Journal of Hydrology 476 (2013) 228–243

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

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

Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling

Vahid Nourani [a,b,c,*,1], Aida Hosseini Baghanam [d,a,1], Jan Adamowski [d], Mekonnen Gebremichael [e]

[a] Dept. of Water Resources Eng., Faculty of Civil Eng., Univ. of Tabriz, Iran
[b] St. Anthony Falls Lab., Univ. of Minnesota, USA
[c] NCED, Dept. of Civil Eng., Univ. of Minnesota, USA
[d] Dept. of Bioresource Engineering, McGill Univ., Quebec, Canada
[e] Dept. of Civil and Environmental Eng., Univ. of Connecticut, Storrs, CT, USA

ARTICLE INFO

Article history:
Received 3 July 2012
Received in revised form 16 September 2012
Accepted 25 October 2012
Available online 16 November 2012
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Fi-John Chang, Associate Editor

Keywords:
Rainfall–runoff modeling
Self-organizing map
Wavelet
Feed-forward neural network
Satellite data
Gilgel Abay watershed

SUMMARY

In this paper, a two-level self-organizing map (SOM) clustering technique was used to identify spatially homogeneous clusters of precipitation satellite data, and to choose the most operative and effective data for a feed-forward neural network (FFNN) to model rainfall–runoff process on a daily and multi-step ahead time scale. The wavelet transform (WT) was also used to extract dynamic and multi-scale features of the non-stationary runoff time series and to remove noise. The performance of the coupled SOM–FFNN model was compared to the newly proposed combined SOM–WT–FFNN model. The performance of these two models was also compared to that of a conventional forecasting method, namely the auto regressive integrated moving average with exogenous input (ARIMAX) model. Daily precipitation data from two satellites and four rain gauges, as well as runoff values recorded from January 2003 to December 2007 in the Gilgel Abay watershed in Ethiopia were used to calibrate and validate the models. Runoff predictions via all of the above methods were investigated for both single-step-ahead and multi-step-ahead lead times. The results indicated that the use of spatial and temporal pre-processed data in the FFNN model led to a promising improvement in its performance for rainfall–runoff forecasting. In the validation phase of single and multi-step-ahead forecasting, it was determined that the SOM–WT–FFNN models provide more accurate forecasts than the SOM–FFNN models (the determination coefficients for validation of the SOM–FFNN and SOM–WT–FFNN models were 0.80 and 0.93, respectively). The proposed FFNN model coupled with the SOM clustering method decreased the dimensionality of the input variables and consequently the complexity of the FFNN models. On the other hand, the application of the wavelet transform to the runoff data increased the performance of the FFNN rainfall–runoff models in predicting runoff peak values by removing noise and revealing the dominant periods.

1. Introduction

Accurate models of the rainfall–runoff process that are embedded with high complexity, non-stationarity, and non-linearity in both spatial and temporal scales can provide important information for watershed management, including for environmental engineering and management of land use, flooding, and water resources. In light of this, numerous models have been developed to simulate this complex process. Due to the large number of ambiguous physical parameters in the rainfall–runoff process, black box (lumped) models are frequently used, since they may have some advantages over fully distributed models (Nourani and Mano, 2007). Conventional black box time series models such as auto regressive integrated moving average (ARIMA), seasonal ARIMA, ARIMA with exogenous input (ARIMAX) and multiple linear regression (MLR), are widely applied to forecast hydrological time series (e.g. Adamowski et al., 2012; Cleaveland and Stahle, 1989; Graumlich, 1987; Hansen and Nelson, 1997; Nourani et al., 2011; Pulido-Calvo and Portela, 2007; Salas et al., 1980; Tankersley et al., 1993; Zhang, 2003; Zhang et al., 2011). These models are linear and assume stationarity of the dataset. Although these models may sometimes be inaccurate because of their inability to handle non-stationarity and non-linearity, such conventional methods are still used both in practice because they are simple to use, and

* Corresponding author at: Dept. of Water Resources Eng., Faculty of Civil Eng., Univ. of Tabriz, Iran. Tel.: +98 914 403 0332; fax: +98 411 334 4287.
E-mail addresses: vnourani@umn.edu, vnourani@yahoo.com (V. Nourani), hosseinibaghanam@gmail.com (A.H. Baghanam), jan.adamowski@mcgill.ca (J. Adamowski), mekonnen@engr.uconn.edu (M. Gebremichael).
[1] Formely.

