Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression

2 authors:

Anteneh Meshesha Belayneh
Carleton University
11 PUBLICATIONS   **499** CITATIONS

Jan Franklin Adamowski
McGill University
347 PUBLICATIONS   **8,128** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Development of a Group Built Coupled Physical – Socio-Economic Modelling Framework View project

Project   UpSWiNG: Understanding game-based approaches for improving sustainable water governance View project

Research Article

Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression

A. Belayneh and J. Adamowski

Department of Bioresource Engineering, Faculty of Agricultural and Environmental Sciences, McGill University,
QC, Canada H9X 3V9

Correspondence should be addressed to J. Adamowski, jan.adamowski@mcgill.ca

Drought forecasts can be an effective tool for mitigating some of the more adverse consequences of drought. Data-driven models are suitable forecasting tools due to their rapid development times, as well as minimal information requirements compared to the information required for physically based models. This study compares the effectiveness of three data-driven models for forecasting drought conditions in the Awash River Basin of Ethiopia. The Standard Precipitation Index (SPI) is forecast and compared using artificial neural networks (ANNs), support vector regression (SVR), and wavelet neural networks (WN). SPI 3 and SPI 12 were the SPI values that were forecasted. These SPI values were forecast over lead times of 1 and 6 months. The performance of all the models was compared using RMSE, MAE, and R^2. The forecast results indicate that the coupled wavelet neural network (WN) models were the best models for forecasting SPI values over multiple lead times in the Awash River Basin in Ethiopia.

1. Introduction

Droughts, a natural occurrence in almost all climatic zones, are a result of the reduction, for an extended period of time, of precipitation from normal amounts. Extended periods of drought can lead to several adverse consequences, which include a disruption of the water supply, low agricultural yields, and reduced flows for ecosystems. Consequently, the ability to forecast and predict the characteristics of droughts, specifically their initiation, frequency, and severity, is important. Effective drought forecasts are an effective tool for water resource management as well as an effective tool for the agricultural industry.

Currently, drought monitoring in Ethiopia is conducted by the National Meteorological Services Agency (NMSA). The NMSA regularly produces a 10-day bulletin that gives an analysis of rainfall based on the long-term average or normal. This bulletin is then circulated to a wide range of users, ranging from local development agents to decision makers at a national level. In addition to rainfall analysis, the normalized vegetation index (NDVI) is provided, which is a satellite-based index widely used to monitor vegetation and drought conditions. The NMSA produces a regular 10-day bulletin regarding NDVI variation that compares the current vegetation condition with normal or conditions of the previous year [1]. However, the NDVI is sensitive to changes in vegetative land cover and may not be effective in areas where vegetation is minimal. In addition, the NMSA of Ethiopia produces medium and seasonal forecasts of precipitation using the aforementioned NDVI.

Unlike other natural hazards, droughts have a slow evolution time [2]. The consequences of droughts take a significant amount of time to come into effect with respect to their inception, and when they are perceived by ecosystems and hydrological systems. Due to this feature, effective mitigation of the most adverse drought impacts is possible, more than in the case of other extreme hydrological events such as floods, earthquakes, or hurricanes, provided

a drought monitoring system, which is able to promptly warn of the onset of a drought and to follow its evolution in space and time, is in operation [3].

A common tool utilized to monitor current drought conditions is a drought index. Several drought indices can be used to forecast the possible evolution of an ongoing drought, in order to adopt appropriate mitigation measures and drought policies for water resources management [4]. This is because a drought index is expressed by a numeric number, which is believed to be far more functional than raw data during decision-making [2]. Several drought indices have been developed around the world in the past based on rainfall as the single variable, including the widely used Deciles [5], Standardized Precipitation Index (SPI) [6], and Effective Drought Index (EDI) [7]. There is also the well-known Palmer Drought Severity Index (PDSI) [8], which considers temperature along with rainfall. The SPI drought index was chosen to forecast drought in this study due to its simplicity, its ability to represent droughts on multiple time scales, and because it is a probabilistic drought index. In addition, the study by Ntale and Gan [9] determined that the SPI is the most appropriate index for monitoring the variability of droughts in East Africa because it is easily adapted to local climate, has modest data requirements, and can be computed at almost any time scale.

Forecasting any hydrologic phenomena can be done using either a physical, conceptual, or data-driven approach. The latter approach is widely used in hydrologic forecasting because data-driven models have low information requirements with respect to the number of variables required for inputs compared to physically based models. Data-driven models also have rapid development times. Unlike physical and conceptual models, data-driven models are not difficult to implement for the purposes of real-time forecasting. Artificial neural networks (ANNs) have been used in several studies as a drought-forecasting tool [10–16]. The most popular type of ANN used for the purposes of drought forecasting is the multilayer perceptron (MLP) that is usually optimized with a back propagation algorithm. However, ANNs are limited in their ability to deal with nonstationarities in the data, a weakness also shared by multiple linear regression (MLR) and autoregressive integrated moving average (ARIMA) models.

This limitation with nonstationary data has led to the recent formation of hybrid models, where data is preprocessed for nonstationary characteristics and then run through a forecasting method such as ANNs to cope with the nonlinearity. Wavelet analysis, an effective tool to deal with nonstationary data, has recently been applied in hydrological forecasting to examine the rainfall-runoff relationship in a Karstic watershed [17], to characterize daily streamflow [18, 19] and monthly reservoir inflow [20], to evaluate rainfall-runoff models [21], to forecast river flow [22–24], to forecast future precipitation values [25], and for the purposes of drought forecasting [26]. The study conducted by Kim and Valdes [26] is the only study that has explored the ability of a wavelet-neural network conjunction model (WN) to forecast a given drought index. However, no studies that assess the ability of WN models to forecast the SPI drought index in particular have been explored.

Support Vector Machines (SVMs) are a relatively new form of machine learning that was developed by Vapnik [27]. The term SVM is used to refer to both classification and regression methods as well as the terms Support Vector Classification (SVC) and Support Vector Regression (SVR), which refer to the problems of classification and regression, respectively [28]. There are several studies where SVRs were used in hydrological forecasting. Khan and Coulibaly [29] found that an SVR model was more effective at predicting 3–12 month lake water levels than ANN models. Rajasekaran et al. [30] used SVR successfully for storm surge predictions, and Kisi and Cimen [31, 32] used SVR to estimate daily evaporation and daily streamflow, respectively. Finally, SVR have been successfully used to predict hourly streamflow by Asefa et al. [33] and were shown to perform better than ANN and ARIMA models for monthly streamflow prediction by Wang et al. [34] and Maity et al. [35], respectively. Yuan and Tan [36] used SVRs as a screening tool to test for drought resistance of rice. However, to date SVRs have not been applied to forecast a given drought index.

This study compared the effectiveness of three data-driven models for forecasting drought conditions in the Awash River Basin of Ethiopia. The Standard Precipitation Index (SPI) was forecasted and compared using artificial neural networks (ANNs), support vector regression (SVR), and wavelet networks (WN). SPI 3 and SPI 12 were forecast over lead times of 1 and 6 months. The forecast lead times were chosen because a 1-month lead time is a typical short-term lead time and a 6-month lead time is representative of the bimodal rainfall pattern in the Awash River Basin. Forecast results of this study are useful for the agricultural water management sector and have the potential to be applied by water resources managers to effectively manage water resources in the region. In addition, accurate forecasts using these data-driven models can complement the forecasts already being used by the NMSA of Ethiopia.

2. Theoretical Development

In the following section, the computation of the SPI is briefly described. In addition to the description of the SPI, this section also describes the data-driven models that were used to forecast the SPI.

2.1. The Standard Precipitation Index (SPI). The Standard Precipitation Index (SPI) was developed by McKee et al. [6]. As mentioned in the previous section, one of the main advantages of the SPI is that it only requires precipitation data as an input, which makes it ideal for areas where data collection is not as extensive (such as in Ethiopia). The fact that the SPI is based solely on precipitation makes its evaluation relatively easy [37]. The SPI is a standardized index. Standardization of a drought index ensures independence from geographical position as the index in question is calculated with respect to the average precipitation in the same place [37].

