Atmósfera 29(2), 121-128 (2016)
doi: 10.20937/ATM.2016.29.02.02

Application of several data-driven techniques to predict a standardized precipitation index

BAHRAM CHOUBIN
Faculty of Natural Resources, Sari University of Agriculture Sciences and Natural Resources, Sari, Iran
Corresponding author; email: Bahram368@gmail.com

ARASH MALEKIAN
Faculty of Natural Resources, University of Tehran, Karaj, Iran

MOHAMMAD GOLSHAN
Faculty of Natural Resources, Sari University of Agriculture Sciences and Natural Resources, Sari, Iran

Received: March 25, 2015; accepted: February 22, 2016

RESUMEN

La modelación y predicción del clima son importantes para la gestión de recursos hidráulicos, especialmente en regiones áridas y semiáridas que con frecuencia sufren escasez de agua. La cuenca de Maharlu-Bakhtegan es una región árida y semiárida de 31 000 km^2 localizada al suroeste de Irán, de modo que la precipitación y escasez de agua en esta zona son muy problemáticas. Este estudio presenta una aproximación a la modelación del índice de sequía con base en índices climáticos de larga duración y el uso del sistema adaptativo de inferencia neurodifusa (ANFIS, por sus siglas en inglés), el árbol de decisión M5P y el modelo perceptrón multicapa (MLP, por sus siglas en inglés). Primero se determinó la mayoría de las señales climáticas a partir de 25 señales climáticas utilizando análisis factorial, y posteriormente se predijo un índice estandarizado de precipitación mediante las técnicas ANFIS, MLP y M5P con anticipación de uno a 12 meses. La evaluación de la aptitud del modelo mediante parámetros de error y diagramas de Taylor demostró que el desempeño del MLP es mejor que el de los otros dos modelos. Los resultados también mostraron que la exactitud de la predicción aumentó de manera considerable cuando se utilizaron índices climáticos del mes previo (t – 1) (RMSE = 0.802, ME = –0.002 y PBIAS = –0.47).

ABSTRACT

Climate modeling and prediction is important in water resources management, especially in arid and semi-arid regions that frequently suffer further from water shortages. The Maharlu-Bakhtegan basin, with an area of 31 000 km^2 is a semi-arid and arid region located in southwestern Iran. Therefore, precipitation and water shortage in this area have many problems. This study presents a drought index modeling approach based on large-scale climate indices by using the adaptive neuro-fuzzy inference system (ANFIS), the M5P model tree and the multilayer perceptron (MLP). First, most of the climate signals were determined from 25 climate signals using factor analysis, and subsequently, the standardized precipitation index (SPI) was predicted one to 12 months in advance with ANFIS, the M5P model tree and MLP. The evaluation of the models performance by error parameters and Taylor diagrams demonstrated that performance of the MLP is better than the other models. The results also revealed that the accuracy of prediction increased considerably by using climate indices of the previous month (t – 1) (RMSE = 0.802, ME = –0.002 and PBIAS = –0.47).

Keywords: Standardized precipitation index (SPI)**,** climate signals**,** multi-layer perceptron (MLP), adaptive neuro-fuzzy inference system (ANFIS), M5P model tree, Taylor diagrams.

1. Introduction

Drought is a climate feature that occurs occasionally. This phenomenon, which affects more people than any other hazard, is considered by many to be the most complex but least understood of all the natural vulnerabilities (Mishra and Desai, 2005). In Iran, arid climate extends to an area of 573 884 km^2 (35.54% of the territory). The Maharlu-Bakhtegan basin is located in this area; therefore, precipitation and water shortage in this region are very problematic. Meteorological drought occurs when the precipitation average is less than the precipitation average during the long-term period. In Iran, we are confronting challenges in many areas that have arid and semi-arid climates and suffer drought events, so it is necessary to pay more attention to precipitation. Pre-knowledge of the possible amount of precipitation is important in planning water recourses, management of agriculture and droughts, etc. Previous studies show that large-scale climate modes (e.g., North Atlantic Oscillation [NAO], South Oscillation Index [SOI]) have an influence on climate and precipitation in different parts of the world (Nazemosadat and Cordey, 2000; Karabörk *et al.*, 2005; Gaughan and Waylen, 2012; Berg *et al.*, 2013; Choubin *et al.*, 2014b).

In this study, we used large-scale climate indices for predicting the standard precipitation index (SPI). Among the several proposed drought monitoring indices, SPI has widespread application for describing and comparing droughts among different time periods and regions with different climatic conditions (Cancelliere *et al.*, 2007). SPI prediction is a critical issue that has attracted much attention in recent decades all over the world in order to carry out hydrological modeling in arid and semi-arid regions (Rezaeian-Zadeh *et al.*, 2012). Today, more non-linear models are applied to prediction. In previous studies, Dahamsheh and Aksoy (2009), Azadi and Sepaskhah (2012), and Rezaeian-Zadeh *et al.* (2012) used artificial neural networks (ANNs), and El-Shafie *et al.* (2011), Sanikhani and Kisi (2012), Jeong *et al.* (2012), and Choubin *et al.* (2014a) successfully applied the adaptive neuro-fuzzy inference system (ANFIS) to predict precipitation. In eastern Australia, Deo and Sahin (2015) investigated the application of the ANN model for the prediction of monthly SPIs using hydrometeorological parameters and climate indices. The results showed that the ANN model is a useful data-driven tool for forecasting monthly SPIs. In the Awash River Basin (Ethiopia), Belayneh *et al.* (2014) forecasted the long term SPI drought using wavelet neural networks. The forecasted results indicated that the coupled wavelet neural network (WA-ANN) models were better than all the other models in this study for forecasting SPI 12 and SPI 24 values. Ruigar and Golian (2016) predicted the precipitation in the Golestan dam watershed using climate indices: their results indicated that the MLP model is capable of accurately predicting monthly maximum precipitation.

In this study we compared the performances of three modeling techniques for predicting drought in a 43-yr period (1967-2009) in the Maharlu-Bakhtegan basin of Iran. We used the M5P model tree in addition to ANFIS and the multilayer perceptron (MLP) network to predict the SPI using large-scale climate indices as input data, over lead times of 1 to 12 months.

2. Methodology

2.1 Study area

The Maharlu-Bakhtegan basin spreads over 31 000 km^2. This area located in southwestern Iran (29º 00′ to 31º 14′ N, 51º 42′ to 54º 31′ W), with annual precipitation of 270 mm, is one of the most important agricultural centers of Iran (Fig. 1). In this paper, precipitation data were collected from the Iranian Water Resource Management Company (TAMAB) for four meteorological stations: Shiraz synoptic station, Dashtbal, Ali Abad Khatr and Dehkade Shahid. First, station data were analyzed and missing data were reconstructed by using the correlation method; then homogeneity and independence of data were evaluated using the run-test method. Homogeneity and dependence were accepted at a high level. We used Thiessen polygons between stations to calculate the average of watershed precipitation.

2.2 Standard precipitation index

The SPI was formulated by McKee *et al.* (1993) in the Colorado Climate Center. It is a relatively new drought index based only on precipitation, which is very important to farmers and responds fairly immediately to rainfall or dryness. This index is the number of standard precipitation deviations that the observed value would deviate from the long-term climatological average. Either a gamma distribution or a Pearson type III distribution is used for its transformation into a normal distribution (Guttman, 1999). It can be