0022-1694/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2012.10.054

because they can be used as 'comparison models' to evaluate newer methods.

During recent decades, artificial neural networks (ANNs) as black box models have increasingly been of interest for modeling complex non-linear hydrological processes such as rainfall–runoff. ANNs are effective for handling high dimensional datasets with non-linear and noisy characteristics, especially in cases where the underlying physical relationships are not clear. Over the course of the last 15 years, one can find numerous examples of rainfall–runoff modeling studies using ANNs (e.g. Antar et al., 2006; Dawson and Wilby, 1998; Hsu et al., 1995; Jain et al., 2004; Lallahem and Maina, 2003; Nourani et al., 2009a, 2011; Sajikumara and Thandaveswara, 1999; Senthil Kumar et al., 2004; Sudheer et al., 2000; Tokar and Johnson, 1999).

One of the concerns regarding the use of ANNs in rainfall–runoff modeling is whether they are capable of multi-step-ahead forecasting, which is important in reducing flood damages. Thus, several researches have been completed to investigate multi-step-ahead predictions using ANNs. For example, Vos and Rientjes (2005) utilized multi-step-ahead discharge predictions to assess the constraints of ANNs for rainfall–runoff modeling. Chang et al. (2007) developed multi-step-ahead ANN models for flood forecasting, where they examined multi-input multi-output (MIMO), multi-input single-output (MISO) and serial-propagated structures and concluded that MISO could lead to more reliable results. More recently, Yonaba et al. (2010) applied multi-step-ahead ANNs to evaluate sigmoid transfer functions in stream flow forecasting.

The performance of an ANN-based rainfall–runoff model relies on the quality and quantity of the input data. Rainfall is conventionally measured using rain gauges, but this method only allows relatively sparse sampling of a watershed. Development in satellite remote-sensing technology has enabled the global distribution of rainfall to be monitored, which may provide a good basis for improving the accuracy of hydrologic predictions (Lábó, 2012). Satellite data have been utilized in a small number of ANN-based rainfall–runoff modeling studies. For example, Artan et al. (2007) established the adequacy of satellite-derived rainfall data for flood and stream flow modeling using a spatially distributed hydrologic model. The results demonstrated the usefulness of remotely sensed precipitation data for hydrologic modeling. Satellite-based data were used to estimate stream flow by Akhtar et al. (2009), Sawunyama and Hunghes (2010), and Shrestha et al. (2008). For example, because of the sparseness of ground-based observations, Akhtar et al. (2009) used satellite precipitation data in an ANN model to forecast river flow after pre-processing by flow length and travel time. It was concluded that incorporating remote sensing data of spatially distributed precipitation as a pre-processing step is a promising alternative for setting up ANN models to forecast river flow.

In spite of the flexibility of ANNs in rainfall–runoff modeling, deficiencies in estimations sometimes arise because of the non-stationarity of signal fluctuations and their seasonalities, which vary over a range of scales from one day to several decades. An ANN model may produce misleading estimations if specific features, seasonality selection, and noise reduction are not carefully taken into account. Therefore, time and/or space pre-processing of data in such situations may be an effective approach to overcome these deficiencies.

The wavelet transform is an appropriate temporal pre-processing method that can be utilized to extract a variety of features from the data, such as short-term and long-term fluctuations, by decomposing the time series into different sub-components. The wavelet decomposition of a non-stationary time series into various scales provides an interpretation of the time series structure and extracts significant information about its history. As a result of these features, the wavelet transform has been applied to time series analysis of non-stationary hydrological signals (e.g. Adamowski, 2008a,b). The use of decomposed sub-signals as inputs into an ANN model helps the ANN model to distinguish the dominant sub-signals by applying relatively strong weights (Nourani et al., 2009a, 2011).

