Global and Planetary Change 144 (2016) 158–170

Contents lists available at ScienceDirect

Global and Planetary Change

journal homepage: www.elsevier.com/locate/gloplacha

Projection of climate change impacts on precipitation using soft-computing techniques: A case study in Zayandeh-rud Basin, Iran

Shapour Kouhestani [a,*], Sayed Saeid Eslamian [a], Jahangir Abedi-Koupai [a], Ali Asghar Besalatpour [b]

[a] Department of Water Engineering, College of Agriculture, Isfahan University of Technology, Isfahan PO. Code 8415683111, Iran
[b] Department of Soil Science, College of Agriculture, Vali-e-Asr University of Rafsanjan, Rafsanjan PO. Code 7718897111, Iran

ARTICLE INFO

Article history:
Received 20 March 2016
Received in revised form 3 July 2016
Accepted 21 July 2016
Available online 26 July 2016

Keywords:
Downscaling
Supervised PCA
Climate change
Supervised learning

ABSTRACT

Due to the complexity of climate-related processes, accurate projection of the future behavior of hydro-climate variables is one of the main challenges in climate change impact assessment studies. In regression-based statistical downscaling processes, there are different sources of uncertainty arising from high-dimensionality of atmospheric predictors, nonlinearity of empirical and quantitative models, and the biases exist in climate model simulations. To reduce the influence of these sources of uncertainty, the current study presents a comprehensive methodology to improve projection of precipitation in the Zayandeh-Rud basin in Iran as an illustrative study. To reduce dimensionality of atmospheric predictors and capture nonlinearity between the target variable and predictors in each station, a supervised-PCA method is combined with two soft-computing machine-learning methods, Support Vector Regression (SVR) and Relevance Vector Machine (RVM). Three statistical transformation methods are also employed to correct biases in atmospheric large-scale predictors. The developed models are then employed on outputs of the Coupled Model Intercomparison Project Phase 5 (CMIP5) multimodal dataset to project future behavior of precipitation under three climate changes scenarios. The results indicate reduction of precipitation in the majority of the sites in this basin threatening the availability of surface water resources in future decades.

1. Introduction

The acceleration of man-made greenhouse gas emissions as a result of economic expansion and population growth is known as the most important contribution to the global warming in the recent decades. The expected increasing trend of global temperature due to the likely continuous release of greenhouse gases in the next decades can potentially affect the hydrological cycle from global to regional scales (Immerzeel et al., 2012; Seager et al., 2013; Sorg et al., 2012; Wu et al., 2012; Yang et al., 2014). These changes are expected to increase the frequency and severity of extreme hydro-climatic events and to threaten the availability of surface water resources in different parts of the world (Du et al., 2014; Fahrenkamp-Uppenbrink, 2015; Hirabayashi et al., 2013; Orlowsky and Seneviratne, 2012; Sunyer et al., 2012). The projection of the impact of climate change on the behavior of hydro-climatic variables is therefore one of the main challenges for water resources managers.

To project the behavior of hydro-climatic variables, Global Circulation Models (GCMs) are considered one of the sources of future climate change information. GCMs are, however, available in coarse spatial scale and unable to project the behavior of hydro-climate variables in finer local scales that is of interest to hydrologists. Due to this limitation, GCMs are restricted to assess the impact of climate change on surface water availability. To address this problem, statistical and dynamical downscaling techniques are developed to bridge large-scale GCMs information to local hydro-climatic variables (Chen et al., 2013; Najafi and Moradkhani, 2015). Requirements such as expensive and complicated computations and the use of biased lateral boundary inputs limit dynamic downscaling techniques to be used in climate change impact assessments. Statistical downscaling methods are, however, popular to hydrologists as they are relatively straightforward to apply (Rocheta et al., 2014). Statistical techniques are based on developing statistical and quantitative relationships between large-scale atmospheric variables, which GCMs provide and local scale hydro-climatic variables. Among the three main categories of the statistical downscaling methods, regression-based techniques are preferred over the other two techniques, including weather typing and transfer function, and weather generation (Najafi and Moradkhani, 2015). The regression-based statistical downscaling techniques are generally carried out in two steps, (i) developing a functional relationship based on historical records of local hydro-climatic and large-scale atmospheric variables, and (ii) projecting local hydro-climatic behavior in future based on the developed model in the first step for climate change scenarios. The procedure relies on the assumption that current empirical models are applicable to the future GCM simulations. In the first step, there are main challenges to develop empirical relationship between simulated

* Corresponding author.
E-mail address: s.kouhestani@ag.iut.ac.ir (S. Kouhestani).

http://dx.doi.org/10.1016/j.gloplacha.2016.07.013
0921-8181/

large-scale atmospheric variables and target hydro-climatic variable. As the first challenge, raw GCMs are unable to represent variability of large-scale climate modes (teleconnections) on large-scale atmospheric variables resulting in uncertainty and biases in projection of hydro-climatic behaviors. On the other hand, an initial pre-processing correction must be implemented based on historical observations to remove the discrepancies between observed and simulated large-scale atmospheric variables. Many techniques have been implemented to address this source of uncertainty (Fowler et al., 2007; Mehrotra and Sharma, 2012, 2015; Ojha et al., 2012). One of the popular and currently used methods is statistical transformations method that aims to adjust the distribution of GCM simulations with the observed climate (Gudmundsson et al., 2012). The main advantage of the method is that it adjusts the entire distributions (including all moments), while maintaining the rank correlation between observations and simulations (Li et al., 2010). This method is such a relatively simple approach that different forms of it have been successfully implemented in different statistical downscaling studies (Boé et al., 2007; Gudmundsson et al., 2012; Johnson and Sharma, 2011; Li et al., 2010).

