The development of statistical downscaling methods for assessing the effects of climate change on the precipitation isotopes concentration

Maryam Mosaffa, Sara Nazif and Youssef Khalaj Amirhosseini

ABSTRACT

In recent years, stable isotopes of the water molecule (oxygen-18 and deuterium) have become a useful tool for tracking the water cycle. The concentration of these tracers changes with variations of water molecules within the water cycle. Due to this feature of isotopes, global large-scale isotope models have been developed. On the other hand, numerous local and global networks have been created in order to monitor the concentration of precipitation isotopes. The main problem with the simultaneous use of these local stations and the large-scale isotope datasets is their temporal and spatial mismatch. To use both isotope databases for monitoring the hydrological cycle in local scale, it is necessary to downscale the large-scale models' outputs. In this research, a downscaling approach is proposed for isotopes' concentrations using three statistical models, including multiple linear regression, generalized linear and weighting least square regression models. The results indicate that the implementation of the statistical downscaling method in the case of information preprocessing based on the seasonal changes, their spatial variations and a suitable method selection is a useful tool for monitoring the climate changes of a region according to the information on the stable oxygen-18 isotope.

Key words | downscaling, stable water isotopes, statistical methods

Maryam Mosaffa
Sara Nazif (corresponding author)
School of Civil Engineering, College of Engineering,
University of Tehran,
Tehran,
Iran
E-mail: *snazif@ut.ac.ir*

Youssef Khalaj Amirhosseini
Water and Environmental Kowsar Research Center,
No. 33, Fatemi St., Tehran,
Iran

INTRODUCTION

Along with all the available methods for identifying the basins and connections among the components of the hydrological cycle, isotope tracing has become an effective tool for a better understanding of hydrological processes through creating new insight into our present and past climate. During recent decades, many researchers have used the water constituent stable isotopes (hydrogen and oxygen) in order to investigate the connection details of the hydrological cycle's components (Bowen & Revenaugh 2003).

The difference in physical properties of water molecules leads to a change in the stable isotope composition of water, due to fractionation during phase changes (Kim & Lee 2011). In general, one can say that the water molecule's isotopologues show different behaviors during the phase transition. The occurring fractionation process on the water isotopologues has led to a significant correlation between the isotopes' concentration and climatic variables (Sutanto *et al.* 2015). Among all the tracers used in hydrology, oxygen-18 isotope concentration is sensitive to phase changes of water during its circulation. Therefore, it has been used as a natural tracer for hydrological cycles since the 1960s because the isotope-related characteristics of water are dependent on phase changes of water in evaporation and condensation as well as geographic and temporal variations. By using the isotope information in precipitation and vapor, one can study atmospheric vapor

doi: 10.2166/wcc.2020.208

cycling processes on various scales, such as large-scale transport and in-cloud processes (Rozanski & Gonfiantini 1990).With the assistance of the World Meteorological Organization (WMO), the International Atomic Energy Agency (IAEA) established the global network of GNIP in 1954 in order to monitor the temporal and spatial variations of oxygen-18, deuterium and tritium contents of precipitation samples, investigate the correlation between the isotopes and climatic variables and simulate the fractionation imposed by the water molecules in a hydrologic cycle. The irregular dispersion of these measurement networks around the globe and lack of data recording at regular intervals has made it impossible to rely on the data recorded at GNIP stations alone for studying the long-term mean changes in the meteorological variables, isotopes and examining the correlation among them. To compensate for this information scarcity, many researchers around the world have tried to create an integrated database with the regular coverage of the entire Earth for the concentration of the existing vapor isotopes in the atmosphere and precipitation by implementing the physical basis of the atmospheric general circulation models (AGCM), identifying the system dynamic of changes in the water molecule's isotopes in this cycle and thermodynamic modeling of the moisture motion in the atmosphere. Attempts have been made to arrive at a novel perspective on climate change by using the water molecule's isotopes via implementing the computational information in the large-scale isotope general circulation models (IsoGCM) (Hoffmann et al. 2000; Bowen & Revenaugh 2003).

In addition to simulating the temporal-spatial distribution of the isotope, the IsoGCM models, which are equipped with the concentration simulation of water molecule's isotopes, can also interpret the effects of kinematic and equilibrium fractionation on the isotopes' concentration during a global hydrological cycle (Haese et al. 2013). Researchers created the IsoGCM models by counting the fractionation process in all transfer phases such as evaporation, condensation, precipitation, air vapor and added them to the GCM models. In the general modeling of the isotopes, the Earth's atmosphere is divided into hundreds of air columns and the parameters of the dynamic differential equations of the atmosphere such as solar radiation, cloud, and convective flow are calculated in these columns and transferred to the next columns. In each transfer, the concentration of isotopes is updated based on the fractionation among the different phases of the air mass. Depending on the conditions and structure of each of the applied IsoGCM models, the use of isotopes also varies in the simulated atmospheric variables (Hoffmann et al. 2000; Dee et al. 2015). Xi (2014) synthesized the recent advances in IsoGCM models. Having introduced nine large-scale models equipped with the simulation of a water molecule's isotopes cycle, Xi (2014) described the temporal and spatial characteristics and the output of each model and stated that there are some discrepancies among model outputs and measurements. This could lead to more model improvements in the future (Xi 2014).

Since 2000, several committees have been established around the globe for isotopic modeling to rectify the results of dynamical and physical mechanisms in isotope hydrology and observations. To evaluate the performance of these models, the output results of the large-scale IsoGCM models were compared with those recorded via ground and air stations (Henderson-Sellers et al. 2006; Zhang et al. 2012; Che et al. 2016). Sutanto et al. (2015) compared the results achieved via the two large-scale isotopic models of GISS and ECHAM with each other and also with those recorded by field measurements. They stated that the results of these two models could not be prioritized as the simulation results of both models were consistent with the ground measurements (Sutanto et al. 2015). Yao et al. (2013) investigated and compared the results of the two large-scale isotopic models LMDZiso and ECHAM5-wiso in the Tibet region. They reported that the isotope concentration in this region during the monsoon precipitations is not consistent with the simulated value. Both of these models successfully indicate the effect of temperature and elevation on the concentration of the water molecule's isotopes in the region (Yao et al. 2013).

Following vapor particles from the ocean surface to land by using the infrared technique was the aim of Zakharov et al. (2004). They illustrated that by monitoring the isotope trackers, the effect of successive fractionations tolerated by water particles during evaporation, distillation, and precipitation can be clearly seen (Yao et al. 2013). The concentration simulation results of the water molecule's isotopes in the LMDZ4 large-scale isotope model have been investigated by Hourdin *et al.* (2006). They indicated that the simulation results in this model are sensitive to temperature and elevation variations, therefore, it can be utilized for modeling the isotopic concentration of the water molecule (Hourdin *et al.* 2006). Lee *et al.* (2007) tried to simulate the concentration distribution of the water molecule's isotopes by improving the NCAR model by adding the evaporation process and precipitation changing rate, further to the temperature. A comparison of their model's outputs with the observed data indicates that the effect of Rayleigh distillation fractionation is diminished (Yoshimura *et al.* 2011). Yoshimura *et al.* (2011) provided a large-scale IsoGSM model. They attempted to simulate the flux of water molecules' isotopes in the rain via adding the evaporation and distillation of a water molecule which undergoes the isotopes fractionation. They compared their simulation results with those of the observation stations and eliminated the difference between the simulation model and observations by two satellite sensors (Yoshimura *et al.* 2011).

Although the isotopic large-scale models are utilized to increase our understanding of the physical relationships between the climate variables and water isotope, the outputs of these models do not have the resolution to assess climate change at the regional scale (Dansgaard 1964). Climate change assessment studies are usually carried out with high spatial resolution in small areas. In this case, it is desirable that the outputs of these models are downscaled. In the meantime, choosing the downscaling technique of the global simulation models' information is one of the primary challenges in using data from these models on a local or regional scale. The regional dynamic models and statistical methods are considered as two common techniques for downscaling the information on the large-scale models.

Yoshimura *et al.* (2008) applied spectral nudging techniques for the downscaling of global reanalysis. They expressed that fractionation during transport is a purely numerical issue, but is clearly non-trivial and persists even in recent isotope schemes (Yoshimura *et al.* 2008). Delavau *et al.* (2011) worked to create a series of empirical relationships among the registered isotope concentrations at 23 stations in Canada by using the data of Canadian Network for Isotopes in Precipitation (CNIP) based on the spatial variables of elevation and latitude, hydrological variables as the predictors of the dependent variable of oxygen-18 isotope concentration. They illustrated that the statistical equations including the temperature and precipitation climate variables have demonstrated better performance (Rahimzadeh *et al.* 2009).

The dynamic downscaling models meet suitable spatial accuracy criteria for the climate assessment but have distinct disadvantages, such as depending on the area of the studied region and model resolution and computational complexity (Akbari *et al.* 2015). On the other hand, the statistical downscaling methods are based on the empirical relationship found between the predicted local variable and the predictor variables (Asadi & Heydari 2011). The statistical downscaling methods are generally classified into three groups, including the regression method, weather typing schemes, and weather generators. Among the above-mentioned methods, the regression method is one of the most widely used data downscaling techniques of the large-scale models (Delavau *et al.* 2011).

Dansgaard (1964) divided the effective factors on the water molecule's isotopes concentration in the cycle into four general categories, including temperature, latitude, seasonal, and continental effects. In this study, while accepting the effects of the above-mentioned factors by Dansgaard (1964) on the isotopes concentration, we attempt to investigate the mechanism of changes in the precipitation isotope concentration in the region by generating a regression numerical model between the concentration of oxygen-18 precipitation isotope and meteorological variables.