A hybrid wavelet–ANN model was first proposed by Aussem et al. (1998) for financial time series forecasting. Different components of the hydrologic cycle, including precipitation, groundwater, river flow, and sedimentation, have since been modeled using the wavelet–ANN approach (e.g. Adamowski and Chan, 2011; Adamowski et al., 2012; Mirbagheri et al., 2010; Nourani et al., 2009a,b, 2011; Partal and Cigizoglu, 2008; Partal and Kisi, 2007; Wang and Ding, 2003). In the field of rainfall–runoff modeling, Anctil and Tape (2004) used the wavelet transform to decompose the time series into three sub-series depicting the rainfall–runoff process, which were short, intermediate and long wavelet periods. Three multi-layer networks were then trained for the sub-series in order to estimate the runoff values. Their results revealed that short-term fluctuations are crucial for any further improvement in ANN-based rainfall–runoff simulation. Nourani et al. (2009a) utilized the combined wavelet–ANN approach for modeling the rainfall–runoff process at daily time scale and showed that the proposed model could predict both short and long-term runoff discharges. Furthermore, Nourani et al. (2011) presented a rainfall–runoff model by taking advantage of wavelet pre-processing to handle the seasonal features of the process.

Hydrologic process data may also be spatially distributed, as well as varying temporally. Therefore, pre-processing of spatial data can improve the efficiency of data-driven methods such as ANNs. Clustering is one suggested method to conduct spatial pre-processing of data, in combination with a temporal pre-processing technique. In the context of ANN-based rainfall–runoff modeling, clustering is usually performed for classification of the data, stations or zones into homogeneous classes (Nourani and Kalantari, 2010), and/or for optimization of the model structure by selecting dominant and relevant inputs (Bowden et al., 2005). Clustering techniques identify structure in an unlabeled data set by objectively arranging data into homogeneous groups, where the within-group-object dissimilarity is minimized, and the between-group-object dissimilarity is maximized. Conventional clustering methods, such as *K*-mean, that proceed according to linear characteristics require the number of clusters to be specified in advance (Hsu and Li, 2010).

The SOM is an unsupervised ANN approach (Kohonen, 1997, 1998) that is used to cluster complex data sets and does not require the number of clusters to be specified in advance. The SOM operates as an effective tool to convert complex, nonlinear, statistical relationships between high-dimensional data items into simple, geometric relationships on a low-dimensional display. Furthermore, the SOM preserves the data space and topology during the clustering process. The SOM maps high-dimensional data sets onto a regular one or two-dimensional grid.

Several studies have reported successful applications of the SOM technique to water resources problems and others in related disciplines. Furundzic (1998) applied SOM to analyze the rainfall–runoff process and the feed-forward multi-layer perceptron (MLP) ANN model for forecasting purposes. Hsu et al. (2002) presented a multivariate ANN procedure entitled self-organizing linear output map (SOLO), which was applied to classic rainfall–runoff problems. The results demonstrated the predominant characteristics of SOLO, such as its ability to rapidly and precisely estimate network structure/parameters and system outputs. Furthermore, SOLO provided features that provided insight into the underlying processes. Toth (2009) utilized SOM for stream flow forecasting and indicated that the SOM method consistently identifies the different parts of the hydrograph, based on the information available in the forecast.

Hsu and Li (2010) applied the SOM to objectively identify spatially homogeneous clusters of precipitation stations on the high-dimensional wavelet-transformed feature space. Lin and Wu (2011) employed SOM to train radial basis function network. SOM was also applied to group design hyetographs (Lin and Wu, 2007) and groundwater levels (Nourani et al., 2012a). Finally, the analysis, modeling and applications of SOM to water resources problems were reviewed by Kalteh et al. (2008) and for meteorology and oceanography by Liu and Weisberg (2011).

In this paper, a novel methodology is proposed that considers the spatiotemporal features of the rainfall and runoff process. An SOM-based clustering method is proposed to select the dominant inputs for the FFNN model. To further enhance the efficiency of the proposed rainfall–runoff model, wavelet-based (WT) temporal pre-processing is also conducted to specify the important features and seasonality of the process. The proposed SOM-WT–FFNN method uses satellite-derived rainfall data, which provide information about the spatial distribution and temporal variability of precipitation. However, satellite data may suffer from inadequate reliability, accuracy, and resolution in space and time (Petty, 1995), which introduces uncertainties in the rainfall–runoff model or may lead to incorrect forecasts. Thus, it is useful to compare model outputs using both satellite and gauge data sets.

