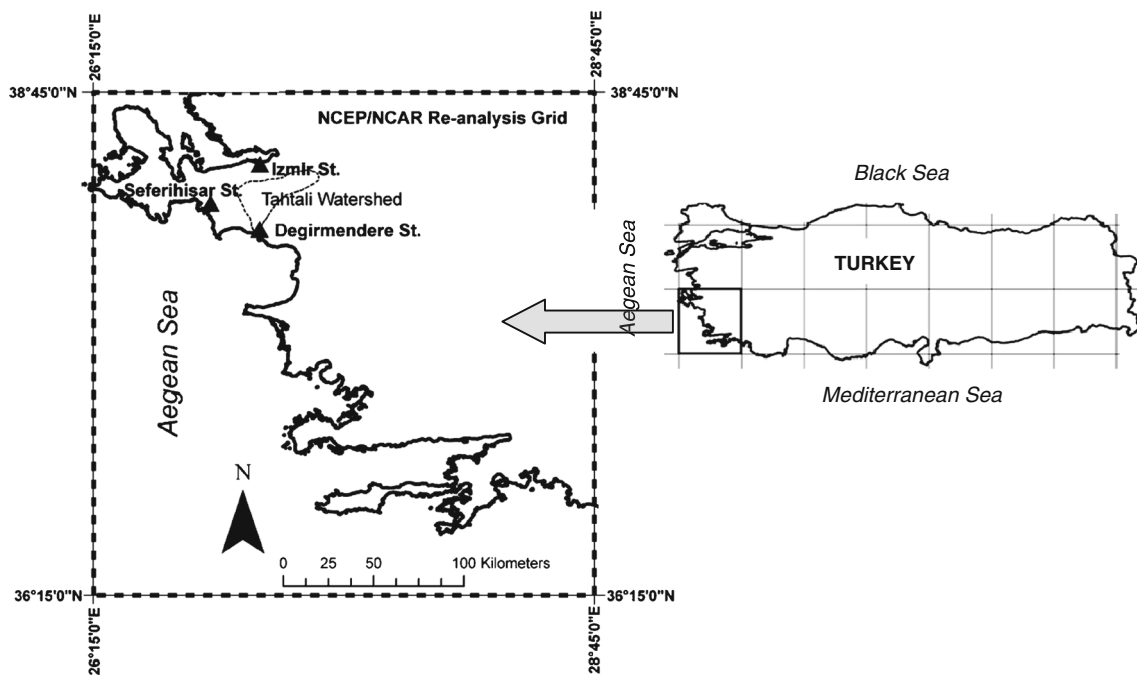
the monthly areal precipitation and temperature of Tahtali watershed, were selected from the National Center for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis data set. In this context, statistical approaches based on the all possible regression method and Mann–Whitney  $U$  homogeneity test were used to determine the effective predictors among the NCEP/NCAR data set. Later, predictors selected from NCEP/NCAR data set were used for training the downscaling models to establish statistical relationships between local-scale variables (observed precipitation and temperature) and climatic variables. The relationships thus obtained were used to project the future precipitation and temperature from CGCM3 (The Third Generation Coupled Global Climate Model) simulations. Following to these statistical downscaling analyses, bias correction procedure was applied. Afterwards, the corrected scenario results were evaluated at watershed scale by means of parametric hydrological model GR2M to observe the probable impacts of temperature and precipitation changes on runoff for near future periods.

The information about study region and climate data, some details about prediction selection, are introduced in the Section 2. Then, downscaling modeling and bias correction process applied in this study and obtained results are presented in Sections 3 and 4, respectively. The application of GR2M model and forecasted runoff results are presented in Section 5. Some conclusions are then made in the final section.

## 2 Study region and data

The study region covers the Tahtali watershed which is located at the Aegean coast of the Turkey (Fig. 1). Tahtali watershed has a total drainage area of 546 km<sup>2</sup> and the annual runoff is in the order of 160 hm<sup>3</sup>. The Tahtali watershed is the major surface water resource for the city of Izmir which is the third largest city in Turkey. The study region has typical Mediterranean climate characteristics. Only three meteorological stations are available around the study area, namely Izmir (17220), Seferihisar (17820), and Degirmendere (6294). The monthly precipitation records for these stations were obtained from the Turkish State Meteorological Service. The monthly mean areal precipitation values, which represent the watershed, are obtained by the Thiessen polygons. The monthly mean temperature values are obtained from Izmir (17220) meteorological station.

The NCEP/NCAR (Kalnay et al. 1996) monthly mean reanalysis data set are selected for the study region as large-scale predictors for the period from 1948 to 2008. The NCEP/NCAR distributes the reanalysis data set of which time resolution varies from hours to month, to present atmospheric conditions at different levels of the atmosphere (Kalnay et al. 1996). The NCEP/NCAR data set have been used in



**Fig. 1** Study region and the meteorological stations in the NCEP/NCAR reanalysis grid

several downscaling applications in different regions over the world, as daily, monthly, and seasonal predictors since reanalysis data are outputs from a high-resolution model operated data from meteorological, upper air, and satellite observation stations (Maheras et al. 2004; Tatli et al. 2004; Anandhi et al. 2008; Fistikoglu and Okkan 2011; Okkan and Fistikoglu 2012). The variables of the selected NCEP/NCAR grid for the study region whose latitudes range from 36.15°N to 38.45°N and longitudes ranges from 26.15°E to 28.45°E at a spatial resolution of 2.5° (Fig. 1) are obtained from the web site <http://www.cdc.noaa.gov/>. The variables extracted from the NCEP/NCAR reanalysis data set include air temperature, relative humidity, and geopotential height at various atmospheric levels and pressure, sea level pressure, and large-scale precipitation.

The monthly climate data used in the study are obtained from the CGCM3 climate model, through the web site <http://esg.llnl.gov:8080/>. The CGCM3 climate model grid is uniform along the longitude with grid box size of 3.75° and roughly uniform along the latitude (nearly 3.75°). CGCM3 climate model was selected in this study because it is one of the frequently used climate model in literature. The monthly climate data for this GCM were compiled for the selected grid. The coordinates of this grid center are 38.97°N and 26.25°E for latitude and longitude, respectively. The nearest grid representing the study area includes air temperature, geopotential height, and relative humidity at different atmospheric levels and surface level variables.

In this study concept, the authors have intended to validate the hydrometeorological forecasts for the present climate and present some evidence of climate change for the near future focusing on a scenario and then discuss possible impacts

based on these projections. Thus, the used data set consist of future climate scenario simulations covering the years 2010s, 2020s, and 2030s and a scenario simulation representing climate of the twentieth century (20C3M) for the historic-based period between 1950 and 1999. For future climate assessment, A1B scenario (a subset of the A1 family), which lies near the high end of the spectrum for future greenhouse gas emissions, particularly through mid-century, has been considered. This scenario projects a future where technology is shared between developed and developing nations to reduce regional economic inequalities.

According to IPCC (2007), a brief description of the scenarios is presented in Table 1. The detailed descriptions of the scenarios have been presented by Anandhi et al. (2008). The atmospheric variables to be provided to the downscaling models for the purpose of the downscaling of the monthly precipitation and temperature series representing the study region consist of determined 12 NCEP/NCAR variables, which common to CGCM3 climate model, including 1948–2008 period and listed in Table 2.

It is often important to determine if the data set is homogenous before any statistical model is applied to it. So, the homogeneity check of climate data is of major importance because some factors make data unrepresentative of the climate variation, and thus the conclusion of hydrometeorological studies are potentially biased (Costa and Soares 2009). The Mann–Whitney  $U$  (M-W) test having greater efficiency than the  $t$  test on nonnormal distributions can be used to determine whether a set of data can be considered homogenous to a certain degree of accuracy. This nonparametric statistical test is used to analyze two comparison groups to

**Table 1** The brief description of the scenarios considered in the study (IPCC 2007; Anandhi et al. 2008)

The name of scenario	Data set	Description of scenario
20C3M	Climate of the twentieth century	Atmospheric CO <sub>2</sub> concentrations and other input data are based on historical records or estimates beginning around the time of the Industrial Revolution
A1B	720 ppm CO <sub>2</sub> maximum	A future world of very rapid economic growth, global population that peaks in mid-century and declines thereafter, and rapid introduction of new and more efficient technologies, with the development balanced across energy sources

identify whether they have the same distribution or not (Mann and Whitney 1947). M-W is based on the bringing together and arranging of two groups. When these group members are lined up, a line number is assigned to each member. The membership status of these members (to which group they belong) is ignored. These line numbers are then summed up. The sum of the members of the first group is  $R_1$  and of the second group is  $R_2$ . The  $U$  values can then be calculated using

$$U_i = N_1N_2 + \frac{N_i(N_i + 1)}{2} - R_i, \quad (i = 1, 2) \tag{1}$$

After the calculation for  $i=1$  and  $i=2$ ,  $U_1$  and  $U_2$  are obtained, and the larger is chosen ( $U^*$ ) to determine the test statistics.

**Table 2** The large-scale common variables selected from the NCEP/NCAR reanalysis and GCM data set

Atmospheric levels	Atmospheric variables
200 hPa	air200 (°C) hgt200 (m)
500 hPa	air500 (°C) hgt500 (m) rhum500 (%)
850 hPa	air850 (°C) hgt850 (m) rhum850 (%)
Surface	air (°C) press (hPa) slp (hPa) prate (kg/m <sup>2</sup> )

*air* mean air temperature (in degree Celsius), *press* pressure (in hectopascal), *slp* sea level pressure (in hectopascal), *prate* precipitation (in kilogram per square meter), *hgt* geopotential height (in meter), *rhum* relative humidity (in percent)

$$z = \left| \frac{U^* - \frac{N_1N_2}{2}}{\sqrt{\frac{N_1N_2(N_1 + N_2 + 1)}{12}}} \right| \tag{2}$$

where  $N_1$  and  $N_2$  are the quantities of data for the groups compared.

For the values of  $z < z_{cr}$ , there is no significant difference between the first group and second group. (For 5 % level of significance,  $z_{cr}=1.96$ ). In this study, NCEP/NCAR variables including the period between 1948 and 2008 (61-year data) and observed local scale monthly precipitation and temperature series were divided into two subgroups as the first group includes 31-year data and the second group includes 30-year data in order to examine the homogeneities of climate data (Table 3).

In literature about statistical downscaling, studies have shown that the explanatory variables could vary from one region to another and any type of predictor can be used for downscaling as long as it has acceptable correlation with the local surface variables (Wilby et al. 1998; Tripathi et al. 2006; Fistikoglu and Okkan 2011; Okkan 2013). For this study, monthly observed precipitation and temperature were selected as the dependent variables, while potential predictors were air200, hgt200, air500, hgt500, rhum500, air850, hgt850, rhum850, air, press, slp, and prate in NCEP/NCAR reanalysis data set given in Table 2. The optimum predictors for both precipitation and temperature downscaling were determined with the help of best regression model structures denoted in Table 4 and Table 5 based on the adjusted determination coefficient and root mean squared error statistics corrected by considering the homogeneity test results presented in Table 3. Generating all possible subset regression models resulted in  $2^{12}-2=4,094$  different models (excluding full model and intercept-only model) for both precipitation and temperature. In this context, MINITAB (version 11) computer package was used to determine best subsets through these different model combinations. Considering the analysis for precipitation, not only data homogeneity but also a satisfactory correlation is provided with large-scale precipitation at surface level (prate), air temperature at 850 hPa (air850), and geopotential height at 200 hPa (hgt200) variables. By using the same procedure, only mean air temperature (air) variable of NCEP/NCAR data set was selected as monthly temperature predictor.