For using isotopic data for climate change studies it is necessary to spatially downscale them to be used beside the weather characteristics. Less attention is paid to this issue in the literature and few models and methods are available for this purpose. The limited amount of isotopic measurements is another issue that causes more difficulty in this regard. Therefore, the aim of this study is to develop a downscaling model considering the limited available data through developing relations between weather and isotopic variables.

THE STUDY AREA

Shiraz, the capital of Fars province with an area of about 8,725 square kilometers, accounts for approximately 7.6%

of Iran's total area and its altitude is 1,486 m above sea level. Shiraz is located 53° 34′ east of the Greenwich meridian and 29° 36′ north of the equator. Based upon the meteorological data during the statistical period of 45 years (1971–2016) for this region, the average annual temperature in this city is 18.04 °C and the total annual precipitation in the province is 320.2 mm. Also, the average long-term annual precipitation of Shiraz is reported as 100.4 mm. The majority of rainfall in the study area is due to the winter masses. The Mediterranean Sea is the main moisture source of the study area during winter and affects the area from the northwest, west, and southwest (Rahimzadeh *et al.* 2009; Akbari *et al.* 2015; Pourtouiserkani *et al.* 2015). The minimum and maximum long-term absolute temperatures of this city are recorded as −14.4 and 43.2 °C, respectively. The long-term monthly average of this station was used for simulating the missing data of precipitation, temperature, and humidity. The main objective of this research is to develop a series of oxygen-18 concentrations in precipitation data on the local scale using large-scale IsoGSM outputs and local weather variables. Preparing these data on a local scale can provide researchers with the opportunity to identify the causes of climate change in a region. In this regard, and to understand the causes of climate change, the basic condition for selecting the study area is that the occurrence of climate change in that area must be proven by other studies. The occurrence of the climate change phenomenon in the Fars province has been proved by many researchers via statistical methods. Since this province is located in a dry and semi-arid climate, it is more vulnerable to the effects of climate change (Asadi & Heydari 2011).

DATASETS

In this study, two datasets are used for the calibration as well as verification of the suggestive statistical downscaling model of the oxygen-18 concentration in Shiraz's monthly precipitation. The first set includes the IsoGCM model data. The second group of data used for model validation are the recorded data in the study area which include rainfall, temperature, humidity, and oxygen-18 concentration in precipitation. Since the number of local measurements is too few, this information is only used to evaluate the function of equations at the regional level. In order to study the effects of climate change here, we have first weighed the isotopes concentration based on the precipitation amount of the model and then used its monthly average for the simulation (Delavau *et al.* 2011). In the following, each of these databases will be introduced in detail.

Local database

To evaluate the efficiency of the proposed downscaling models in the study area, the statistical information of the meteorological station located in Shiraz has been used. The meteorological station information includes temperature, precipitation, wind speed, and relative humidity on a daily basis. The observation period of the meteorological data at this station is from 1983 to 2017. The daily information of these data has been provided from Iran Meteorological Organization. To use this information, a monthly time scale has been selected. The average monthly temperature recorded at this station is 17.5 °C and its total annual precipitation is 312 mm. Along with the meteorological station data, 12 precipitation samples of Shiraz city were evaluated in terms of the isotopic concentration of deuterium and oxygen-18 during the data downscaling period. The isotope concentrations of the precipitation samples are indicated with δ notation relative to the VSMOW (Vienna Standard Mean Ocean Water). The isotopic abundance ratios are expressed as parts per mill of their deviations as given by VSMOW and δ is calculated according to the following relation:

$$\delta(‰) = \frac{\alpha_{sample} - \alpha_{standard}}{\alpha_{standard}} \times 1000 \qquad (1)$$

where, α represents the concentration ratio of heavy to light isotope in the measured and standard samples. Currently, VSMOW is declared as the standard defining the water isotopic composition by IAEA. The concentrations of oxygen-18 and deuterium isotopes of the samples have been measured using the Los Gatos Research (LGR) setup. Table 1 lists the results of the isotopic experiments and precipitation sampling time in Shiraz city.

Table 1 | Data corresponding to the precipitation samples in Shiraz

Date	Season[a]	δ^{18}O (‰)	δD (‰)	Pre. (mm)	Temp. (°C)	RH (%)
5/31/1989	MAM	−2.39	−0.2	0.25	30	49.76
5/29/1989		−2.59	−1.2	0.2	30.2	45
3/16/1990		−6.1	−25	10.4	10.85	56.35
3/20/1989		−6.4	−36	3	9.3	64
1/14/1991	DJF	−5.53	−18.3	145	3.05	74.5
1/16/1992		−4.3	−20.6	65	4.3	91.87
9/1/1991	SON	1.65	28.3	7.9	20.3	14.35
9/2/1991		1.95	27.9	0.3	20.5	35.25
1/24/2016	DJF	−3.89	−24.29	8	7.65	78.9
1/25/2016		−4.1	−22.29	4	8	82.13
1/6/2018		−5.66	−14.29	5.9	8.2	52.2
1/7/2018		−6.73	−19.33	6.8	5	48

δ^{18}O and δD concentration of oxygen-18 and deuterium in precipitation, respectively. Pre., Temp., and RH correspond to precipitation, temperature, and relative humidity, respectively.
[a]Spring season MAM (March, April, May), fall season (September, October, November) and winter season (December, January, February).

The large-scale isotopic model's database

In this research, the concentration of the stable isotope oxygen-18 (δ^{18}O) in precipitation and meteorological parameters is extracted from the IsoGSM model which contains data from the period 1979–2016. IsoGSM stands for Isotopes-incorporated Global Spectral Model. Improving the data of the NCEP/DOE large-scale model based upon the spectral nudging technique, Yoshimura et al. (2010) have simulated the existing isotope's data in the precipitation of this model (Yoshimura et al. 2008, 2010; Kim & Lee 2011). This model includes 19 atmospheric variables along with the concentrations of the water molecule's isotopes in precipitation and available vapor in the air, which are located in networks with approximate dimensions of 200 km on the horizon. The air column in this model is divided into 17 sections up to the pressure level between 500 and 800 hPa. The information from this model is available with a 6-hour time resolution and can be accessed online at https://data.giss.nasa.gov/swing2. Table 2 lists the 19 simulated variables by this model, among which, the independent variables for the concentration prediction of oxygen-18 isotope in precipitation are deduced. Since the present research will attempt to investigate climate change, data will be used monthly. The deduced concentration of oxygen-18 isotope from this model is first weighted by the daily precipitation amount on the basis of Equation (2) and then its weighted average evaluated:

Table 2 | The selection of large-scale climate factors for the statistical downscaling methods in this study

No.	Predictor	Abbreviation	Unit
1	Precipitable water	PWAT	mm
2	Oxygen – 18 for PWAT	PWAT^{18}O	mm/SMOW
3	HDO for PWAT	PWATHDO	mm/SMOW
4	Total precipitation	PRATE	kg/m^2/s
5	Oxygen – 18 for PRATE	PRATE^{18}O	kg/m^2/s/SMOW
6	HDO for PRATE	PRATEHDO	kg/m^2/s/SMOW
7	Convective precipitation	CPRAT	kg/m^2/s
8	Oxygen – 18 for CPRAT	CPRAT^{18}O	kg/m^2/s/SMOW
9	HDO for CPRAT	CPRATHDO	kg/m^2/s/SMOW
10	Latent heat flux	LHTFL	W/ m^2
11	Oxygen – 18 for LHTFL	LHTFL^{18}O	W/m^2/SMOW
12	HDO for LHTFL	LHTFLHDO	W/m^2/SMOW
13	Specific humidity at 2 m	SPFH2 m	kg/kg
14	Oxygen – 18 for SPFH2 m	SPFH2 m^{18}O	Kg/kg/SMOW
15	HDO for SPFH2 m	SPFH2 m HDO	Kg/kg/SMOW
16	Air temperature at 2 m	TMP2 m	K
17	Relative humidity at 2 m	RH2 m	%
18	Surface pressure	PRESsfc	Pa
19	500 hPa geopotential height	HGT500	m

$$\delta^{18}O = \frac{\sum P_i \times \delta^{18}O_i}{\sum P_i} \tag{2}$$

where, i, P_i, and $\delta^{18}O_i$ define the month indicator, average monthly precipitation, and monthly isotopic concentration of oxygen-18, respectively.

In the data downscaling process of the IsoGSM model, 70% of the studied time data have been randomly selected in order to train the regression model and the remaining 30% used for validating the model. The IsoGSM data output scale is a 6-hour timeframe and their monthly average is used in the present study. Figure 1 indicates the study area and network location of the IsoGSM large-scale model.

Alijani & Harmant (1985) have determined the path of Iran's moisture fronts. Based on the maps provided by them (wind direction and movement path of the moisture fronts), four points of this large-scale model have been selected around the studied area (Alijani & Harman 1985). According to the findings of Dansgaard (1964), the geographic location of the studied area is one of the most significant factors affecting the concentration of precipitation isotopes (Dansgaard 1964), but according to Liebminger *et al.* (2006) and Argiriou & Lykoudis (2008), finding the effect of this parameter can be ignored in atmospheric datasets which have finer resolution range. Therefore, based on the location of the study area, which is situated among these four points of the large-scale model, two different scenarios are defined for selecting the IsoGSM grids to be used in developing the downscaling model. In the first scenario, due to the fact that the local sampling point and Point 1 and Point 3 are approximately located at the same latitude, only the information from these points (Point 1 and Point 3) was used for creating the statistical downscaling model. However, in the second scenario, taking into account the directions of wind and moisture sources, the information for all four nodes around the region has been used for developing the downscaling model.