This study proposes, for the first time, a new method of rainfall–runoff modeling based on the coupling of three techniques, namely SOM, wavelet transform, and FFNN. The proposed method is tested for single and multi-step ahead forecasting in the Gilgel Abay watershed in Ethiopia. In the next sections of the paper, the concepts of the forecasting methods (i.e. FFNN and ARIMAX), the wavelet transform, and SOM are reviewed. The following sections describe the study area and data sources, and the methodology for model development, as well as the evaluation criteria. The results obtained using the proposed methodology are then presented and discussed, followed by some concluding remarks.

2. Forecasting models

2.1. Feed forward neural network (FFNN)

The FFNN is widely applied in hydrology and water resource studies as a forecasting tool. It has already been demonstrated that an FFNN model trained by the back-propagation (BP) algorithm with three layers is satisfactory for forecasting and simulating hydrological engineering problems (ASCE, 2000; Hornik et al., 1989). Three-layered FFNNs, which have usually been used in forecasting hydrologic time series, provide a general framework for representing nonlinear functional mapping between a set of input and output variables (Fig. 1). They are based on a linear combination of the input variables, which are transformed by a nonlinear activation function. The term "feed-forward" means that a neuron connection only exists from a neuron in the input layer to other neurons in the hidden layer or from a neuron in the hidden layer to neurons in the output layer. The neurons within a layer are not interconnected.

In Fig. 1 i, j and k denote input layer, hidden layer and output layer neurons, respectively, and w is the applied weight by the neuron. The explicit expression for an output value of a three-layered FFNN is given by Nourani et al. (2012b):

$$\hat{y}_k = f_o\left[\sum_{j=1}^{M_N} w_{kj}\cdot f_h\left(\sum_{i=1}^{N_N} w_{ji}x_i + w_{jo}\right) + w_{ko}\right] \quad (1)$$

where w_{ji} is a weight in the hidden layer connecting the ith neuron in the input layer and the jth neuron in the hidden layer, w_{jo} is the bias for the jth hidden neuron, f_h is the activation function of the hidden neuron, w_{kj} is a weight in the output layer connecting the jth neuron in the hidden layer and the kth neuron in the output layer, w_{ko} is the bias for the kth output neuron, f_o is the activation function for the output neuron, x_i is ith input variable for input layer and $\hat{y}_k$ and y are computed and observed output variables, respectively. N_N and M_N are the number of the neurons in the input and hidden layers, respectively. The weights are different in the hidden and output layers, and their values can be changed during the network training process.

In any FFNN-based modeling, there are two important points to which attention must be paid: firstly, the architecture, i.e. the number of neurons in the input and hidden layers, and secondly, the training iteration (epoch) number. Appropriate selection of these two parameters improves model efficiency in both the training and testing steps. Furthermore, a high epoch number and poor quality or quantity of data could cause the network to over fit during the training step. If this occurs, the model cannot adequately generalize new data outside of the training set.

2.2. Auto regressive integrated moving average with exogenous input (ARIMAX)

The auto regressive integrated moving average (ARIMA) methodology has the ability to identify complex patterns in data and generate forecasts (Box and Jenkins, 1976). It is used to predict

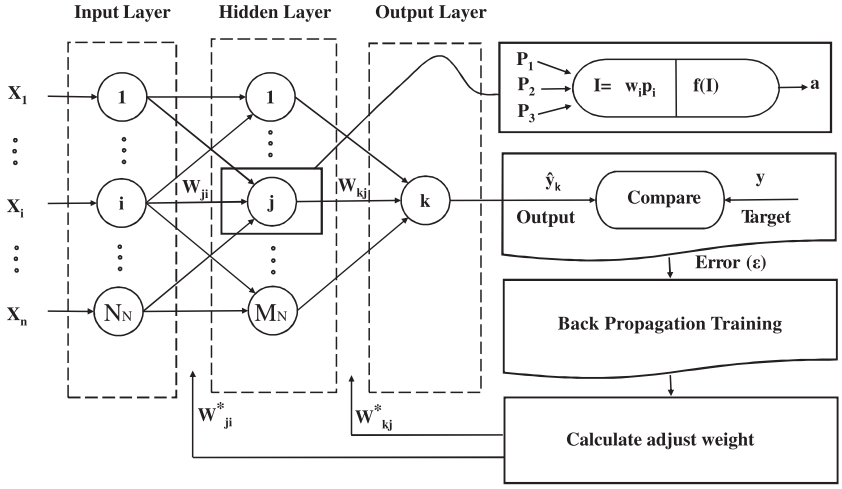

Fig. 1. A three layered feed-forward neural network with BP training algorithm.