TABLE 1: Drought classification based on SPI [6].

SPI values	Class
>2	Extremely wet
1.5–1.99	Very wet
1.0–1.49	Moderately wet
−0.99 to 0.99	Near normal
−1 to −1.49	Moderately dry
−1.5 to −1.99	Very dry
<−2	Extremely dry

The computation of the SPI drought index for any location is based on the long-term precipitation record (at least 30 years) cumulated over a selected time scale [38]. This long-term precipitation time series is then fitted to a gamma distribution, which is then transformed through an equal probability transformation into a normal distribution [38, 39]. Positive SPI values indicate wet conditions with greater than median precipitation, and negative SPI values indicate dry conditions with lower than median precipitation [38]. Table 1 below indicates SPI drought classes.

In most cases, the probability distribution that best models observational precipitation data is the Gamma distribution [37]. The density probability function for the Gamma distribution is given by the expression [37]:

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad \text{for } x > 0, \tag{1}$$

where $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter, and $x > 0$ is the amount of precipitation. $\Gamma(\alpha)$ is the value taken by the standard mathematical function known as the Gamma function, which is defined by the integral [37]:

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy. \tag{2}$$

In general, the Gamma function is evaluated either numerically or using the values tabulated depending on the value taken by parameter α.

In order to model the data observed with a gamma distributed density function, it is necessary to estimate parameters α and β appropriately. Different methods have been suggested in the literature for the estimate of these two parameters. For example, the Thom [40] approximation is used for maximum probability in Edwards and McKee [41]:

$$\hat{\alpha} = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right),$$
$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}}, \tag{3}$$

where for n observations

$$A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n}. \tag{4}$$

The estimate of the parameters can be further improved by using the interactive approach suggested in Wilks [42].

After estimating coefficients α and β the density of probability function $g(x)$ is integrated with respect to x and we obtain an expression for cumulative probability $G(x)$ that a certain amount of rain has been observed for a given month and for a specific time scale [37]:

$$G(x) = \int_0^x g(x) dx = \frac{1}{\hat{\beta}\Gamma(\hat{\alpha})} = \int_0^x x^{\hat{\alpha}-1} e^{-x/\beta} dx. \tag{5}$$

The Gamma function is not defined by $x = 0$, and since there may be no precipitation, the cumulative probability becomes [37]

$$H(x) = q + (1 - q)G(x), \tag{6}$$

where q is the probability of no precipitation. $H(x)$ is the cumulative probability of precipitation observed. The cumulative probability is then transformed into a normal standardized distribution with null average and unit variance from which we obtain the SPI index.

The above approach, however, is neither practical nor numerically simple to use if there are many grid points of many stations on which to calculate the SPI index. In this case, an alternative method is described in Edwards and McKee [41] using the technique of approximate conversion developed in Abramowitz and Stegun [43] that converts the cumulative probability into a standard variable Z. The SPI index is then defined as

$$Z = \text{SPI}$$
$$= \begin{cases} -\left(t - \dfrac{c_0 + c_1 t + c_2 t_2}{1 + d_1 t + d_2 t_2 + d_3 t_3}\right), & \text{for } 0 < H(x) \le 0.5, \\ +\left(t - \dfrac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right), & \text{for } 0.5 < H(x) < 1, \end{cases} \tag{7}$$

where

$$t = \begin{cases} \sqrt{\ln\left[\dfrac{1}{(H(x))^2}\right]}, & \text{for } 0 < H(x) \le 0.5, \\ \sqrt{\ln\left[\dfrac{1}{(1 - H(x))^2}\right]}, & \text{for } 0.5 < H(x) < 1, \end{cases} \tag{8}$$

where x is precipitation, $H(x)$ is the cumulative probability of precipitation observed, and $c_0, c_1, c_2, d_0, d_1, d_2$ are constants with the following values:

$$c_0 = 2.515517, \quad c_1 = 0.802853, \quad c_2 = 0.010328,$$
$$d_0 = 1.432788, \quad d_1 = 0.189269, \quad d_2 = 0.001308. \tag{9}$$

2.2. Artificial Neural Networks (ANNs).

Artificial neural networks (ANNs) are flexible computing frameworks that resemble the structure of a nerve system. ANNs have been used to model a broad range of hydrologic time series over the past two decades. The main advantage of using ANNs is that there is no need to define the physical processes between the inputs

and outputs [11]. This feature makes ANNs suitable for the purposes of drought forecasting, where all the variables that may cause a drought are not fully understood.

In this paper, the multilayer perceptron (MLP) feed-forward network was used to forecast the SPI time series. Figure 1 is an illustration of a typical feed-forward neural network. ANN models in this study were trained with the Levenberg Marquardt (LM) back propagation algorithm. MLPs have been used extensively in hydrologic forecasting studies [10, 12, 23, 26, 44, 45] due to their simplicity. In terms of their architecture, MLPs consist of an input layer, one or more hidden layers, and an output layer. The hidden layer contains the neuron-like processing elements that connect the input and output layers and is given by [26]

$$y'_k(t_s) = f_0 \left[\sum_{j=1}^{m} w_{kj} . f_n \left(\sum_{i=1}^{n} w_{ji} x_i(t_s) + \left(w_{j0} \right) \right) + w_{k0} \right], \tag{10}$$

where n is the number of input variables; m is the number of hidden neurons; $x_i(t)$ = the ith input variable at time step t_s; 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; $y'_k(t_s)$ is the forecasted kth output at time step t_s [26].

2.3. Support Vector Regression.
Support vector machines (SVM) were developed by Vapnik [27] as a tool for classification and regression. SVMs embody the structural risk minimization principle, while neural networks embody the empirical risk minimization principle. In contrast to ANNs that seek to minimize training error, SVMs attempt to minimize the generalization error. SVMs have two components: support vector classification (SVC) and support vector regression (SVR). Since the main objective of this study is to forecast the SPI, the SVR was used.

Support vector regression (SVR) is used to describe regression with SVMs [27]. In regression estimation with SVR, the purpose is to estimate a functional dependency $f(\vec{x})$ between a set of sampled points $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_l\}$ taken from R^n and target values $Y = \{y_1, y_2, \ldots, y_l\}$ 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). Assuming that these samples have been generated independently from an unknown probability distribution function $P(\vec{x}, y)$ and a class of functions [27]:

$$F = \left\{ f \mid f(\vec{x}) = \left(\vec{W}, \vec{x} \right) + B : \vec{W} \in R^n, R^n \longrightarrow R \right\}, \tag{11}$$

where $\vec{W}$ and B are coefficients that have to be estimated from the input data. The main objective is to find a function $f(\vec{x}) \in F$ that minimizes a risk functional [46]:

$$R[f(\vec{x})] = \int l(y - f(\vec{x}), \vec{x}) dP(\vec{x}, y), \tag{12}$$

where l is a loss function used to measure the deviation between the target, y, and estimate $f(\vec{x})$, values. As the probability distribution function $P(\vec{x}, y)$ is unknown, one cannot minimize the risk functional directly, but can only compute the empirical risk function as [46]

$$R_{\text{emp}}[f(\vec{x})] = \frac{1}{N} \sum_{i=1}^{N} l(y_1 - f(\vec{x}_i)), \tag{13}$$

where N is the number of samples. This traditional empirical risk minimization is not advisable without any means of structural control or regularization. To avoid this issue a regularized risk function with the smallest steepness among the functions that minimize the empirical risk function can be used as [46]

$$R_{\text{reg}}[f(\vec{x})] = R_{\text{emp}}[f(\vec{x})] + \gamma \left\| \vec{W} \right\|^2, \tag{14}$$

where γ is a constant ($\gamma \geq 0$). This additional term reduces the model space and thereby controls the complexity of the solution resulting in the following form of this expression [46, 47]:

$$R_{\text{reg}}[f(\vec{x})] = C_c \sum_{x_i \in X} l_\varepsilon (y_1 - f(\vec{x}_i)) + \frac{1}{2} \left\| \vec{W} \right\|^2, \tag{15}$$

where C_c is a positive constant that has to be selected beforehand. The constant C_c that influences a trade-off between or an approximation error and the regression (weight) vector $\| \vec{W} \|$ is a design parameter. The loss function in this expression, which is called an ε-insensitive loss function (l_ε), has the advantage that it will not need all the input data for describing the regression vector $\| \vec{W} \|$ and can be written as [46]

$$l_\varepsilon(y_1 - f(\vec{x}_i)) = \begin{cases} 0, & \text{for } |y_1 - f(\vec{x}_i)| < \varepsilon \\ y_1 - f(\vec{x}_i), & \text{otherwise.} \end{cases} \tag{16}$$