Two scenarios are defined for using IsoGSM data in this study area, as follows:

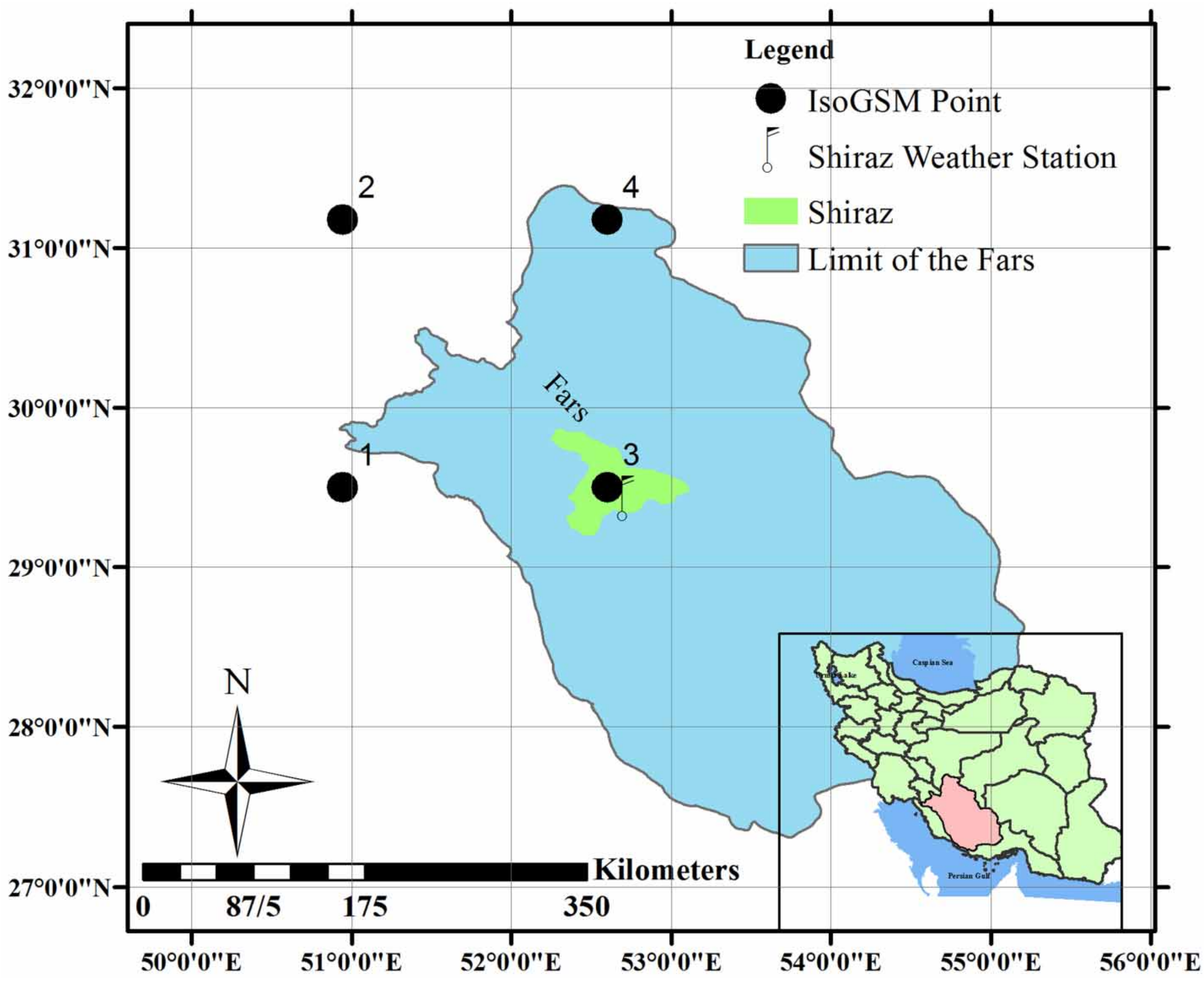

Figure 1 | Location map of IsoGSM grid points around the study area, local sampling point, and weather stations in the study area.

(1) using Point 1 and Point 3 information in developing the downscaling model because the local sampling point and the two points of 1 and 3 are approximately located at the same latitude;
(2) using Point 1 to Point 4 information in developing the downscaling model considering sources of moisture in the study area and wind direction.

Table 3 illustrates the geographical coordinates of the large-scale model's points and sampling site for the precipitation isotopes in the study area.

METHODOLOGY

The research implementation method is summarized in four steps. The first step is the information classification of the IsoGSM large-scale model according to the temporal and spatial changes (step A). The second step concludes the selection of dominant independent variables among the simulated variables by the large-scale isotopic model for predicting the concentration of oxygen-18 isotope in precipitation. In this step, the correlation between the predictand variable of the oxygen-18 isotope concentration in precipitation is assessed with other variables in the large-scale IsoGSM (step B). In the third step, attempts will be made to calibrate the statistical model for the selective nodes' conditions in the two above-mentioned scenarios using the multiple linear regression (MLR), generalized linear model (GLM), and weighting least square regression (WLSR) methods together with the creation of a linear relation between the predictand variable (the oxygen-18 isotope concentration in precipitation) and predictor (variables selected in step B) (step C). The fourth step deals with verifying the obtained equations in step C and evaluating their performance in predicting the concentration of oxygen-18 isotope in precipitation (step D). Figure 2 depicts the flowchart of the methodology process used in the present study. In the following, efforts will be made to briefly introduce the steps outlined above.

Table 3 | Location of IsoGSM model points around the study area

Station	Long.	Lat.	Elev. (m)
Point 1	50.625	29.523	53.5
Point 2	50.625	31.428	1,267.4
Point 3	52.500	29.523	1,485.2
Point 4	52.500	31.428	2,376.5
Shiraz	52.600	29.534	1,488

Step A

Since a method is being developed for estimating the concentration of oxygen-18 isotope in rainfall, the days in which rainfall has not occurred or the amount of precipitation is small, must be deleted from the database. In this study, the boundary value of 0.5 mm of precipitation per day has been used. The basis for choosing this boundary value for the wet days is according to the suggestion of Chen *et al.* (2010).

On the other hand, before choosing the predictor variables of the oxygen-18 isotope concentration, it is necessary to first assess the conditions of the region by reviewing the Dansgaard rules. Dansgaard (1964) divided the effective parameters on the concentration of water molecules' isotopes into four general categories including the temperature changes, latitude, seasonal, and continental effects. Therefore, to consider these cases, there is a need to adopt two different attitudes to the data: first, spatial classification of the data to take into account the latitude and continental effects and second, the data temporal grouping in order to consider the seasonal and temperature effects. In the spatial classification, as stated earlier, the information is introduced to the simulation via two different scenarios.

A review of technical literature shows that several geostatistical methods are developed that generate gridded sets of isotopic composition. They use several IsoGCM models and GNIP station data with different resolution. The results show that the regression coefficient of distance is negligible versus the other variables such as temperature and precipitation in atmospheric datasets which have a finer resolution range, from $0.5^\circ \times 0.5^\circ$ down to $5^\circ \times 5^\circ$ (Liebminger *et al.* 2006; Argiriou & Lykoudis 2008; Delavau *et al.* 2011). Besides, the distances (distances between the nodes are 200 km) among the four nodes surrounding the region are such that one can ignore the latitudinal and continental effects among them. In the temporal classification, the

