Journal of Hydrology 508 (2014) 418–429

Contents lists available at ScienceDirect

Journal of Hydrology

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

Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models

A. Belayneh [a], J. Adamowski [a,*], B. Khalil [a], B. Ozga-Zielinski [b]

[a] Department of Bioresource Engineering, Faculty of Agricultural and Environmental Sciences, McGill University, 21 111 Lakeshore, Ste Anne de Bellevue, Quebec H9X 3V9, Canada
[b] Department of Environment Protection and Development, Environmental Engineering Faculty, Warsaw University of Technology, ul. Nowowiejska 20, 12 Warsaw 00-653, Poland

ARTICLE INFO

Article history:
Received 14 May 2013
Received in revised form 29 October 2013
Accepted 30 October 2013
Available online 7 November 2013
This manuscript was handled by Geoff Syme, Editor-in-Chief, with the assistance of Bellie Sivakumar , Associate Editor

Keywords:
ANN
Support vector regression
SPI
Drought forecasting
Wavelet transforms
Africa

SUMMARY

Long-term drought forecasts can provide valuable information to help mitigate some of the consequences of drought. Data driven models are suitable forecast tools due to their minimal information requirements and rapid development times. This study compares the effectiveness of five data driven models for forecasting long-term (6 and 12 months lead time) drought conditions in the Awash River Basin of Ethiopia. The Standard Precipitation Index (SPI 12 and SPI 24) was forecasted using a traditional stochastic model (ARIMA) and compared to machine learning techniques such as artificial neural networks (ANNs), and support vector regression (SVR). In addition to these three model types, wavelet transforms were used to pre-process the inputs for ANN and SVR models to form WA-ANN and WA-SVR models; this is the first time that WA-SVR models have been explored and tested for long-term SPI forecasting. The performances of all models were compared using RMSE, MAE, R^2 and a measure of persistence. The forecast results indicate 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 over lead times of 6 and 12 months in the Awash River Basin.

1. Introduction

Drought, which occurs when there is a deficit in precipitation compared to the long-term average, has many impacts. Of all the extreme climate induced events, droughts have the most complex consequences, mainly due to the difficulty in identifying their inception and their end (Wilhite, 1993). In addition, droughts impact a wide range of geographic areas. Globally, 22% of the economic damage caused by natural disasters and 33% of the damage in terms of the number of persons affected can be attributed to drought (Keshavarz et al., 2013). The impacts of drought are more severe in sub-Saharan Africa, where rain-fed agriculture comprises 95% of all agriculture. This dependence of rain-fed agriculture leaves sub-Saharan Africa vulnerable to the impacts of drought. For example, in 2009, poor rainfall led to increased droughts and an increase of 53 million food insecure people in the region (Husak et al., 2013).

Due to a slow evolution in time, drought is often a phenomenon whose consequences take a significant amount of time with respect to its inception to be perceived by the socio-economic sector (Cancelliere et al., 2007b). This feature allows for the mitigation of some of the impacts of drought, provided there is an effective monitoring system to provide information to decision makers. Effective monitoring of droughts can significantly help early warning systems. In sub-Saharan Africa, effective rainfall monitoring contributes to the allocation of aid during periods of drought (Husak et al., 2013). An important aspect of drought monitoring and the development of an early warning system is the ability to effectively forecast future drought events. Forecasting future drought events in a region is very important for finding sustainable solutions to water management and risk assessment of drought occurrences (Bordi and Sutera, 2007).

Drought forecasts can be done using either physical/conceptual or data driven models. While physical/conceptual models are good at providing insight into catchment processes, they have been criticised for being difficult to implement for forecasting applications, requiring many different types of data and resulting in models that are overly complex (Beven, 2006). In contrast, data driven models have minimum information requirements, rapid development times, and have been found to be accurate in various hydrological forecasting applications (Adamowski, 2008).

Data driven stochastic models have traditionally been used for drought forecasting. Autoregressive integrated moving average

* Corresponding author. Tel.: +1 514 398 7786; fax: +1 514 398 8387.
E-mail address: jan.adamowski@mcgill.ca (J. Adamowski).

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.10.052

models (ARIMA) (Mishra and Desai, 2005,2006; Mishra et al., 2007; Han et al., 2010) have been the most widely used stochastic models for hydrologic drought forecasting. Stochastic models are linear models and are limited in their ability to forecast non-linear data. To effectively forecast non-linear data, researchers in the last two decades have increasingly begun to utilize artificial neural networks (ANNs) to forecast hydrological data. ANNs have been used in a number of studies as a drought forecasting tool (Mishra and Desai, 2006; Morid et al., 2007; Bacanli et al., 2008; Barros and Bowden, 2008; Cutore et al., 2009; Karamouz et al., 2009; Marj and Meijerink, 2011; Mishra and Nagarajan, 2012).

Along with ANN models, Support Vector Machines (SVMs) are a machine learning technique that has become increasingly popular in hydrologic forecasting. SVMs can be divided into two main techniques, the Support Vector Classification (SVC) and Support Vector Regression (SVR), which address problems of classification and regression, respectively (Gao et al., 2001). There are several studies where SVR was used in hydrological forecasting. Khan and Coulibaly (2006) found that a SVR model performed better than ANNs in 3–12 month predictions of lake water levels. Kisi and Cimen (2009) used SVRs to estimate daily evaporation. A study by Belayneh and Adamowski (2012) used SVR models to forecast the SPI in the Awash River Basin. The research presented in the current paper complements the Belayneh and Adamowski (2012) study by forecasting the SPI over a larger selection of stations in the same area, and by coupling, for the first time, SVR models with wavelet transforms.

The ability of these machine learning techniques (i.e. ANNs and SVRs) to forecast non-stationary data is limited. To overcome this limitation, wavelet analysis has begun to be explored in hydrologic forecasting. Wavelet transforms are often used as a data pre-processing tool in order to reveal aspects of a time series that other signal processing techniques cannot. Wavelet analysis has been applied to examine rainfall-runoff relationships in Karstic watersheds (Labat et al., 1999), to evaluate rainfall-runoff models (Lane, 2007), to forecast river flow (Adamowski, 2008; Adamowski and Sun, 2010), to forecast groundwater levels (Adamowski and Chan, 2011), to forecast urban water demand (Adamowski et al., 2012) and for the purposes of drought forecasting (Kim and Valdes, 2003; Ozger et al., 2012; Mishra and Singh, 2012; Belayneh and Adamowski, 2012).

The main objective of the current study is to compare the effectiveness of machine learning methods (in this case the ANN and SVR methods) as well as machine learning methods pre-processed with wavelet transforms (WA-ANN and WA-SVR models), for long-term drought forecasting in arid regions (in this case the Awash River Basin of Ethiopia). All these methods were compared to a traditional stochastic method (ARIMA model), which was used for comparative purposes (to compare the performance of the newer methods with a widely used traditional method). The standardized precipitation index (SPI), a meteorological drought index, was the drought index forecasted in this study, as it is a good indicator of the variability of East African droughts (Ntale and Gan, 2003). SPI 12 and SPI 24 were forecast for lead times of 6 and 12 months; SPI 12 and SPI 24 are good indicators of long-term drought conditions. The coupling of wavelet transforms with SVR models for the purpose of forecasting the SPI has not been explored to date in the literature.

2. The Standardized Precipitation Index

The Standardized Precipitation Index (SPI) was developed by McKee et al. (1993). The SPI is a meteorological drought index as it is based solely on precipitation data. Two main advantages arise from the use of the SPI index. First, as the index is based on precipitation alone its evaluation is relatively easy (Cacciamani et al., 2007). Second, the index makes it possible to describe drought on multiple time scales (Tsakiris and Vangelis, 2004; Mishra and Desai, 2006; Cacciamani et al., 2007). The SPI can only be computed when a sufficiently long (at least 30 years) and possibly continuous time-series of monthly precipitation data is available (Cacciamani et al., 2007). SPI calculation begins by selecting a suitable probability density function to describe the precipitation data (Cacciamani et al., 2007). The cumulative probability of an observed precipitation amount is computed after an appropriate density function is chosen. The inverse normal (Gaussian) function is then applied to the probability (Cancelliere et al., 2007a). For each rainfall gauge in this study the gamma distribution function was selected to fit the rainfall data. Alternatively, a lognormal or an exponential distribution can be used to model the precipitation. In this paper, we followed Edossa et al. (2010) on selecting the gamma distribution function to fit the rainfall data. Detailed computation of the SPI can be found in Cacciamani et al. (2007).

The SPI is a z-score and represents an event departure from the mean, expressed in standard deviation units. The SPI is a normalized index in time and space and this feature allows for the comparison of SPI values among different locations. SPI values can be categorized according to classes (Cacciamani et al., 2007). Normal conditions are established from the aggregation of two classes: $-1 < \mathrm{SPI} < 0$ (mild drought) and $0 \leqslant \mathrm{SPI} \leqslant 1$ (slightly wet). SPI values are positive or negative for greater or less than mean precipitation, respectively. Variance from the mean is a probability indication of the severity of the flood or drought that can be used for risk assessment (Morid et al., 2006). The more negative the SPI value for a given location, the more severe the drought. The time series of the SPI can be used for drought monitoring by setting application-specific thresholds of the SPI for defining drought beginning and ending times. Accumulated values of the SPI can be used to analyze drought severity. In this study, the SPI_SL_6 program developed by the National Drought Mitigation Centre, University of Nebraska-Lincoln, was used to compute time series of drought indices (SPI) for each station in the basin and for each month of the year at different time scales. A drought event occurs at the time when the value of the SPI is continuously negative; the event ends when the SPI becomes positive. Even though the SPI is a meteorological drought index, it can be used to interpret different aspects of drought. This study only forecast SPI 12 and SPI 24, which are tied to long-term drought conditions, which is the focus of this study. This study was interested in providing information about drought conditions that affect streamflow, groundwater or other hydrological systems within the Awash River basin. Shorter SPI runs, such as SPI 3 and SPI 6, are representative of agricultural drought conditions, which are measured by a deficit in soil moisture content (Mishra and Desai, 2006).