A similar methodology which has been applied to monthly precipitation downscaling in an earlier work (Fistikoglu and Okkan 2011) has been improved for this study concept at the same basin scale. In both papers, the precipitation is down-scaled with Levenberg–Marquardt algorithm-based feed forward neural networks method for a present climatic period while in this paper future projections are also produced.

**Table 3** Mann–Whitney  $U$  homogeneity test results both for NCEP/NCAR variables and observed monthly precipitation/temperature

			Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
prate	H.M.N	Mann–Whitney $U$	224	347	319	418	380	426	431.5	363	455.5	454	369	404
	10	$Z$	3.477	<b>1.702</b>	2.106	<b>0.678</b>	<b>1.226</b>	<b>0.563</b>	<b>0.495</b>	<b>1.533</b>	<b>0.137</b>	<b>0.159</b>	<b>1.385</b>	<b>0.880</b>
slp	H.M.N	Mann–Whitney $U$	278	292	331.5	247	275.5	266.5	256	256.5	233.5	271	334.5	293
	2	$Z$	2.698	2.496	<b>1.926</b>	3.145	2.734	2.864	3.015	3.008	3.340	2.799	<b>1.883</b>	2.481
press	H.M.N	Mann–Whitney $U$	267	297	322.5	243.5	263	240	242	239.5	217.5	255	335	283.5
	1	$Z$	2.856	2.424	2.056	3.195	2.914	3.246	3.217	3.253	3.571	3.030	<b>1.876</b>	2.618
air	H.M.N	Mann–Whitney $U$	407.5	406	371.5	379.5	376.5	302.5	352	337.5	353	372.5	458.5	455.5
	11	$Z$	<b>0.830</b>	<b>0.851</b>	<b>1.349</b>	<b>1.234</b>	<b>1.277</b>	2.344	<b>1.630</b>	<b>1.839</b>	<b>1.616</b>	<b>1.334</b>	<b>0.094</b>	<b>0.137</b>
air200	H.M.N	Mann–Whitney $U$	422	303.5	408.5	431.5	440	445.5	407	445.5	454	463	260	344.5
	10	$Z$	<b>0.636</b>	2.367	<b>0.828</b>	<b>0.498</b>	<b>0.372</b>	<b>0.285</b>	<b>0.881</b>	<b>0.299</b>	<b>0.163</b>	<b>0.030</b>	3.074	<b>1.775</b>
air500	H.M.N	Mann–Whitney $U$	260	465	398	427	392	197	445	457	335	335.5	440	439.5
	8	$Z$	3.024	<b>0.000</b>	<b>0.990</b>	<b>0.572</b>	<b>1.099</b>	4.226	<b>0.303</b>	<b>0.125</b>	2.022	1.962	<b>0.369</b>	<b>0.377</b>
air850	H.M.N	Mann–Whitney $U$	398.5	396	413.5	382.5	415	321.5	343.5	391	321	365.5	430.5	447.5
	10	$Z$	<b>0.977</b>	<b>1.007</b>	<b>0.752</b>	<b>1.209</b>	<b>0.739</b>	2.184	<b>1.799</b>	<b>1.091</b>	2.128	<b>1.454</b>	<b>0.506</b>	<b>0.258</b>
hgt200	H.M.N	Mann–Whitney $U$	196	413	307	366.5	340.5	218.5	414	366	261	289	390	355
	7	$Z$	3.881	<b>0.750</b>	2.279	<b>1.421</b>	<b>1.796</b>	3.556	<b>0.736</b>	<b>1.428</b>	2.943	2.539	<b>1.082</b>	<b>1.587</b>
hgt500	H.M.N	Mann–Whitney $U$	215	398	323	322	305	113	221	221.5	202	251	401	332
	3	$Z$	3.607	<b>0.967</b>	2.049	2.063	2.308	5.078	3.520	3.513	3.794	3.087	<b>0.923</b>	<b>1.919</b>
hgt850	H.M.N	Mann–Whitney $U$	243	314.5	296.5	248	249	196	196.5	167	206.5	211.5	350.5	262
	1	$Z$	3.203	2.171	2.431	3.131	3.116	3.881	3.874	4.299	3.729	3.657	<b>1.652</b>	2.929
rhum500	H.M.N	Mann–Whitney $U$	109	102.5	165.5	227.5	237.5	116	280	306.5	202	264.5	235	256.5
	0	$Z$	5.142	5.237	4.328	3.438	3.289	5.046	2.679	2.297	3.808	2.898	3.323	3.015
rhum850	H.M.N	Mann–Whitney $U$	307.5	406.5	290.5	337	330	359	459	404	413	434	402.5	399
	10	$Z$	2.277	<b>0.846</b>	2.524	<b>1.850</b>	<b>1.952</b>	<b>1.533</b>	<b>0.087</b>	<b>0.882</b>	<b>0.752</b>	<b>0.448</b>	<b>0.904</b>	<b>0.954</b>
Observed temperature	H.M.N	Mann–Whitney $U$	436	411	405.5	354.5	372.5	202.5	278	326.5	314	356.5	399.5	403.5
	8	$Z$	<b>0.419</b>	<b>0.779</b>	<b>0.931</b>	<b>1.595</b>	<b>1.335</b>	3.792	2.700	2.000	2.181	<b>1.566</b>	<b>0.945</b>	<b>0.888</b>
Observed precipitation	H.M.N	Mann–Whitney $U$	385	391	447	435	377.5	316.5	397.5	305	317	394	317	440
	8	$Z$	<b>1.154</b>	<b>1.068</b>	<b>0.260</b>	<b>0.433</b>	<b>1.262</b>	2.143	<b>0.994</b>	2.344	2.139	<b>1.024</b>	2.135	<b>0.361</b>

Bold values denote homogeneous months.  $Z_{cr}$ , 1.96

*H.M.N* homogeneous month numbers

The common belief is based on selecting the subset regression model structures with highest  $R^2$  and  $Adj.R^2$  and with the smallest root mean squared error as the best one. In the previous study, Fistikoglu and Okkan (2011) have selected nine predictors in precipitation downscaling modeling considering this common belief. On the other hand, non-homogenous variables may reduce the statistical downscaling model efficiency. With redundant variables, there is also a potential of overfitting the data. In the study presented, we were able to reduce the number of precipitation predictors using proposed procedure involving both homogeneity test and all possible regression method in order to efficiently create a statistical downscaling model. Moreover, it can be said that changes on adjusted determination coefficient ( $Adj.R^2$ ) and root mean squared error (RMSE) statistics are no longer significant when the number of predictors exceeds three. In other words, performance of the regression-based

model with three predictors is nearly the same as that of the full model with 12 variables.

### 3 Developed downscaling strategy

In this study, statistical relationships based on an ANN algorithm were developed between large-scale climatic variables and local surface observed variables, with selected reanalysis data as predictors and observed precipitation and temperature as predictands. These relationships were assessed to model the future precipitation and temperature using CGCM3 outputs.

ANN, which has been recently used in many fields, could be defined as a black box model producing outputs against inputs and is used as one of the applied frequently used statistical downscaling techniques (Fistikoglu and Okkan 2011). Among ANN algorithms, the feed forward neural



**Table 4** Summary of the best results derived from all possible regression analyses of monthly precipitation predictors

Model number	$R^2$	Adj $R^2$	RMSE (mm)	air850	hgt850	rhum850	air500	hgt500	rhum500	air200	hgt200	air	press	slp	prate
1	0.563	0.563	53.639												•
2	0.612	0.611	50.626	•											•
3	0.632	0.630	49.325	•							•				•
4	0.634	0.632	49.208	•	•						•				•
5	0.648	0.646	48.296	•				•		•	•				•
6	0.650	0.647	48.205	•				•		•	•	•			•
7	0.651	0.648	48.125	•				•		•	•		•	•	•
8	0.654	0.650	48.009	•	•		•			•	•	•	•		•
9	0.656	0.652	47.885	•	•		•			•	•	•	•	•	•
10	0.658	0.653	47.797	•	•		•		•	•	•	•	•	•	•
11	0.659	0.654	47.716	•	•		•	•	•	•	•	•	•	•	•
12	0.660	0.654	47.699	•	•	•	•	•	•	•	•	•	•	•	•

*air* mean air temperature (in degree Celsius), *press* pressure (in hectopascal), *slp* sea level pressure (in hectopascal), *prate* precipitation (in kilogram per square meter), *hgt* geopotential height (in meter), *rhum* relative humidity (in percent)

networks (FFNN) type is the most frequently preferred algorithm. There are many literatures which provide a detailed description of the FFNN (Hagan and Menhaj 1994; Ham and Kostanic 2001; Fistikoglu and Okkan 2011), and hence only a brief description of FFNN is given here.

The basic concept of the FFNN is that they are typically made up of neurons. And in FFNN, the neurons are organized in the form of layers. The first and last layer of FFNN is called the input and the output layer, respectively. The input layer does not perform any computations but only serves to feed the input data to the hidden layer which is between the input and output layers. In general, there can be any number of hidden layers in the FFNN structures; however, from practical applications, only one or two hidden layers are used. In addition to this, the number of hidden layers and also the number of