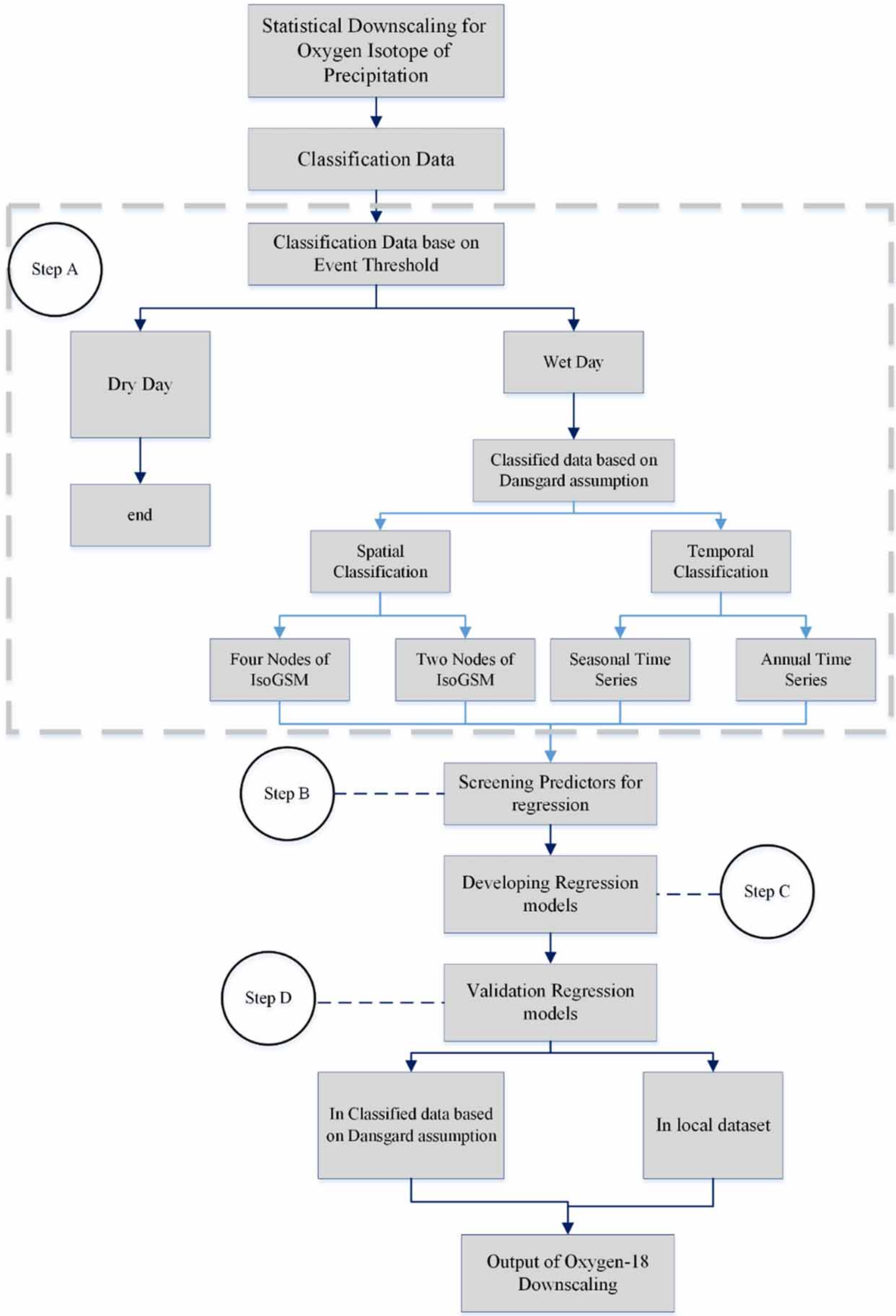

Figure 2 | The proposed algorithm for downscaling isotope concentration.

effects of season and temperature on the concentration of precipitation isotopes will be considered by categorizing the data according to the seasons.

Step B

The results of numerous studies in the field of downscaling have shown that using all the output variables of a large-scale model to simulate the predictor variable does not necessarily meet with good results. The importance of this section is due to the fact that the characteristics of the model and results are directly influenced by the selected independent variables. The independent variables selected from the existing ones are called the dominant variables. In most cases, there is a correlation among the selected variables for the predictor group. To this end, this correlation is investigated by means of the variance inflation factor (VIF). The dominant variables must be physically related to the dependent ones and meet a high degree of statistical correlation with them. In this research, the stepwise regression technique is implemented in order to examine the correlation between the predicted variable (oxygen-18) and predictor ones (output of the large-scale IsoGSM model) and also to reduce the variables and their optimum selection (Silhavy *et al.* 2017).

In this approach, the independent variables are added to the equation one by one and they are eliminated if they do not play a significant role in the regression. In the stepwise regression model, the correlation analysis is used for examining the effect of each variable and an independent variable with less correlation is eliminated in each step. The correlation evaluation criteria in this method are the correlation (R^2) and determination (R) coefficients (Karamouz *et al.* 2014). This method provides an automatic procedure for selecting the variables which should be entered into the regression models by conducting a series of t-tests at the 95% confidence level (Delavau *et al.* 2011).

Step C

Here, three simple regression methods have been used for simulating and downscaling the concentration of oxygen-18 isotope, including GLM, MLR, and WLSR. Usually, the regression equations rarely show all the climatic changes but they are suitable to illustrate the generalities of the behavior occurring in the region and the average behavior of climate variables is well represented by them. Statistical downscaling is a suitable choice compared to the dynamic one which is computationally costly and complex and does not show enough flexibility for convenient use by users. Statistical methods have more advantages and capabilities when there is a need for quicker assessment of the factors affecting the climate change (Kigawa 2014). Each of the methods will be briefly explained below.

C-1) MLR

The multiple linear downscaling method is based upon the empirical relations between the predictors and predictand variable. From the strengths of this model, one can point out the ease of use for predicting and accessibility of the equations for the user. In this study, the MLR technique is implemented for predicting the concentration of oxygen-18 isotope in precipitation according to the following relation:

$$y = \beta_0 + \sum_{j=1}^{n} \beta_i X_i \tag{3}$$

where, y and X_i denote the value of dependent variable (δ^{18}O) and selected predictor variables in step B (precipitation, temperature, and relative humidity), respectively. β_i stands for the regression coefficient being estimated by means of the least squares method. In the MLR approach, it is assumed that there is no multicollinearity among the predictor variables. This assumption is assessed using the VIF coefficient which can be obtained as below:

$$VIF_i = \frac{1}{(1 - R_i^2)} \tag{4}$$

where, R_i^2 is the coefficient of determination of the model in which the independent variable of X_i is regressed on the remaining ones.

This factor indicates how much the variance of the estimated coefficients is inflated compared to the state in which the estimated variables are not linearly correlated. As an empirical rule, for VIF values greater than 5, the multicolinearity is high. If the colinearity level among the independent variables is high, it indicates correlation existence among them and the selection of the predictor variables should be reconsidered. Since the independent and dependent data have different measurement units and the range of numbers varies for each variable, the data must be standardized before the analysis according to the following relation:

$$Nx_i = \frac{x_i - \mu}{\sigma} \tag{5}$$

where, μ and σ stand for the average and standard deviation of the data, respectively. Also, Nx_i is the standardized value with average and variance values of 0 and 1, respectively.

C-2) GLM

The GLM model is a method used to simulate the linear relationships between several independent variables with an independent one and it is assumed in this method that all the investigated data are subjected to the normal distribution. The GLM regression model is defined as:

$$y = \beta_0 + \sum_{j=1}^{n} \beta_i Nx_i + \varepsilon_i \tag{6}$$

in which, ε_i is the random error with normal distribution as $\varepsilon_i \sim N(0, \sigma^2)$. Since the prediction of the relation between the dependent and independent variables is not accurate, the considered random variable as the model error prevents bias in simulations made by Equation (6). Nx_i indicates the standardized predictor variables selected in step B. The other variables and parameters are defined as in Equation (3).

In this research, the Kolmogorov–Smirnov test is implemented due to the large amount of input data. When examining the data normality, the zero assumption is tested at a 5% error rate. If the test statistic is greater than 0.05, then there is no reason to reject the zero assumption based on the normal distribution of data.

C-3) WLSR

The WLSR method used here for the information downscaling is very similar to the MLR one, with the difference that in this regression approach, a weight W is assigned to each of the independent variables. This weight is equal to the inverse square of the linear regression variance fitted on the independent variables by stepwise regression method. The WLSR method is used to weigh the independent variables of a regression when the error variance is constant. The data weight according to the least squares indicates the error's random behavior in the simulation model. If the parameters of a regression model are estimated using the LSM, the error terms will not depend on each other and will have the same variance. This is because when less weight is assigned to the data that are more accurate and higher to those with less accuracy, this makes the data behave better compared to the overall variance. Therefore, weighing independent variables in the regression equation reduces the number of deviated information from the time series. The equation governing the WLSR is defined as:

$$y = \beta_0 + \sum_{j=1}^{n} w_i(\beta_i Nx_i + \varepsilon_i) \tag{7}$$

where w_i denotes the weight of ith variable given by:

$$w_i = \begin{pmatrix} \frac{1}{\sigma_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sigma_n} \end{pmatrix} \tag{8}$$

Step D

In order to evaluate and analyze the performance of estimation and forecasting models, there are various performance indicators. In the following, a brief explanation is given for the indices used in this research. The resultant equations are verified for 30% of the remaining data from the IsoGSM model. Comparison of the standardized residual values, box plot, probability density function, and modeling performance indicators are from the methods used in this research for evaluating the results. The utilized modeling performance indices include the root mean square error, Nash–Sutcliffe efficiency coefficient, ratio index of simulation standard deviation to observations standard, and coefficient of determination.

The root mean square error is used as a scale to indicate the difference between the simulated and measured values. This criterion, which is defined by Equation (9), is considered as the most commonly used error index. Equation (10) describes the Nash–Sutcliffe efficiency (NSE) coefficient. If the NSE value is equal to 1, there is a complete proportion between the observational and simulation data. In fact, this criterion indicates the relative importance of the simulated values variance in comparison to that of the observational ones. The ratio index of simulation standard deviation to observations standard is another way of assessing the observations (Equation (11)).

Equation (12) represents the coefficient of determination for evaluating the simulation results. This coefficient indicates how much the dependent variable's changes are affected by the independent variable and its remaining variation is related to the other factors:

$$RMSE = \frac{\sqrt{\sum_{k=1}^{n} (O_{io} - O_{pi})^2}}{n} \tag{9}$$

$$E_{NS} = 1 - \left(\frac{\sum_{i=1}^{n} (O_{io} - O_{pi})^2}{\sum_{i=1}^{n} (O_{io} - \overline{O_{io}})^2} \right) \tag{10}$$

$$RS = \frac{S_{sim}}{S_{obs}} \tag{11}$$

$$R^2 = \frac{\sum_{k=1}^{n} O_{io} O_{pi}}{\sum_{k=1}^{n} (O_{io})^2 \sum_{k=1}^{n} (O_{pi})^2} \tag{12}$$

In the above equations, O_{io} defines the simulated concentration of oxygen-18 isotope in the large-scale model of IsoGSM or the observed one in the Shiraz station. Also, O_{pi} stands for the simulated concentration of this isotope via the regression equations described in step C and n is the number of samples.