3. Awash River Basin

In this study, the SPI was forecast in the Awash River Basin of Ethiopia (Fig. 1). Forecasts were made and compared for twelve stations within the basin. Drought is a common occurrence in the Awash River Basin (Edossa et al., 2010), and with approximately 90% of the population engaged in agricultural activities the area is especially vulnerable to the effects of drought (Desalegn et al., 2006). The heavy dependence of the population on rain-fed agriculture has made the people and the country's economy extremely vulnerable to the impacts of droughts. Current monthly and seasonal drought forecasts are done using the normalized difference vegetation index (NDVI). While the NDVI is an effective drought index, it is sensitive to changes in vegetation and has limitations in areas where vegetation is minimal. Forecasts of SPI

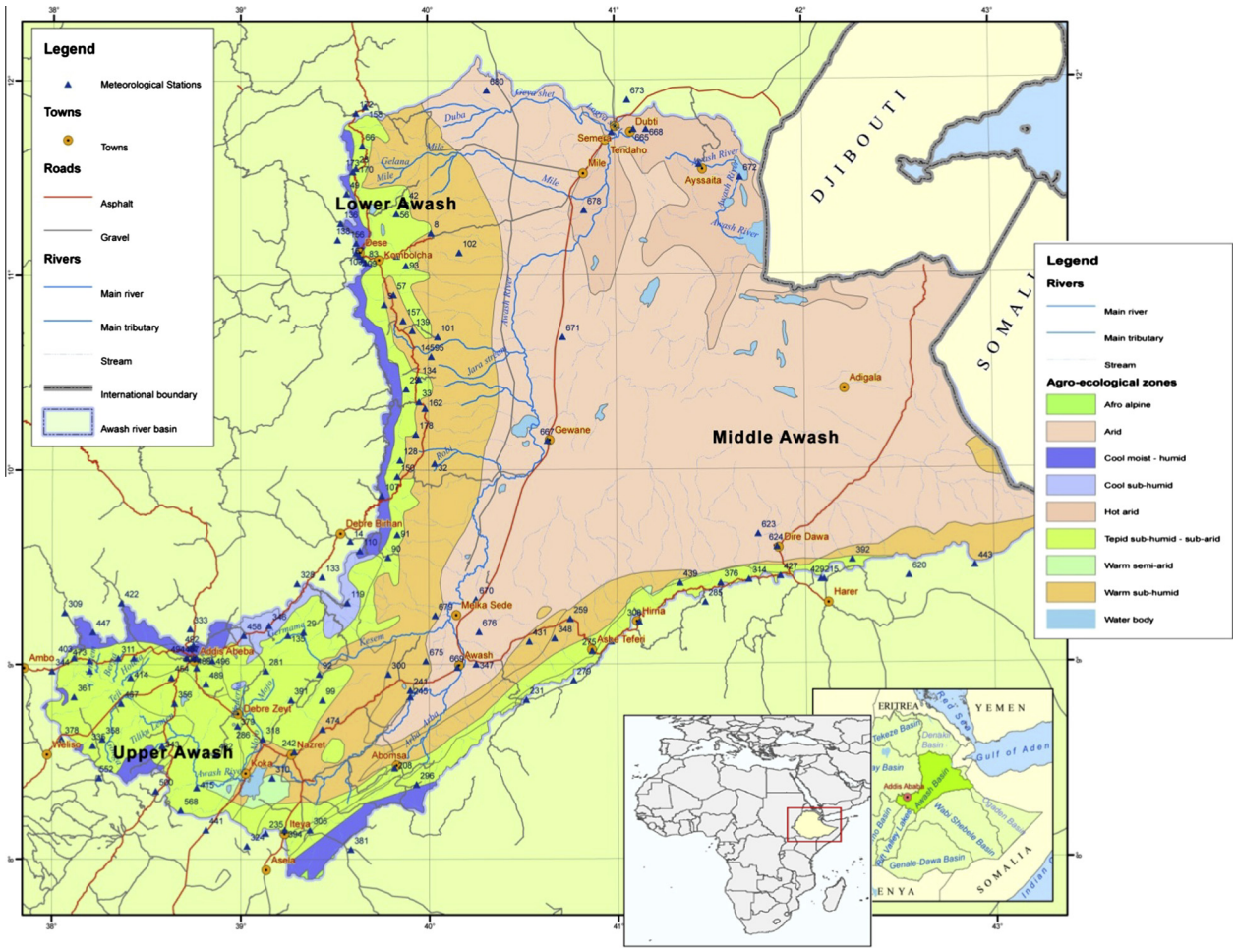

Fig. 1. The Awash River Basin. Ministry of Water Resources, Ethiopia. Agricultural Water Management Information System. http://www.mowr.gov.et/AWMISET/images/Awash_agroecologyv3.pdf Accessed 06-June-2013.

in this case SPI 12 and SPI 24) are not dependent on vegetative cover and can be used as another tool for drought forecasts within the basin and the country as a whole to complement the NDVI forecasts.

The mean annual rainfall of the basin varies from about 1600 mm in the highlands to 160 mm in the northern point of the basin. The total amount of rainfall also varies greatly from year to year, resulting in severe droughts in some years and flooding in others. The total annual surface runoff in the Awash Basin amounts to some 4900×10^6 m^3 (Edossa et al., 2010). Effective forecasts of the SPI can be used for mitigating the impacts of hydrological drought that manifests as a result of rainfall shortages in the area. The climate of the Awash River Basin varies between a mild temperate climate in the Upper Awash sub-basin to a hot semi-arid climate in both the Middle and Lower sub-basins. The Awash River Basin supports farming, from the growth of lowland crops such as maize and sesame to pastoral farming practices. Rainfall records from 1970 to 2005 were used to generate SPI 12 and SPI 24 time series from each station. Descriptive statistics for precipitation at the rainfall stations is shown in Table 1.

The normal ratio method, recommended by Linsley et al. (1988), was used to estimate the missing rainfall records at any stations that had incomplete precipitation records. With this method, rain depths for missing data are estimated from observations at three stations as close to, and as evenly spaced around the station with missing records, as possible. The distance matrix was established for all rain gauge stations in the basin based on their geographic locations in order to assess the proximity of stations with each other. All data sets were normalized using:

$$X_n = \frac{X_0 - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

where X_0 and X_n represent the original and normalized data respectively, while $X_{\min}$ and $X_{\max}$ represent the minimum and maximum value among the original data.

4. Model development

4.1. ARIMA models

ARIMA models were developed based on the Box and Jenkins approach and consist of three steps: model identification, parameterization and validation. The general non-seasonal ARIMA model may be written as (Box and Jenkins, 1976):

$$z_t = \frac{\theta(B)a_t}{\phi(B)\nabla^d} \tag{2}$$

$$\phi B = (1 - \phi_t B - \phi_2 B^2 - \cdots \phi_p B^p \tag{3}$$

and

Table 1
Descriptive Statistics of the Awash River Basin.

Basin	Station	Mean annual precipitation (mm)	Max annual precipitation (mm)	Standard deviation (mm)
Upper Awash Basin	Bantu Liben	91	647	111
	Ginchi	97	376	90
	Sebeta	111	1566	172
	Ejersalele	67	355	75
	Ziquala	100	583	110
Middle Awash Basin	Modjo	76	542	92
	Wolenchiti	76	836	95
	Gelemsso	77	448	75
	Dire Dawa	51	267	54
Lower Awash Basin	Eliwuha	44	374	57
	Dubti	87	449	89
	Bati	26	268	40

$$\theta B = (1 - \theta_1 B - \theta_2 B^2 - \cdots \theta_q B^q \tag{4}$$

where z_t is the observed time series and B is a back shift operator. $\phi(B)$ and $\theta(B)$ are polynomials of order p and q, respectively. The orders p and q are the order of non-seasonal auto-regression and the order of non-seasonal moving average, respectively. Random errors a_t are assumed to be independently and identically distributed with a mean of zero and a constant variance. ∇^d describes the differencing operation to data series to make the data series stationary and d is the number of regular differencing.

The first step in developing ARIMA models was determining the stationarity of a time series. To determine stationarity, the NUMxl Excel add-on was used. Once non-stationarity is removed the autocorrelation (ACF) and partial autocorrelation functions (PACF) were used to determine the correlation structure of the data. Once the significant lags were determined using the ACF and PACF, different combinations were used to determine the optimal model structure. For instance, if the ACF and PACF for SPI 12 showed significant lags at 5 and 2 lags respectively, different combinations for p and q were used with intervals between 1 and 5 for p and 1 and 2 for q. The selection of ARIMA models was determined based on both the accuracy and precision of models. The combinations that provided the most accurate forecast models as measured by the MAE and the most precise models as measured by the RMSE were chosen. The details on the development of ARIMA models for SPI time series can be found in the works of Mishra and Desai (2005) and Mishra et al. (2007). For all ARIMA models, the data was partitioned so that 90% of the data was a calibration set and 10% of the data was a validation set.

4.2. ANN models

The advantage of using ANNs is their parsimonious data requirements, rapid execution time and ability to produce models where the relationship between inputs and outputs are not fully understood. The ANN models used in this study have a feed forward Multi-layer perceptron (MLP) architecture which was trained with the Levenberg–Marquardt (LM) back propagation algorithm. MLPs have often been used in hydrologic forecasting due to their simplicity. MLPs consist of an input layer, one or more hidden layers, and an output layer (Kim and Valdes, 2003):

$$y'_k(t) = f_0\left[\sum_{j=1}^{m} w_{kj} \cdot f_n\left(\sum_{i=1}^{N} w_{ji} x_i(t) + (w_{j0}) + w_{k0}\right)\right] \tag{5}$$

where N is the number of samples, m is the number of hidden neurons, $x_i(t)$ = the ith input variable at time step t; w_{ji} = weight that connects the ith neuron in the input layer and the jth neuron in the hidden layer; w_{j0} = bias for the jth hidden neuron; f_n = activation function of the hidden neuron; w_{kj} = weight that connects the jth neuron in the hidden layer and kth neuron in the output layer; w_{k0} = bias for the kth output neuron; f_0 = activation function for the output neuron; and $y'_k(t)$ is the forecasted kth output at time step t (Kim and Valdes, 2003).