the future according to the past input and/or output data of the process. ARIMA models can be used to analyze and forecast univariate time series data. The three steps to develop ARIMA models are identification, estimation, and forecasting. The ARIMA model function is represented by (p, d, q), with p representing the number of autoregressive terms, d the number of non-seasonal differences, and q the number of lagged forecast errors in the prediction equation. A model described as (2, 1, 3) signifies that it contains two autoregressive (p) parameters, and three moving average (q) parameters, which were computed for the series after it was differenced (d) once. ARIMA (p, d, q) models are defined as follows (Box and Jenkins, 1976):

$$W_t = \mu + \frac{\theta_q(B)}{\psi_p(B)} a_t \tag{2}$$

where $W_t = (1-B)^d Y_t$ is the difference of Y_t which is the univariate time series under investigation, B is the backshift operator; that is, $BX_t = X_{t-1}$, μ is a constant value, $\psi_p(B)$ is the auto regressive polynomial of B of order p: $\psi_p(B) = 1 - \psi_1(B) - \cdots - \psi_p(B)^p$, $\theta_q(B)$ is the moving average polynomial of B of order q: $\theta_q(B) = 1 - \theta_1(B) - \cdots - \theta_q(B)^q$, and d is the order of differencing used for the regular or non-seasonal part of the series, and a_t is a random perturbation or white noise (zero mean, constant variant, and zero covariance).

When an ARIMA model includes more than one time series as input variables, the model is known as an auto regressive integrated moving average with exogenous inputs (ARIMAX) model. Because more information is used in forecasting when exogenous inputs are considered, results usually show improvement in comparison to univariate ARIMA models. In the current study, the precipitation and antecedents of runoff data are considered as exogenous inputs, which are used to predict future runoff values as the output.

The ARIMAX model, which is an extension of the ARIMA model, is written as (Pankratz, 1991):

$$W_t = \mu + \sum_i \frac{\omega_i(B)}{\delta_i(B)} B^{k_i} X_{i,t} + \frac{\theta(B)}{\psi(B)} a_t \tag{3}$$

where $X_{i,t}$ is the ith input time series or a difference of the ith input series at time t, k_i is the pure time delay for the effect of the ith input series, $\omega_i(B)$ is the numerator polynomial of the transfer function for the ith input series, and $\delta_i(B)$ is the denominator polynomial of the transfer function for the ith input series. The model can also be written more compactly as:

$$W_t = \mu + \sum_i \Psi_i(B) X_{i,t} + \eta_t \tag{4}$$

where $\Psi_i(B)$ is the transfer function weights for the ith input series modeled as a ratio of the ω and δ polynomials: $\Psi_i(B) = (\omega_i(B)/\delta_i(B))B^{k_i}$, and η_t is the noise series: $\eta_t = (\theta(B)/\psi(B)a_t$.

This model expresses the prediction series as a combination of past values of the exogenous input series and noise series. The output and input series can be considered as the dependent and independent series, respectively.

3. Wavelet transform

The wavelet transform has increased in usage and popularity in recent years since its inception in the early 1980s, but it is still not as widely used as the Fourier transform. Fourier analysis has a serious drawback, namely that time information is lost in transforming to the frequency domain so that it is impossible to tell from a Fourier transform of a signal when a particular event took place. However, wavelet analysis allows the use of long time intervals to give more precise low-frequency information, as well as shorter regions to give high-frequency information. The time-scale localization of processes derives from the compact support of the wavelet transform's basic function, as opposed to the classical trigonometric function of Fourier analysis. The wavelet transform searches for correlations between the signal and the wavelet function. This calculation is done at different scales of a and locally around the time of b. The result is a wavelet coefficient ($T(a,b)$) contour map known as a scalogram. Grossmann and Morlet (1984) introduced wavelet transform applications in the field of earth sciences, particularly on geo-physical seismic signals. A comprehensive literature survey of wavelet applications in geosciences can be found in Foufoula-Georgiou and Kumar (1995); the most recent contributions have been cited by Labat (2005). Wavelet theory is discussed thoroughly in Labat et al. (2000) and Mallat (1998).