RESULTS AND DISCUSSION

In this section, attempts are made to show the results of the statistical downscaling of the oxygen-18 isotope concentration according to the steps described in the previous sections.

Step A

In the first step, by first applying the boundary value of 0.5 mm of rainfall per day, the days of the year are divided into two categories of wet and dry. On dry days, due to the absence of precipitation, isotope concentrations cannot be monitored. The abundance of oxygen-18 in precipitation is temporally and spatially variable. Major isotope variations are linked to air temperature changes, especially seasonal variation (creates temporal variations in long-term isotope records). Attitude and continental (distance from the coast) changes cause spatial variation in the isotopic composition of the precipitation. To consider these effects, data are divided into two sets of seasonal and annual ones based on the time and two spatial scenarios according to the geographical location of the large-scale IsoGSM model's nodes. Regarding the effect of temperature variations on the concentration of the water molecule's isotopes, the outputs of IsoGSM are classified into four categories based on the seasons. In addition to the seasons categorization, the results of the annual downscaling are also evaluated in this study so that one can be able to state whether Dansgaard rules can be ignored in the region or not through comparing these equations. To consider the effects of geographical changes on the isotope concentration, the large-scale model's data have been divided into two sets based on the location: first, setting up a database for creating regression models by using the information of Point 1 and Point 3 which are located in the same latitude as the study area (first scenario) and second, creating another database by using the information of all four nodes which are located around the study area (second scenario).

Step B

The IsoGSM large-scale model has 19 outputs as defined in Table 2. One of the most important steps in the statistical downscaling is to select the dominant variables used in the model. Since the purpose of this study is to achieve statistical downscaling methods of the oxygen-18 isotope concentration based on the meteorological variables, only the meteorologically measurable variables have been selected for correlation analysis among the above model's outputs. These variables include the precipitation amount, relative humidity, specific humidity, and temperature. The correlation between the predicted variable and predictors has been investigated according to the stepwise regression method.

It is necessary to mention that climate change can impact the type of precipitation and the fraction of convective and stratiform rainfalls (Aggarwal *et al.* 2016; Tharammal *et al.* 2017). Based on Aggarwal *et al.* (2016), this issue can be well investigated using isotopic data. The downscaled data in this study can be used for such purposes which is of high value. In this study, because of limitations in data accessibility, the effect of precipitation type is not investigated.

Table 4 illustrates the steps of the stepwise regression model. In each step, a variable is added to the regression model and correlation (R) and determination (R^2) coefficients are estimated as well. The results show that in both scenarios, the combination of three independent variables of precipitation, temperature, and relative humidity meet the highest correlation with the dependent variable (oxygen-18 isotope concentration). The values of VIF and correlation among the selected combination's variables are listed in Tables 5 and 6 for both scenarios. As can be seen from the results of Table 5, the value of this coefficient in two scenarios is lower than 2.5, therefore, one can ignore the correlation among the independent variables.

Table 4 | Stepwise regression model's performance summary for selecting appropriated variables from IsoGSM (a) in scenario 1 and (b) in scenario 2

No.	Predictors	R	R^2	Sig.
(a) Scenario 1				
1	RH2 m	0.465	0.218	0.000
2	RH2 m, TMP2 m	0.468	0.219	0.000
3	RH2 m, TMP2 m, PRATE	0.471	0.220	0.000
4	RH2 m, TMP2 m, PRATE, SPFH2 m	0.449	0.201	0.000
5	TMP2 m, PRATE, SPFH2 m	0.435	0.189	0.068
(b) Scenario 2				
1	RH2 m	0.429	0.184	0.000
2	RH2 m, TMP2 m	0.430	0.185	0.000
3	RH2 m, TMP2 m, PRATE	0.431	0.185	0.000
4	RH2 m, TMP2 m, PRATE, SPFH2 m	0.428	0.183	0.000
5	TMP2 m, PRATE, SPFH2 m	0.425	0.180	0.078

Table 5 | Multicollinearity test result on each scenario and independent variables

Scenario	Independent variable	VIF
PO.1 & PO.3	RH2 m	2.33
	TMP2 m	1.82
	PRATE	1.57
PO.1 TO PO.4	RH2 m	2.27
	TMP2 m	1.78
	PRATE	1.53

Step C

The downscaling results of the oxygen-18 isotope concentration via MLR (Equation (3)), GLM (Equation (6)), and WLSR (Equation (7)) models are given in Table 7 based upon the seasonal and annual classifications in two different spatial scenarios. In all seasons and scenarios, the relative humidity has a negative effect on the concentration of oxygen-18 isotope in precipitation and its increment means the precipitation is depleted from the light isotopes. Apart from in the GLM method, the precipitation amounts possess negative effects on the concentration of oxygen-18 isotope in precipitation. During the autumn and winter seasons, temperature increase has a positive effect on the concentration of isotopes and enriches the precipitation of heavy isotopes.

Verification of the downscaling regression models' results

Verification of the model results with those of the large-scale IsoGSM model

As mentioned before, in the data downscaling process of the IsoGCM model, 70% of the studied time data have been

Table 6 | Correlation coefficient between variables selected in each scenario of the stepwise regression model

Correlations

PO.1 TO PO.4					PO.1 & PO.3				
		T	**RH**	**P**			**T**	**RH**	**P**
P	Pearson correlation	0.034	0.433	1	P	Pearson correlation	0.029	0.408	1
	Sig.	0.132	0			Sig.	0.072	0	
	N	1,987	1,987	1,987		N	3,832	3,832	3,832
RH	Pearson correlation	−0.474	1	0.433	RH	Pearson correlation	−0.425	1	0.408
	Sig.	0		0		Sig.	0		0
	N	1,987	1,987	1,987		N	3,832	3,832	3,832
T	Pearson correlation	1	−0.474	0.034	T	Pearson correlation	1	−0.425	0.029
	Sig.		0.000	0.132		Sig.		0	0.072
	N	1,987	1,987	1,987		N	3,832	3,832	3,832

Table 7 | Regression coefficient for each scenario at 5% significant level

Class	**Scenario**	**Methods**	β_0	**NRH**	**NP**	**NT**
Annual simulation	PO.1 & PO.3	MLR	0.00	−0.435	−0.036	0.028
		GLM	0.628	−0.107	−0.001	−0.058
		WLSR	−0.003	−0.392	−0.053	0.041
	PO.1 TO PO.4	MLR	0.00	−0.406	−0.022	0.023
		GLM	0.576	−0.095	−0.009	−0.028
		WLSR	−0.005	−0.299	−0.053	0.155
Winter season	PO.1 & PO.3	MLR	0.00	−0.244	−0.17	0.272
		GLM	0.650	−0.10	0.013	0.027
		WLSR	0.005	−0.389	−0.038	0.092
	PO.1 TO PO.4	MLR	0.00	−0.275	−0.127	0.244
		GLM	0.536	−0.114	0.019	0.027
		WLSR	0.002	−0.295	−0.101	0.192
Autumn season	PO.1 & PO.3	MLR	0.00	−0.465	−0.036	0.028
		GLM	0.781	−0.206	−0.04	0.042
		WLSR	0.008	−0.482	−0.018	0.002
	PO.1 TO PO.4	MLR	0.00	−0.412	−0.053	0.041
		GLM	0.651	−0.186	−0.042	0.067
		WLSR	−0.004	−371	−0.007	−0.019
Spring season	PO.1 & PO.3	MLR	0.00	−0.635	−0.041	0.028
		GLM	0.583	−0.117	0.026	−0.018
		WLSR	−0.043	−0.293	−0.151	0.085
	PO.1 TO PO.4	MLR	0.00	−0.582	−0.028	0.028
		GLM	0.553	−0.102	0.022	−0.008
		WLSR	−0.098	−0.381	−0.121	0.015

randomly selected in order to train the regression model and the remaining 30% used for the model validation. In this section, we evaluate the regression models introduced in step C for simulating the data of the large-scale IsoGSM model using the appropriate statistical methods.

