Remote Sensing of Environment 211 (2018) 229–247

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A fusion-based methodology for meteorological drought estimation using remote sensing data

Mohammad Reza Alizadeh, Mohammad Reza Nikoo*

Department of Civil and Environmental Engineering, Shiraz University, Shiraz, Iran

A R T I C L E I N F O

Keywords:
Data fusion
Remote sensing data
Ordered weighted averaging
Nonparametric standardized precipitation index (nonparametric-SPI)
K-nearest neighbors algorithm (KNN)

A B S T R A C T

An effective planning and management to deal with potential impacts of drought requires accurate estimation and analysis of this natural complex phenomenon. Application of new fusion approaches using high-resolution satellite-based products, unlike ground-based observations, can provide accurate drought analysis. This study examines three advanced fusion-based methodologies including Ordered Weighted Averaged (OWA) approach based on ORNESS weighting method (ORNESS-OWA) and ORLIKE weighting method (ORLIKE-OWA) as well as K-nearest neighbors algorithm (KNN) to fuse estimations by five individual estimator models using different remotely sensed data products. The precipitation data from Global Precipitation Climatology Project (GPCP), CPC Merged Analysis of Precipitation (CMAP), CICS High-Resolution Optimal Interpolation Microwave Precipitation from Satellites (CHOMPS), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR), Tropical Rainfall Measuring Mission (TRMM), The second Modern-Era Retrospective analysis for Research and Applications (MERRA-2) and Global Land Data Assimilation System Version-2 (GLDAS-2) products is utilized in estimating nonparametric-SPI as a meteorological drought index versus ground-based observations analysis. To achieve more accurate drought estimation, ground-based observations are classified in different clusters based on K-means clustering algorithm. Five individual Artificial Intelligence (AI) models including Multi-Layer Perceptron (MLP), Adaptive Neuro-Fuzzy Inference System (ANFIS), M5P model tree, Group Method of Data Handling (GMDH) and Support Vector Regression (SVR) are developed for each cluster and their best results are used in fusion process. In addition, the Genetic Algorithm (GA) optimization model is utilized to determine optimal weights in weighting methods. Estimation performance of all models are evaluated using statistical error indices of Mean Absolute Relative Error (MARE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2). Application of proposed methodology is verified over Fars province in Iran and the results are compared. Results showed that ORNESS-OWA method with lowest estimation error (MARE of 2.51% and R^2 of 95%) had the superb performance in comparison with all other individual AI and fusion-based models. Also, the proposed framework based on remotely sensed precipitation data and fusion-based models demonstrated an effective proficiency in drought estimation.

1. Introduction

Drought as a part of natural climate variability have had significant social, economic and environmental impacts on human life in many parts of the world (Sheffield et al., 2012). Accurate estimation of drought provides useful information for many decision makers such as hydrologists, agriculturalists, water managers, and industrialists in future planning and management. In last decades, many researches have been performed to monitor and analyze droughts using ground-based observations or interpolated grids (e.g., Shen and Tabios, 1996; Henriques and Santos, 1999; Santos et al., 2010). In general, ground-based observations from synoptic and rain gauge stations are very limited in many areas such as over the oceans, mountains, scattered islands and inaccessible land areas (Ragab and Prudhomme, 2002). In addition, sampling error, the lack of sufficient historical record of observations and limited monitoring stations for drought-related variables such as precipitation have created many obstacles for an accurate drought estimation (Easterling, 2013). In the past decades, the emergence of different types of satellites has provided alternative monitoring methods based on remote-sensing satellite techniques. These remote-sensing techniques have opened a new window in drought monitoring from different perspectives (e.g., meteorological, agricultural,

* Corresponding author.
E-mail address: nikoo@shirazu.ac.ir (M.R. Nikoo).

https://doi.org/10.1016/j.rse.2018.04.001
Received 25 July 2017; Received in revised form 4 March 2018; Accepted 2 April 2018

hydrological, and ecological) (Wardlow et al., 2012; Zhang et al., 2017a). New remote sensing products provide useful datasets and information including precipitation, snow, soil moisture, land surface temperature, evaporation, total water storage, vegetation, and land cover in order to carry out a precise drought monitoring and analysis in many parts of the world (NASA, 2010; Lillesand et al., 2014). Satellite-based information is often consistent and with high spatial resolution which is available over the internet (Morgan, 1989; Sahoo et al., 2015; Holgate et al., 2016). Many researchers have sought to take advantage of satellite remote sensing as an efficient tool for measuring and monitoring of drought-related variables on a global scale (e.g., Krajewski et al., 2006; Sorooshian et al., 2011; Wardlow et al., 2012; Sadegh et al., 2017). While the application of remotely sensed products provide many advantages for drought analysis, there are major challenges to enhance the accuracy, reliability and interoperability of estimations and analyses using different methods and models.