After correcting inherent biases in GCMs, two other challenges remain, including i) the identification of proper and relevant atmospheric predictors, and ii) the development of an appropriate nonlinear functional relationship between target hydro-climate variable and GCM predictors. In term of selecting appropriate projectors conveying relevant information associated with the target variable, a predictor domain in terms of spatial extent and geographical location is chosen. In this procedure several related projectors from surrounding GCMs with respect to the location of the study site are extracted. Projecting a target hydro-climatic variable from a set of high-dimensional predictors in regression-based statistical downscaling leads to inadequate performance accuracy. The reason is due to the existence of correlation among projectors which results into redundancy and collinearity (Sarhadi et al., 2016). Despite its importance, this challenge has not gained enough attention in statistical downscaling studies. The majority of studies use conventional dimensionality reduction methods, including Principal Component Analysis (PCA) (Shashikanth et al., 2014; Tripathi et al., 2006), Canonical Correlation Analysis (CCA) (Joshi et al., 2013; Wójcik, 2015), and different linear types of multivariate linear regression methods (Ghosh and Mujumdar, 2008; Tisseuil et al., 2010). Sarhadi et al., (2016) introduced a kernelized version of Supervised-PCA for dimensionality reduction in statistical downscaling. The technique takes into account the response variable and seeks a sequence of principal components having maximum dependency with the target variable. The method is also able to capture nonlinear variabilities between the target variable and atmospheric projectors. These advantages address the above-mentioned drawbacks of dimensionality and results in enhancing the performance accuracy of the statistical downscaling process.

Another challenge in regression-based statistical downscaling is the development of reliable empirical function to capture the complex nonlinear relationships between target variable and predictors. Because of complex nonlinear relationship between hydro-climatic variables and atmospheric predictors, conventional linear approaches fail to capture the nonlinearity. To address this challenge, in recent years, research on nonlinear-based soft computing methods have become popular. Among all methods, machine-learning methods have gained more popularity in statistical downscaling. Among these types of methods, Support Vector Regression (SVR) method has been widely used in hydrology for nonlinear stochastic modeling (Chen et al., 2010; Nasseri et al., 2013). In spite of many advantages in modeling, the SVR suffers from overfitting in prediction in statistical downscaling. A Spare Bayesian Learning (SBL) algorithm known as Relevance Vector Machine (RVM) has covered all the drawbacks of the SVR method by using fewer kernel functions, which avoid over-fitting. This method has recently been used in statistical downscaling studies and shown to improve the performance of the regression-based statistical downscaling processes (Joshi et al., 2013).

Overall, the present study focuses on reducing uncertainty arising from different aspects of the regression-based statistical modeling to improve the reliability of hydro-climate variable projections. Therefore, different sources of uncertainty, including biases in GCM simulations, high-dimensionality, and nonlinear transfer functions are addressed in this study by introducing a comprehensive combined framework. Using the comprehensive methodology helps water managers better assess the future impact of climate change on the availability of surface-water resources in regions of interest.

The rest of the paper is organized as follows: Section 2 introduces the study area and the dataset used for statistical downscaling modeling. Section 3 describes the background of the comprehensive methodology employed for the regression-based statistical modeling in the present study. The results and discussion are presented in Section 4. Conclusions are given in Section 5.

2. Study area and dataset

The Zayandeh-Rud basin, as one of main basins in central Iran is selected as the domain for the present study. Originated from the Zagrous Mountains (Fig. 1), the Zayandeh-Rud River with an average flow of 1400 million cubic meters (MCM) per year is the most important available water resource for domestic, industrial and agriculture consumptions in the basin. Currently, more than 3.7 million people are living in the basin, making it the second most populated water basin in Central Iran. The population growth accompany with the occurrence of severe drought events have negatively influenced surface water resources availability for different sectors, including agricultural (as the main water consumer), industrial and drinking-water sectors in the area. A number of engineering-based alternatives, including developing multi-purpose reservoirs and water transfer projects have been implemented to address the water shortage in the basin in recent years. Such projects have, however, led to water conflict between neighboring watersheds and provinces.

The results of projections based on previous research using the Fourth Assessment Report (AR4) of Intergovernmental Panel on Climate Change (IPCC) show that the Zayandeh-Rud river basin will experience more severe water shortage under a changing climate (Elmahdi et al., 2009; Gohari et al., 2013; Zareian et al., 2014).

In the current study, the GCMs outputs from the newest version of the Fifth Assessment Report (AR5) of Intergovernmental Panel on Climate Change (IPCC) are used to project the impact of climate change on the availability of surface water resources in the Zayandeh-Rud basin. For this purpose, the recorded monthly precipitation time series in eighteen meteorological stations are employed in the current study (location of the sites is illustrated in Fig. 1). The precipitation data provided by the Iran Meteorological Organization (IRIMO) cover the time span of 1951 to 2010 (the record length varies in different stations). As a proxy of observed large-scale atmospheric predictors, National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis data is employed. The reanalysis data are available from 1948 to present and cover observations from 1948 to present with spatial resolution of $2.5° \times 2.5°$. In the pre-processing step of the statistical downscaling, NCEP/NCAR data are used as predictors to correct systematic biases in historical simulated GCMs. It is therefore important to select the best predictor variables from the reanalysis data. Selected predictors should follow three implicit assumptions, including i) predictors are relevant to the host GCMs and can simulated by them realistically, ii) the empirical relationship developed under present climate conditions is also valid for future altered climate conditions, iii) selected predictors are able to represent climate change signals (Hewitson and Crane, 1992). According to these assumptions, the main identified atmospheric predictors are precipitation, mean, maximum, and minimum air temperature variables, mean sea level pressure, surface specific and relative humidity, and geopotential height at three pressure levels (1000, 500, 250 hPa). All of these predictors are