Computers and Electronics in Agriculture 144 (2018) 164–173

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

# Application of artificial intelligence models for the prediction of standardized precipitation evapotranspiration index (SPEI) at Langat River Basin, Malaysia

Y.W. Soh, C.H. Koo, Y.F. Huang*, K.F. Fung

*Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Jalan Bandar Sg. Long, Bandar Sg. Long, 43000 Kajang, Selangor, Malaysia*

ARTICLE INFO

*Keywords:*
ANFIS
ANN
ARIMA
Drought
Wavelet

ABSTRACT

Drought forecasting is a vital for mitigating the impact of drought events on the economy, tourism, agriculture and water resource systems. This paper adopts the proposed Wavelet-ARIMA-ANN (WAANN) model and the latest Wavelet-Adaptive Neuro-Fuzzy Inference System (WANFIS) model to predict the Standardized Precipitation Evapotranspiration Index (SPEI) at the Langat River Basin for different time scales (1-month, 3-months and 6-months). Model input data pre-processing with wavelet decomposition for improving the performance of the models was carried out apriori. The historical SPEI from 1976 to 2007 were used in the WAANN and WANFIS models for predicting the SPEI for the test period from 2008 to 2015. The Adjusted Coefficient of Determination ($R^2_{adj}$), Root-Mean-Square-Error (RMSE), Mean Absolute Error (MAE), Willmott's Index of Agreement (d) and the Nash-Sutcliffe Coefficient of Efficiency (E) were used to assess the models. It was found that the prediction accuracy of the two models improved with time scale length. For the prediction of SPEI-1 (1-month), the errors associated with both models were considered relatively high. Based on the performance measures and graphical plots, the WAANN model is better for the prediction of SPEI-3 and SPEI-6. The WANFIS model had satisfactory prediction of the mid-term drought forecasting for all stations. The WAANN model developed in this study however, gives better accuracy for both, the short-term and mid-term drought forecasting.

## 1. Introduction

On the basis of the long-run average precipitation (normal precipitation) for a particular basin, the declining trend of precipitation indicates the initiation of droughts (Jalalkamali et al., 2015). Low relative humidity, temperatures, high wind velocity, rainfall characteristics including intensity, duration of precipitation and the distribution of rainfall during the crop growing seasons are important features of the droughts (Mishra and Singh, 2010). It has been reported in researches that the drought-induced conditions developed seasonally in the event of El Nino (also known as the warm phase of El Nino Southern Oscillation, ENSO). El Nino is induced by the reduction of trade winds and, in tandem with the increase of earth surface temperatures, appears to occur on the average, every 3–4 years (Paz et al., 2007).

Even though Malaysia receives an average of 2800 mm of precipitation annually, the country however is still subjected to prolonged dry spells; especially at the Langat River Basin, where the rapid urbanisation of the Kuala Lumpur City area has resulted in an increased demand for the freshwater supply (Pour et al., 2014). The ENSO profoundly affects the condition of climate in Malaysia and in Southeast Asia. Yusof et al. (2012) applied the Kriging method to analyse the upward and downward trends during the occurrence of droughts in Peninsular Malaysia. Their results showed that the major regions of West Malaysia are subjected to an upward trend throughout the dry season, particularly in the eastern and western regions. It is extremely imperative for the water resources department to predict the drought intensity, severity and duration. With prior awareness of the onset of droughts, appropriate actions to mitigate the consequential damages can be considered. Precise and representative drought index series about the onset, extent and the end of the drought event allows the proper drought contingency plans to be established (Subash et al., 2011).

Drought monitoring and early warning are important phases to manage droughts (Bachmir et al., 2016). Among the approaches for drought forecasting, the use of artificial intelligence (AI) shows outstanding performance and accuracy (Masinde, 2014; Ozger et al., 2011; Belayneh et al., 2014). The flexibility and adaptability of AI is useful in predicting the occurrence of droughts that poses varying durations,

* Corresponding author.
*E-mail address:* huangyf@utar.edu.my (Y.F. Huang).

https://doi.org/10.1016/j.compag.2017.12.002
Received 26 May 2017; Received in revised form 12 September 2017; Accepted 3 December 2017