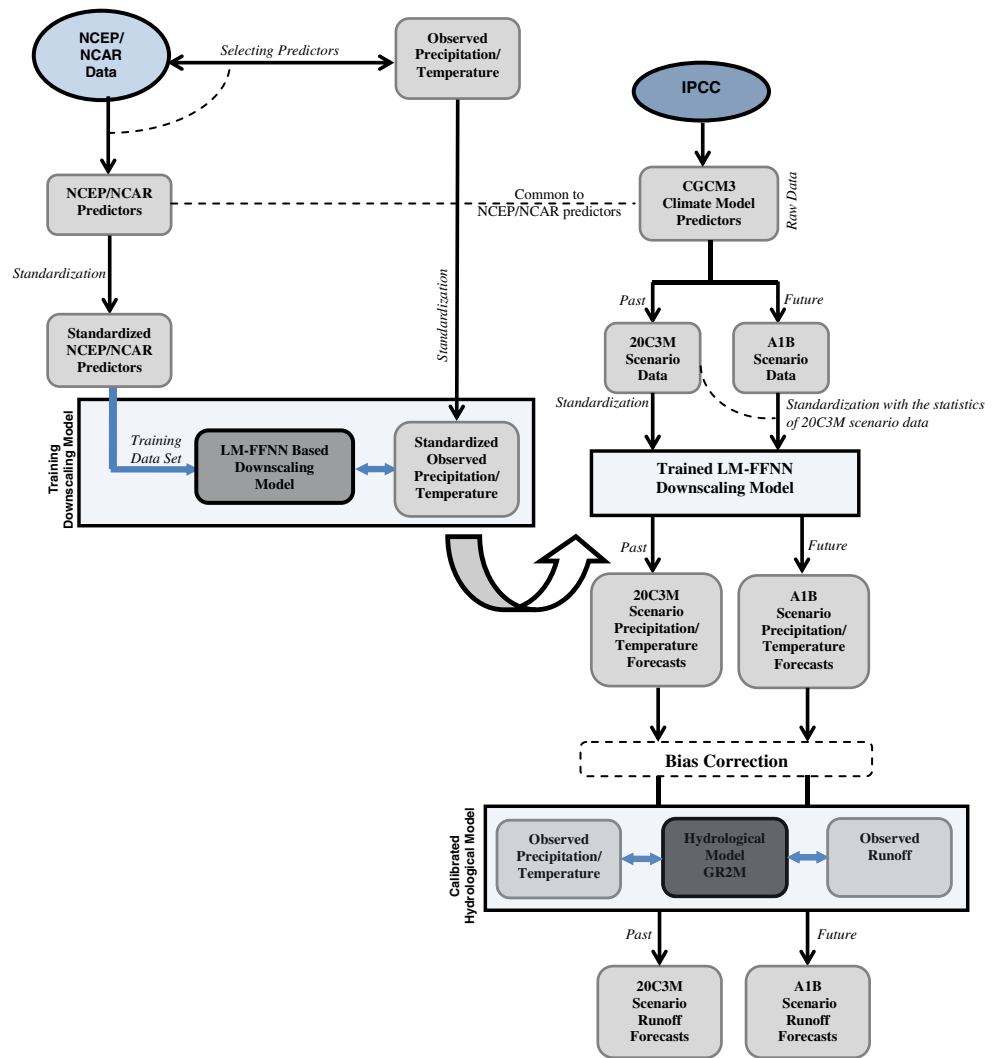
neurons of hidden layers can be determined by the trial and error (Ham and Kostanic 2001). There are also three important components of an FFNN structure: weights, summing function, and activation function. The importance and functionality of the inputs on ANN models are obtained with weights ( $W$ ). So, the success of the model depends on the precise and correct determination of weight values. The summing function (*net*) acts to add all outputs; that is, each neuron input is multiplied by the weights and then summed. After computing the sum of weighted inputs for all neurons, the activation function  $f(\cdot)$  serves to limit the amplitude of these values. The activation functions are usually continuous, non-decreasing, and bounded functions. Various types of the activation function are possible but generally log-sigmoid function is preferred in applications (Ham and Kostanic 2001).

**Table 5** Summary of the best results derived from all possible regression analyses of monthly temperature predictors

Model number	$R^2$	Adj $R^2$	RMSE (°C)	air850	hgt850	rhum850	air500	hgt500	rhum500	air200	hgt200	air	press	slp	prate
1	0.986	0.986	0.822									•			
2	0.988	0.988	0.777									•		•	
3	0.989	0.989	0.741									•	•	•	
4	0.989	0.989	0.728		•			•			•	•			
5	0.990	0.990	0.712		•			•		•	•	•			
6	0.990	0.990	0.703	•					•		•	•	•	•	
7	0.990	0.990	0.699	•	•				•		•	•	•	•	
8	0.990	0.990	0.696	•	•				•		•	•	•	•	
9	0.990	0.990	0.692	•	•			•	•	•	•	•	•	•	
10	0.990	0.990	0.692	•	•		•	•	•	•	•	•	•	•	
11	0.990	0.990	0.692	•	•	•		•	•	•	•	•	•	•	•
12	0.990	0.990	0.692	•	•	•	•	•	•	•	•	•	•	•	•

*air* mean air temperature (in degree Celsius), *press* pressure (in hectopascal), *slp* sea level pressure (in hectopascal), *prate* precipitation (in kilogram per square meter), *hgt* geopotential height (in meter), *rhum* relative humidity (in percent)

**Fig. 2** Proposed modeling strategy in the scope of the study



This activation function generates outputs between 0 and 1 as the input signal goes from negative to positive infinity.

$$f(\cdot) \cong \frac{1}{1 + e^{-\cdot}} \tag{3}$$

In addition to the structure and its components of FFNN, the running procedure is also important which involves typically two phases: forward computing and backward computing.

In forward computing, each layer uses a weight matrix ( $W^{(v)}$ , for  $v=1, 2$ ) associated with all the connections made from the previous layer to the next layer. The hidden layer has the weight matrix  $W^{(1)} \in R^{h \times n}$ , the output layer's weight matrix is  $W^{(2)} \in R^{m \times h}$ . Given the network input vector  $x \in R^{n \times 1}$ , the output of the hidden layer  $x_{out,1} \in R^{h \times 1}$  can be written as

$$x_{out,1} = f^{(1)}[net^{(1)}] = f^{(1)}[W^{(1)}x] \tag{4}$$

which is the input to the output layer. The output of the output layer, which is the response (output) of the network  $y = x_{out,2} \in R^{m \times 1}$ , can be written as

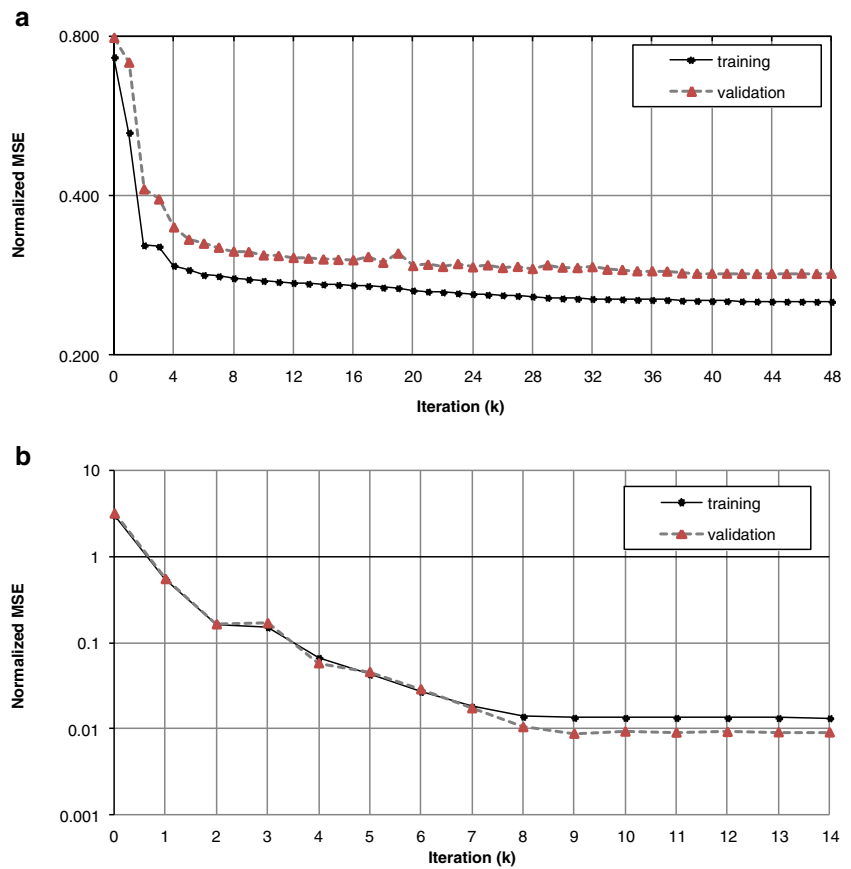
$$y = x_{out,2} = f^{(2)}[net^{(2)}] = f^{(2)}[W^{(2)}x_{out,1}] \tag{5}$$

Substituting (Eq. 4) into (Eq. 5) for  $x_{out,1}$  gives the final output  $y = x_{out,2}$  of the network as

$$y = f^{(2)}[W^{(2)}f^{(1)}[W^{(1)}x]] \tag{6}$$

After the phase of forward computing, backward computing which depending on the algorithms to adjust weights is used in the ANN. The process of adjusting these weights to minimize the differences between the actual and the desired output values is called training or learning the network. If these differences (error) are higher than the desired values, the errors are passed backwards through the weights of the

**Fig. 3** Error performances of training and validation sets for **a** precipitation downscaling model and **b** temperature downscaling model



network. In ANN terminology, this phase is also called the backpropagation. Once the comparison error is reduced to an acceptable level for the whole training set, the training period ends, and the network is also tested for another known input and output data set in order to evaluate the generalization capability of the ANN (Ham and Kostanic 2001).

Depending on the techniques to train FFNN models, different backpropagation algorithms have been used for modeling studies. These modeling studies generally include the standard backpropagation (BP) algorithms such as gradient-descent, gradient-descent with momentum rate, conjugate gradient, etc. As the BP algorithms have some disadvantages relating to the time requirement and slow convergency in training, Levenberg–Marquardt algorithms, which are alternative approaches to standart BP algorithms, were used in some applications (Fistikoglu and Okkan 2011; Okkan 2011).

In this study, Levenberg–Marquardt algorithm (LM-FFNN) was used for training. This algorithm is a second-order nonlinear optimization technique that is usually faster and more reliable than any other standard back propagation techniques. It represents a simplified version of Newton's method (Marquardt 1963) applied to the training FFNN (Hagan and Menhaj 1994).

Considering FFNN structure, the running of the network training can be viewed as finding a set of weights that minimized

the error ( $e_p$ ) for all samples in the training set ( $Q$ ). If the performances function is a sum of squares of the errors as

$$E(W) = \frac{1}{2} \sum_{p=1}^P (d_p - y_p)^2 = \frac{1}{2} \sum_{p=1}^P (e_p)^2, P = mT \quad (7)$$

where  $Q$  is the total number of training samples,  $m$  is the number of output layer neurons,  $W$  represents the vector containing all the weights in the network,  $y_p$  is the actual network output, and  $d_p$  is the desired output.