Accordingly, many efforts have been made to increase the accuracy of drought predictions using different models, among them hydrological models (Brown et al., 2015), Markov Chain models (e.g., Paulo et al., 2005; Paulo and Pereira, 2008; Rahmat et al., 2017) and autoregressive integrated moving average models (ARIMA) (e.g., Mishra and Desai, 2005, 2006; Mishra et al., 2007; Han et al., 2010). Since these models are linear models, they are not very effective in forecasting nonlinearities associated with hydrologic data. In this regard, many researchers have applied data-driven models in drought analysis. For instance, Artificial Neural Networks (ANNs) have been used to forecast drought in several studies (e.g., Morid et al., 2007; Barros and Bowden, 2008; Bacanli et al., 2009; Karamouz et al., 2009; Marj and Meijerink, 2011). However, like other stochastic models, ANNs are incapable to deal with non-stationarities in drought estimations.

In the past years, the need for increased accuracy and precision in data-driven models has motivated the researchers to develop new approaches such as multi-model fusion-based methods. The main concept of data fusion is combining or amalgamating data derived from fused information in order to provide more accurate estimations in comparison with using single-source data alone (Dasarathy, 1997). Data fusion has been recently used by many researchers in many fields of research such as hydrological engineering (Shu and Burn, 2004), river-level forecasting (See and Abrahart, 2001), flood analysis (Shu and Burn, 2004). At more complicated features such as drought analysis, in particular, fusion-based estimations could be applied to integrate strengths of individual estimation models. Many studies have suggested that data fusion and aggregation concepts can be more accurate and reliable to derive drought indices (e.g., Azmi et al., 2010; Park et al., 2017). Table 1 presents a summary of relevant literature on drought analysis studies using data-driven models, specifically those based on fusion techniques using ground-based and remotely sensed data.

Reviewing the past studies, different Artificial Intelligence (AI) models have been used for drought forecasting by individual drought indices such as SPI (e.g., Belayneh and Adamowski, 2013; Jalalkamali et al., 2015; Deo et al., 2017). On the other side, some researchers have tried to derive new drought indices using different methods such as fusion concepts instead of individual drought indices alone (e.g., Barua et al., 2012; Zhang and Jia, 2013; Azmi et al., 2016). However, previous fusion-based methods in drought analysis still suffer from some deficiencies such as process of combinations, relying on input parameters selection and determination of optimal weights for fusion models. Obviously, there is no study that have considered the application of fusion-based methods directly in fusing multiple data from different data-driven models and combine them simultaneously for drought analysis based on remote sensing products. In this study, we describe a new methodology based on three advanced data fusion methods and remotely sensed data from different high-resolution satellite products to estimate nonparametric Standardized Precipitation Index (np-SPI) as the meteorological drought index. Each type of data-driven techniques like AI models is often robust to grab some aspects of hydrological behavior in drought analysis better than other models. The fusion of these models could exploit the strengths of each individual approach through combining information from multiple data sources to improve estimations that are more accurate. By fusing a set of individual data driven models, this study develops a framework to analyze drought conditions by objectively combining estimation from different AI models. The fusion-based methodology makes use of advanced statistical and optimization methods (Advanced weighting methods such as Ordered Weighted Averaged and Genetic Algorithm), and also considers the two weighting method characteristics (ORNESS and ORLIKE) to fuse the ultimate estimations by each data-driven models at each time scale. In order to test the ability of the new approach to fuse a range of AIs, five different data driven models, each developed with remotely sensed data, are considered. The three main objectives of this study are:

(i) To calculate np-SPI in drought estimations using different high resolution satellite-based products and compare the results with ground-based SPIs;
(ii) To compare five individual AI models including Multi-Layer Perceptron (MLP) neural network, Adaptive Neuro-Fuzzy Inference System (ANFIS), M5P model tree, Group Method of Data Handling (GMDH) and Support Vector Regression (SVR) in estimating np-SPI based on remotely sensed data;
(iii) To evaluate three advanced fusion-based approaches, namely, Ordered Weighted Averaged approach based on ORNESS weighting method (ORNESS-OWA) and ORLIKE weighting method (ORLIKE-OWA) and K-nearest neighbors (KNN) algorithm in data fusion for improving drought estimations.

These considered objectives provide the structure for the following Methods, Results and Discussion sections.

2. Study site and materials

2.1. Study site

Fars province as an agricultural hub is located between 27° and 31° east longitude in southwestern parts of Iran. Fig. 1 shows the location of Fars province and ground-based stations over the study area.

Fars province has a total area of 122,400 Km^2 (almost 8.1% of the country's area) with about three distinct climatic regions including the mountainous area of the north and northwest with moderate winters and summers, the central regions with relatively rainy temperate winters and hot dry summers and the area located in the south and southeast with cold winters and hot summers (FARS Meteorological Organization, 2017). Some climatology features of Fars province are shown in Fig. 2.

The average annual temperature and rainfall depth in this plain are 18.2 °C and 330 mm, respectively. According to the recorded data during the past forty years, the average annual rainfall in Fars province has been 300 mm. In last decade, this region has experienced a decreased between 5 and 65% of rainfall (FARS Meteorological Organization, 2017). Especially in recent years, this region has been severely affected by the drought condition, which has led to many water crises.

2.2. Materials

Due to availability of high quality and consistent recorded data with sufficient data range in order to calculate np-SPIs, 24 ground-based observations were selected for drought analysis in this study. The precipitation data for 24 stations over 30-year period from 1981 to 2011 were collected from the Fars Meteorological Organization and Fars Regional Water Organization. The climate information for ground-based stations are presented in Table 2.