This function behaves as a biased estimator when it is combined with the regularization term ($\gamma \| \vec{W} \|^2$). The loss is equal to 0 if the difference between the predicted and observed value is less than ε. The nonlinear regression function is described by the following expression [27, 46, 48]:

$$f(x) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K(x, x_i) + B, \tag{17}$$

where $\alpha_i, \alpha_1^* \geq 0$ are the Lagrange multipliers, B is a bias term, and $K(x, x_i)$ is the Kernel function which is based upon Reproducing Kernel Hilbert Spaces [32]. The Kernel function enables operations to be performed in the input space as opposed to the potentially high-dimensional feature space. Several types of functions are treated by SVR such as polynomial functions, Gaussian radial basis functions, exponential radial basis functions, multilayer perception functions, and functions with splines and so forth [32].
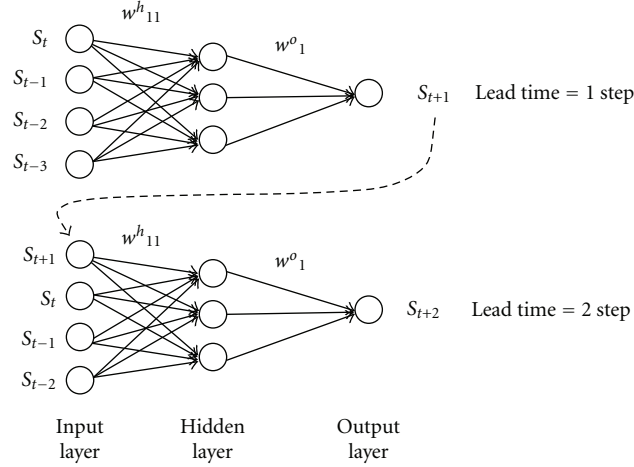
FIGURE 1: Typical Feed-forward Neural Network.

2.4. Wavelet Transforms. Wavelet transforms are mathematical functions that can be used for the analysis of time-series that contain nonstationarities. Wavelet transforms allow for the use of long time intervals for low frequency information and shorter intervals for high frequency information. They are capable of revealing aspects of data like trends, breakdown points, and discontinuities that other signal analysis techniques might miss [26]. Another advantage of wavelet analysis is the flexible choice of the mother wavelet according to the characteristics of the investigated time series [45].

An important step in the use of wavelet transforms is the choice of a mother wavelet (ψ). The continuous wavelet transform (CWT) is defined as the sum over all time of the signal multiplied by scale and shifted versions of the wavelet function ψ [26]:

$$W(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - \tau}{s} \right) dt, \quad (18)$$

where s is the scale parameter; τ is the translation and $*$ corresponds to the complex conjugate [26]. The CWT produces a continuum of all scales as the output. Each scale corresponds to the width of the wavelet; hence, a larger-scale means that more of a time series is used in the calculation of the coefficient than in smaller scales. The CWT is useful for processing different images and signals; however, it is not often used for forecasting because its computation is complex and time consuming. As an alternative, in forecasting applications, the discrete wavelet transform (DWT) is used, due to its simplicity and shorter computation time. DWT scales and positions are usually based on powers of two (dyadic scales and positions). This is achieved by modifying the wavelet representation to [49]

$$\psi_{j,m}(m) = \frac{1}{\sqrt{\left| s_0^j \right|}} \sum_k \psi \left(\frac{k - m\tau_0 s_0^j}{s_0^j} \right) x(k), \quad (19)$$

where j and m are integers that control the scale and translation, respectively, while $s_0 > 1$ is a fixed dilation step and

τ_0 is a translation factor that depends on the aforementioned dilation step. The effect of discretizing the wavelet is that the time-space scale is now sampled at discrete levels. The DWT operates two sets of functions: high-pass and low-pass filters. The original time series is passed through high-pass and low-pass filters, and detailed coefficients and approximation series are obtained.

One of the inherent challenges of using the DWT for forecasting applications is that if we change values at the beginning of our time series, all of the wavelet coefficients will subsequently change. To overcome this problem, a redundant algorithm, known as the à trous algorithm can be used, given by [50]

$$C_{i+1}(k) = \sum_{l=-\infty}^{+\infty} h(l) c_i \left(k + 2^i l \right), \quad (20)$$

where h is the low pass filter and the finest scale is the original time series. To extract the details, $w_i(k)$, that were eliminated in (21), the smoothed version of the signal is subtracted from the coarser signal that preceded it, given by [51]

$$w_i(k) = c_{i-1}(k) - c_i(k), \quad (21)$$

where $c_i(k)$ is the approximation of the signal and $c_{i-1}(k)$ is the coarser signal. Each application of (20) and (21) creates a smoother approximation and extracts a higher level of detail. Finally, the nonsymmetric Haar wavelet can be used as the low pass filter to prevent any future information from being used during the decomposition [52].

3. The Awash River Basin

This study forecasted the SPI in the Awash River Basin of Ethiopia. The mean annual rainfall of the basin varies from about 1,600 mm in the highlands north east of Addis Ababa, to 160 mm in the northern point of the basin [53]. 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 $4,900 \times 106\,\mathrm{m}^3$ [54].

The Awash River Basin (Figure 2) was separated into three smaller basins for the purpose of this study on the basis of various factors such as location, altitude, climate, topography, and agricultural development. A study conducted by Edossa et al. [54] separated the Awash Basin in a similar fashion. The subbasins were called the Upper, Middle, and Lower Awash Basins, respectively. The reasoning behind the use of these three subbasins was to ensure the methods used in this study were effective in forecasting short-term drought in different conditions. The characteristics of each sub-basin are briefly described in the following sections.

3.1. Upper Awash Basin. The Upper Awash Basin has a temperate climate with annual mean temperatures ranging between 15–22°C and an annual precipitation of between 500–2000 mm [54]. Rainfall distribution in the Upper Awash Basin is unimodal. Seven rainfall gauges located in the Upper Awash River Basin were chosen for this study (Table 2). These stations were chosen because their precipitation records from 1970–2005 were either complete or relatively complete. Any station, which had over 10% of their records missing was not selected.

3.2. Middle Awash Basin. The Middle Awash Basin is in the semiarid climatic zone with a long hot summer and a short mild winter. Annual rainfall varies between 200–1500 mm [54]. The rainfall distribution is bimodal in this subbasin. Minor rains normally occur in March and April and major rains from July to August. Eight rainfall gauges located in the Middle Awash Basin were selected using the same criteria as in the Upper Awash Basin and are shown in Table 2.

3.3. Lower Awash Basin. The Lower Awash River Basin has a hot, semi-arid climate. The annual mean temperature of the region ranges between 22 and 32°C with average annual precipitation between 500 and 700 mm [54]. Five rainfall gauges were selected form the Lower Awash Basin using the same criteria used in the two other sub-basins and are shown in Table 2.

4. Methodology

The methodology section of this paper describes how the SPI was calculated and then forecast over two separate lead times using ANN, WN, and SVR models.