For ANN model development, the determination of the architecture of the model is very important. The optimal number of neurons in the input layer was determined by trial and error. As the SPI requires only precipitation for its computation it was the only input used. The SPI was lagged to generate several neurons in the input layer and the number of neurons that provided the lowest RMSE values was chosen as the appropriate number. Traditionally, the number of hidden neurons for ANN models is selected via a trial and error method. However a study by Wanas et al. (1998) empirically determined that the best performance of a neural network occurs when the number of hidden neurons is equal to log (N), where N is the number of training samples. Another study conducted by Mishra and Desai (2006) determined that the optimal number of hidden neurons is $2n + 1$, where n is the number of input neurons. In this study, the optimal number of hidden neurons was determined to be between log(N) and ($2n + 1$). For example, if using the method proposed by Wanas et al. (1998) gave a result of 4 hidden neurons and using the method proposed by Mishra and Desai (2006) gave 7 hidden neurons, the optimal number of hidden neurons was assumed to be between 4 and 7; thereafter the optimal number was chosen via trial and error. These two methods helped establish an upper and lower bound for the number of hidden neurons.

The ANN models used to forecast the SPI were recursive models. A recursive ANN model is similar to an ARIMA model in terms of the forecasting approach, and forecasts one time step ahead. For subsequent forecasts the network is applied recursively, using the previous predictions as inputs. Recursive ANN models have been shown to be effective for forecasts of the SPI for long lead times (e.g., Mishra and Nagarajan, 2012). In addition, preliminary analysis conducted for each station in our study indicated that recursive models were more accurate than direct multistep models, and as such the recursive forecasting approach was used in this study. For all the ANN models, 80% of the data was used to train the models, while the remaining 20% of the data was divided into a testing and validation set with each set comprising 10% of the data.

4.3. SVR models

SVR models adhere to the structural risk minimization principle as opposed to the empirical risk minimization principle used by conventional neural networks (Vapnik, 1995). As a result, these models reduce the generalization error as opposed to the training error. With SVR, the purpose is to estimate a functional dependency $f(\vec{x})$ between a set of sampled points $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ taken from R^n and target values $Y = \{y_1, y_2, \ldots, y_n\}$ with $y_i \in R$ (the input and target vectors (x_i's and y_i's) refer to the monthly records of the SPI index). Detailed descriptions of SVR model development can be found in Cimen (2008).

All SVR models were created using the OnlineSVR software created by Parrella (2007), which can be used to build support vector machines for regression. The data was partitioned into two sets: a calibration set and a validation set. 90% of the data was partitioned into the calibration set while the final 10% of the data was used as the validation set. Unlike neural networks the data can only be partitioned into two sets with the calibration set being equivalent to the training and testing sets found in neural networks. All inputs and outputs were standardized between 0 and 1.

All SVR models used the non-linear radial basis function (RBF) kernel. As a result, each SVR model consisted of three parameters that were selected: gamma (γ), cost (C), and epsilon (ε). The γ parameter is a constant that reduces the model space and controls the complexity of the solution, while C is a positive constant that is a capacity control parameter, and ε is the loss function that describes the regression vector without all the input data (Kisi and Cimen, 2011). These three parameters were selected based on a trial and error procedure. The combination of parameters that produced the lowest RMSE and MAE values for the calibration data sets were selected.

4.4. Wavelet decomposition

The wavelet transform is a mathematical tool that provides a time–frequency representation of a signal in the time domain (Partal and Kisi, 2007). In addition, wavelet analysis can often compress or de-noise a signal (Kim and Valdes, 2003) and thus, is an effective method for dealing with local discontinuities in a given time series. The continuous wavelet transform (CWT) of a signal $x(t)$ is defined as (Nason and Von Sachs, 1999):

$$W(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-\tau}{s}\right)dt \tag{6}$$

where s is the scale parameter; τ is the translation, ψ is the mother wavelet and * corresponds to the complex conjugate (Kim and Valdes, 2003).

CWT is not often used for forecasting due to its complexity and long computation times. Instead, the wavelet is discretized in forecasting applications to simplify the numerical calculations. The discrete wavelet transform (DWT) requires less computation time and is simpler to implement (Nason and Von Sachs, 1999):

$$\psi_{j,k}(t) = \frac{1}{\sqrt{|s_0^j|}}\psi\left(\frac{t - k\tau_0 s_0^j}{s_0^j}\right) \tag{7}$$

where j and k are integers that control the scale and translation respectively, while $s_0 > 1$ is a fixed dilation step (Cannas et al., 2006) and τ_0 is a translation factor that depends on the aforementioned dilation step.

One of the inherent limitations of using the DWT for forecasting applications is that it is not shift invariant (i.e. if we change values at the beginning of our time series, all of the wavelet coefficients will change). To overcome this problem, a redundant algorithm, known as the à trous algorithm, can be used and is given by (Mallat, 1998):

$$C_{i+1}(k) = \sum_{l=-\infty}^{+\infty} h(l)c_i(k + 2^i l) \tag{8}$$

where h is the low pass filter and $C_{i+1}(k)$ is the original time series. To extract the details, $w_i(k)$, that were eliminated in Eq. (8), the smoothed version of the signal is subtracted from the coarser signal that preceded it, given by (Murtagh et al., 2003):

$$w_i(k) = c_{i-1}(k) - c_i(k) \tag{9}$$

where $c_i(k)$ is the approximation of the signal and $c_{i-1}(k)$ is the coarser signal. Each application of Eqs. (8) and (9) results in a smoother approximation and extracts a higher level of detail. Finally, the non-symmetric Haar wavelet can be used as the low pass filter for the à trous algorithm to prevent any future information from being used during the decomposition (Renaud et al., 2002).

The selection of the appropriate wavelet transform for an application requires a prior understanding of the attributes of the candidate wavelet. The shift invariant property of the 'a trous' Haar wavelet makes it the most suitable wavelet for forecasting applications (Maheswaran and Khosa, 2012b), and this is why it was used in this study. Furthermore, a study by Kim and Valdes (2003) found the 'a trous' wavelet to be more effective for forecasting studies than the Morlet or Daubechies wavelets. The aim of the 'a trous' wavelet is to fill any gaps with redundant information that is obtained from the original series. The redundant information provides a basis for enhanced forecasting accuracy (Maheswaran and Khosa, 2012a). The Haar wavelet, which is a low-pass filter, is concentrated over the narrowest support band, and therefore has good localization properties. This attribute makes the Haar wavelet more suitable for dynamic time series where it is effective at detecting changes within the time series (Maheswaran and Khosa, 2012b). The 'a trous' wavelet is appropriate for forecasting due to its localization capability in both the time and frequency domains, and the Haar wavelet as a low-pass filter is appropriate for detecting any changes that may occur within the 35 year SPI record in this study due to its localization properties as a result of its energy being concentrated over a narrow support band.

When conducting wavelet analysis, the number of decomposition levels that is appropriate for the data must be carefully selected. In this study, each SPI time series was decomposed between 1 and 8 levels. Fig. 2 depicts the SPI 12 time series of the Eliwuha station and the approximation series at different decomposition levels. The results were compared at all decomposition levels to determine the appropriate decomposition level. Fig. 2 indicates that the higher the level of decomposition, the less likely the transformed signal represents the original time series. Hence, the transformed times series was chosen after decomposing the original time series to level three or level four. The results from a decomposition of level three or four provided the most accurate forecast results as measured by the performance measures used in this study. Once a given time series was successfully decomposed, it was used as an input for either ANN or SVR models. Instead of using the original SPI data and its subsequent lags, a new wavelet transformed time series was used. After pre-processing, the generation of ANN and SVR models, including data partition was done in exactly the same way as the ANN and SVR models without wavelet transforms.

4.5. Performance measures

The following measures of goodness of fit were used to evaluate the forecast performance of all the aforementioned models:

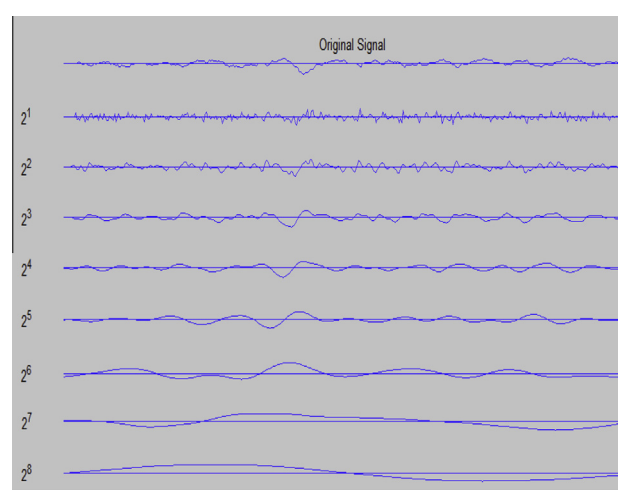

Fig. 2. The SPI 12 time series at the Eliwuha station and the approximation time series at different decomposition levels. The first time series is the original signal followed by decompositions a level 1–8.

$$\text{The coefficient of determination}(R^2) = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \quad (10)$$

$$\text{where } \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \quad (11)$$

where $\bar{y}$ is the mean value taken over N, y_i is the observed value, $\hat{y}_i$ is the forecasted value and N is the number of data points. The coefficient of determination measures the degree of association among the observed and predicted values.

$$\text{The Root Mean Squared Error(RMSE)} = \sqrt{\frac{SSE}{N}} \quad (12)$$

where SSE is the sum of squared errors, and N is the number of samples used. SSE is given by:

$$SSE = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (13)$$

with the variables already having been defined.

$$\text{The Mean Absolute Error(MAE)} = \sum_{i=1}^{N}\frac{|\hat{y}_i - y_i|}{N} \quad (14)$$

The MAE is used to measure how close forecasted values are to the observed values. It is the average of the absolute errors.

The results in this section were also compared via the persistence index (*PERS*)

$$PERS = 1 - \frac{SSE}{SSE_{naive}} \quad (15)$$

$$\text{where } SSE_{naive} = \sum (y_i - y_{i-L})^2 \quad (16)$$

As mentioned above, SSE is the sum of squared errors. y_{1-L} is the estimate from a persistence model that takes the last observation (at time 1 minus the lead time (L)) (Tiwari and Chatterjee, 2010). A value of *PERS* smaller or equal to 0 indicates that the model under study performs worse or no better than the easy to implement naïve model. A *PERS* value of 1 is obtained when the model under study provides exact estimates of observed values.