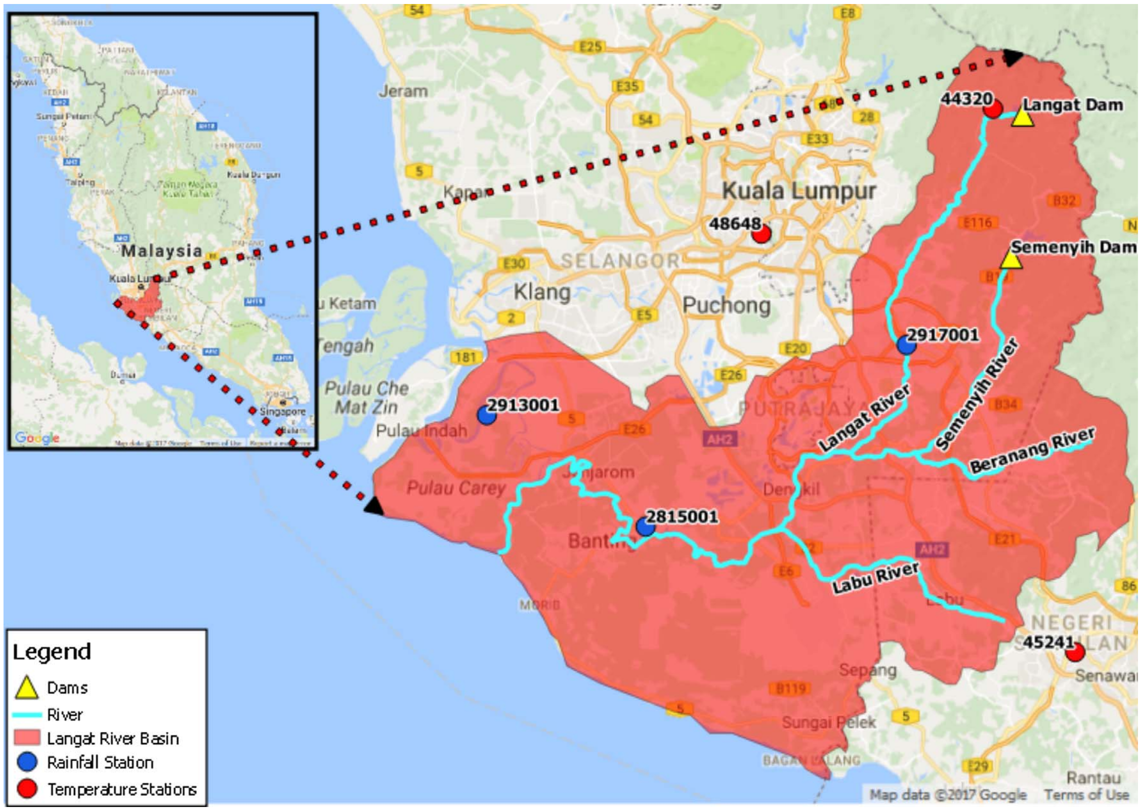


Fig. 1. Location of Langat River Basin.

frequencies and intensities; characteristics that are not easily determined effectively using empirical relations. Artificial Neural Network (ANN), Fuzzy Logic (FL) and Support Vector Regression (SVR) are the examples of AI models that can be used to forecast the time series modelling based on historical data (Djerbouai and Souag-Gamane, 2016). Another popular type of data-driven model is the stochastic method. The older Auto Regressive Integrated Moving Average Model (ARIMA) and Seasonal Auto Regressive Moving Integrated Average Model (SARIMA) are the most widely adopted stochastic approaches, as they are relatively simple yet providing excellent results (Mossad and Alazba, 2015; Bazrafshan et al., 2015).

In order to improve the prediction accuracy in the time series forecasting, different types of hybrid models were developed to minimize the error between the actual and predicted values. For example, both Choubin et al. (2014, 2016) explored the effectiveness of ANFIS model in forecasting the Standardized Precipitation Index (SPI) for different regions. Apart from that, ARIMA-ANN, Wavelet-ANN (WANN) and WANFIS are also effective forecasting tools among the hybrid models (Belayneh et al., 2014; Shabri, 2014). The wavelet transform had been proven to decompose the incoming signal into different frequencies (Djerbouai and Souag-Gamane, 2016; Belayneh and Adamowski, 2012, 2013; Özger et al., 2011, 2012) and thus, providing a more accurate output for the data driven models. By combining the strength of different models, the hybrid models can give a higher degree of accuracy.

In 2013, the WANFIS model was first used in drought forecasting because the wavelet transform has the ability to improve the accuracy of the models. The performance of the WANFIS model was compared with the WANN, ANN and ANFIS. WANFIS model results were more precise than other models in the study and the wavelet transform was particularly good in improving meteorological drought forecasting. Shabri (2014) also proposed to use the WANFIS model to forecast the drought events in Malaysia. Subsequently, the WANFIS model was widely recognised as an effective tool to forecast the time series of drought index. However, the performance of the WANFIS highly relied on the modelling of the ANFIS model. Seo et al. (2016) had pointed the drawbacks of using the ANFIS modelling that would lead to difficulty in the identification of network parameters and the number and type of membership functions. Moreover, the development of the ANFIS model is computationally expensive and complex.

The aim of this paper is to improve the drought forecasting procedure at the Langat River Basin, using hybrid modelling. An alternative model, the WAANN was proposed in this study since the wavelet transform is able to reduce the complexity of the time series and the ARIMA-ANN model was proven as an appropriate tool in the time series forecasting (Babu and Reddy, 2014; Khashei and Bijari, 2011). The latest hybrid model WANFIS in drought forecasting is used for comparison. The expected results are the SPEI series predicted with both models and the accompanying performance evaluations.

## 2. Methodology

### 2.1. Study area and data acquisition

The catchment area of the Langat River Basin in Selangor, Malaysia is about 2400 $km^2$, supplying approximately 65% of water usage in the state of Selangor. Two reservoirs in the Langat River Basin, namely the Semenyih dam (area of 56.6 $km^2$) and the Hulu Langat dam (area of 41.0 $km^2$) supply water to the state. The recorded daily precipitation and temperature for each station in the basin were retrieved from the Department of Irrigation and Drainage (DID) Malaysia and the Malaysian Meteorology Department (MMD). After determining the suitability of the data, the records from six stations were used. A set of daily meteorological data, including rainfall and temperature were obtained for these stations for the period of 1976–2015; with the exception of station 44320, where only the daily data from 1985 to 2015 is available. The location of the Langat River Basin is illustrated in Fig. 1 and the details of selected hydrometric stations in Langat River