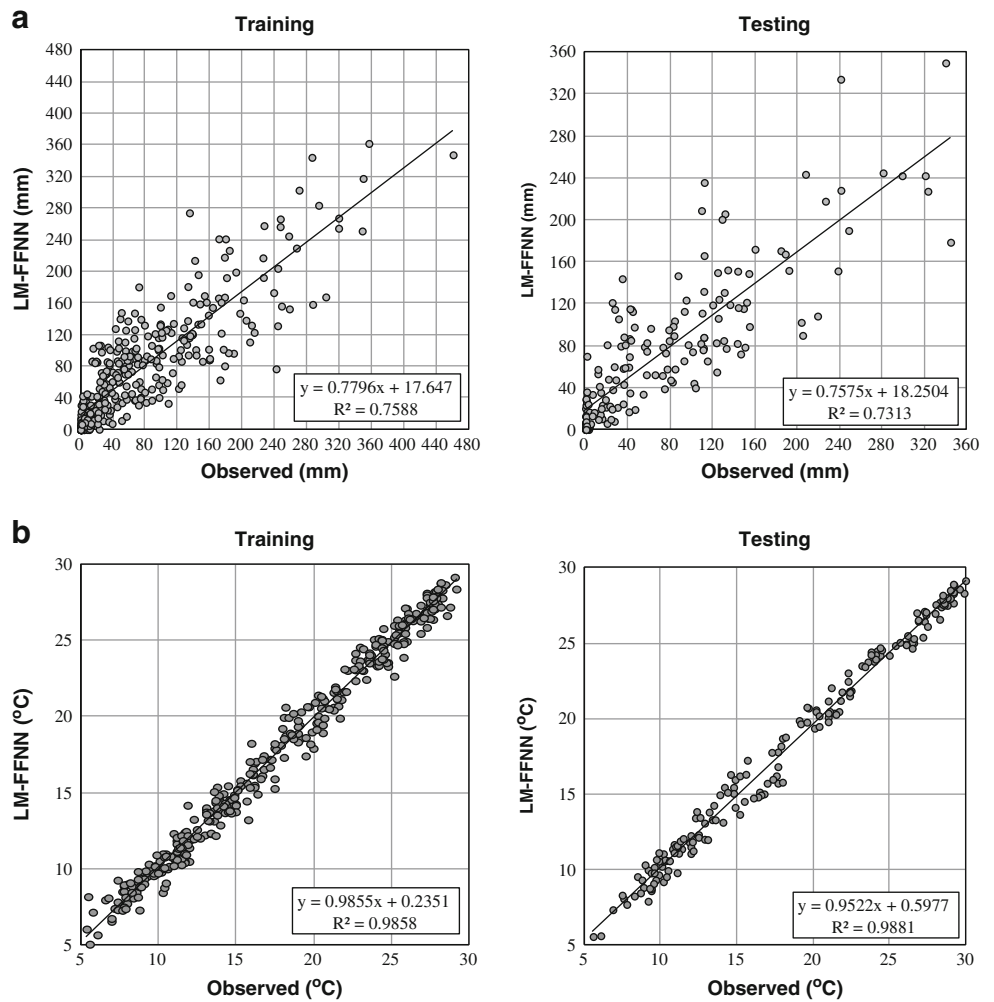
When training with the Levenberg–Marquardt optimization algorithm, the changing of weights  $\Delta W$  can be computed as follows

$$\Delta W_k = -[J_k^T J_k + \mu_k I]^{-1} J_k^T e_k \quad (8)$$

where  $J$  is the Jacobian matrix,  $I$  is the identify matrix,  $\mu$  is the Marquardt parameter which is to be updated using the decay rate  $\beta$  depending on the outcome. In particular,  $\mu$  is multiplied by the decay rate  $\beta$  ( $0 < \beta < 1$ ) whenever  $E(W)$  decreases, while  $\mu$  is divided by  $\beta$  whenever  $E(W)$  increases in a new step ( $k$ ).



**Fig. 4** Scatter plots of precipitation downscaling model (a) and temperature downscaling model (b) for training and testing periods



The LM-FFNN training process can be illustrated in the following pseudo codes,

1. Initialize the weights and  $\mu$  ( $\mu=0.001$  is appropriate).
2. Compute the sum of squared errors over all inputs,  $E(W)$ .
3. Compute the Jacobian matrix  $J$ .
4. Solve Eq. 8 to obtain the changing of weights  $\Delta W$ .
5. Recompute the sum of squared errors  $E(W)$  using

$$W_{(k+1)} = W_{(k)} - [J_k^T J_k]^{-1} J_k^T e_k$$

as the trial  $W$ , and judge

IF trial  $E(W) < E(W)$  in Step 2, THEN

$$W_{(k+1)} = W_{(k)} - [J_k^T J_k]^{-1} J_k^T e_k$$

$$\mu_{(k+1)} = \mu_k \beta (\beta = 0.1)$$

go back to Step 2.

ELSE

$$\mu_{(k+1)} = \mu_k / \beta$$

go back to Step 4.

END IF

In this study, the flowchart of proposed downscaling strategy, which was constructed by a MATLAB code, summarized in Fig. 2 is considered. According to this flowchart, observed precipitation/temperature data and the selected NCEP/NCAR predictors were turned into standardized series before being presented to the LM-FFNN-based downscaling model. Wilby et al. (2004) have been emphasized that standardization procedure is used prior to downscaling to reduce biases in the mean and variances of GCM outputs relative to the observations or NCEP/NCAR variables. This procedure involves subtraction of mean and division by standard deviation of the related variable. The variables for CGCM3 related to A1B future scenario were also standardized using the mean and standard deviation statistics of 20C3M scenario variables including 1950–1999 baseline period.

For developing the downscaling model between standardized precipitation/temperature and standardized NCEP/NCAR predictors, data set were divided in three stages including training (50 %), validation (25 %), and testing (25 %) under certain proportions. In the training of LM-FFNN, the early stopping method is used and overtraining of the network was aimed to be prevented.

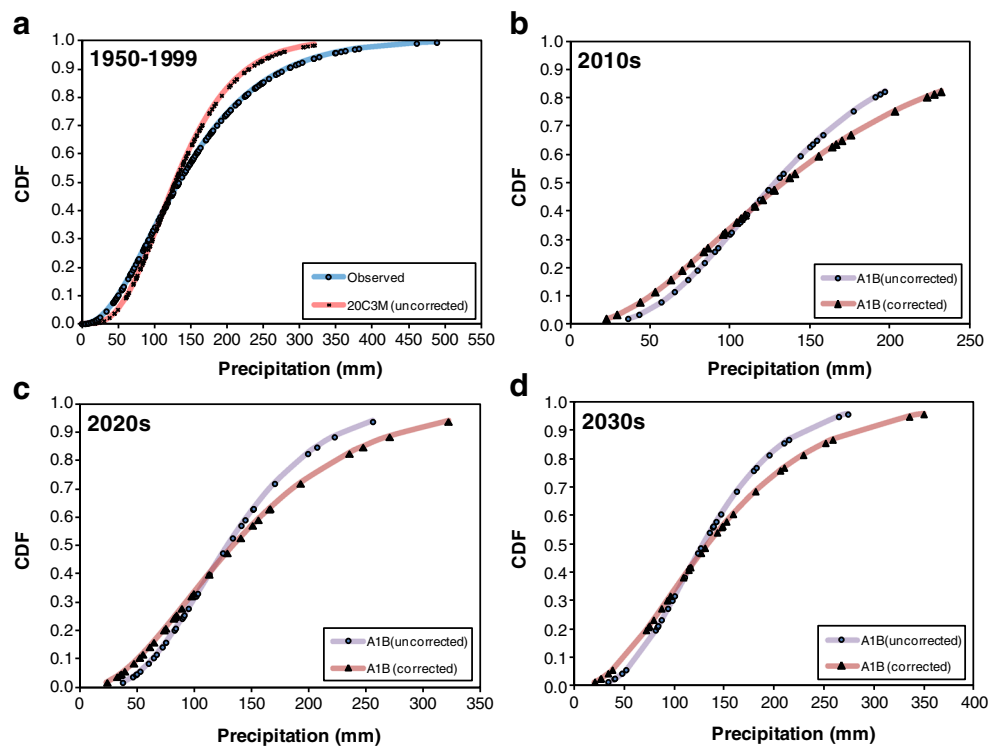
**Table 6** Performance measures of developed precipitation downscaling model and temperature downscaling model

Precipitation downscaling model								
Training								
	MSE (mm <sup>2</sup> )	RMSE (mm)	WMAPE	R <sup>2</sup>	Adj.R <sup>2</sup>	NS	Min. (mm)	Max. (mm)
Observed	–	–	–	–	–	–	0.00	461.09
LM-FFNN	<b>1,548.6620</b>	<b>39.3530</b>	<b>0.3911</b>	<b>0.7588</b>	<b>0.7568</b>	<b>0.7572</b>	<b>0.00</b>	<b>362.22</b>
MLR	2,195.2419	46.8534	0.4952	0.6579	0.6551	0.6559	–41.39	320.93
Testing								
	MSE (mm <sup>2</sup> )	RMSE (mm)	WMAPE	R <sup>2</sup>	Adj. R <sup>2</sup>	NS	Min. (mm)	Max. (mm)
Observed	–	–	–	–	–	–	0.00	344.14
LM-FFNN	<b>1,662.4714</b>	<b>40.7734</b>	<b>0.3969</b>	<b>0.7313</b>	<b>0.7267</b>	<b>0.7300</b>	<b>0.00</b>	<b>350.03</b>
MLR	2,233.0729	47.2554	0.4950	0.6377	0.6315	0.6374	–31.85	212.67
Temperature downscaling method								
Training								
	MSE (°C <sup>2</sup> )	RMSE (°C)	WMAPE	R <sup>2</sup>	Adj.R <sup>2</sup>	NS	Min. (°C)	Max. (°C)
Observed	–	–	–	–	–	–	5.40	29.20
LM-FFNN	<b>0.6591</b>	<b>0.8119</b>	<b>0.0357</b>	<b>0.9858</b>	<b>0.9857</b>	<b>0.9858</b>	5.05	<b>29.19</b>
MLR	0.7919	0.8899	0.0398	0.9839	0.9838	0.9829	<b>5.47</b>	30.68
Testing								
	MSE (°C <sup>2</sup> )	RMSE (°C)	WMAPE	R <sup>2</sup>	Adj.R <sup>2</sup>	NS	Min. (°C)	Max. (°C)
Observed	–	–	–	–	–	–	5.60	30.00
LM-FFNN	<b>0.7582</b>	<b>0.8708</b>	<b>0.0400</b>	<b>0.9881</b>	<b>0.9880</b>	<b>0.9853</b>	<b>5.55</b>	<b>29.15</b>
MLR	1.0567	1.0279	0.0478	0.9876	0.9875	0.9796	6.21	31.71