5. Results and discussion

In this study, the proposed forecast models for SPI 12 and SPI 24 are presented for forecast lead times of 6 and 12 months. A SPI 12 forecast of 6 months lead time represents a 6 month warning time for SPI 12, and a 12 month lead time represents a 12 month warning time and shows the variation in precipitation from year to year. Table 2 shows the inputs used for the proposed data driven models at the Eliwuha station (6 months lead time). Table 2 shows how the models are applied recursively to reach the final target of a forecast of 6 months lead time. The table is applicable for the input structure of both SPI 12 and SPI 24 (6 month lead time). The performance results of the proposed models for each station are presented in Table 3 through Table 6. As mentioned earlier, models that have a persistence index between 0 and 1 perform better than a naïve model. All the data driven models had a persistence index greater than 0. ARIMA models had a *PERS* of 0.36, ANN models had a *PERS* of 0.46, SVR models had a *PERS* of 0.41, WA-ANN models had a *PERS* of 0.58 and WA-SVR models had a *PERS* of 0.55 respectively. The results presented are based on the validation data sets. Table 7 shows the performance results for the training/calibration set for SPI 12 forecasts (6 months lead time) at the Eliwuha station. The table shows that the forecast accuracy shown in the validation sets is consistent with the results shown in the training or calibration sets (see Table 8).

Table 2
The decompositions of the SPI 12 series at the Eliwuha station with the correlation and variation of these decompositions to the original SPI 12 series.

Decomposition level	R^2	RMSE
Level 1	0.8408	0.0742
Level 2	0.8408	0.0742
Level 3	0.9727	0.0412
Level 4	0.9458	0.0490
Level 5	0.8408	0.0742
Level 6	0.7033	0.1068
Level 7	0.6671	0.1300
Level 8	0.6129	0.1446

The models were forecast one time step ahead and the subsequent result was used as an input in another model and forecast one time step ahead. The lead time in this study refers to the amount of time between the original time series and the final predicted time series. The final output time series was either 6 or 12 months ahead of the original time series. For each rainfall station, forecasts of SPI 12 and SPI 24 were made for 6 and 12 months lead times. These forecasts were made using the five model types outlined above. The results presented in the following sections are from forecasts of SPI 12 (6 months lead time) at the Eliwuha station. Only these results were presented in detail as it would be very difficult to present all the results from each station in detail. The forecast results for all the stations are presented in Tables 3–6, but these are not presented in detail.

5.1. ARIMA model results

The parameters for the ARIMA models were selected based on the ACF and PACF of the time series in question. Once the significant lags were determined from the ACF and PACF, ARIMA models with different combinations were developed and the model that had the lowest RMSE and MAE value was chosen. Figs. 3a and 3b are ACF and PACF figures for the Eliwuha station. The autocorrelation of the time series is significant for lags greater than 10 in Fig. 3a; however no lag greater than 5 was selected for the purpose of parsimony. For this station a combination of parameters from 1 to 5 were tested for p, while a value of 1 was tested for q. The combination that provided the lowest RMSE and MAE values was chosen as the model parameters. The results from the ARIMA models as shown in Tables 3 and 5 are significantly less accurate than the forecast results of the other model types (likely because ARIMA models are linear models).

5.2. ANN models

Fig. 4 provides the ANN model results at the Eliwuha station for SPI 12 at 6 months lead time. Fig. 4 indicates that the model has good generalization ability for SPI 12 and very little time shift error as the extreme events for the predicted values correspond to the extreme values of the observed values. With respective extreme drought or extreme precipitation, the model does underestimate

Table 3
Model input structure for the Eliwuha station (6 months lead time).

Recursive order	Input structure	Target
1st step	SPI(t−4), SPI(t−3), SPI(t−2), (SPI t−1), SPI (t)	SPI(t+1)
2nd step	SPI(t−3), SPI(t−2), SPI (t−1), SPI (t), SPI(t+1)	SPI(t+2)
3rd step	SPI(t−2), SPI(t−1), SPI(t), (SPI t+1), SPI (t+2)	SPI(t+3)
4th step	SPI(t−1), SPI(t), SPI (t+1), SPI (t+2), SPI(t+3)	SPI(t+4)
5th step	SPI(t), SPI(t+1), SPI(t+2), (SPI t+3), SPI (t+4)	SPI(t+5)
Test step	SPI(t+1), SPI(t+3), SPI (t+4), SPI (t+4), SPI(t+5)	SPI(t+6)

Table 4
The best ARIMA, ANN and SVR models for 6 and 12 month forecasts of SPI 12. Column 3 is the ANN architecture detailing the number of nodes in the input, hidden and output layers respectively. In column 11 the parameters of the SVR models are given.

Basin	Station	ANN models	R^2	RMSE	MAE	ARIMA (p,d,q)	R^2	RMSE	MAE	SVR (γ,C,ε)	R^2	RMSE	MAE
6 Month lead time													
Upper	Bantu Liben	5-4-1	0.7013	0.3737	0.3051	(5,0,0)	0.5912	0.8757	0.7714	0.4, 99, 0.008	0.7343	0.2427	0.2063
	Ginchi	5-4-1	0.7120	0.3834	0.3592	(5,1,1)	0.5391	0.8851	0.7459	0.4, 94, 0.008	0.7028	0.2535	0.1984
	Sebeta	5-4-1	0.6993	0.3784	0.3516	(4,0,0)	0.5827	0.8789	0.7407	0.4, 96, 0.008	0.7320	0.2805	0.2007
	Ejersalele	5-4-1	0.7103	0.3751	0.3458	(3,1,0)	0.5261	0.8846	0.7539	0.5, 90, 0.006	0.7127	0.2997	0.1984
	Ziquala	6-4-1	0.7201	0.3765	0.3271	(4,0,0)	0.5713	0.8623	0.7436	0.3, 89, 0.005	0.7382	0.2690	0.2052
Middle	Modjo	5-4-1	0.7064	0.3673	0.3281	(3,110)	0.5329	0.8600	0.7333	0.6, 88, 0.008	0.7326	0.2917	0.2044
	Wolenchiti	6-4-1	0.7139	0.3722	0.3481	(5,0,0)	0.5537	0.8665	0.7388	0.9, 90, 0.008	0.7044	0.2884	0.1987
	Gelemsso	5-4-1	0.7045	0.3619	0.3581	(4,0,0)	0.5572	0.8751	0.7459	0.8, 91, 0.007	0.7150	0.2891	0.2147
	Dire Dawa	5-4-1	0.7123	0.3781	0.3440	(5,0,1)	0.5928	0.8731	0.7443	0.4, 92, 0.005	0.7329	0.2838	0.1885
Lower	Dubti	6-4-1	0.7077	0.3409	0.3103	(5,0,0)	0.5472	0.8928	0.7607	0.8, 93, 0.008	0.7346	0.3054	0.1851
	Eliwuha	5-4-1	0.7565	0.3296	0.3019	(4,0,1)	0.5534	0.8857	0.7798	0.8, 95, 0.008	0.7653	0.3136	0.1953
	Bati	6-4-1	0.7210	0.3434	0.3097	(5,0,0)	0.5237	0.8928	0.7940	0.5, 87, 0.009	0.7547	0.2831	0.2162
12 Month lead time													
Upper	Bantu Liben	5-4-1	0.5129	0.4120	0.3742	(4,0,0)	0.4421	0.9556	0.7577	0.4, 90, 0.01	0.6087	0.3723	0.2582
	Ginchi	6-4-1	0.5346	0.4309	0.3781	(4,0,1)	0.4472	0.9015	0.7343	0.5, 91, 0.007	0.6078	0.3794	0.2352
	Sebeta	5-4-1	0.5456	0.4456	0.3981	(5,0,0)	0.4638	0.9123	0.7295	0.6, 97, 0.003	0.6123	0.3912	0.2807
	Ejersalele	6-4-1	0.5422	0.4021	0.3448	(5,0,0)	0.4537	0.9286	0.7416	0.6, 100, 0.01	0.6078	0.3466	0.2051
	Ziquala	7-4-1	0.5437	0.4358	0.3891	(4,0,1)	0.4682	0.9526	0.7321	0.8, 99, 0.03	0.6234	0.3556	0.2599
Middle	Modjo	6-4-1	0.5273	0.3996	0.3568	(4,1,0)	0.4572	0.9725	0.7755	0.5, 96, 0.008	0.6350	0.3482	0.2034
	Wolenchiti	7-4-1	0.5358	0.4180	0.3689	(3,0,0)	0.4589	0.9545	0.7785	0.8, 96, 0.008	0.6267	0.3114	0.1963
	Gelemsso	5-4-1	0.5441	0.4322	0.3889	(3,0,1)	0.4578	0.9592	0.7752	0.6, 94, 0.009	0.6290	0.3300	0.2115
	Dire Dawa	5-4-1	0.5291	0.3900	0.3770	(3,0,0)	0.4280	0.9678	0.7690	0.5, 96, 0.008	0.6443	0.3466	0.2016
Lower	Dubti	6-4-1	0.5127	0.3976	0.3782	(2,0,2)	0.4627	0.9285	0.7904	0.6, 96, 0.004	0.6239	0.3359	0.2212
	Eliwuha	4-4-1	0.5371	0.4304	0.3623	(5,0,0)	0.4632	0.9204	0.7837	0.8, 94, 0.008	0.6219	0.3087	0.1958
	Bati	7-4-1	0.5230	0.4032	0.3703	(4,0,0)	0.4547	0.9890	0.7837	0.5, 96, 0.008	0.6387	0.3783	0.2426

Table 5
The best WA-ANN and WA-SVR models for 6 and 12 month forecasts of SPI 12. Column 3 is the ANN architecture detailing the number of nodes in the input, hidden and output layers respectively. In column 7 the parameters of the SVR models are given.