Investigating the standardized residual values. Figure 3 depicts the standardized residuals in the verification data (30% of the whole data separated by season) for each regression model and spatial scenario. All graphs represent the randomness of the residuals' pattern. This randomness

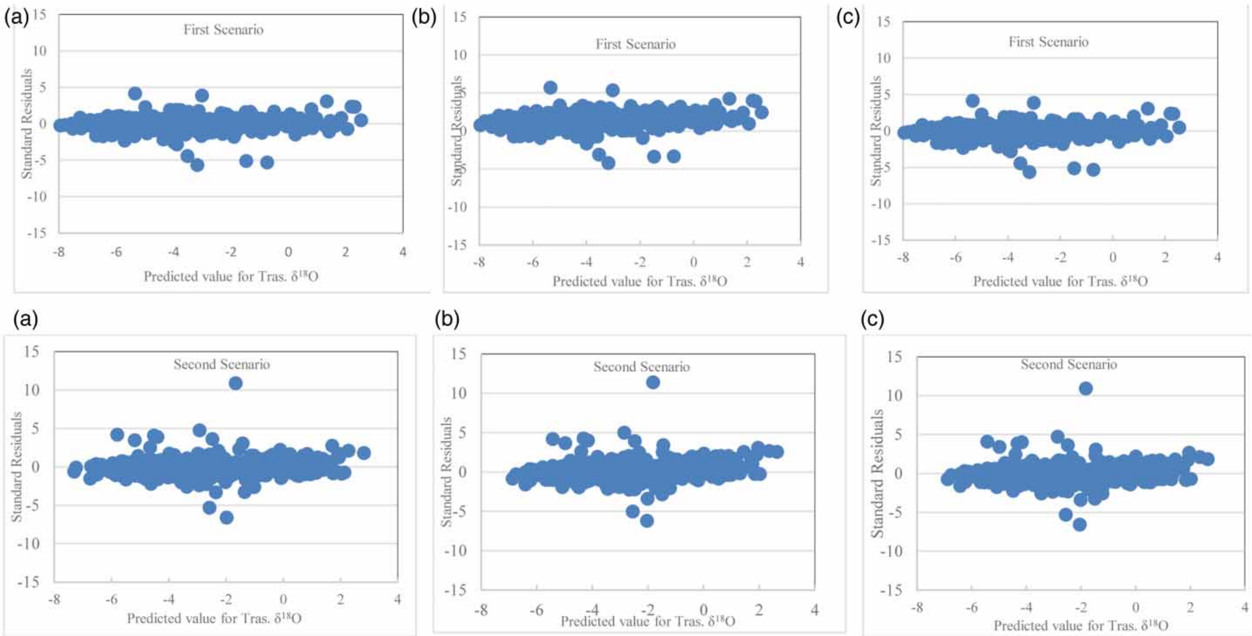

Figure 3 | Standardized residuals from downscaling models of (a) MLR, (b) GLM, and (c) WLSR against the predictor variable in the first and second scenarios, respectively.

of the residuals' pattern, indicates the proper fitting of the linear models to the simulated data of the large-scale IsoGSM model. On the other hand, the standardized residual error around the x-axis is approximately symmetrically distributed. This dispersion in the first spatial scenario (Points 1 and 3) has a more symmetric distribution than the second counterpart.

Box plot investigation. The simulated concentration value of the oxygen-18 isotope in winter using different models is plotted in Figure 4 for the first spatial scenario (Points 1 and 3).

Figure 4 shows a sample of variations of models' results in comparison with the IsoGSM model simulations in winter for the first spatial scenario (Points 1 and 3). As can be observed, the extreme values are not well simulated by all developed models while the simulation of long-term average value of isotopic concentrations is acceptable. This is because of the nature of regression-based models which are commonly weak in extreme values' simulation. The weakest results belong to the GLM model because the GLM model is not sensitive to extreme value. The performance of MLR and WLSR models are almost the same but better than the GLM.

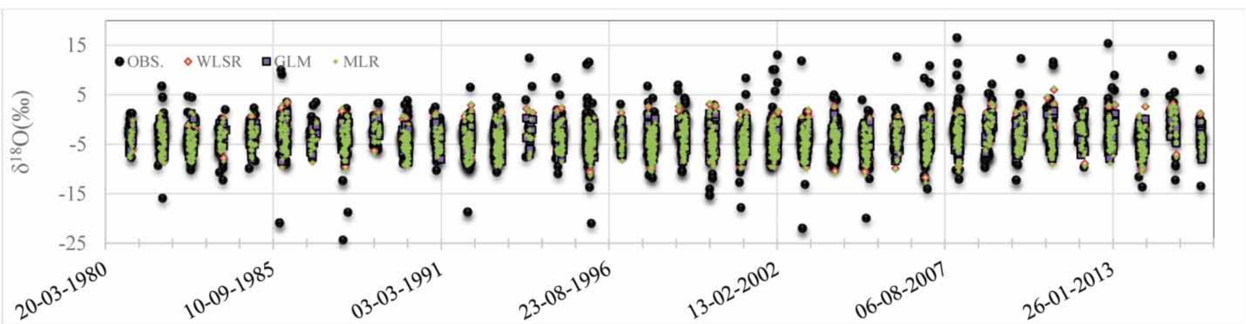

Figure 4 | Total isotope concentration time series of the three downscaling models compared to the IsoGSM model during the time study period in the first scenario.

Figure 5 shows the box plot associated with the simulation results of oxygen-18 isotope concentration in precipitation. It seems that the size of error bars in the MLR approach is very close to the observations' error. In the MLR model, the average simulated isotope concentration remains relatively stable for both scenarios. The amount of simulated isotope concentration using validation data is not accurately modeled in the second scenario with the help of the MLR model. The WLSR model indicates a reduction in the first and third quartiles rather than the observational data in the large-scale IsoGSM model's nodes. However, in the first scenario, this method illustrates decrement and increment in the third and first quartiles, respectively. The simulation error by the GLM model relative to the observations is illustrated in this graph as well. In the box plot of this technique, the average simulated isotope concentration for the first scenario is less than the observational one. However, the trend becomes reversed for the second scenario and the average simulated value is higher than the observational one. This method has not been successful even in simulating the training data. Also, the results of Figure 4 clarify that the maximum difference between the simulated and observational oxygen-18 isotope concentration occurs in

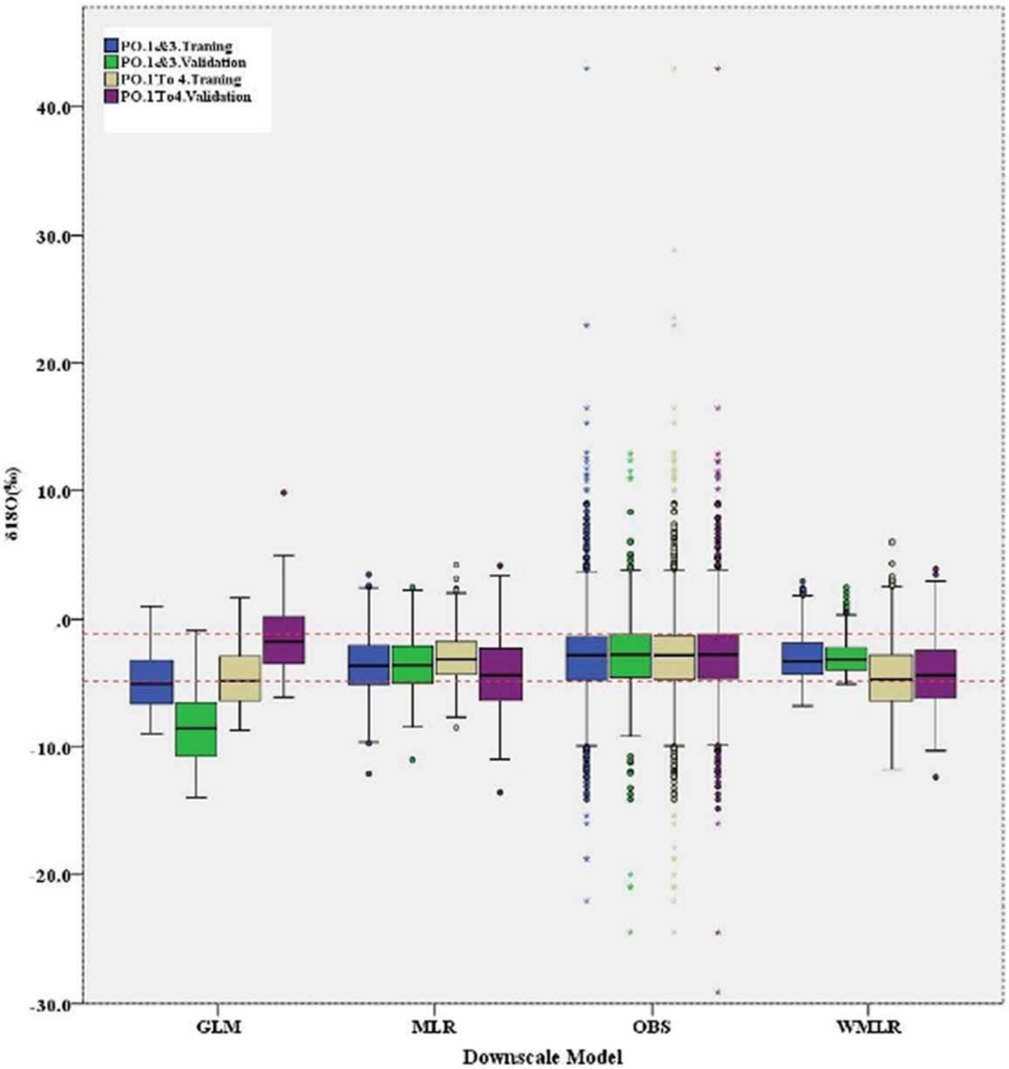

Figure 5 | JFD box plot of isotope concentration for the different modeling methods.

the GLM model; whereas, the results via the two models of MLR and WLSR are closer to the observational value in the large-scale IsoGSM.

Probability density function (pdf) investigation. The pdf investigation is considered as another test for evaluating the simulation performance of the modeling methods. Figure 6 depicts the pdf associated with each regression method used for downscaling the concentration data of oxygen-18 isotope in precipitation and two spatial scenarios during the winter. The pdf shape in the downscaling methods has a small skewness to the left. However, the pdf of the simulated data by the IsoGSM model shows a tendency to the normal distribution function. Among the three regression methods, the skewness amount in the pdf of the MLR technique is more similar to that of the IsoGSM model. In the first scenario, the GLM model has indicated the worst performance. According to Figures 4 and 5, this method underestimates the oxygen-18 isotope concentration compared to the real values. As shown by pdf forms, none of the methods are able to simulate the limit values in accordance with the observational function of the IsoGSM model. In both scenarios, the GLM and MLR models have been able to partially simulate the minimum limit values.