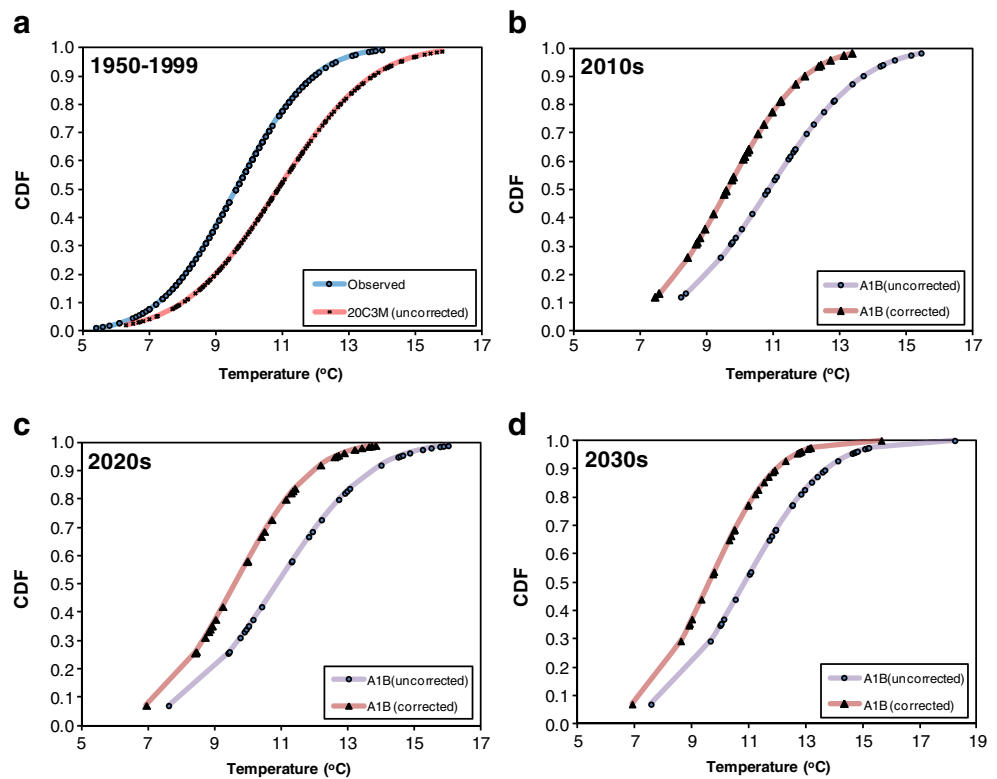
MSE mean squared error, RMSE root mean squared error, WMAPE weighted mean absolute percentage error, R<sup>2</sup> determination coefficient; Adj.R<sup>2</sup> adjusted determination coefficient, NS Nash–Sutcliffe coefficient

The bold values denote the best results

**Fig. 5** Correction for bias in downscaling GCM simulations for winter precipitation



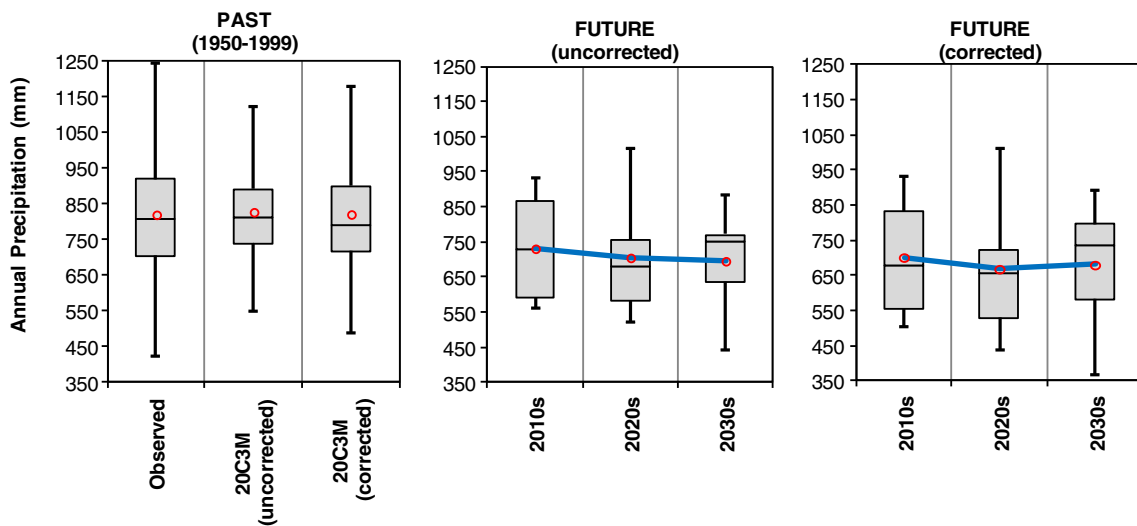
**Fig. 6** Correction for bias in downscaling GCM simulations for winter temperature



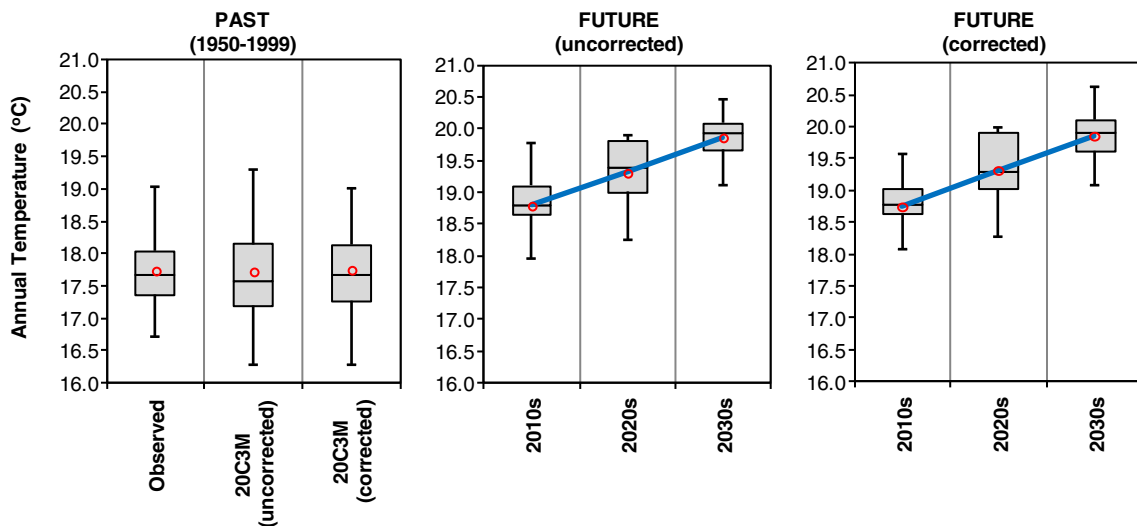
**4 Downscaling results corresponding to climate change scenarios**

In the training of the LM-FFNN models, the numbers of neurons in the hidden layer were determined by trial and error approach. The number of neurons in the hidden layer making the statistical performance of training and testing sets highest for the precipitation and temperature downscaling modeling were determined as

9 and 3, respectively, their  $\mu_0$  parameters were selected as 0.001, and their  $\beta$  parameters chosen as 0.1. Various types of the activation function are possible for feed forward neural networks but sigmoid activation function and linear activation were used for hidden and output layers, respectively. When the early stopping method was considered, the training of precipitation and temperature downscaling models were finalized in 48 and 14 iterations, respectively (Fig. 3).



**Fig. 7** Uncorrected and corrected precipitation forecasts produced from the trained downscaling model for annual periods (The horizontal lines in the middle of the boxes and the circles represent the median and mean values, respectively. The blue lines joining the circles denote the precipitation trends)



**Fig. 8** Uncorrected and corrected temperature forecasts produced from the trained downscaling model for annual periods (The *horizontal lines in the middle of the boxes* and the *circles* represent the median and mean values, respectively. The *blue lines joining the circles* denote the temperature trends)

Some statistical approaches are suggested for modeling accuracy evaluation according to literature related to training and testing of models. In this study, five statistical performance measures were considered (Nash and Sutcliffe 1970; Krause et al. 2005; Nayak et al. 2005). LM-FFNN model with optimum parameters and structure provided the best training result in terms of the minimum RMSE, weighted mean absolute percentage error (WMAPE), and the maximum determination coefficient ( $R^2$ ), the maximum adjusted determination coefficient ( $Adj.R^2$ ), and Nash–Sutcliffe efficiency (NS) were also employed for the testing period. RMSE statistics evaluates the residual between desired and output data, and WMAPE measures the weighted mean absolute percentage error of the prediction.  $R^2$  and  $Adj.R^2$  evaluate the power of regression relation between desired and output data, while NS evaluates the capability of the model in simulating output data away from the mean statistics. The detailed descriptions of

these statistical performance measures were presented by Okkan and Serbes (2012).

The scatter plots of the LM-FFNN models are presented in Fig. 4, statistical performance measures are presented in Table 6. It can be observed in Table 6 that LM-FFNN models have good performances during both training and testing periods, and they outperform multiple linear regression models (MLR) in terms of the different performance measures. Thus, by using precipitation at surface (prate), air temperature at 850 hPa (air850), and geopotential height at 200 hPa (hgt200) as an input of LM-FFNN-based downscaling model, determination coefficient ( $R^2$ ) values of training and testing period were obtained as 75.88 and 73.13 %; RMSE were 39.35 and 40.77 mm, respectively. Moreover, by using mean air temperature at surface level (air) as an input of LM-FFNN-based downscaling model, the determination coefficient ( $R^2$ ) values of training and testing period were

**Table 7** Before and after bias correction, the annual descriptive statistics of precipitation and temperature

	Precipitation					
	Mean (mm/year)	Std. Dev. (mm/year)	Cv (%)	Upper whisker (mm/year)	Lower whisker (mm/year)	Range (mm/year)
Observed	820.03	188.89	23.0	1,244.46	421.67	216.49
20C3M (uncorrected)	827.31	138.15	16.7	1,122.93	546.93	153.80
20C3M (corrected)	820.97	168.51	20.5	1179.68	488.35	185.10
	Temperature					
	Mean (°C/year)	Std. Dev. (°C/year)	Cv (%)	Upper whisker (°C/year)	Lower whisker (°C/year)	Range (°C/year)
Observed	17.74	0.51	2.9	19.04	16.71	0.67
20C3M (uncorrected)	17.72	0.71	4.0	19.30	16.28	0.96
20C3M (corrected)	17.75	0.56	3.2	19.02	16.29	0.87

*Std. Dev* standard deviation, *Cv* variation coefficient, *Range* third quartile–first quartile

**Table 8** The mean statistics (in millimeter) of corrected scenario forecasts and examining statistically significant changes (in percent) on precipitation for future periods ( $\alpha=5\%$ ; italic values indicate the non-homogeneous future periods in terms of medians, and bold values show

that means of the forecasts representing the future climate are significantly different from means of 20C3M scenario representing the climate of the past period)

Scenario	Period	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
Mean statistics (in millimeter) of corrected scenario forecasts														
20C3M	1950–1999	152.14	122.14	93.92	59.21	26.71	13.08	0.24	0.61	8.85	48.49	109.59	185.99	820.97
A1B	2010s	113.31	99.31	89.39	62.78	29.35	7.01	0.24	0.48	4.62	28.35	115.48	151.37	701.68
	2020s	137.75	87.05	81.89	55.59	40.83	10.30	0.79	4.19	5.39	23.97	104.61	116.22	668.56
	2030s	153.08	98.83	72.53	55.33	27.52	1.63	4.97	6.76	7.45	16.80	78.67	156.63	680.20
Statistically significant changes (in percent) on precipitation for future periods														
A1B	2010s	-25.52	<b>-18.70</b>	-4.83	6.03	9.85	<i>-46.42</i>	-0.35	<b>-21.14</b>	<i>-47.77</i>	<i>-41.53</i>	5.37	<b>-18.61</b>	<b>-14.53</b>
	2020s	-9.46	<i>-28.73</i>	<i>-12.81</i>	<i>-6.12</i>	<i>52.83</i>	<b>-21.21</b>	<i>231.51</i>	<i>588.63</i>	<i>-39.04</i>	<i>-50.57</i>	<i>-4.55</i>	<i>-37.51</i>	<i>-18.56</i>
	2030s	0.62	<b>-19.09</b>	<i>-22.78</i>	<i>-6.55</i>	3.03	<i>-87.56</i>	<i>1,992.29</i>	<i>1,012.08</i>	<b>-15.79</b>	<i>-65.35</i>	<i>-28.21</i>	<b>-15.78</b>	<i>-17.15</i>

obtained as 98.58 and 98.81 %; RMSE were obtained as 0.81 and 0.87 °C, respectively. When the statistical performance measures are considered, proposed statistical downscaling models are found satisfactory to simulate the scenario results of CGCM3.