Basin	Station	WA-ANN	R^2	RMSE	MAE	WA-SVR	R^2	RMSE	MAE
6 Month lead time									
Upper	Bantu Liben	5-6-1	0.8696	0.2023	0.1933	0.4, 99, 0.008	0.8507	0.2409	0.1934
	Ginchi	5-5-1	0.8276	0.2095	0.1831	0.4, 94, 0.008	0.8144	0.2314	0.2136
	Sebeta	5-5-1	0.8815	0.2013	0.1833	0.4, 96, 0.008	0.8380	0.2343	0.2214
	Ejersalele	7-5-1	0.9090	0.2066	0.1821	0.5, 90, 0.006	0.8363	0.2221	0.2054
	Ziquala	6-5-1	0.8616	0.2057	0.1830	0.3, 89, 0.005	0.8330	0.2413	0.2138
Middle	Modjo	5-4-1	0.8953	0.2083	0.1828	0.6, 88, 0.008	0.8732	0.2245	0.2054
	Wolenchiti	7-4-1	0.9332	0.2015	0.1892	0.9, 90, 0.008	0.8644	0.2518	0.2210
	Gelemsso	6-5-1	0.9204	0.2000	0.1845	0.8, 91, 0.007	0.8726	0.2167	0.2017
	Dire Dawa	7-6-1	0.9129	0.2001	0.1928	0.4, 92, 0.005	0.8968	0.2212	0.2036
Lower	Dubti	4-4-1	0.9231	0.2060	0.1847	0.8, 93, 0.008	0.8640	0.2185	0.2084
	Eliwuha	5-4-1	0.9326	0.2088	0.1804	0.8, 95, 0.008	0.8671	0.2440	0.2213
	Bati	6-4-1	0.9005	0.2183	0.1997	0.5, 87, 0.009	0.8441	0.2467	0.2341
12 Month lead time									
Upper	Bantu Liben	5-6-1	0.8034	0.2235	0.2115	0.4, 90, 0.01	0.7535	0.2484	0.2228
	Ginchi	6-5-1	0.8261	0.2416	0.2128	0.5, 91, 0.007	0.7533	0.2455	0.2030
	Sebeta	5-5-1	0.8049	0.2314	0.2126	0.6, 97, 0.003	0.7148	0.2734	0.2622
	Ejersalele	8-6-1	0.8162	0.2208	0.2128	0.6, 100, 0.01	0.7342	0.2645	0.2406
	Ziquala	7-5-1	0.8300	0.2206	0.2128	0.8, 99, 0.03	0.7604	0.2922	0.2711
Middle	Modjo	6-4-1	0.8292	0.2334	0.2242	0.5, 96, 0.008	0.7643	0.2645	0.2449
	Wolenchiti	8-5-1	0.8046	0.2365	0.2247	0.8, 96, 0.008	0.7843	0.2817	0.2733
	Gelemsso	7-6-1	0.8272	0.2354	0.2216	0.6, 94, 0.009	0.7721	0.2996	0.2717
	Dire Dawa	8-6-1	0.8219	0.2362	0.2117	0.5, 96, 0.008	0.7813	0.2674	0.2241
Lower	Dubti	5-5-1	0.8549	0.2406	0.2052	0.6, 96, 0.004	0.7443	0.2745	0.2620
	Eliwuha	6-4-1	0.8473	0.2719	0.2534	0.8, 94, 0.008	0.7641	0.3012	0.2639
	Bati	8-4-1	0.8437	0.2350	0.2248	0.5, 96, 0.008	0.7134	0.2611	0.2418

some instances where the observed SPI value corresponds to extreme conditions (−2 and 2) respectively. Fig. 5 is a scatter plot of the observed and ANN forecast results for the Eliwuha station. The scatter plot in Fig. 5 shows several points significantly below the trend line indicating a certain level of underestimation in the ANN model results. The proposed ANN model from the Eliwuha station had an R^2 of 0.7564, an RMSE of 0.3296 and an MAE of 0.3019, respectively. The R^2 values of 0.7564 shows a good correlation between observed and predicted results at 6-months lead time.

5.3. WA-ANN models

The results of all the proposed WA-ANN models can be found in Table 4 for both forecasts of 6 and 12 months lead time. The forecast results of WA-ANN models are more accurate than ANN models according to the performance measures used. The use of wavelets as a pre-processing tool resulted in more accurate and precise models as seen in Tables 4 and 6. These model results have a higher level of correlation between the observed and predicted time series and the RMSE and MAE values are lower compared to

Table 6
The best ARIMA, ANN and SVR models for 6 and 12 month forecasts of SPI 24. Column 3 is the ANN architecture detailing the number of nodes in the input, hidden and output layers respectively. In column 11 the parameters of the SVR models are given.

Basin	Station	ANN models	R^2	RMSE	MAE	ARIMA (p,d,q)	R^2	RMSE	MAE	SVR (γ,C,ε)	R^2	RMSE	MAE
6 Month lead time													
Upper	Bantu	4-4-1	0.7832	0.2775	0.2404	(5,1,0)	0.6486	0.5867	0.5735	0.8, 93, 0.08	0.7754	0.3192	0.2967
	Liben	5-4-1	0.7949	0.3302	0.3238	(4,1,0)	0.6302	0.5853	0.5707	0.8, 94, 0.06	0.7581	0.3027	0.2883
	Ginchi	5-4-1	0.7682	0.3421	0.3325	(5,0,1)	0.6682	0.5844	0.5688	0.9, 93, 0.08	0.7961	0.3338	0.2991
	Sebeta	6-4-1	0.7881	0.3347	0.3131	(5,1,0)	0.6427	0.5832	0.5665	0.8, 93, 0.08	0.7515	0.3546	0.3029
	Ejersalele	4-4-1	0.7947	0.3619	0.3421	(5,0,1)	0.6362	0.5822	0.5644	0.5, 99, 0.05	0.7862	0.2627	0.2534
	Ziquala												
Middle	Modjo	4-4-1	0.7915	0.2836	0.2629	(5,0,0)	0.6526	0.6046	0.5421	0.7, 92, 0.06	0.7637	0.3115	0.2951
	Wolenchiti	4-4-1	0.7844	0.3302	0.3026	(4,0,1)	0.5721	0.5927	0.5721	0.6, 94, 0.08	0.7619	0.3143	0.3065
	Gelemsso	5-4-1	0.7916	0.2262	0.2202	(5,0,0)	0.6676	0.5821	0.5626	0.45, 96, 0.01	0.7738	0.3498	0.3431
	Dire Dawa	5-4-1	0.8042	0.2662	0.2632	(3,1,1)	0.6742	0.5780	0.5581	0.4, 89, 0.03	0.7887	0.2933	0.2788
Lower	Dubti	4-4-1	0.8041	0.3397	0.3187	(5,0,0)	0.6248	0.5819	0.5355	0.4, 94, 0.07	0.7983	0.3141	0.2874
	Eliwuha	5-4-1	0.7965	0.3129	0.3017	(5,0,0)	0.6574	0.5827	0.5282	0.5, 97, 0.06	0.7831	0.3149	0.2932
	Bati	5-4-1	0.7914	0.2786	0.2645	(5,1,0)	0.6340	0.5595	0.5295	0.8, 98, 0.05	0.7681	0.2958	0.2804
12 Month lead time													
Upper	Bantu	5-4-1	0.7122	0.3545	0.3422	(4,1,1)	0.5442	0.7470	0.6946	0.55, 88, 0.08	0.7259	0.3256	0.3179
	Liben	6-4-1	0.7294	0.3620	0.3467	(5,1,0)	0.5078	0.7414	0.6379	0.65, 96, 0.08	0.7219	0.3282	0.3074
	Ginchi	6-4-1	0.7164	0.3632	0.3474	(5,010)	0.5939	0.7377	0.6200	0.85, 99, 0.06	0.7083	0.3387	0.3188
	Sebeta	7-4-1	0.7231	0.3640	0.3493	(5,1,0)	0.5025	0.7330	0.5876	0.8, 88, 0.09	0.7296	0.3667	0.3594
	Ejersalele	5-4-1	0.7017	0.3692	0.3508	(5,0,1)	0.5145	0.7288	0.5804	0.6, 91, 0.06	0.7139	0.3786	0.3683
	Ziquala												
Middle	Modjo	5-4-1	0.7129	0.3816	0.3710	(4,0,1)	0.4780	0.8637	0.7777	0.6, 91, 0.05	0.7233	0.3783	0.3684
	Wolenchiti	5-4-1	0.7064	0.3705	0.3657	(5,0,0)	0.4471	0.8762	0.8186	0.45, 100, 0.09	0.7235	0.3630	0.3529
	Gelemsso	6-4-1	0.7215	0.3534	0.3290	(4,1,1)	0.5659	0.8889	0.8300	0.6, 98, 0.08	0.7242	0.3768	0.3710
	Dire Dawa	6-4-1	0.7146	0.3412	0.3148	(5,1,0)	0.5216	0.9035	0.7709	0.65, 92, 0.05	0.7023	0.3687	0.3433
Lower	Dubti	5-4-1	0.7179	0.3716	0.3580	(5,0,0)	0.5158	0.8854	0.7244	0.5, 88, 0.09	0.7253	0.3398	0.3158
	Eliwuha	6-4-1	0.7189	0.3715	0.3471	(4,0,1)	0.4265	0.8643	0.7275	0.25, 100, 0.1	0.7018	0.3420	0.3261
	Bati	6-4-1	0.7136	0.3708	0.3600	(5,1,0)	0.4915	0.8312	0.7729	0.88, 95, 0.07	0.7061	0.3387	0.2916

Table 7
The best WA-ANN and WA-SVR models for 6 and 12 month forecasts of SPI 24. Column 3 is the ANN architecture detailing the number of nodes in the input, hidden and output layers respectively. In column 7 the parameters of the SVR models are given.