Table 8 shows the comparison of the data recorded in the large-scale IsoGSM model with the results of three methods applied for simulating the oxygen-18 isotope concentration in precipitation under the two defined scenarios. In the Sig column, there are no significant differences between the groups' average in the first scenario under the two modeling methods of MLR and WLSR; that is, the isotope modeling behavior is homogeneous relative to the information recorded in the large-scale model. However, the concentration simulation of oxygen-18 isotopes via the GLM model differs from the recorded data of the large-scale one. In the second scenario, only the WLSR method has no significant difference with the data recorded in the large-scale model.

Investigation of the performance indicators. Table 9 lists the performance indicator values associated with each regression model in two different spatial scenarios. As illustrated by the results, the two models of WLSR and MLR achieve better performance. These two models have the highest RS performance index for simulating the concentration of oxygen-18 isotopes in precipitation, respectively. The value of E_{NS} index in the regression model of GLM is equal to 0. This value indicates that the GLM model can successfully simulate the values close to the observational time series.

Validation of the model's results by means of the local measured data

As stated before, investigating the long-term changes in the rainfall's isotope concentrations is a suitable method for assessing the changes in the source of precipitation and occurrence of the climate change in a region. It has been attempted in this section to investigate the behavior of the selective downscaling models in the studied area. At this stage, the concentration of oxygen-18 was first calculated using the equations presented in Table 7 and using the weather data of the region and then compared with the results of the observational data (Table 1). Figure 7 shows the comparison of the regression models' achievements with the regional observational data in both spatial scenarios. As can be seen from this figure, all annual downscaling models (except the GLM method in the second scenario), estimate the concentration of oxygen-18 isotope to be less than the observational values. Further to this, with the separation of data and time separation, the simulation results seem to be somewhat satisfactory with the help of the seasonal equations. On the other hand, simulation results are closer to the observational values with the help of the first scenario. This means that the information classification according to Dansgaard rules has resulted in better simulation results.

CONCLUSION

Using long-term data of the large-scale isotopic models for climate investigation at local and regional levels is the primary challenge in climate studies. On the one hand, the use of information from these models at the local level requires the use of a downscaling process. On the other hand, it is difficult and costly to measure the concentration of isotope molecules in water, and expensive and

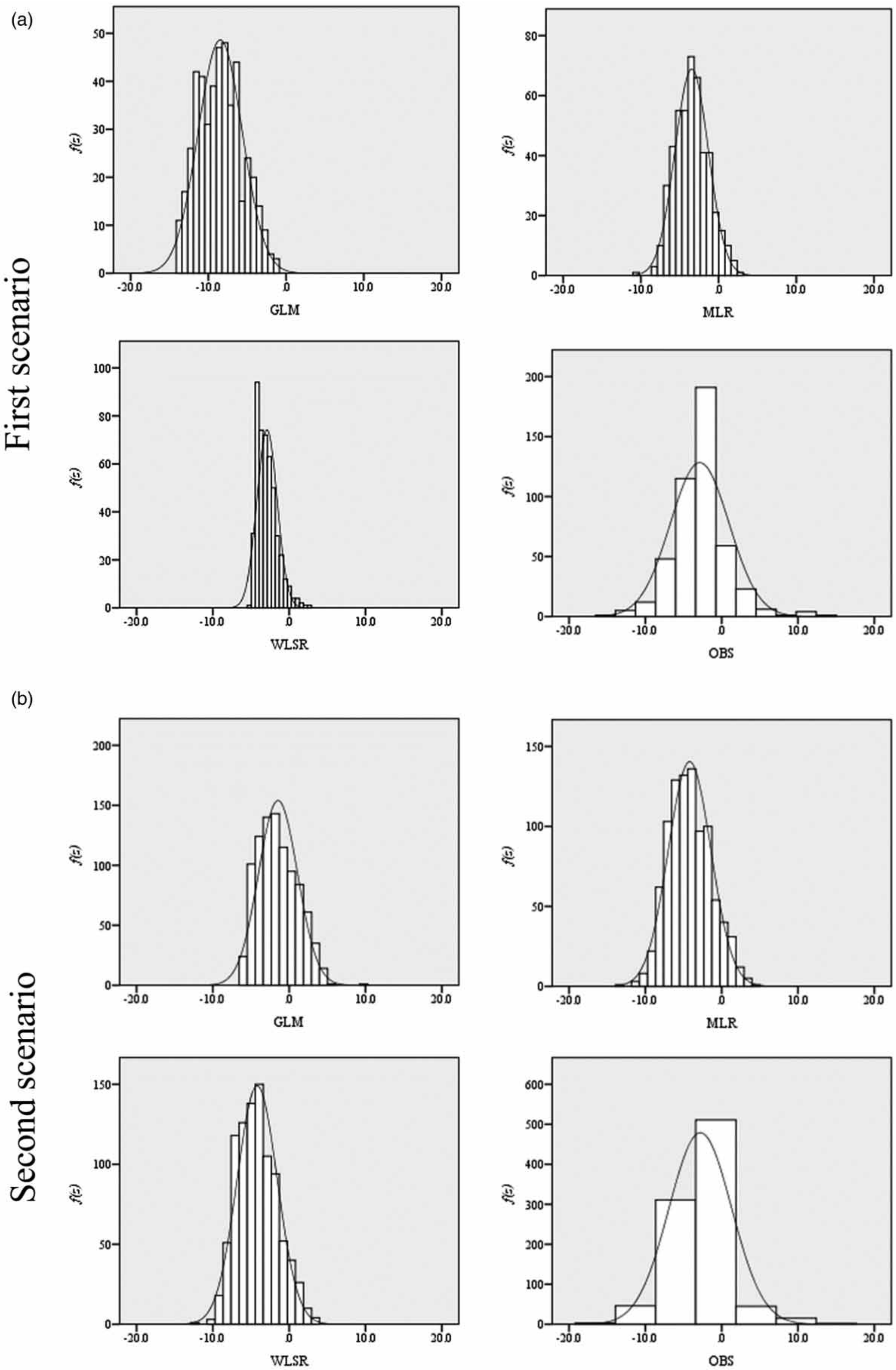

Figure 6 | Probability density function of $\delta^{18}O$ values: (a) first scenario and (b) second scenario in verification samples of IsoGSM database in JFD.

Table 8 | Paired-samples t-test results in JFD

Scenario		Mean	Std. error mean	95% Confidence interval of the difference		t	df	Sig.
				Lower	Upper			
First scenario	OBS-WLSR	−0.01	3.61	0.12	−0.24	0.22	935	0.945
	OBS-GLM	−1.33	3.384	0.13	−1.58	−1.09	937	0.000
	OBS-MLR	0.00	3.6	0.12	−0.23	0.23	935	1.000
Second scenario	OBS-WLSR	0.02	3.52	0.162	−0.30	0.34	469	0.906
	OBS-GLM	5.59	3.8	0.175	5.25	5.94	469	0.000
	OBS-MLR	0.58	3.41	0.157	0.27	0.88	469	0.000

Table 9 | Statistics of downscaling results on winter season (JFD months)

		Calibration				Validation			
		RSME (‰)	E_{NS}	RS	R^2	RSME	E_{NS}	RS	R^2
PO.1 & PO.3	MLR	0.02	0.22	0.47	0.22	0.03	0.18	0.47	0.23
	GLM	0.00	0.19	0.44	0.19	0.00	−0.19	0.37	0.17
	WLSR	0.03	0.91	0.44	0.37	0.03	0.18	0.44	0.22
PO.1 TO PO.4	MLR	0.01	0.18	0.43	0.21	0.02	0.22	0.43	0.22
	GLM	0.00	0.16	0.40	0.18	0.00	−0.09	0.64	0.16
	WLSR	0.00	0.95	1.94	0.26	0.02	0.22	0.44	0.23

delicate equipment is required. Also, the dispersion of measuring stations for isotopic concentration of precipitation near meteorological ones is very limited for regional studies. Therefore, the present study attempted to create a regular database from the time series of oxygen-18 isotope concentration at the local scale through downscaling the concentration of the precipitation isotopes which have been simulated in global large-scale models. This study dealt with the development and evaluation of three statistical downscaling models in order to predict the concentration of oxygen-18 isotope in precipitation. These models are capable of annually and seasonally presenting oxygen-18 isotope concentration in precipitation using a combination of meteorological parameters. Two spatial scenarios have also been defined here to consider the geographical changes on the isotopes' concentration. In the data downscaling process of the IsoGCM model, 70% of the studied time data have been randomly selected in order to train the regression model and the remaining 30% used for validating the model.

The above methodologies have their limitations. The main limitation of this study is the small sample size of isotopic data which is mostly because of the high cost of these kinds of measurements. Using a bigger sample of isotopic data will definitely improve the study results and outputs' accuracy. Furthermore, other weather variables such as geopotential height and surface pressure can be taken into account in developing the downscaling model. This is not considered in this study because of limitations in data accessibility.