4.1. SPI Calculation. In order to calculate the SPI, a probability density function that adequately describes the precipitation data must be determined. The gamma distribution function was selected to fit the raw rainfall data from each station in this study. 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. SPI values can be categorized according to classes. In this study, the near normal class is established from the aggregation of two classes: $-1 < \mathrm{SPI} < 0$ (mild drought)

and $0 \leq \mathrm{SPI} \leq 1$ (slightly wet). The departure from the mean is a probability indication of the severity of the wetness or drought that can be used for risk assessment. 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.

In each sub-basin, for each station, SPI 3 and SPI 12 were computed. These SPI values were subsequently forecast over lead times of 1 and 6 months. A 3-month SPI compares the precipitation for that period with the same 3-month period over the historical record. For example, a 3-month SPI at the end of September compares the precipitation total for the July–September period with all the past totals for that same period. A 3-month SPI indicates short and medium term trends in precipitation and is still considered to be more sensitive to conditions at this scale than the Palmer Index. A 3-month SPI can be very effective in showing seasonal trends in precipitation and is a good indicator of agricultural drought. SPI 12 reflects long-term precipitation patterns. SPI 12 is a comparison of the precipitation for 12 consecutive months with the same 12 consecutive months during all the previous years of available data and is a good indicator of long-term drought conditions. Because these time scales are the cumulative result of shorter periods that may be above or below normal, the longer SPIs tend toward zero unless a specific trend is taking place. Forecast lead times of 1 and 6 months were chosen because 1 month is the shortest possible monthly lead time and 6 months is representative of the bimodal rainfall pattern in parts of the Awash River Basin discussed in Section 3.2.

4.2. Wavelet Decomposition. In the proposed WN model, the SPI data for each of the rainfall stations was decomposed into subseries of approximations and details (DWs). The process consists of a number of successive filtering steps. The original SPI time series is first decomposed into an approximation and accompanying detail signal. The decomposition process is then iterated, with successive approximation signals being decomposed in turn. As a result the original SPI time series is broken down into many lower resolution components.

When conducting wavelet analysis, the number of decomposition levels that is appropriate for the data must be chosen. A commonly used method to determine the number of decomposition levels is based on the signal length [55] and is given by $L = \mathrm{int}[\log(N)]$, where L is the level of decomposition and N is the length of the signal. The training set in this study comprised between 1290 and 3017 samples (samples varied depending on the number of inputs for each rainfall station). Thus, the decomposition level was selected as $L = 3$.

As discussed in Section 2.4, the "a trous" wavelet algorithm with a low pass Haar filter was used to create four

TABLE 2: Descriptive statistics for the Awash River Basin.

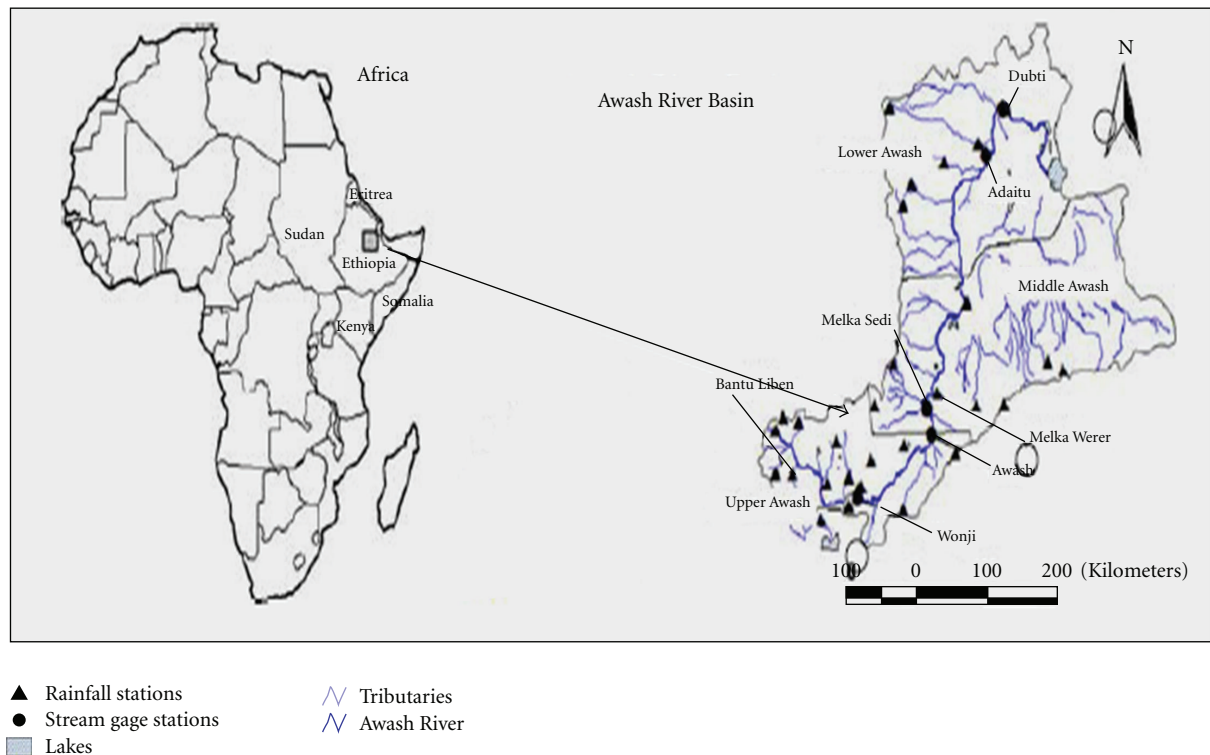| Basin	Station	Mean annual precipitation (mm)	Max annual (1970–2005) precipitation (mm)	Standard deviation (mm)
Upper Awash Basin	Bantu Liben	91	647	111
	Tullo Bullo	94	575	114
	Ginchi	97	376	90
	Sebeta	111	1566	172
	Ejersalele	67	355	75
	Ziquala	100	583	110
	Debre Zeit	73	382	81
Middle Awash Basin	Koka	97	376	90
	Modjo	76	542	92
	Nazereth	73	470	85
	Wolenchiti	76	836	95
	Gelemsso	77	448	75
	Hirna	78	459	86
	Dire Dawa	51	267	54
	Meisso	61	361	61
Lower Awash Basin	Dubti	15	192	23
	Eliwuha	44	374	57
	Mersa	87	449	89
	Mille	26	268	40
	Bati	73	357	80

FIGURE 2: Awash River Basin (Source: [54]).

sets of wavelet subseries. These four sub-series included a low frequency component (the approximation) used to uncover the trend of each signal and a set of three high frequency components (the details) used to uncover the periodicity of the signal. All decomposed sub-series were added together to generate one time series and used as an input to the ANN models. Using the sum of all the sub-series as an input in this study provided more accurate results than using certain sub-series or sub-series that exhibited the highest correlations with the original time series.

4.3. ANN Models. All the ANN models were created with the MATLAB (R.2010a) ANN toolbox. The hyperbolic tangent sigmoid transfer function was the activation function for the hidden layer, while the activation function for the output layer was a linear function. All the ANN models in this study were trained using the LM back propagation algorithm. The LM back propagation algorithm was chosen because of its efficiency and reduced computational time in training models [45].

In this study, there were between 4–8 input neurons for each ANN model. The optimal number of input neurons for each station was selected using a trial and error procedure. The data-driven models were recursive models, where a model is forecast one lead time ahead, and the subsequent forecasts include the output from the previous forecast as an input. Hence, a forecast of 6 months lead time will have the outputs from forecasts of lead times of 1–5 months. Recursive models were used because it was determined that it would be simpler to use an ANN with one output neuron. Mishra and Desai [10] compared recursive ANN models and ANN models with more than one output neuron (direct ANN models) and found the results to be comparable for forecasting the SPI. The inputs and outputs were normalized between 0 and 1. A study by Wanas et al. [56] empirically determined that the best performance of a neural network occurs when the number of hidden nodes is equal to $\log(T)$, where T is the number of training samples. Another study conducted by Mishra and Desai [10] determined that the optimal number of hidden neurons is $2n + 1$, where n is the number of input layers. In this study the optimal number of hidden neurons was determined to be between $\log(T)$ and $(2n + 1)$. For example, if using the method proposed by Wanas et al. [56] gave a result of 4 hidden neurons and using the method proposed by Mishra and Desai [10] gave 6 hidden neurons, the optimal number of hidden neurons was between 4 and 6, thereafter the optimal number was determined using trial and error. These two methods helped establish an upper and lower bound for the number of hidden neurons.