Basin	Station	WA-ANN	R^2	RMSE	MAE	WA-SVR	R^2	RMSE	MAE
6 Month lead time									
Upper	Bantu Liben	5-6-1	0.9665	0.1968	0.1803	0.8, 93, 0.08	0.8832	0.2461	0.2108
	Ginchi	5-5-1	0.9254	0.2850	0.2671	0.8, 94, 0.06	0.8740	0.2475	0.2148
	Sebeta	5-5-1	0.8864	0.1821	0.1723	0.9, 93, 0.08	0.8742	0.2581	0.2331
	Ejersalele	7-5-1	0.8791	0.1778	0.1632	0.8, 93, 0.08	0.8683	0.2287	0.2023
	Ziquala	6-5-1	0.8978	0.2546	0.2395	0.5, 99, 0.05	0.8869	0.2192	0.2133
Middle	Modjo	5-4-1	0.9166	0.2561	0.2481	0.7, 92, 0.06	0.8938	0.2407	0.2281
	Wolenchiti	7-4-1	0.9014	0.2115	0.2049	0.6, 94, 0.08	0.8892	0.2218	0.2106
	Gelemsso	6-5-1	0.9407	0.2258	0.2054	0.45, 96, 0.01	0.8935	0.2773	0.2694
	Dire Dawa	7-6-1	0.9215	0.2335	0.2149	0.4, 89, 0.03	0.8865	0.2487	0.2124
Lower	Dubti	5-4-1	0.8953	0.2618	0.2531	0.4, 94, 0.07	0.8938	0.3094	0.2870
	Eliwuha	5-4-1	0.9122	0.2190	0.2015	0.5, 97, 0.06	0.9024	0.2100	0.1993
	Bati	6-4-1	0.9450	0.2236	0.2159	0.8, 98, 0.05	0.8953	0.2386	0.2114
12 Month lead time									
Upper	Bantu Liben	5-6-1	0.8372	0.3342	0.3025	0.55, 88, 0.08	0.8171	0.3531	0.3388
	Ginchi	6-5-1	0.8637	0.3373	0.3191	0.65, 96, 0.08	0.8284	0.2743	0.2414
	Sebeta	5-5-1	0.8331	0.3325	0.3282	0.85, 99, 0.06	0.8398	0.3548	0.3302
	Ejersalele	8-6-1	0.8277	0.3372	0.3362	0.8, 88, 0.09	0.8420	0.3353	0.3209
	Ziquala	7-5-1	0.8316	0.3245	0.3083	0.6, 91, 0.06	0.8554	0.3560	0.3422
Middle	Modjo	6-4-1	0.8419	0.3061	0.2931	0.6, 91, 0.05	0.8013	0.3591	0.3396
	Wolenchiti	8-5-1	0.8134	0.2948	0.2844	0.45, 100, 0.09	0.8400	0.3159	0.2760
	Gelemsso	7-6-1	0.8178	0.2879	0.2753	0.6, 98, 0.08	0.8223	0.2777	0.2697
	Dire Dawa	8-6-1	0.8659	0.2853	0.2762	0.65, 92, 0.05	0.8548	0.3410	0.3163
Lower	Dubti	5-5-1	0.8303	0.2827	0.2749	0.5, 88, 0.09	0.8472	0.3051	0.2827
	Eliwuha	6-4-1	0.8462	0.2874	0.2697	0.25, 100, 0.1	0.8494	0.3184	0.3073
	Bati	8-4-1	0.8086	0.2819	0.2620	0.88, 95, 0.07	0.8538	0.3019	0.2794

ANN models. For forecasts of SPI 12 (6 months lead time) the results are very similar across all stations. Fig. 6 shows the forecast results of the Eliwuha station for SPI 12. The figure indicates that the WA-ANN forecasts closely mirror the observed results and do not underestimate extreme events of precipitation or drought. The model does not accurately depict the inception of the drought period at the end of 1970. While the observed SPI series indicates a slight drought, the forecasts indicate a moderate period of precipitation.

Fig. 7 is a scatter plot of the observed and predicted SPI values from the WA-ANN model. The figure shows that the points are closer to the trend line and there are no points of significant overestimation or underestimation. Fig. 8 illustrates the effects of the approximation series on the SPI time series for the Eliwuha station.

Table 8
Results for the calibration/training data set of the Eliwuha station (SPI 12, 6 months lead time).

Model	R^2	RMSE	MAE
ARIMA	0.5837	0.8903	0.8773
ANN	0.7241	0.3348	0.3180
SVR	0.7528	0.3062	0.2381
WA-ANN	0.9335	0.2140	0.1906
WA-SVR	0.8631	0.2481	0.2301

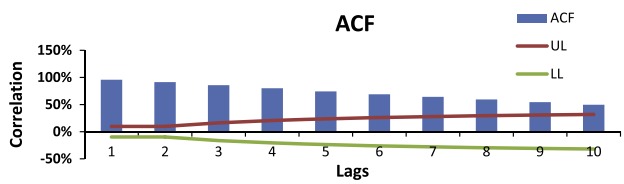

Fig. 3a. Autocorrelation of the SPI 12 time series for the Eliwuha station.

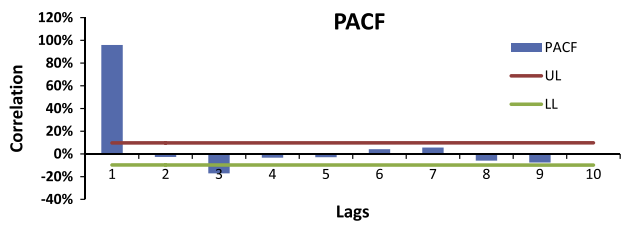

Fig. 3b. Partial autocorrelation of the SPI 12 time series for the Eiwuha station.

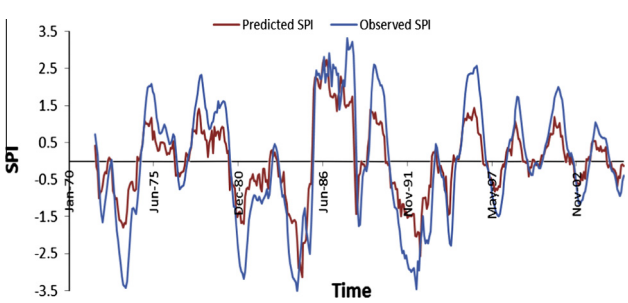

Fig. 4. Observed vs. predicted SPI 12 at Eliwuha station for the ANN model.

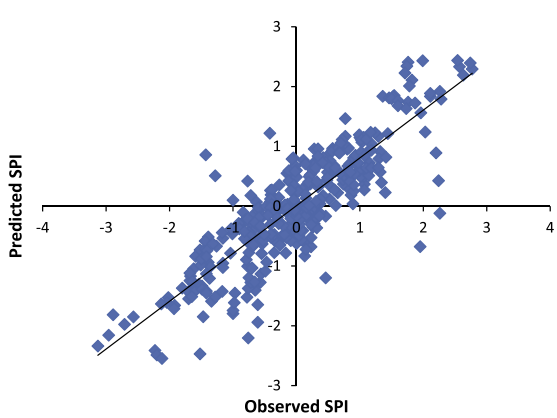

Fig. 5. Scatter plot for SPI values at Eliwuha station for the ANN model.

This study found that the approximation series after wavelet decomposition is useful to use in forecasting SPI time series. For the model in Fig. 8, the approximation series is from wavelet

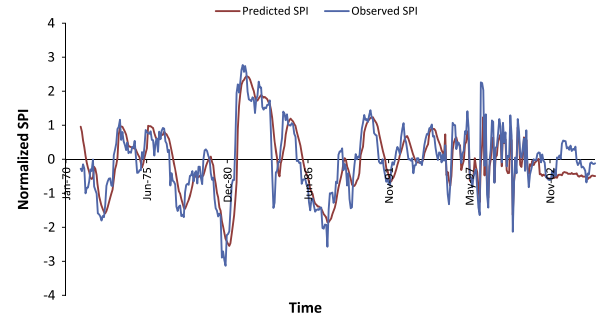

Fig. 6. Observed vs. predicted SPI at the Eliwuha station for the WA-ANN model.

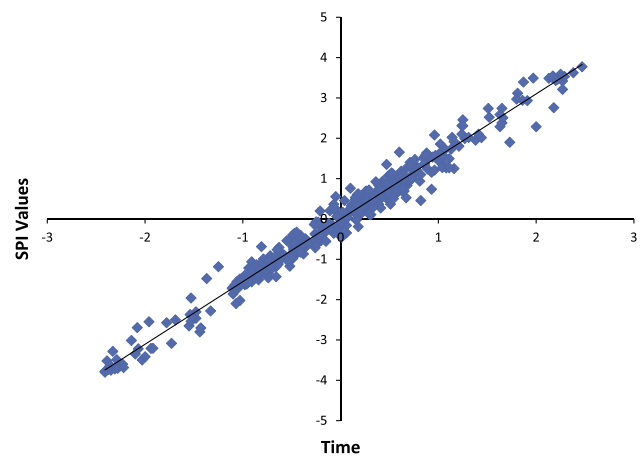

Fig. 7. Scatter plot from WA-ANN forecasts of SPI 12 at the Eliwuha station.

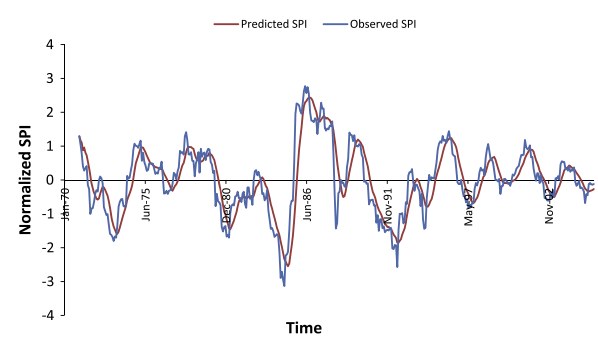

Fig. 8. Time series showing the observed SPI values against the approximation series after wavelet decomposition for the Eliwuha station.

decomposition at level three. The approximation series closely mirrors the periods of precipitation and drought exhibited by the SPI series. This approximation time series was subsequently used as an input in WA-ANN or WA-SVR models. As seen in Tables 3–6, the forecast results for WA-ANN and WA-SVR models are improved compared to models without any wavelet pre-processing.

5.4. SVR models

Fig. 9 illustrates the forecast results of the best proposed SVR model at the Eliwuha station. The SVR model had trouble predicting the inception of drought at the end of 1970. While the model does accurately predict events of drought there seems to be a slight time shift error. The predicted values are slightly lagged compared

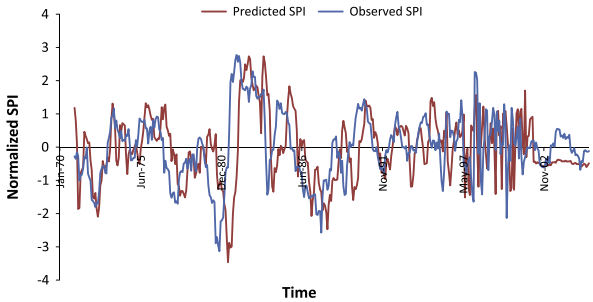

Fig. 9. Observed vs. predicted SPI 12 from a SVR model at the Eliwuha Station.