After the training of the downscaling models using NCEP/NCAR data set, CGCM3 variables including precipitation at surface level (prate), air temperature at surface level (air), air temperature at 850 hPa (air850) level, and geopotential height at 200 hPa (hgt200) level both for the past and the future periods were compiled. According to the downscaling strategy summarized in Fig. 2, the compiled scenario variables of CGCM3 were used as the new inputs of the trained statistical downscaling models. Therefore, the precipitation and temperature forecasts at basin scale for 20C3M (1950–1999) scenario representing the climate of the past period and A1B scenario representing the future climate were obtained.

For validation purpose, seasonal precipitation and temperature were evaluated for the baseline period of years 1950–1999 with downscaled CGCM3 20C3M scenario outputs. In this stage, several parametric probability distributions were fitted to the seasonal precipitation and temperature and the best probability density functions were chosen among them with the help of Anderson–Darling tests. Considering these goodness of fit tests, Gamma distribution was found suitable for both seasonal precipitation and temperature. The cumulative distribution functions (CDFs) obtained from observed data and GCM outputs (downscaled 20C3M scenario outputs), using probability plotting formulas of selected distributions, are presented for only winter seasons due to space limitation (Fig. 5a and Fig. 6a).

According to CDF results, downscaled GCM outputs for 20C3M scenario have significant deviations from the observed data and such biases are detected for each season (bias=

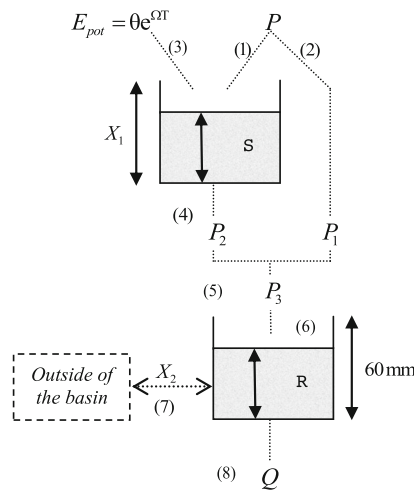
**Table 9** The mean statistics (in centigrade) of corrected scenario forecasts and examining statistically significant changes ( $\Delta t=t_2-t_1$ ) on temperature for future periods ( $\alpha=5\%$ ; italic values indicate the non-homogeneous future periods in terms of medians, and bold values show

that means of the forecasts representing the future climate are significantly different from means of 20C3M scenario representing the climate of the past period)

Scenario	Period	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
Mean statistics (in centigrade) of corrected scenario forecasts														
20C3M	1950–1999	8.84	8.34	11.96	15.27	20.49	24.96	27.26	28.09	22.97	19.57	13.53	11.68	17.75
A1B	2010s	9.56	9.28	13.23	17.63	22.28	25.63	27.96	28.78	24.36	20.30	13.88	12.11	18.75
	2020s	9.74	9.51	14.66	18.21	23.27	25.81	28.07	28.83	24.42	21.16	15.15	13.00	19.32
	2030s	10.42	9.87	16.05	18.80	24.25	26.36	28.34	29.02	25.23	21.87	15.05	13.02	19.86
Statistically significant changes on temperature for future periods														
A1B	2010s	0.72	<i>0.94</i>	1.27	<i>2.36</i>	<i>1.79</i>	<i>0.67</i>	<i>0.70</i>	<i>0.69</i>	<i>1.39</i>	0.73	0.35	0.43	<i>1.00</i>
	2020s	<b>0.90</b>	<i>1.17</i>	<i>2.70</i>	<i>2.94</i>	<i>2.78</i>	<i>0.85</i>	<i>0.81</i>	<i>0.74</i>	<i>1.45</i>	<i>1.59</i>	<i>1.62</i>	<i>1.32</i>	<i>1.57</i>
	2030s	<i>1.58</i>	<i>1.53</i>	<i>4.09</i>	<i>3.53</i>	<i>3.76</i>	<i>1.40</i>	<i>1.08</i>	<i>0.93</i>	<i>2.26</i>	<i>2.30</i>	<i>1.52</i>	<i>1.34</i>	<i>2.11</i>



**Fig. 9** Diagram of the hydrological model GR2M (Mouelhi et al. 2006)



$$(1) \quad S_1 = \frac{S + X_1 \phi}{1 + \phi \frac{X_1}{S}}; \quad \phi = \tanh\left(\frac{P}{X_1}\right)$$

$$(2) \quad P_1 = P + S - S_1$$

$$(3) \quad S_2 = \frac{S_1(1 - \psi)}{1 + \psi \left(1 - \frac{S_1}{X_1}\right)}; \quad \psi = \tanh\left(\frac{E_{pot}}{X_1}\right)$$

$$(4) \quad S = \frac{S_2}{\left[1 + \left(\frac{S_2}{X_1}\right)^3\right]^{1/3}}; \quad P_2 = S_2 - S;$$

$$(5) \quad P_3 = P_1 + P_2$$

$$(6) \quad R_1 = R + P_3$$

$$(7) \quad R_2 = X_2 R_1$$

$$(8) \quad Q = \frac{R_2^2}{R_2 + 60}; \quad R = R_2 - Q$$

observed data–simulated GCM output). The biases may be caused by partial ignorance about geophysical process, assumptions for numerical modeling, and parameterization. In the phase of modeling hydrometeorological variables under the climate change scenarios, such biases should be taken into consideration; otherwise, it will propagate in the computations for future periods (Ghosh and Mujumdar 2008). The standardization procedure before statistical downscaling can reduce the biases in the mean and standard deviations of the predictors but the biases in large-scale patterns of atmospheric circulation in GCMs or unrealistic intervariable relationships may not be removed by only standardization procedure (Wilby and Dawson 2004; Ghosh and Mujumdar 2008).

Ghosh and Mujumdar (2008) have proposed a methodology after downscaling process to remove such biases from downscaled outputs. The proposed methodology consists of following steps for this study content:

- CDFs are obtained with observed and downscaled CGCM3 20C3M scenario outputs for the 1950–1999 period using the determined probability potting position formulas.
- For a given value of generated output under 20C3M scenario (seasonal uncorrected precipitation/temperature), the value of CDF is computed.

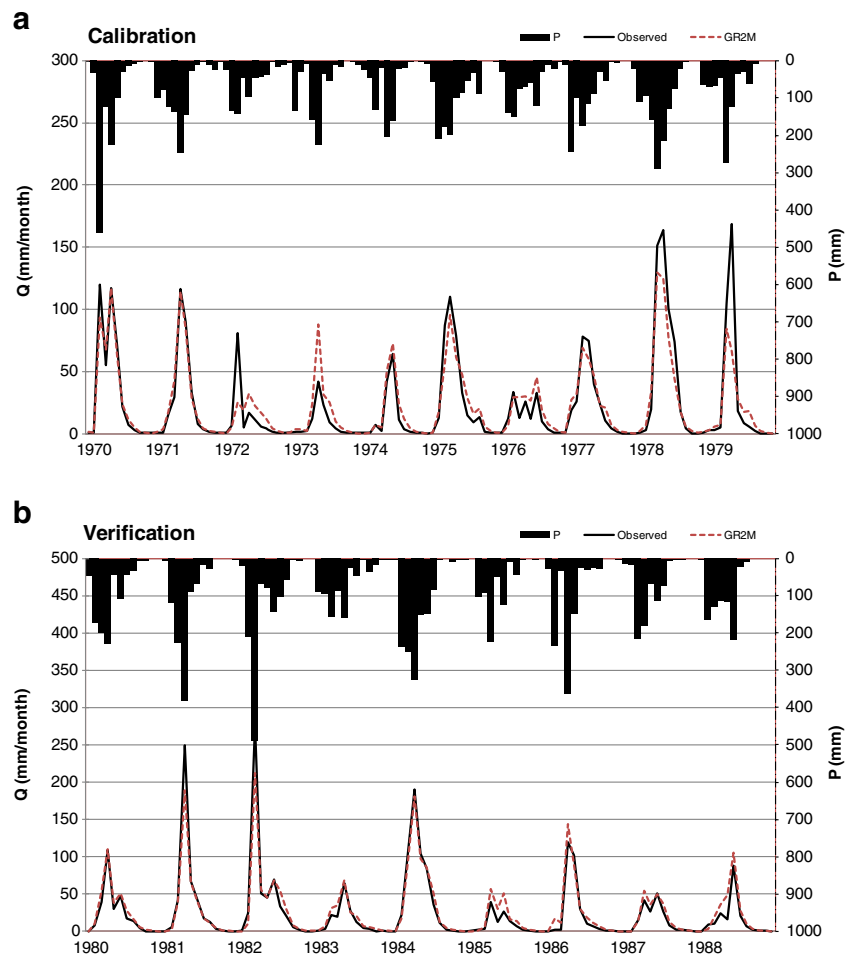
- Corresponding to CDF value of used GCM the observed precipitation/temperature values are obtained from the CDFs of observed data.
- The CGCM3 generated precipitation/temperature are replaced by the observed data, thus computed, having the same CDF values.
- The CDFs of CGCM3 generated and observed precipitation/temperature, obtained for the 1950–1999 periods, act as reference, and based on them the corrections are applied to precipitation/temperature values obtained from CGCM3 model for future periods.

Thus, corrected 20C3M and A1B scenario outputs were produced for both seasonal precipitation and temperature. The box plot graphs of corrected downscaled annual total precipitation and annual mean temperature are presented in Fig. 7 and Fig. 8. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the extreme data points, not considered as outliers, are indicated. After bias correction, descriptive statistics of precipitation and temperature are presented in Table 7. For example, the annual mean and standard deviation of observed precipitation are 820.03 and 188.89 mm and those of CGCM3 20C3M scenario precipitation before bias correction were 827.31 and 138.15 mm. After bias correction process for precipitation, the annual mean and standard deviation statistics are 820.97 and 168.51 mm, respectively, which show bias has been significantly reduced. Similarly, after bias correction for temperature, obtained statistics prove that bias has been reduced. Other statistics involving upper/lower whisker

**Table 10** Calibrated parameters of hydrological model for Tahtali River

$X_1$ (mm)	$X_2$	$\theta$	$\Omega$
484.778	0.705	6.422	0.096

**Fig. 10** Observed and modeled runoff of Tahtali river for (a) calibration period and (b) verification period



and range values are as much as possible corrected. The CDFs projected near future precipitation and temperature of winter season are presented for 10-year time slices 2010s, 2020s, and 2030s in Fig. 5b–d and Fig. 6b–d.