For all the ANN models the cross validation technique [57] was used to partition the data sets; 80% of the data was used to train the models, while the remaining 20% of the data was used to test and validate the models, with 10% used for testing and 10% used for validation. The training set was used to compute the error gradient and to update the network weights and biases. The error from the validation set was used to monitor the training process. If the network overfits the data, the error in the validation set will begin to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned. The testing data set is an independent data set and is used to verify the performance of the model.

4.4. WN Models. The WN models were trained in the same way as the ANN models, with the exception that the inputs were made up from the wavelet decomposed subseries. In this study, the significant wavelets (approximation and detail series) were summed together once the insignificant coefficients were excluded, similar to what was done by Partal [58] and Kisi and Cimen [32]. In this study, the summed sub-series provided better results than using the individual wavelet coefficients as inputs.

For WN models, an input layer with 4–8 neurons, a single hidden layer composed of 4–6 neurons, and one output layer consisting of one neuron were developed. The number of neurons was determined in the same way as for the traditional ANN models. All the ANN models that had wavelet decomposed subseries as their inputs were also partitioned in a similar manner to the traditional ANN models.

4.5. SVR Models. All SVR models were developed using the OnlineSVR software created by Parrella [59]. OnlineSVR is a technique used to build support vector machines for regression. The OnlineSVR software partitions the data into only two sets: a training set and a testing set. The SVR models were partitioned in a similar manner to the ANN and WN models.

All SVR models used the nonlinear 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, 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 [32]. These three parameters were selected based on a trial and error procedure. The combination of parameters that produced the lowest RMSE values for the training data sets was selected.

4.6. Performance Measures. The performance of the forecasts resulting from the data-driven models was evaluated by the following measures of goodness of fit:

$$\text{The coefficient of determination } (R^2) = \frac{\sum_{i=1}^{N}\left(\hat{y}_i - \overline{y}_i\right)}{\sum_{i=1}^{N}\left(y_i - \overline{y}_i\right)^2},$$

$$\overline{y}_i = \frac{1}{N}\sum_{i=1}^{N} y_i,$$

$$(22)$$

where $\overline{y}_i$ 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. The higher the value of R^2 (with 1 being the highest possible value), the better the performance of the model

$$\text{The Root Mean Squared Error (RMSE)} = \sqrt{\frac{\text{SSE}}{N}}, \quad (23)$$

where SSE is the sum of squared errors and N is the number of data points used. SSE is given by

$$\text{SSE} = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (24)$$

with the variables already having been defined. The RMSE evaluates the variance of errors independently of the sample size

$$\text{The Mean Absolute Error (MAE)} = \sum_{i=1}^{N} \frac{|y_i - \hat{y}_i|}{N}. \quad (25)$$

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

5. Results and Discussion

For each subbasin of the Awash River Basin, the station that showed the best performance results for each data driven model are presented below. In this study, SPI 3 and SPI 12 were forecast over lead times of 1 and 6 months to determine the effectiveness of the data-driven models over short- and long-term lead times.

As shown in Table 3(a), the best data-driven model in the Upper Awash Basin for forecasts of SPI 3 and 12 is the WN model. All the models exhibited better results for forecasts of a 1-month lead time (L1) compared to forecasts of 6-months lead time (L6). Forecasts of SPI 12, for all the data-driven models, had better performance results than forecasts of SPI 3 in terms of R^2, RMSE, and MAE, regardless of forecast lead time. The best 1-month lead time WN forecast of SPI 12 had results of 0.9534, 0.0600, and 0.0536 in terms of R^2, RMSE, and MAE, respectively. The second best results were from ANN models with results of 0.9451, 0.0610, and 0.0603 in terms of R^2, RMSE and MAE, respectively. Figures 3 and 4 show the ANN and WN 1-month forecast results for SPI 12 at the Ejersalele station.

The performance of both these models is quite similar, as indicated by Figures 3 and 4. Both models adequately represent the periods of abundant and acute precipitation as indicated by the peaks and valleys in the figures.

Similar to the results for the Upper Awash Basin, the best forecast results in the Middle Awash Basin were from WN models. The WN models had the best results for both SPI 3 and SPI 12, for forecast lead times of 1 and 6 months, respectively (Table 3(b)). The forecast results of all the data-driven models deteriorated when the forecast lead time was increased from 1 to 6 months.

Figure 5 illustrates the relationship between the observed SPI 12 and the predicted SPI 12 from the ANN model at
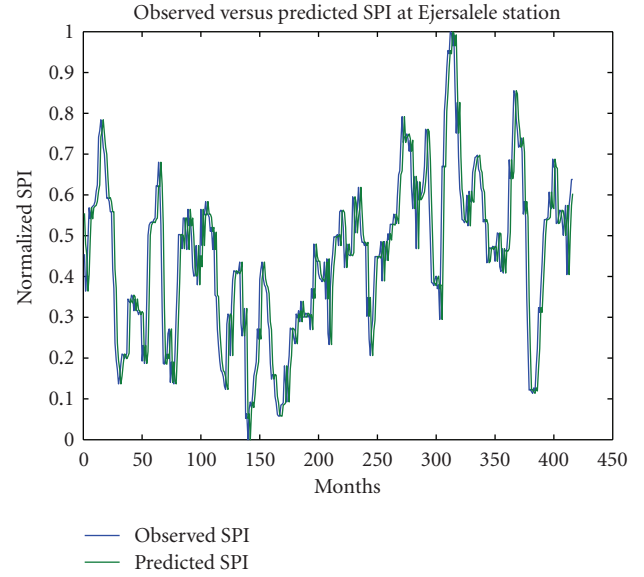

FIGURE 3: SPI 12 forecast results for the best ANN model at the Ejersalele station (1-month lead time).
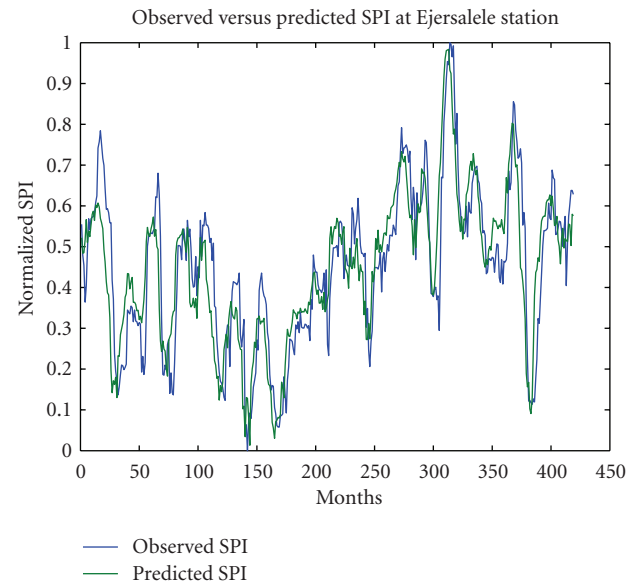

FIGURE 4: SPI 12 forecast results for the best WA-ANN model at the Ejersalele station (1-month lead time).

the Nazereth station. The ANN model underestimates the severity of the drought period at 112 months. In contrast, the WN model for SPI 12 at the Nazereth station displays improved results with respect to the drought period at 112 months (Figure 6).

In the Lower Awash Basin, the forecast results exhibited the same trend shown in the Upper and Middle sub-basins. The WN models had the best results for both SPI 3 and SPI 12, for forecast lead times of 1 and 6 months, respectively. Figures 7 and 8 illustrate the best SPI 12 forecasts at the Dubti station where both ANN and WN models predict

TABLE 3: (a) Performance results for the Ejersalele station, Upper Awash Basin, (b) Performance results of Nazereth Station, Middle Awash Basin, (c) Performance results of Dubti Station, Lower Awash Basin.