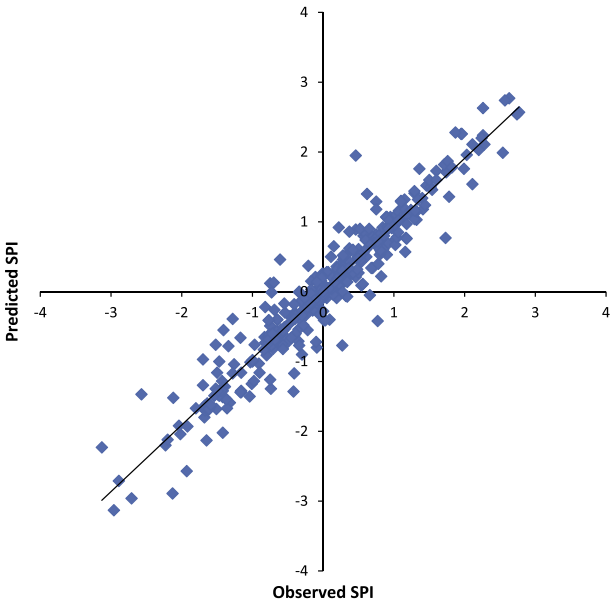

Fig. 10. Scatter plot of SPI 12 values for a SVR model at the Eliwuha Station.

to the observed values, which may make applications of SVR models for drought forecasting problematic. The SVR model also had a good correlation between the observed and predicted values as exhibited by the R^2 value of 0.7382. The scatter plot in Fig. 10 illustrates the correlation between the observed and predicted SPI values. The observations in the scatter plot are symmetrical around the trend line.

5.5. WA-SVR models

The use of wavelet analysis improved the forecasting ability of SVR models with respect to SPI 12 and SPI 24. As shown in Tables 4 and 6, the use of wavelet analysis improves the performance measures across all the stations. The R^2 values for the WA-SVR models are between 0.8144 and 0.8968; these values are all greater than any SVR model for SPI 12, indicating a greater level of correlation between observed and predicted values. For forecasts of 12 months lead time for both SPI values refer to Tables 3–6. As with other models an increase in lead time results in a deterioration of the forecast accuracy.

The forecast results from WA-SVR models are more accurate than SVR models, according to the performance measures. The inputs for the WA-SVR models were generated from the approximation series after wavelet decomposition of the SPI time series. The fact that the use of just the approximation series improved the results significantly indicates that the approximation series

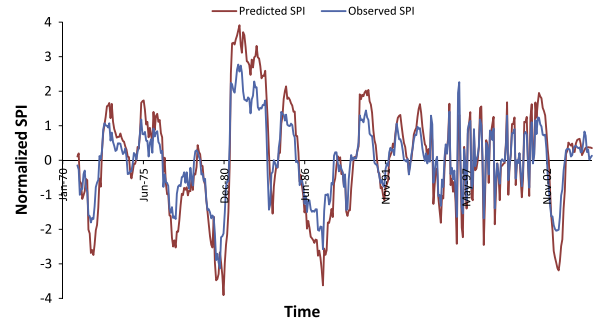

Fig. 11. Observed vs. predicted SPI 12 from a WA-SVR model at the Eliwuha Station.

adequately de-noises the data and avoids any discontinuities within the SPI time series.

Fig. 11 indicates that the WA-SVR model at the Eliwuha station is good at predicting the inception of a drought as shown in the figure around the end of 1970. Fig. 11 also indicates that the WA-SVR model tends to overestimate periods of extreme precipitation or drought. The overestimation is not a mis-categorization of the event. SPI values of −2 and −3 are both representative of extreme events. In general, the predicted values mirror the trends observed with the original time series. For the Eliwuha station, the WA-SVR model is effective at predicting the end of the SPI time series. The results from Figs. 6 and 9 show the models predicting slight to moderate drought while the observed time series indicates slight precipitation. The WA-SVR model was better at predicting the events at the end of the time series. The scatter plot in Fig. 12 shows the good level of correlation between the observed and predicted SPI values at the Eliwuha station.

6. Discussion

This study has shown that data driven models can be an effective means of forecasting drought at forecast lead times of 6 and 12 months in the Awash River Basin. The results indicate that machine learning techniques (ANNs and SVR) are more effective than a traditional stochastic model such as an ARIMA model in forecasting SPI 12 and SPI 24 at the aforementioned lead times. This is likely due to the fact that ANN and SVR models are effective in modeling non-linear components of time series data. Furthermore, the use of wavelet analysis as a pre-processing tool improved the forecast results for both ANN and SVR models. As might be expected, the results also indicate that as the forecast lead time is increased the correlation between observed and predicted values, as measured by R^2, decreases considerably. While the RMSE and the MAE increase with increasing forecast lead time, their increase is not as pronounced. An increase in forecast lead time from 6 to 12 months did not result in poor results, especially when wavelet analysis was used, which highlights the effectiveness of this pre-processing method for ANN and SVR models in predicting the SPI. The input structure of the models does not change with the increase of forecast lead time, which makes the models convenient for operational purposes.

The results from all the data driven models generally show that SPI 24 forecasts were more accurate than SPI 12 forecasts. Both SPI 12 and SPI 24 are long-term SPI and each new month has less impact on the period of sum precipitation (McKee et al., 1993) compared to short-term precipitation. As a result, monthly variation in precipitation has a smaller impact for both these SPI compared to short-term SPI. However, as SPI 24 is a longer term SPI its sensitivity to changes in precipitation within the time series is less than that of SPI 12. This lack of sensitivity may explain why the models are slightly more effective at generalizing SPI 24 better than SPI 12.

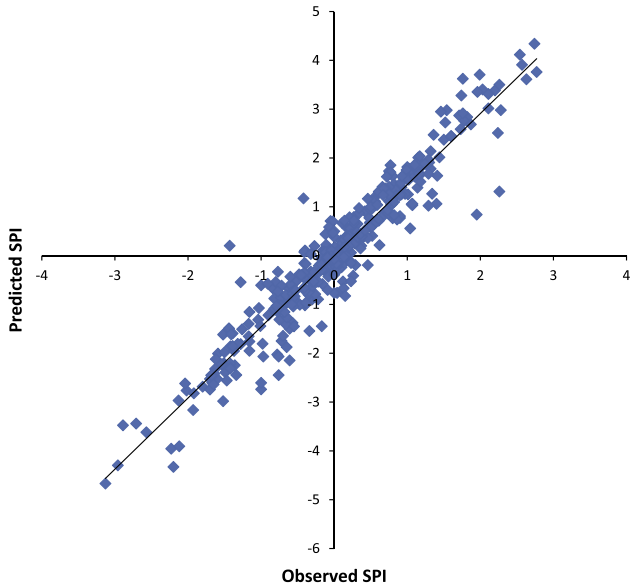

Fig. 12. Scatter plot from WA-SVR forecasts of SPI 12 at the Eliwuha station.

The forecast accuracy of the proposed models does not differ significantly between each of the basins. The lack of a significant difference in terms of the forecast accuracy indicates that the different conditions within the three sub-basins do not appreciably affect the forecast of the SPI. In general, the performances of SVR and ANN models were comparable. The use of wavelets improved the results of both machine learning techniques, with the forecast measures indicating that WA-ANN models slightly outperformed WA-SVR models. Theoretically, SVR models should perform better than ANN models because they adhere to the structural risk minimization principle instead of the empirical risk minimization principle. They should, in theory, not be as susceptible to local minima or maxima. However, there have been studies that have shown that the performance of SVR and ANN models are comparable. Lima et al. (2013) found that SVR models were more effective at precipitation forecasts when MAE was the performance measure and ANN models were more effective when MSE was the performance measure. A study by Shin et al. (2005) and Chevalier et al. (2011) found that the application of ANN models in time series forecasting was comparable to those of SVR models especially as the size of the training set was increased. The study by Chevalier et al. (2011) also found that SVR models were superior in the training phase, while ANN models were superior in the evaluation phase. Witten et al. (2011) found that ANN models are comparable to SVR models because they can learn to ignore irrelevant attributes. Witten et al. (2011) also state that there is no universally superior learning method.

From the figures illustrating all the forecasts using all of the data driven models, it is apparent that there is not much time step error in the forecasts of SPI 12. However, the SVR model at the Eliwuha station for SPI 12 did show a lag between observed and predicted events. This time shift error is unique to the SVR models in this study. This time shift error associated with the SVR models is indicated by the delayed drought forecasts. Time shift error is caused by the autoregressive components related to the selection of input variables (Abrahart et al., 2007). The use of past data in the forecasts of SPI values at long lead times introduces time shift error. Long lead time forecasts that do not possess time shift error usually result in noisier forecasts (Chua and Wong, 2011). In an operational setting, the presence of a time shift error would likely compromise the ability of planners to implement an effective drought warning system. Forecasts of SPI 12 at the Eliwuha station using the SVR model showed that the inception of the drought (according to the observed time series) was not accurately predicted. However, the WA-SVR model at the same station did adequately predict the inception of drought but was relatively noisier than the SVR model.

7. Conclusion

This study investigated the ability of data driven models to forecast drought. This study also proposed and evaluated, for the first time, the use of the WA-SVR method for long-term drought forecasting.

Overall, coupled wavelet-neural network (WA-ANN) models were found to provide better results than the other model types used for forecasts of SPI 12 and SPI 24 in the Awash River Basin, especially for SPI 24. WA-ANN models showed a higher coefficient of determination between observed and predicted SPI, as well as lower RMSE and MAE values compared to simple ANNs, ARIMA, SVR and WA-SVR models. Wavelet coupled models also consistently showed lower values of RMSE and MAE compared to the other data driven models. The coupled models provide more accurate results because pre-processing the original SPI time series with wavelet decompositions improves the forecast results over time series that do not use wavelet decompositions. Wavelet analysis de-noises the SPI time series and subsequently allows the ANN and SVR model to model the main signal without the noise. Wavelet analysis also reduces the sensitivity to changes in monthly precipitation within the SPI time series. In the case of SVR models, the use of wavelets also reduces the lag between forecasts and the observed SPI values.

