Atmospheric Research 171 (2016) 21–30

Contents lists available at ScienceDirect

Atmospheric Research

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

Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology

Roslan Hashim [a,b,*], Chandrabhushan Roy [a], Shervin Motamedi [a,b], Shahaboddin Shamshirband [c], Dalibor Petković [d], Milan Gocic [e], Siew Cheng Lee [a]

[a] Department of Civil Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Institute of Ocean and Earth Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[d] Faculty of Mechanical Engineering, Department for Mechatronics and Control, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia
[e] Faculty of Civil Engineering and Architecture, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form 2 December 2015
Accepted 4 December 2015
Available online 15 December 2015

Keywords:
Rainfall
Forecasting
Meteorological data
Anfis
Variable selection

ABSTRACT

Rainfall is a complex atmospheric process that varies over time and space. Researchers have used various empirical and numerical methods to enhance estimation of rainfall intensity. We developed a novel prediction model in this study, with the emphasis on accuracy to identify the most significant meteorological parameters having effect on rainfall. For this, we used five input parameters: wet day frequency (d_{wet}), vapor pressure ($\bar{e}_a$), and maximum and minimum air temperatures (T_{max} and T_{min}) as well as cloud cover (cc). The data were obtained from the Indian Meteorological Department for the Patna city, Bihar, India. Further, a type of soft-computing method, known as the *adaptive-neuro-fuzzy inference system* (ANFIS), was applied to the available data. In this respect, the observation data from 1901 to 2000 were employed for testing, validating, and estimating monthly rainfall via the simulated model. In addition, the ANFIS process for variable selection was implemented to detect the predominant variables affecting the rainfall prediction. Finally, the performance of the model was compared to other soft-computing approaches, including the artificial neural network (ANN), support vector machine (SVM), extreme learning machine (ELM), and genetic programming (GP). The results revealed that ANN, ELM, ANFIS, SVM, and GP had R^2 of 0.9531, 0.9572, 0.9764, 0.9525, and 0.9526, respectively. Therefore, we conclude that the ANFIS is the best method among all to predict monthly rainfall. Moreover, d_{wet} was found to be the most influential parameter for rainfall prediction, and the best predictor of accuracy. This study also identified sets of two and three meteorological parameters that show the best predictions.

1. Introduction

Increasing population and industrialization have put tremendous pressures on the natural processes (e.g., Bawa and Seidler, 2015; Brown, 2015; Gupta, 2014; Heald and Spracklen, 2015; Lemaire et al., 2014). In addition, in recent years, the frequency of natural hazards (i.e., droughts, floods, hurricanes, cyclones and tsunamis) have increased (Gong and Forrest, 2014; Guan et al., 2015; Kumar et al., 2015; Sowmya et al., 2015). Therefore, temporal prediction of natural processes, such as rainfall, is helpful in appropriate planning and management of urban areas. For example, rainfall determines the ground water level, which in turn supports the water supply to urban and rural populations (Dhakal and Sullivan, 2014; Van Eekelen et al., 2015). In addition, rainfall endures various human ecology interactions such as agriculture, thermal comfort, and evaporation (Arvor et al., 2014; de Abreu-Harbich et al., 2015; Müller et al., 2014; Singh et al., 2014).

Rainfall is not easy to predict, because it depends on the space and time scales. Researchers consider rainfall as a stochastic process (Kundu et al., 2014; Ramesh et al., 2013; Schleiss et al., 2014). In this regard, various empirical and numerical models have been developed to investigate the non-linear trend of rainfall (Buzzi et al., 2014; Dai et al., 2015; Kumar et al., 2014; Nicholson, 2014; Nielsen et al., 2014). Recently, accuracy of conventional prediction models have been improved through the meteorological and satellite observations (Mehran and AghaKouchak, 2014; Zhou et al., 2014). The efficiency of soft-computing techniques was investigated for non-linear natural processes, such as rainfall (Nastos et al., 2014), drought (Deo and Şahin, 2015), oceanic wave (Sayemuzzaman et al., 2015), and sediment transport (Makarynskyy et al., 2015). In this view, use of the soft-computing methods for predicting rainfall can be beneficial.

Rainfall models that employ soft-computing techniques can be classified into two groups. The first group uses historical time series data of a station to train the model and predict rainfall (e.g., Ramana et al., 2013;

* Corresponding author.
E-mail address: roslan@um.edu.my (R. Hashim).

http://dx.doi.org/10.1016/j.atmosres.2015.12.002
0169-8095/

Shamshirband et al., 2015a,b). Second set of rainfall models uses multiple meteorological parameters to make prediction (e.g., Ortiz-García et al., 2014; Nastos et al., 2014).