(a)

Model-Lead time	SPI 3			SPI 12		
	R^2	RMSE	MAE	R^2	RMSE	MAE
ANN-L1	0.7694	0.1574	0.1433	0.9451	0.0610	0.0603
ANN-L6	0.6232	0.1744	0.1567	0.8614	0.1011	0.0885
WN-L1	0.8829	0.0700	0.0352	0.9534	0.0600	0.0536
WN-L6	0.6433	0.1070	0.0356	0.8731	0.0790	0.0662
SVR-L1	0.7219	0.1046	0.0915	0.7611	0.1312	0.1129
SVR-L6	0.6647	0.1118	0.1042	0.6941	0.1341	0.1247

(b)

Model-Lead time	SPI 3			SPI 12		
	R^2	RMSE	MAE	R^2	RMSE	MAE
ANN-L1	0.7319	0.1170	0.1016	0.9158	0.1003	0.0911
ANN-L6	0.6546	0.1240	0.1142	0.7542	0.1104	0.0919
WN-L1	0.9483	0.0510	0.0441	0.9167	0.0753	0.0629
WN-L6	0.8641	0.0727	0.0512	0.8012	0.1072	0.0802
SVR-L1	0.7114	0.1216	0.1114	0.7713	0.1147	0.1130
SVR-L6	0.6540	0.1320	0.1217	0.7326	0.1244	0.1215

(c)

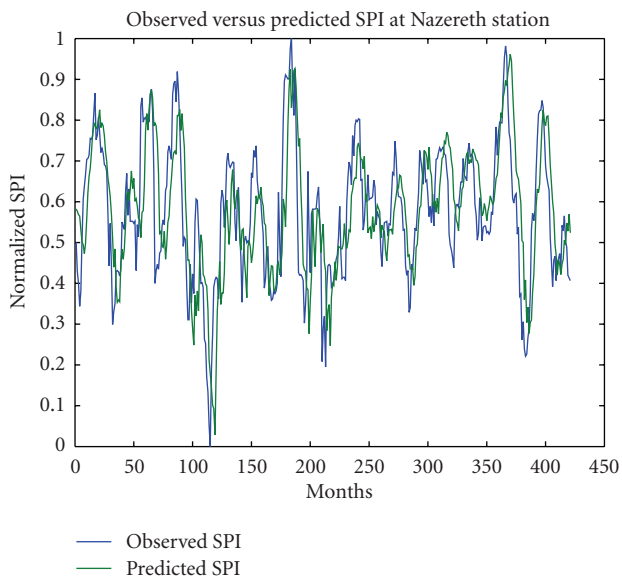| Model-Lead time	SPI 3			SPI 12		
	R^2	RMSE	MAE	R^2	RMSE	MAE
ANN-L1	0.7368	0.1175	0.1095	0.9188	0.0710	0.0648
ANN-L6	0.6806	0.1302	0.1147	0.7135	0.0938	0.0836
WN-L1	0.9018	0.0652	0.0581	0.9473	0.0648	0.0560
WN-L6	0.8119	0.0706	0.0642	0.8641	0.0846	0.0747
SVR-L1	0.6990	0.1146	0.1022	0.7041	0.1102	0.1009
SVR-L6	0.6331	0.1309	0.1242	0.6705	0.1107	0.1025

FIGURE 5: SPI 12 forecast results for the best ANN model at the Nazereth station (1 month lead time).
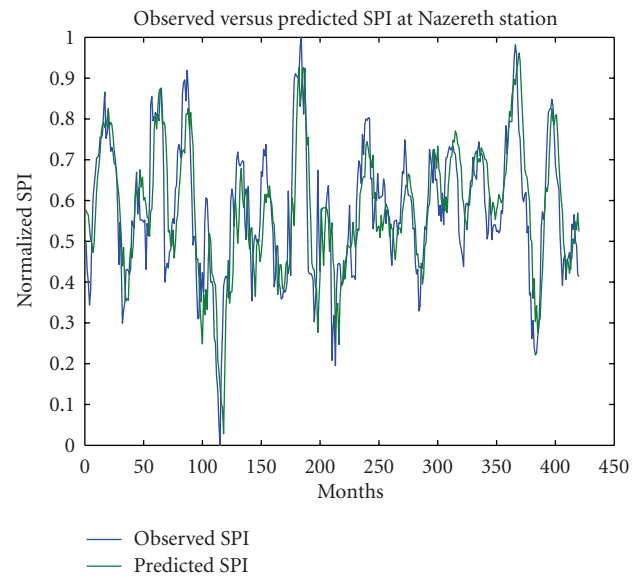

FIGURE 6: SPI 12 forecast results for the best WN model at the Nazereth station (1-month lead time).
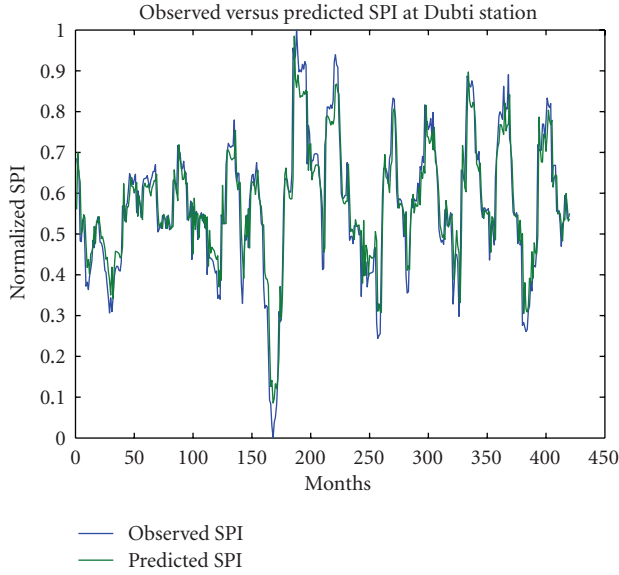
FIGURE 7: SPI 12 forecast results for the best ANN model at the Dubti station (1-month lead time).

FIGURE 8: SPI 12 forecast results for the best WN model at the Dubti station (1-month lead time).

the periods of abundant and acute precipitation quite well. When the forecast lead time was increased, the performance of all the models deteriorated, especially with respect to R^2. Data-driven models in the Upper and Lower Awash basins exhibited their best results for forecasts of SPI 12, indicating that data-driven models are more effective in predicting long-term drought conditions in those two basins, while in the Middle Awash Basin most models also exhibited their best results for forecasts of SPI 12 except WN models, which exhibited their best results for forecasts of SPI 3. This trend could be due to the fact that long-term SPI, which is a cumulative of short-term time scales, tend toward zero unless a specific trend is taking place. The exception regarding the WN models in the Middle Awash Basin may be due to the fact that the precipitation record at this station is relatively stable, meaning there are not many changes from one month to the next and the SPI 3 is not sensitive to those changes.

Overall, all three data-driven models forecast SPI 3 and SPI 12 well for forecast lead times of 1 and 6 months. The results indicate that ANN models are more effective than SVR models at forecasting in this study. The use of wavelet analysis improved the forecast results of ANN models, specifically in predicting extreme events as shown in Figure 6. Indeed, using a measure for peak relative error as shown by

$$Z = 100 \left| \frac{q_s(\text{peak}) - q_o(\text{peak})}{q_o} \right|, \qquad (26)$$

it was determined that the relative error of the ANN model, 95%, was reduced to 88% when a WN model was used.