The results of the three basic statistical downscaling models have been compared with the information of the large-scale isotopic (information used for the validation) one and the observational results in terms of the statistical downscaling ability of oxygen-18 isotope concentration in precipitation of Shiraz, southwest Iran. Comparison of the simulation results with the data via the global and regional large-scale data shows that the first spatial scenario utilizing data from nodes 1 and 3 has better performance than the second scenario which contains the information from four nodes around the study area. The standardized residual

Figure 7 | Comparison of the downscaling results of isotope concentration with the observational data in precipitation: (a) MLR, (b) WLSR and (c) GLM.

error around the x-axis in the first spatial scenario (Points 1 and 3) is more symmetric than the second counterpart. This is mainly due to the fact that the two nodes of 1 and 3 have approximate latitudes as wide as the study area. Comparison of the time separation information also shows the effect of season and time variations on the concentration of isotopes. Simulating oxygen-18 isotope concentration using the seasonal equations is more accurate than their annual simulation. Therefore, it can be concluded that the use of the statistical downscaling method in the case of information pre-processing based on their seasonal and spatial changes and the selection of a suitable method is a useful tool for monitoring climate change of a region based on stable oxygen-18 isotope information.

The maximum difference between the simulated and observational oxygen-18 isotope concentration occurs in the GLM model. Whereas, the results via the two models of MLR and WLSR are closer to the observational value in the large-scale IsoGSM model. Furthermore, the MLR and WLSR models have the highest RS performance index for simulating the concentration of oxygen-18 isotopes in precipitation. In contrast to the WLSR and GLM methods, the skewness amount in the pdf of the MLR technique is more similar to that of the IsoGSM model. Also, comparing the results obtained via the improved statistical simulation models with the observational data and those of the isotopic global large-scale model (information used for the verification), indicates that the WLSR and MLR models have better performance of the three selected statistical methods.

Using the equations presented in this research, one is able to draw the monthly concentration distribution map of oxygen-18 precipitation isotopes in Shiraz using meteorological variables (precipitation, temperature, and humidity) and utilize them for long-term ecological, hydrological, and hydrogeological studies by means of isotope trackers.

A review of technical literature shows that the isotopic concentration is affected by the type of precipitation. For further studies, downscaled isotopic data can be used for the investigation of climate change impacts on the type of rainfall.

REFERENCES

Aggarwal, P. K., Romatschke, U., Araguas-Araguas, L., Belachew, D., Longstaffe, F. J., Berg, P., Schumacher, C. & Funk, A. 2016 Proportions of convective and stratiform precipitation revealed in water isotope ratios. *Nature Geoscience* **9** (8), 624–629.

Akbari, H., Rakhshandehroo, G. R., Afrooz, A. H. & Pourtouiserkani, A. 2015 Climate change impact on intensity-duration-frequency curves in Chenar-Rahdar river basin. In *Watershed Management Symposium, 2015*, Reston, VA, USA.

Alijani, B. & Harman, J. R. 1985 Synoptic climatology of precipitation in Iran. *Annals of the Association of American Geographers* **75** (3), 404–416.

Argiriou, A. & Lykoudis, S. P. 2008 Improving the spatial coverage of stable isotope data. In: *New Topics in Water Resources Research and Management* (H. M. Andreassen, ed.). Nova Science Publishers, New York, USA, pp. 335–347.

Asadi, A. & Heydari, A. 2011 Analysis of temperature and precipitation changes of Shiraz during the period of 1951–2005. *Geography Environmental Planning (University Isfahan)* **22**, 37–40.

Bowen, G. J. & Revenaugh, J. 2003 Interpolating the isotopic composition of modern meteoric precipitation. *Water Resources Research* **39** (10), 1–13.

Che, Y., Zhang, M., Wang, S., Wang, J., Liu, Y. & Zhang, F. 2016 Stable water isotopes of precipitation in China simulated by SWING2 models. *Arabian Journal of Geosciences* **9** (19), 732.

Chen, S. T., Yu, P. S. & Tang, Y. H. 2010 Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *Journal of Hydrology* **385** (1–4), 13–22.

Dansgaard, W. 1964 Stable isotopes in precipitation. *Tellus* **16** (4), 436–468.

Dee, S., Noone, D., Buenning, N., Emile-Geay, J. & Zhou, Y. 2015 SPEEDY-IER: a fast atmospheric GCM with water isotope physics. *Journal of Geophysical Research: Atmospheres* **120** (1), 73–91.

Delavau, C., Stadnyk, T. & Birks, J. 2011 Model based spatial distribution of oxygen-18 isotopes in precipitation across Canada. *Canadian Water Resources Journal/Revue Canadienne des Ressources Hydriques* **36** (4), 313–330.

Haese, B., Werner, M. & Lohmann, G. 2013 Stable water isotopes in the coupled atmosphere–land surface model ECHAM5-JSBACH. *Geoscientific Model Development* **6**, 1463–1480.

Henderson-Sellers, A., Fischer, M., Aleinov, I., McGuffie, K., Riley, W. J., Schmidt, G. A., Sturm, K., Yoshimura, K. & Irannejad, P. 2006 Stable water isotope simulation by current land-surface schemes: results of iPILPS phase 1. *Global and Planetary Change* **51** (1–2), 34–58.

Hoffmann, G., Jouzel, J. & Masson, V. 2000 Stable water isotopes in atmospheric general circulation models. *Hydrological Processes* **14** (8), 1385–1406.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J. L., Fairhead, L., Filiberti, M. A., Friedlingstein, P., Grandpeix, J. Y. & Krinner, G. 2006 The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics* **27** (7–8), 787–813.

Karamouz, M., Zahmatkesh, Z., Nazif, S. & Razmi, A. 2014 An evaluation of climate change impacts on extreme sea level variability: coastal area of New York City. *Water Resources Management* **28** (11), 3697–3714.

Kigawa, S. 2014 Techniques of precipitation analysis and prediction for high-resolution precipitation nowcasts. *Weather Service Bulletin* **81**, 22.

Kim, K. & Lee, X. 2011 Isotopic enrichment of liquid water during evaporation from water surfaces. *Journal of Hydrology* **399** (3–4), 364–375.

Lee, J. E., Fung, I., DePaolo, D. J. & Henning, C. C. 2007 Analysis of the global distribution of water isotopes using the NCAR atmospheric general circulation model. *Journal of Geophysical Research: Atmospheres* **112** (D16), 1–14.

Liebminger, A., Haberhauer, G., Varmuza, K., Papesch, W. & Heiss, G. 2006 Modeling the oxygen 18 concentration in precipitation with ambient climatic and geographic parameters. *Geophysical Research Letters* **33** (5), 1–5.

Pourtouiserkani, A., Rakhshandehroo, G. R. & Akbari, H. 2015 Comparison of statistical downscaling methods in projecting climate change impact on extreme runoff events in Chenar Rahdar River basin (Iran). *Watershed Management* **2015**, 23–35.

Rahimzadeh, F., Asgari, A. & Fattahi, E. 2009 Variability of extreme temperature and precipitation in Iran during recent decades. *International Journal of Climatology* **29** (3), 329–343.

Rozanski, K. & Gonfiantini, R. 1990 Isotopes in climatological studies. *IAEA Bulletin* **32** (4), 9–15.

Silhavy, R., Silhavy, P. & Prokopova, Z. 2017 Analysis and selection of a regression model for the use case points method using a stepwise approach. *Journal of Systems and Software* **125**, 1–14.

Sutanto, S. J., Hoffmann, G., Worden, J., Scheepmaker, R. A., Aben, I. & Roeckmann, T. 2015 Atmospheric processes governing the changes in water isotopologues during ENSO events from model and satellite measurements. *Journal of Geophysical Research: Atmospheres* **120** (13), 6712–6729.

Tharammal, T., Bala, G. & Noone, D. 2017 Impact of deep convection on the isotopic amount effect in tropical precipitation. *Journal of Geophysical Research: Atmospheres* **122** (3), 1505–1523.

Xi, X. 2014 A review of water isotopes in atmospheric general circulation models: recent advances and future prospects. *International Journal of Atmospheric Sciences* **2014**, 1–16.

Yao, T., Masson-Delmotte, V., Gao, J., Yu, W., Yang, X., Risi, C., Sturm, C., Werner, M., Zhao, H., He, Y. & Ren, W. 2013 A review of climatic controls on $\delta 18O$ in precipitation over the Tibetan Plateau: observations and simulations. *Reviews of Geophysics* **51** (4), 525–548.

Yoshimura, K., Kanamitsu, M., Noone, D. & Oki, T. 2008 Historical isotope simulation using reanalysis atmospheric data. *Journal of Geophysical Research: Atmospheres* **113** (D19), 1–15.

Yoshimura, K., Kanamitsu, M. & Dettinger, M. 2010 Regional downscaling for stable water isotopes: a case study of an atmospheric river event. *Journal of Geophysical Research: Atmospheres* **115** (D18), 1–13.

Yoshimura, K., Frankenberg, C., Lee, J., Kanamitsu, M., Worden, J. & Röckmann, T. 2011 Comparison of an isotopic atmospheric general circulation model with new quasi-global satellite measurements of water vapor isotopologues. *Journal of Geophysical Research: Atmospheres* **116** (D19), 1–15.

Zakharov, V. I., Imasu, R., Gribanov, K. G., Hoffmann, G. & Jouzel, J. 2004 Latitudinal distribution of the deuterium to hydrogen ratio in the atmospheric water vapor retrieved from IMG/ADEOS data. *Geophysical Research Letters* **31** (12), 1–4.

Zhang, X., Sun, Z., Guan, H., Zhang, X., Wu, H. & Huang, Y. 2012 GCM simulations of stable isotopes in the water cycle in comparison with GNIP observations over East Asia. *Acta Meteorologica Sinica* **26** (4), 420–437.

First received 25 September 2019; accepted in revised form 1 April 2020. Available online 16 May 2020