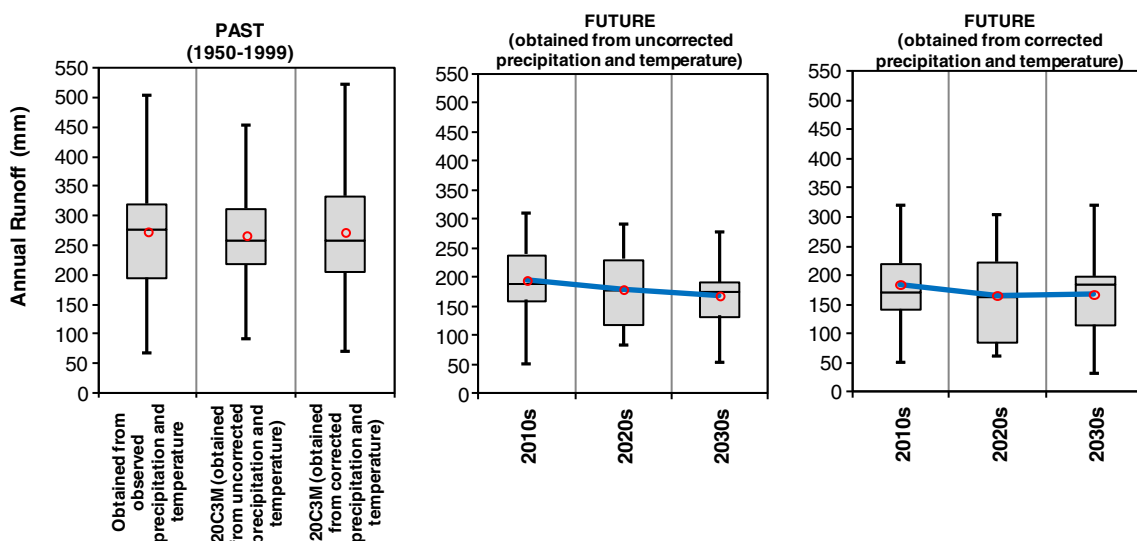
After bias correction analyses, some statistical tests including M-W homogeneity and two-sample  $t$  test of equality of mean statistics were applied to examine the significances of computed variations. For all projection periods (2010s, 2020s,

and 2030s), these tests are applied at 5 % level of significance to check whether means and homogeneities of the forecasted series at basin scale for the A1B scenario representing the future climate are significantly different from the statistics for 20C3M (1950–1999) representing the past periods (Table 8 and Table 9). The downscaled precipitation results were evaluated on scenario basis, and they were interpreted in terms of their mean statistics and homogeneities as follows. According

**Table 11** Performance measures of calibrated hydrological model GR2M for the calibration and verification periods

	MSE (mm <sup>2</sup> )	RMSE (mm)	WMAPE (%)	$R^2$ (%)	Adj. $R^2$ (%)	NS (%)	Mean (mm)	Std. Dev. (mm)	Min. (mm)	Max. (mm)
Calibration										
Observed	–	–	–	–	–	–	22.52	37.47	0.00	168.45
Modeled	218.95	14.80	33.93	85.77	85.53	84.27	22.82	30.12	0.29	129.32
Verification										
Observed	–	–	–	–	–	–	24.98	45.86	0.02	276.86
Modeled	135.23	11.63	23.36	94.52	94.41	93.51	26.96	40.45	0.14	211.81

*MSE* mean squared error, *RMSE* root mean squared error, *WMAPE* weighted mean absolute percentage error,  $R^2$  determination coefficient, *Adj.  $R^2$*  adjusted determination coefficient, *NS* Nash–Sutcliffe coefficient



**Fig. 11** Simulated runoff forecasts obtained from uncorrected/corrected precipitation and temperature for annual periods (The horizontal lines in the middle of the boxes and the circles represent the median and mean values, respectively. The blue lines joining the circles denote the runoff trends)

to the annual statistics, there are foreseen decreases for all projection periods. In the 2020s, it is foreseen that there will be decrease at the levels of 18.6%. Significant changes on annual precipitation are not foreseen during the 2010s. Similar analyses were carried out for annual variance statistics by applying *f* test and, in general, significant changes were not seen in this statistics for future. After assessing precipitation statistics, similar tests have been applied at 5% level of significance to check whether the future temperature statistics are significantly different from the 20C3M scenario statistics. It is found that, for all projection periods, the homogeneities and mean statistics of future and 20C3M scenario temperature series are significantly different so that they do not belong to the same populations. According to the downscaled temperature results

obtained under the A1B scenario, it can be concluded that statistically significant increases of 1.00, 1.57, and 2.11 °C in annual mean temperature may be expected for the 2010s, 2020s, and 2030s, respectively. Moreover, in 2020s and 2030s, there are foreseen significant increases for nearly all months. The equalities of variances have been also tested by *f* test and annual variance statistics of temperature does not display significant changes in future periods.

### 5 GR2M for climate change scenarios

In this section, the downscaled climate change scenarios of the previous step are evaluated at watershed scale by means

**Table 12** The mean statistics (in millimeter) of corrected scenario forecasts and examining statistically significant changes (in percent) on runoff for future periods ( $\alpha=5\%$ ; italic values indicate the non-homogeneous future periods in terms of medians, and bold values show that means of

the forecasts representing the future climate are significantly different from means of 20C3M scenario representing the climate of the past period)

Scenario	Period	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
Mean statistics (in millimeter) of corrected scenario forecasts														
20C3M	1950–1999	60.33	58.14	47.63	31.12	15.09	6.55	2.31	0.91	0.46	0.89	7.09	42.20	272.71
A1B	2010s	31.01	37.65	38.33	28.11	13.88	5.23	1.83	0.72	0.33	0.24	3.55	24.16	185.05
	2020s	37.38	32.23	31.26	23.19	13.81	5.26	1.86	0.79	0.36	0.27	2.64	17.17	166.22
	2030s	37.06	39.45	32.50	20.98	11.01	3.54	1.47	0.64	0.30	0.20	0.95	19.80	167.92
Statistically significant changes (in percent) on runoff for future periods														
<b>A1B</b>	2010s	-48.60	-35.24	-19.53	-9.66	-8.03	-20.21	-20.54	-20.44	<b>-27.28</b>	-72.63	-49.92	-42.75	-32.14
	2020s	-38.05	-44.56	-34.36	-25.47	-8.50	-19.63	-19.27	-12.48	-21.99	-69.58	-62.76	-59.32	-39.05
	2030s	-38.57	-32.14	-31.77	-32.58	<b>-27.01</b>	-46.01	-36.16	-29.05	-33.50	-77.42	-86.62	-53.07	-38.43

of the GR2M (Génie Rural à 2 paramètres au pas de temps Mensuel) parametric conceptual hydrological model to observe the impacts of temperature and precipitation changes on runoff regime in the study region. Makhoul and Michel (1994) reported this model for French watersheds, which originated from a daily rainfall–runoff model. Despite having only two free parameters, the GR2M has been shown to perform well when compared to similar models; on a benchmark test consisting of 410 basins in the world, it shows the best performance among several models, some of them counting five free model parameters (Mouelhi et al. 2006).

Mouelhi et al. (2006) provide a detailed description of the GR2M model, and hence only a summarized description of model is presented here. The two free parameters of GR2M model are  $X_1$ , the soil moisture storage maximum capacity, and  $X_2$ , the water exchange term with neighboring catchments. The internal state variables consist of soil moisture accounting store ( $S$ ) and quadratic reservoir ( $R$ ). The model operates on a monthly basis with precipitation ( $P$ ) and potential evapotranspiration ( $E_{\text{pot}}$ ) as input variables. Users should note that the parameter  $X_1$ , the soil moisture store capacity, controls the model response to precipitation event, and to a certain extent, the variability of the modeled runoff. As  $X_1$  increases, the modeled runoff depends less on the current precipitation and more on the store level, itself dependent on past precipitation. For small  $X_1$ , more precipitation is directed as excess precipitation and directly routed as output runoff (Mouelhi et al. 2006; Huard and Mailhot 2008). In this study, monthly potential evapotranspiration ( $E_{\text{pot}}$ ) is defined as an exponential function ( $E_{\text{pot}} = \theta e^{\Omega T}$ ) of monthly mean temperature (Ozkul 2009), using two additional parameters  $\theta$  and  $\Omega$ . Thus, two-parameter GR2M model was turned into a four-parameter model for this study. Figure 9 shows a sketch of the GR2M model and its running procedure.

The calibration of the GR2M is based on the maximization of the value of the Nash–Sutcliffe efficiency (NS). For GR2M model, using of this measure decreases rapidly as the magnitude of random errors over precipitation increases, but much more slowly in the case of random errors over potential evapotranspiration (Huard and Mailhot 2008). In the calibration process, Newton's method-based program, which was constructed by MS-EXCEL, was used.

The reservoir of the Tahtali dam is fed by Tahtali River, and runoff values of the river were observed by Derebogazi station for the period between 1970 and 1988. For the study region, the monthly data sets of runoff data were collected from the records of II. Regional Directorate of State Hydraulic Works of Turkey. The calibration of model is carried out with the observed monthly runoff series of 1970–1979 water year period, and verification is carried out for the water year period 1980–1988. The parameters obtained from calibration phase are given in Table 10, and the modeled and observed runoff

series of the calibration and verification periods are presented in Fig. 10. The performance measures of calibrated GR2M model for the calibration and verification periods are given in Table 11.

When the computed performance measures are investigated, calibrated GR2M model is found successful and can be applied to simulate the runoff series under the future climate conditions. Although the structure of GR2M model is simple, it has an ability to model complex and nonlinear relations. In addition, other parametric hydrological models that may be much superior are used as alternative ways to GR2M for monthly rainfall–runoff modeling.

After the building of the dam, the Derebogazi stream gauging station was closed. The runoff values may be obtained by using water budget equations in monthly reservoir operation reports. It is widely accepted that the use of this approach in obtaining missing runoff values is insufficient. Being able to obtain missing runoff series will give us information about how watershed behaved after the building of the dam. To simulate missing runoff series of 1950–1999 common period, we ran the calibrated GR2M model, which is able to obtain the better prediction accuracy in terms of different performance measures during the planning period of dam, with the observed precipitation and temperature inputs of 1950–1999 period.

In the last step of the study, the calibrated GR2M model was operated with downscaled precipitation and temperature series under the 20C3M and A1B climate change scenarios. Then, the runoff series obtained from both uncorrected and corrected precipitation and temperature forecasts are examined. The box plot graphs for annual total runoff are presented in Fig. 11. When Fig. 11 is considered, it can be shown that runoff operated with corrected precipitation and temperature can represent the observed period statistics. For example, the annual mean and standard deviation statistics of simulated runoff obtained from observed precipitation and temperature are 273.72 and 106.16 mm and those of 20C3M uncorrected runoff were 267.23 and 75.64 mm. The annual mean and standard deviation statistics of runoff generated with corrected precipitation and temperature are 272.71 and 94.19 mm, respectively, which display bias correction procedure at the end of precipitation and temperature downscaling affirmatively affects the runoff modeling under the climate change scenarios.