The fact that wavelet analysis is an effective tool at revealing local discontinuities helps explain why it was more effective in predicting the extreme events in the Middle Awash Basin. Wavelet analysis may help de-noise the original
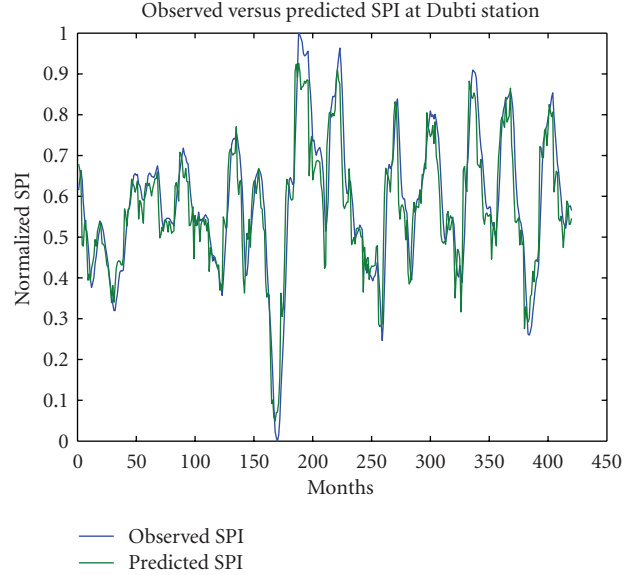
SPI time series compared to a traditional ANN model. The forecast of this de-noised signal may further explain the fact that extreme events are forecast better using wavelet analysis.

An increase in forecast lead time results in a deterioration of performance in all the models. However, this deterioration does not result in poor models, indicating the stability of these data-driven models in predicting the SPI. The results in terms of RMSE and MAE do not deteriorate drastically with an increase of lead time. For example, for the Dubti station, the RMSE and MAE of SVR models deteriorate by 0.05 and 0.26%, respectively.

There is variability with regards to the best forecasts of both SPI 3 and SPI 12 amongst the three subbasins. For example, the best forecast of SPI 3 at a 1-month lead time occurred in the Middle Awash Basin (WN model), while the best forecast of SPI 12 at a 1-month lead time occurred in the Upper Awash Basin (WN model). While each subbasin has a different climatology, there does not seem to be a clear trend linking climatology with forecast accuracy. It seems that the reason behind the best models for each data-driven method being in various subbasins is linked with the characteristics of the individual station and not the characteristics of the subbasin as a whole.

In addition, the forecast results for SPI 12 are better than the forecast results for SPI 3 in almost all cases. For SPI 3 and other short-term SPI, each new month has a large impact on the period sum of precipitation [6]. As a result, the SPI 3 is sensitive to any change in precipitation from one month to another. In the case of SPI 12, each individual month has less impact on the total and the index is not as sensitive to changes in precipitation from one month to the next. The fact that SPI 3 is more sensitive to changes in precipitation results in less accurate forecast results than SPI 12. However, the effects of wavelet analysis are more significant for SPI 3 than for SPI 12, especially for forecast lead times of 6

months. As stated previously, the ANN forecasts of SPI 12 are not as sensitive to changes in precipitation and thus good results are obtained. The ability of wavelet analysis to improve these results exists as shown but is not as high as the improvement seen in SPI 3 forecasts because ANN forecasts of SPI 3 suffer due to the sensitivity of SPI 3 to slight changes in precipitation over the long-term record.

All three subbasins had a different climatology. The forecast results have all shown that WN models are the most effective at forecasting the SPI in all the sub-basins in terms of R^2, RMSE and MAE. Whether this is the case in all climatic zones needs to be explored in future studies.

6. Conclusion

This study tried to determine the most effective data-driven model for forecasts of the SPI drought index in the Awash River Basin of Ethiopia. WN models were shown to be the most effective model for forecasts of SPI 3 and 12 in all three subbasins. WN models showed greater correlation between observed and predicted SPI compared to simple ANNs and SVR models. WN models also consistently showed lower values of RMSE and MAE compared to the other data driven models explored in this study. All the data-driven models showed increased forecast results for SPI 12 compared to SPI 3. Forecast results deteriorated as the forecast lead time increased for all the models. Of the two machine learning techniques, ANNs are more effective in forecasting the SPI compared to SVR models. This trend occurs in all three subbasins and should be studied in other regions to determine if ANNs are more effective tools for drought forecasting compared to SVR models. It is thought that WN models provide more accurate results because preprocessing the original SPI time series with wavelet decompositions "denoises" the data. Future studies should attempt to explore WSVR models, ensemble WN and WSVR models, and explore SPI forecasts using these new methods in other regions with different characteristics. Future studies should also attempt to quantify time shift error as it is a part of forecasting problems with regression models.

Acknowledgments

References

[1] E. Mersha and V. K. Boken, "Agricultural drought in Ethiopia," in *Monitoring and Predicting Agricultural Drought: A Global Study*, V. K. Boken, A. P. Cracknell, and R. L. Heathcote, Eds., Oxford University Press, 2005.

[2] A. K. Mishra and V. P. Singh, "A review of drought concepts," *Journal of Hydrology*, vol. 391, no. 1-2, pp. 202–216, 2010.

[3] T. Ross and N. Lott, "A climatology of 1980–2003 extreme weather and climate events," National Climatic Data Center Technical Report No. 2003-01. NOAA/ NESDIS, National Climatic Data Center, Asheville, NC, USA.

[4] A. Cancelliere, G. di Mauro, B. Bonaccorso, and G. Rossi, "Stochastic forecasting of drought indices," in *Methods and Tools For Drought Analysis and Management*, G. Rossi, T. Vega, and B. Bonaccorso, Eds., Springer, 2007.

[5] W. J. Gibbs and J. V. Maher, *Rainfall Deciles as Drought Indicators*, vol. 48 of *Bulletin (Commonwealth Bureau of Meteorology, Australia)*, Bureau of Meteorology, Melbourne, Australia, 1967.

[6] T. B. McKee, N. J. Doesken, and J. Kleist, "The relationship of drought frequency and duration to time scales," in *Proceedings of the 8th Conference on Applied Climatology*, American Meteorological Society, Anaheim, Calif, USA, 1993.

[7] H. R. Byun and D. A. Wilhite, "Objective quantification of drought severity and duration," *Journal of Climate*, vol. 12, no. 9, pp. 2747–2756, 1999.

[8] W. Palmer, "Meteorological drought," Tech. Rep. 45, U.S. Weather Bureau, Washington, DC, USA, 1965.

[9] H. K. Ntale and T. Y. Gan, "Drought indices and their application to East Africa," *International Journal of Climatology*, vol. 23, no. 11, pp. 1335–1357, 2003.

[10] A. K. Mishra and V. R. Desai, "Drought forecasting using feed-forward recursive neural network," *Ecological Modelling*, vol. 198, no. 1-2, pp. 127–138, 2006.

[11] S. Morid, V. Smakhtin, and K. Bagherzadeh, "Drought fore-casting using artificial neural networks and time series of drought indices," *International Journal of Climatology*, vol. 27, no. 15, pp. 2103–2111, 2007.

[12] U. G. Bacanli, M. Firat, and F. Dikbas, "Adaptive Neuro-Fuzzy inference system for drought forecasting," *Stochastic Environmental Research and Risk Assessment*, vol. 23, no. 8, pp. 1143–1154, 2009.

[13] A. P. Barros and G. J. Bowden, "Toward long-lead operational forecasts of drought: an experimental study in the Murray-Darling River Basin," *Journal of Hydrology*, vol. 357, no. 3-4, pp. 349–367, 2008.

[14] P. Cutore, G. Di Mauro, and A. Cancelliere, "Forecasting palmer index using neural networks and climatic indexes," *Journal of Hydrologic Engineering*, vol. 14, no. 6, pp. 588–595, 2009.

[15] M. Karamouz, K. Rasouli, and S. Nazif, "Development of a hybrid Index for drought prediction: case study," *Journal of Hydrologic Engineering*, vol. 14, no. 6, pp. 617–627, 2009.

[16] A. F. Marj and A. M. J. Meijerink, "Agricultural drought forecasting using satellite images, climate indices and artificial neural network," *International Journal of Remote Sensing*, vol. 32, no. 24, pp. 9707–9719, 2011.