The soft-computing methods can effectively recognize relationships between dependent and independent variables for non-linear natural processes, such as rainfall (Datta et al., 2014; Geetha and Nasira, 2014; Maheswaran and Khosa, 2014; Sharma and Bose, 2014; Valverde et al., 2014). For example, Ramana et al. (2013) used the wavelet technique with the ANN to satisfactorily predict the rainfall time series. They used monthly rainfall gauge data for Darjeeling, India, to calibrate the developed model. They reported that the wavelet neural network models were superior to the ANN models. A recent study compared the performance of the ANFIS and the support vector regression (SVR)-based rainfall prediction models (Shamshirband et al., 2015a,b). The results show that the ANFIS better predicts rainfall than the SVR. This study employed only the historical monthly rainfall data of 29 stations in Serbia, measured from the year 1946 to 2012, as input to train the models. They suggested that even with one input, the ANFIS and SVR can make accurate monthly predictions. A similar study used the historical daily rainfall data from the years 1987 to 2001 in three stations in Turkey to estimate one-day ahead rainfall using the wavelet-neuro-fuzzy technique. The results showed R^2 of 0.9 for the predicted rainfall using the wavelet-neuro-fuzzy network, while the conventional neuro-fuzzy method yielded a very low R^2 (0.1). The authors concluded that the results of the wavelet-neuro-fuzzy compare well with other conventional soft-computing techniques, such as ANN and multi-linear regression (MLR) methods.

In addition to the time series rainfall data, many soft-computing studies attempted to predict rainfall using multiple meteorological parameters. For example, Ortiz-García et al. (2014) compared rainfall estimation by the SVM, multi-layer perceptron, extreme learning machine (ELM), and other classical algorithms, such as K-nearest neighbor classifiers and decision tree. They used parameters, such as total perceptible water, air temperature, humidity, wind speed, and wind direction. The results indicated that the SVM outperforms other classification techniques. Nastos et al. (2014) used the monthly total rain and corresponding rain days (d_{wet}, wet day) in the ANN model to predict rainfall intensity in Athens, Greece. They used meteorological data for 11 years, attained from the *National Observatory of Athens*. The rainfall predictions of the 4-month observation was in good agreement with the field data. Furthermore, the RMSE and R^2 proved the robustness of the models, with a statistical significance of $P < 0.01$.

Based on the above discussion, it can be concluded that the soft-computing models satisfactorily predict rainfall. In addition, previous research works have employed multiple meteorological parameters, but until now, no study has investigated the effect of input parameters on the accuracy of rainfall prediction.

In this study, we developed a prediction model using the ANFIS to identify the most influential parameters that accurately estimate rainfall. The process, which is called variable selection, includes a number of ways to discover a subset of the total recorded parameters that show good capability of prediction. The ANFIS network was used to perform a variable search and thereafter, it was used to examine how 5 input parameters with influence on rainfall prediction (monthly cloud cover (cc), average monthly vapor pressure ($\overline{e}_a$), frequency of monthly wet day (d_{wet}), and monthly maximum and minimum air temperatures (T_{max} and T_{min})). This study obtained the average monthly meteorological observations from the year 1901 to 2002 for the district Patna in the state of Bihar, India. The collected data w used to calibrate and estimate monthly rainfall using the model.

2. Methodology

This section describes the study area and provides details of the input parameters, observation data and the ANFIS method. Section 2.1 discusses the observation area and Section 2.2 presents the input and output meteorological data. Also, section 2.2 presents the ANFIS model that was used for predicting rainfall. The schematic flow of this research work is depicted in Fig. 1.

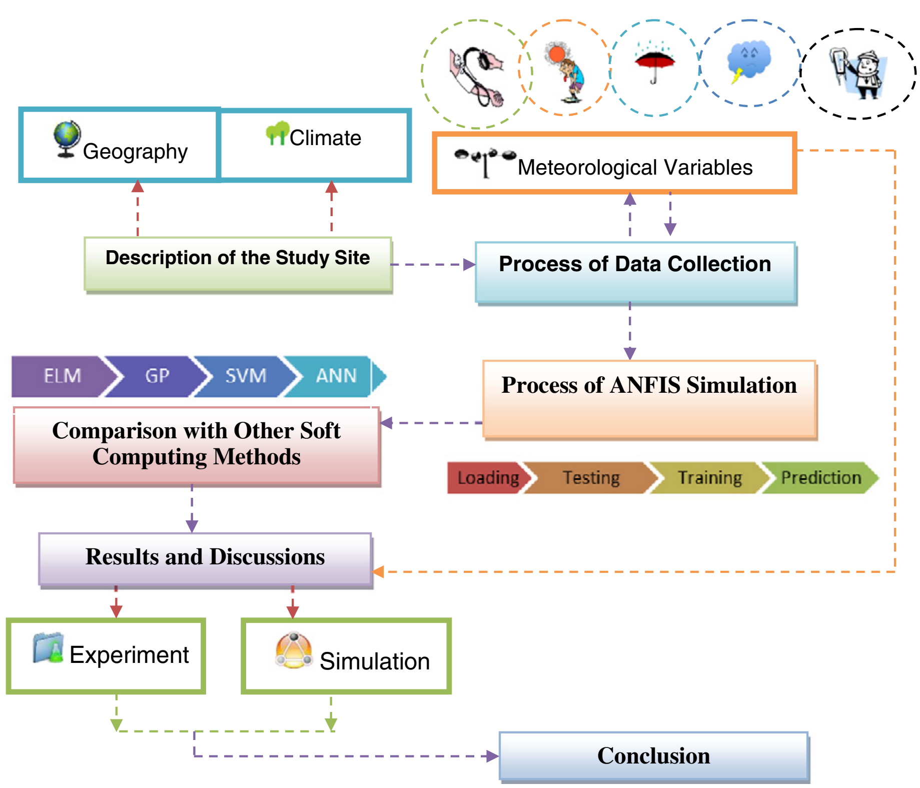

Fig. 1. Flow of research works in this study.