This study focused on long-term drought forecasts of SPI 12 and SPI 24 in the Awash River Basin. Further studies need to be done to determine which of these data driven models is suitable for forecasting long-term SPI values in other locations with different climates and different physical characteristics. Considering the fact that the Middle and Lower Awash sub-basins have a very similar climate, studies of areas with different climates should be conducted to determine whether there is a significant link between forecast accuracy and climate. This study found that the different characteristics and climatology of the sub-basins had no discernible effect on forecast accuracy. Future studies should also attempt to couple data driven drought forecasting models with uncertainty analysis, such as bootstrapping or boosting ensembles.

Acknowledgements

This research was funded by an NSERC Discovery Grant and a CFI Grant held by Jan Adamowski. The data was obtained from National Meteorological Agency of Ethiopia. Their help is greatly appreciated.

References

Abrahart, R.J., Heppenstall, A.J., See, L.M., 2007. Timing error correction procedure applied to neural network rainfall0runoff modelling. Journal of Hydrologic Sciences 52 (3), 414–432.

Adamowski, J., 2008. Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. Journal of Hydrology 353, 247–266.

Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology 407, 28–40.

Adamowski, J., Sun, K., 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. Journal of Hydrology 390, 85–91.

Adamowski, J., Chan, E., Prasher, S., Ozga-Zielinski, B., Sliusareve, A., 2012. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network and wavelet artificial neural network methods for urban water demand forecasing in Montreal, Canada. Water Resources Research 48, W01528. http://dx.doi.org/10.1029/2010WR009945.

Bacanli, U.G., Firat, M., Dikbas, F., 2008. Adaptive neuro-fuzzy inference system for drought forecasting. Stochastic Environmental Research and Risk Assessment 23 (8), 1143–1154.
Barros, A., Bowden, G., 2008. Toward long-lead operational forecasts of drought: an experimental study in the Murray-Darling River Basin. Journal of Hydrology 357 (3–4), 349–367.
Belayneh, A., Adamowski, J., 2012. Standard Precipitation Index drought forecasting using neural networks, wavelet neural networks, and support vector regression. Applied Computational Intelligence and Soft Computing. Article ID 794061, 13 pages.
Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology 320, 18–36.
Bordi, I., Sutera, A., 2007. Drought monitoring and forecasting at large scale. In: Rossi, G. et al. (Eds.), Methods and Tools for Drought Analysis and Management. Springer.
Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.
Cacciamani, C., Morgillo, A., Marchesi, S., Pavan, V., 2007. Monitoring and forecasting drought on a regional scale: Emilia-Romagna region. Water Science and Technology Library 62 (1), 29–48.
Cancelliere, A., Di Mauro, G., Bonaccorso, B., Rossi, G., 2007a. Stochastic forecasting of drought indices. In: Rossi, G. et al. (Eds.), Methods and Tools for Drought Analysis and Management. Springer.
Cancelliere, A., Di Mauro, G., Bonaccorso, B., Rossi, G., 2007b. Drought forecasting using the standard precipitation index. Water Resource Management 21, 801–819.
Cannas, B., Fanni, A., Sias, G., Tronci, S., Zedda, M.K., 2006. River flow forecasting using neural networks and wavelet analysis. Proceedings of the European Geosciences Union.
Chevalier, R.F., Hoogenboom, G., McClendon, R.W., Paz, J.A., 2011. Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. Neural Computing and Applications 20, 151–159.
Chua, I.H.C., Wong, T.S.W., 2011. Runoff forecasting for an asphalt plane by artificial neural networks and comparisons with kinematic wave and autoregressive moving average models. Journal of Hydrology 397, 191–201.
Cimen, M., 2008. Estimation of daily suspended sediments using support vector machines. Hydrological Sciences Journal 53 (3), 656–666.
Cutore, P., Di Mauro, G., Cancelliere, A., 2009. Forecasting palmer index using neural networks and climatic indexes. Journal of Hydrologic Engineering 14, 588–595.
Desalegn, C., Babel, M.S., Das Gupta, A., Seleshi, B.A., Merrey, D., 2006. Farmers' perception about water management under drought conditions in the Awash River Basin, Ethiopia. International Journal of Water Resources Development 22 (4), 589–602.
Edossa, D.C., Babel, M.S., Gupta, A.D., 2010. Drought analysis on the Awash River Basin, Ethiopia. Water Resource Management 24, 1441–1460.
Gao, J.B., Gunn, S.R., Harris, J., Brown, M., 2001. A probabilistic framework for SVM regression and error bar estimation. Machine Learning 46, 71–89.
Han, P., Wang, P., Zhang, S., Zhu, D., 2010. Drought forecasting with vegetation temperature condition index. Wuhan Daxue Xuebao (Xinxi Kexue Ban)/ Geomatics and Information Science of Wuhan University 35 (10), 1202–1206+1259.
Husak, G.J., Funk, C.C., Michaelsen, J., Magadzire, T., Goldsberry, K.P., 2013. Developing seasonal rainfall scenarios for food security early warning. Theoretical and Applied Climatology. http://dx.doi.org/10.1007/s00704-013-0838-8.
Karamouz, M., Rasouli, K., Nazil, S., 2009. Development of a hybrid index for drought prediction: case study. Journal of Hydrologic Engineering 14, 617–627.
Keshavarz, M., Karami, E., Vanclay, F., 2013. The social experience of drought in rural Iran. Journal of Land Use Policy 30, 120–129.
Khan, M.S., Coulibaly, P., 2006. Application of support vector machine in lake water level prediction. Journal of Hydrologic Engineering 11 (3), 199–205 (American Society of Civil Engineering).
Kim, T., Valdes, J.B., 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. Journal of Hydrologic Engineering 8, 319–328.
Kisi, O., Cimen, M., 2009. Evapotranspiration modelling using support vector machines. Hydrological Science Journal 54 (5), 918–928.
Kisi, O., Cimen, M., 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. Journal of Hydrology 399, 132–140.
Labat, D., Ababou, R., Mangin, A., 1999. Wavelet analysis in Karstic hydrology 2nd Part: rainfall–runoff cross-wavelet analysis. Comptes Rendus de l'Academie des Sciences Series IIA Earth and Planetary Science 329, 881–887.
Lane, S.N., 2007. Assessment of rainfall–runoff models based upon wavelet analysis. Hydrological Processes 21, 586–607.
Lima, A.R., Cannon, A.J., Hsieh, W.W., 2013. Nonlinear regression in environmental sciences by support vector machines combined with evolutionary strategy. Computers and Geosciences 50, 136–144.
Linsley, R.K., Kohler, M.A., Paulhus, J.L.H., 1988. Hydrology for Engineers, International Edition. McGraw-Hill, Singapore.
Maheswaran, R., Khosa, R., 2012a. Comparative study of different wavelets for hydrologic forecasting. Computers and Geosciences 46, 284–295.
Maheswaran, R., Khosa, R., 2012b. Wavelet-volterra coupled model for monthly streamflow forecasting. Journal of Hydrology 450, 320–335.
Mallat, S.G., 1998. A Wavelet Tour of Signal Processing. Academic, San Diego, p. 577.
Marj, A.F., Meijerink, A.M., 2011. Agricultural drought forecasting using satellite images, climate indices and artificial neural network. International Journal of Remote Sensing 32 (24), 9707–9719.
McKee, T.B., Doesken, N.J., Kleist, J., 1993. The Relationship of Drought Frequency and Duration to Time Scales, Paper Presented at 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA.
Mishra, A.K., Desai, V.R., 2005. Drought forecasting using stochastic models. Stochastic Environmental Research and Risk Assessment 19 (5), 326–339.
Mishra, A.K., Desai, V.R., 2006. Drought forecasting using feed-forward recursive neural network. Ecological Modelling 198 (1–2), 127–138.
Mishra, S.S., Nagarajan, R., 2012. Forecasting drought in Tel River Basin using feed-forward recursive neural network. International Conference on Environmental, Biomedical and Biotechnology.
Mishra, A.K., Singh, V.P., 2012. Simulating hydrological drought properties at different spatial units in the United States based on wavelet-bayesian approach. Earth Interactions 16 (17), 1–23.
Mishra, A.K., Desai, V.R., Singh, V.P., 2007. Drought forecasting using a hybrid stochastic and neural network model. Journal of Hydrologic Engineering 12, 626–638.
Morid, S., Smakhtin, V., Moghaddasi, M., 2006. Comparison of seven meteorological indices for drought monitoring in Iran. International Journal of Climatology 26 (7), 971–985.
Morid, S., Smakhtin, V., Bagherzadeh, K., 2007. Drought forecasting using artificial neural networks and time series of drought indices. International Journal of Climatology 27 (15), 2103–2111.
Murtagh, F., Starcj, J.L., Renaud, O., 2003. On neuro-wavelet modeling. Decision Support Systems 37, 475–484.
Nason, G.P., Von Sachs, R., 1999. Wavelets in time-series analysis. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 357 (1760), 2511–2526.
Ntale, H.K., Gan, T.Y., 2003. Drought indices and their application to East Africa. International Journal of Climatology 23 (11), 1335–1357.
Ozger, M., Mishra, A.K., Singh, V.P., 2012. Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. Journal of Hydrometeorology 13 (1), 284–297.
Parrella, F., 2007. Online Support Vector Regression. Master Thesis, University of Genoa.
Partal, T., Kisi, O., 2007. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. Journal of Hydrology 342 (1–2), 199–212.
Renaud, O., Starck, J., Murtagh, F., 2002. Wavelet-based Forecasting of Short and Long Memory Time Series. Department of Economics, University of Geneve.
Shin, K., Lee, T.S., Kim, H., 2005. An application of support vector machines in bankruptcy prediction model. Expert Systems with Applications 28, 127–135.
Tiwari, M.K., Chatterjee, C., 2010. Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach. Journal of Hydrology 394, 458–470.
Tsakiris, G., Vangelis, H., 2004. Towards a drought watch system based on spatial SPI. Water Resources Management 18 (1), 1–12.
Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer Verlag, New York, USA.
Wanas, N., Auda, G., Kamel, M.S., Karray, F., 1998. On the optimal number of hidden nodes in a neural network. Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering., 918–921.
Wilhite, D.A. (Ed.), 1993. Drought Assessment, Management, and Planning: Theory and Case Studies. Natural Resource Management and Policy Series. Kluwer Publishers, Boston, MA.
Witten, I.H., Frank, E., Hall, M.A., 2011. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufman, Burlington, MA.