[17] D. Labat, R. Ababou, and A. Mangin, "Wavelet analysis in karstic hydrology. 2nd part: rainfall-runoff cross-wavelet analysis," *Comptes Rendus de l'Academie de Sciences*, vol. 329, no. 12, pp. 881–887, 1999.

[18] P. Saco and P. Kumar, "Coherent modes in multiscale vari-ability of streamflow over the United States," *Water Resources Research*, vol. 36, no. 4, pp. 1049–1067, 2000.

[19] L. C. Smith, D. L. Turcotte, and B. L. Isacks, "Stream flow characterization and feature detection using a discrete wavelet transform," *Hydrological Processes*, vol. 12, no. 2, pp. 233–249, 1998.

[20] P. Coulibaly, F. Anctil, and B. Bobée, "Daily reservoir inflow forecasting using artificial neural networks with stopped training approach," *Journal of Hydrology*, vol. 230, no. 3-4, pp. 244–257, 2000.

[21] S. N. Lane, "Assessment of rainfall-runoff models based upon wavelet analysis," *Hydrological Processes*, vol. 21, no. 5, pp. 586–607, 2007.

[22] J. F. Adamowski, "Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis," *Journal of Hydrology*, vol. 353, no. 3-4, pp. 247–266, 2008.

[23] J. Adamowski and K. Sun, "Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds," *Journal of Hydrology*, vol. 390, no. 1-2, pp. 85–91, 2010.

[24] M. Özger, A. K. Mishra, and V. P. Singh, "Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas," *Journal of Hydrometeorology*, vol. 13, no. 1, pp. 284–297, 2012.

[25] T. Partal and Ö. Kişi, "Wavelet and neuro-fuzzy conjunction model for precipitation forecasting," *Journal of Hydrology*, vol. 342, no. 1-2, pp. 199–212, 2007.

[26] T. W. Kim and J. B. Valdes, "Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks," *Journal of Hydrologic Engineering*, vol. 8, no. 6, pp. 319–328, 2003.

[27] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.

[28] J. B. Gao, S. R. Gunn, C. J. Harris, and M. Brown, "A probabilistic framework for SVM regression and error bar estimation," *Machine Learning*, vol. 46, no. 1–3, pp. 71–89, 2002.

[29] M. S. Khan and P. Coulibaly, "Application of support vector machine in lake water level prediction," *Journal of Hydrologic Engineering*, vol. 11, no. 3, pp. 199–205, 2006.

[30] S. Rajasekaran, S. Gayathri, and T.-L. Lee, "Support vector regression methodology for storm surge predictions," *Journal of Ocean Engineering*, vol. 35, no. 16, pp. 1578–1587, 2008.

[31] O. Kisi and M. Cimen, "Evapotranspiration modelling using support vector machines," *Hydrological Sciences Journal*, vol. 54, no. 5, pp. 918–928, 2009.

[32] O. Kisi and M. Cimen, "A wavelet-support vector machine conjunction model for monthly streamflow forecasting," *Journal of Hydrology*, vol. 399, no. 1-2, pp. 132–140, 2011.

[33] T. Asefa, M. Kemblowski, M. McKee, and A. Khalil, "Multi-time scale stream flow predictions: the support vector machines approach," *Journal of Hydrology*, vol. 318, no. 1–4, pp. 7–16, 2006.

[34] W. C. Wang, K. W. Chau, C. T. Cheng, and L. Qiu, "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series," *Journal of Hydrology*, vol. 374, no. 3-4, pp. 294–306, 2009.

[35] R. Maity, P. P. Bhagwat, and A. Bhatnagar, "Potential of support vector regression for prediction of monthly streamflow using endogenous property," *Hydrological Processes*, vol. 24, no. 7, pp. 917–923, 2010.

[36] Z. M. Yuan and X. S. Tan, "Nonlinear screening indicators of drought resistance at seedling stage of rice based on support vector machine," *Acta Agronomica Sinica*, vol. 36, no. 7, pp. 1176–1182, 2010.

[37] C. Cacciamani, A. Morgillo, S. Marchesi, and V. Pavan, "Monitoring and forecasting drought on a regional scale: emilia-romagna region," *Water Science and Technology Library*, vol. 62, part 1, pp. 29–48, 2007.

[38] I. Bordi and A. Sutera, "Drought monitoring and forecasting at large-scale," in *Methods and Tools For Drought Analysis and Management*, G. Rossi, T. Vega, and B. Bonaccorso, Eds., pp. 3–27, Springer, New York, NY, USA, 2007.

[39] N. B. Guttman, "Accepting the standardized precipitation index: a calculation algorithm," *Journal of the American Water Resources Association*, vol. 35, no. 2, pp. 311–322, 1999.

[40] H. C. S. Thom, "A note on gamma distribution," *Monthly Weather Review*, vol. 86, pp. 117–122, 1958.

[41] D. C. Edwards and T. B. McKee, "Characteristics of 20th century drought in the United States at multiple scales," Atmospheric Science Paper 634, 1997.

[42] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences an Introduction*, Academic Press, San Diego, Calif, USA, 1995.

[43] M. Abramowitz and A. Stegun, Eds., *Handbook of Mathematical Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, NY, USA, 1965.

[44] S. Morid, V. Smakhtin, and M. Moghaddasi, "Comparison of seven meteorological indices for drought monitoring in Iran," *International Journal of Climatology*, vol. 26, no. 7, pp. 971–985, 2006.

[45] J. Adamowski and H. F. Chan, "A wavelet neural network conjunction model for groundwater level forecasting," *Journal of Hydrology*, vol. 407, no. 1–4, pp. 28–40, 2011.

[46] M. Çimen, "Estimation of daily suspended sediments using support vector machines," *Hydrological Sciences Journal*, vol. 53, no. 3, pp. 656–666, 2008.

[47] A. J. Smola, *Regression Estimation with Support Vector Learning Machines [M.S. thesis]*, Technische Universitat Munchen, Munich, Germany, 1996.

[48] S. Gunn, "Support vector machines for classification and regression," ISIS Technical Report, Department of Electronics and Computer Science, University of Southampton, 1998.

[49] B. Cannas, A. Fanni, G. Sias, S. Tronci, and M. K. Zedda, "River flow forecasting using neural networks and wavelet analysis," in *Proceedings of the European Geosciences Union*, 2006.

[50] S. G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 1998.

[51] F. Murtagh, J. L. Starck, and O. Renuad, "On neuro-wavelet modeling," *Decision Support Systems*, vol. 37, no. 4, pp. 475–484, 2004.

[52] O. Renaud, J. Starck, and F. Murtagh, *Wavelet-Based Forecasting of Short and Long Memory Time Series*, Department of Economics, University of Geneve, 2002.

[53] C. E. Desalegn, M. S. Babel, A. Das Gupta, B. A. Seleshi, and D. Merrey, "Farmers' perception of water management under drought conditions in the upper Awash Basin, Ethiopia," *International Journal of Water Resources Development*, vol. 22, no. 4, pp. 589–602, 2006.

[54] D. C. Edossa, M. S. Babel, and A. D. Gupta, "Drought analysis in the Awash River Basin, Ethiopia," *Water Resources Management*, vol. 24, no. 7, pp. 1441–1460, 2010.

[55] M. K. Tiwari and C. Chatterjee, "Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach," *Journal of Hydrology*, vol. 394, no. 3-4, pp. 458–470, 2010.

[56] N. Wanas, G. Auda, M. S. Kamel, and F. Karray, "On the optimal number of hidden nodes in a neural network," in *Proceedings of the 11th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE '98)*, pp. 918–921, May 1998.

[57] J. C. Principe, N. R. Euliano, and W. Curt Lefebvre, *Neural and Adaptive Systems*, John Wiley & Sons, 2000.

[58] T. Partal, "Modelling evapotranspiration using discrete wavelet transform and neural networks," *Hydrological Processes*, vol. 23, no. 25, pp. 3545–3555, 2009.

[59] F. Parrella, *Online support vector regression [M.S. thesis]*, University of Genoa, 2007.