After runoff simulations, the annual mean statistics were evaluated by using statistical tests including M-W homogeneity,  $f$  test of equality of variances, and  $t$  test of equality of mean statistics. For all projection periods, statistical tests are applied at 5 % level of significance ( $\alpha$ ) to check whether means, variances, and homogeneities of the simulated series at basin scale for the scenarios representing that the future climate are significantly different from 20C3M scenario statistics representing past climate (Table 12).

The simulated runoff series were evaluated on scenario basis, and they were interpreted in terms of their mean statistics and homogeneities as follows. For all months, there are foreseen decreases under the A1B scenario. Considering annual mean runoff statistics, it is determined that statistically significant decreases are respectively foreseen for the 2010s, 2020s, and 2030s as 32, 39, and 38 %. In addition, in 2030s, the forecasted decreases are significant for nearly all months. Similar assessments were carried out for annual variances by using  $f$  test and, in general, significant changes were not foreseen.

## 6 Conclusions

When the studies about climate change effects on river basins are examined, it can be observed that almost all studies have been carried out by researchers of developed countries (e.g., Arnell et al. 2001; Bates et al. 2008). Those studies are performed in those countries even though they expect to have an increase in their water potentials. On the contrary, Mediterranean countries including Turkey who are expecting to face a decrease on runoff are quite poor in regard to the number of studies performed and published in the literature. Having its motivation from the facts emphasized above, the presented study intends to demonstrate the effects of climate change on the runoff of Tahtali Dam which is the most important water resource in the Aegean Region having Mediterranean climate characteristic.

According to the presented assessment by statistical downscaling, bias correction at the end of statistical downscaling, and GR2M rainfall-runoff model, in the 2010s, which represents the near future, the runoff feeding the reservoir will decrease as 32 %, and thus there will be an important lack of supply. If we take the increase in the population growth and water demand into consideration, it is clear that in order to be able to compensate the lack of supply, additional water resources may be needed. Although there are several methods such as water transferring between neighbor basins and purification of seawater, our foresight is that the optimum solution is an additional reservoir capable of storing water and thus regulating the fluctuations in the runoff regime.

As is known, climate scenarios for the future are based on the predictions of GCMs. In this study, climate projections of CGCM3 were evaluated. In the literature, the differences between the atmospheric and surface variables of different climate models thus the inner uncertainties of them have been emphasized (Khan et al. 2006; Okkan and Fistikoglu 2012; Okkan 2013). These uncertainties between the projections for different GCMs constitute a current issue studied by the climate researchers. Organizing these studies, IPCC conducted some activities studying the uncertainties between the projections for different climate models within the scope

of Fifth Assessment Report (AR5) which is intended to be published ([www.IPCC.CH](http://www.IPCC.CH)). Therefore, it is considerably important to use the new GCMs to be published within the scope of AR5 with their new projections and to update the results in terms of minimizing the uncertainties.

If we are to make an overall evaluation, negative changes in the climate will have an effect on the basins which are used to supply drinking water and the ones used in agriculture, along with the other basins, affecting their water supplies and runoff regimes. In addition, in the drought period, there can be an increase in the frequency, duration, and impact area of the desertification and forest fires and thus problems will occur in the drinking water supply and in agriculture. It is clear that immediate measures must be taken in a national scale against these problems caused by climate change. We hope that results derived from this study will be of assistance in the struggle against the negative effects of climate change on Izmir-Tahtali freshwater basin in Turkey.

**Acknowledgments** The authors feel responsible to thank the reviewers of *Theoretical and Applied Climatology* and Gul Inan (from Middle East Technical University) for their valuable comments and contributions to the revision of this study.

## References

- Anandhi A, Srinivas VV, Nanjundiah SR, Kumar ND (2008) Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine. *Int J Climatol* 28:401–420
- Arnell NW, Liu R, Compagnucci L, da Cunha K, Hanaki C, Howe G et al (2001) Hydrology and water resources. *Climate Change 2001: impacts, adaptation and vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, McCarthy, J.J., Canziani, O.F., Leary, N.A., Dokken, D.J. and White, K.S. (Eds.). Cambridge University Press, Cambridge, 191–234
- Bardossy A, Bogardi I, Matyasovszky I (2005) Fuzzy rule-based downscaling of precipitation. *Theor Appl Climatol* 82:116–119
- Bates BC, Kundzewicz ZW, Wu S, Palutikof JP (eds) (2008) *Climate Change and water. Technical Paper of the Intergovernmental Panel on, Climate Change*, 210
- Crane RG, Hewitson BC (1998) Doubled CO<sub>2</sub> precipitation changes for the Susquehanna basin: downscaling from the GENESIS general circulation model. *Int J Climatol* 18:65–76
- Costa AC, Soares A (2009) Homogenization of climate data: review and new perspectives using geostatistics. *Math Geosci* 41:291–305
- Fistikoglu O, Okkan U (2011) Statistical downscaling of monthly precipitation using NCEP/NCAR reanalysis data for Tahtali River basin in Turkey. *J Hydrol Eng* 16(2):157–164
- Frei C, Scholl R, Fukutome S, Schmidli J, Vidale PL (2006) Future change of precipitation extremes in Europe: an intercomparison of scenarios from regional climate models. *J Geophys Res-Atmos* 111, D06105. doi:10.1029/2005JD005965
- Gedney N, Cox PM, Betts RA, Boucher O, Huntingford C, Stott PA (2006) Detection of a direct carbon dioxide effect in continental river runoff records. *Nature* 439(7078):835–838



- Ghosh S, Mujumdar PP (2008) Statistical downscaling of GCM simulations to streamflow using relevance vector machine. *Adv Water Resour* 31(1):132–146
- Graham LP, Hagemann S, Jaun S, Beniston M (2007) On interpreting hydrological change from regional climate models. *Clim Chang* 81: 97–122
- Hagan MT, Menhaj MB (1994) Training feed forward techniques with the Marquardt algorithm. *IEEE Trans Neural Netw* 5(6):989–993
- Ham F, Kostanic I (2001) Principles of neurocomputing for science and engineering (1st ed.). Macgraw-Hill, New York
- Holt CP, Jones AA (1996) Equilibrium and transient global warming scenario implications for water resources in Wales. *Water Resour Bull* 32:711–722
- Huard D, Mailhot A (2008) Calibration of hydrological model GR2M using Bayesian uncertainty analysis. *Water Resour Res* 44, W02424. doi:10.1029/2007WR005949
- IPCC (2007) Climate Change 2007: the scientific basic. Contribution of Working Group I to the fourth assessment report of the intergovernmental panel on climate change. Summary for policy makers
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L et al (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–471
- Khan MS, Coulibaly P, Dibike Y (2006) Uncertainty analysis of statistical downscaling methods. *J Hydrol* 319:357–382
- Krause P, Boyle DP, Base F (2005) Comparison of different efficiency criteria for hydrological model assessment. *Adv Geosci* 5:89–97
- Landman WA, Mason SJ, Tyson PD, Tennant WJ (2001) Statistical downscaling of GCM simulations to streamflow. *J Hydrol* 252(1–4):221–236
- Lettenmaier DP, Wood AW, Palmer RN, Wood EF, Stakhiv EZ (1999) Water resources implications of global warming: A U.S. regional perspective. *Clim Chang* 43:537–579
- Leung RL, Mearns LO, Giorgi F, Wilby RL (2003) Regional climate research: needs and opportunities. *Bull Am Meteorol Soc* 84: 89–95
- Maheras P, Tolika K, Anagnostopoulou C, Vafiadis M, Patrikas I, Flocas H (2004) On the relationships between circulation types and changes in rainfall variability in Greece. *Int J Climatol* 24:1695–1712
- Makhlouf Z, Michel C (1994) A two-parameter monthly water balance model for French watersheds. *J Hydrol* 162:299–318
- Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* 18: 50–60
- Marquardt D (1963) An algorithm for least squares estimation of non-linear parameters. *J Soc Ind Appl Math* 11(2):431–441
- Milly PCD, Dunne KA, Vecchia AV (2005) Global pattern of trends in streamflow and water availability in a changing climate. *Nature* 438(7066):347–350
- Mouelhi S, Michel C, Perrin C, Andreassian V (2006) Stepwise development of a two-parameter monthly water balance model. *J Hydrol* 318:200–214
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models, Part I—a discussion of principles. *J Hydrol* 10:282–290
- Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS (2005) Short-term flood forecasting with a neurofuzzy model. *Water Resour Res* 41, W04004. doi:10.1029/2004WR003562
- Okkan U (2011) Application of Levenberg–Marquardt optimization algorithm based multilayer neural networks for hydrological time series modeling. *Int J Optim Control Theor Appl* 1(1):53–63
- Okkan U, Fistikoglu O (2012) Downscaling of precipitation to Tahtali watershed in Turkey for climate change scenarios, 10th International Congress on Advances in Civil Engineering
- Okkan U, Serbes ZA (2012) Rainfall–runoff modeling using least squares support vector machines. *Environmetrics* 23:549–564
- Okkan U (2013) Assessment of climate change effects on river flows. PhD Thesis, Department of Civil Engineering, Dokuz Eylul University, Izmir, Turkey
- Ozkul S (2009) Assessment of climate change effects in Aegean river basins: the case of Gediz and Buyuk Menderes Basins. *Clim Chang* 97:253–283
- Phillips ID, McGregor GR, Wilson CJ, Bower D, Hannah DM (2003) Regional climate and atmospheric circulation controls on the discharge of two British rivers, 1974–1997. *Theor Appl Climatol* 76(3):141–164
- Prudhomme C, Reynard N, Crooks S (2002) Downscaling of global climate models for flood frequency analysis: where are we now? *Hydrol Processes* 16(6):1137–1150
- Schoof JT, Pryor SC, Robeson SM (2007) Downscaling daily maximum and minimum temperatures in the midwestern USA: a hybrid empirical approach. *Int J Climatol* 27(4):439–454
- Struglia MV, Mariotti A, Filogrosso A (2004) River discharge into the mediterranean sea: climatology and aspects of the observed variability. *J Clim* 17(24):4740–4751
- Tatli H, Dalfes HN, Mentés S (2004) A statistical downscaling method for monthly total precipitation over Turkey. *Int J Climatol* 24(2):161–180
- Tripathi S, Srinivas VV, Nanjundiah RS (2006) Downscaling of precipitation for climate change scenarios: a support vector machine approach. *J Hydrol* 330:621–640
- Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: a comparison of methods. *Water Resour Res* 34: 2995–3008
- Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO (2004) The guidelines for use of climate scenarios developed from statistical downscaling methods. Supporting material of the Intergovernmental Panel on Climate Change (IPCC), prepared on behalf of Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA)
- Wilby RL, Dawson CW, Barrow EM (2002) SDSM—a decision support tool for the assessment of climate change impacts. *Environ Model Softw* 17:147–159
- Wilby RL, Dawson CW (2004) Using SDSM version 3.1 A decision support tool for the assessment of regional climate change impacts, User Manual
- Wilby RL, Tomlinson OJ, Dawson CW (2003) Multi-site simulation of precipitation by conditional resampling. *Clim Res* 23:183–194