The time-scale wavelet transform of a continuous time signal, $x(t)$, is defined as (Mallat, 1998):

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} g^*\left(\frac{t-b}{a}\right) x(t) \cdot dt \tag{5}$$

where g^* corresponds to the complex conjugate and $g(t)$ is the wavelet function or mother wavelet. The parameter a acts as a dilation factor, while b corresponds to a temporal translation of the function $g(t)$, which allows the study of the signal around b.

For practical applications in hydrology, discrete time signals are usually available, rather than continuous time signal processes. A discretization of Eq. (5) based on the trapezoidal rule is the simplest discretization of the continuous wavelet transform. This transform produces N^2 coefficients from a data set of length N; hence, redundant information is locked up within the coefficients, which may or may not be a desirable property (Addison et al., 2001). To overcome this redundancy, logarithmic uniform spacing can be used for the a scale discretization with correspondingly coarser resolution of the b locations, which allows for N transform coefficients to completely describe a signal of length N. Such a discrete wavelet has the form (Mallat, 1998):

$$g_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} g\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \tag{6}$$

where m and n are integers that control the wavelet dilation and translation, respectively; a_0 is a specified dilation step greater than 1; and b_0 is the location parameter, which must be greater than zero. The most common and simplest choice for parameters are $a_0 = 2$ and $b_0 = 1$.

This power of two logarithmic scaling of the translation and dilation is known as the dyadic grid arrangement. The dyadic wavelet can be written in more compact notation as (Mallat, 1998):

$$g_{m,n}(t) = 2^{-m/2} g(2^{-m}t - n) \tag{7}$$

Discrete dyadic wavelets of this form are commonly chosen to be orthonormal; i.e. (Mallat, 1998):

$$\int_{-\infty}^{+\infty} g_{m,n}(t) g_{m',n'}(t) dt = \delta_{m,m'} \delta_{n,n'} \tag{8}$$

where δ is the Kronecker delta. This allows for the complete regeneration of the original signal as an expansion of a linear combination of translation and dilation orthonormal wavelets.

For a discrete time series, x_i, the dyadic wavelet transform becomes (Mallat, 1998):

$$T_{m,n} = 2^{-m/2} \sum_{i=0}^{N-1} g(2^{-m}i - n) x_i \tag{9}$$

where $T_{m,n}$ is the wavelet coefficient for the discrete wavelet of scale $a = 2^m$ and location $b = 2^m n$. Eq. (6) considers a finite time series, x_i,

$i = 0, 1, 2, \ldots, N-1$; and N is an integer power of 2: $N = 2^M$. This gives the ranges of m and n as, respectively, $0 < n < 2^{M-m} - 1$ and $1 < m < M$. At the largest wavelet scale (i.e. 2^m where $m = M$) only one wavelet is required to cover the time interval, and only one coefficient is produced. At the next scale (2^{m-1}), two wavelets cover the time interval, hence two coefficients are produced, and so on down to $m = 1$. At $m = 1$, the a scale is 2^1, i.e. 2^{M-1}or $N/2$ coefficients are required to describe the signal at this scale. The total number of wavelet coefficients for a discrete time series of length $N = 2^M$ is then $1 + 2 + 4 + 8 + \cdots + 2^{M-1} = N - 1$.

In addition to this, a signal-smoothed component, $\overline{T}$, is left, which is the signal mean. Thus, a time series of length N is broken into N components, i.e. with zero redundancy. The inverse discrete transform is given by (Mallat, 1998):

$$x_i = \overline{T} + \sum_{m=1}^{M} \sum_{n=0}^{2^{M-m}-1} T_{m,n} 2^{-m/2} g(2^{-m} i - n) \tag{10}$$

or in a simple format as (Mallat, 1998):

$$x_i = \overline{T}(t) + \sum_{m=1}^{M} W_m(t) \tag{11}$$

where $\overline{T}(t)$ is the approximation sub-signal at level M and $W_m(t)$ are detail sub-signals at levels $m = 1, 2, \ldots, M$. The wavelet coefficients, $W_m(t)$ ($m = 1, 2, \ldots, M$), provide the detail signals, which can capture small features of interpretational value in the data. The residual term $\overline{T}(t)$ represents the background information of data.

Because of the simplicity of $W_1(t)$, $W_2(t)$, …, $W_M(t)$, $\overline{T}(t)$, some interesting characteristics, such as period (time taken to see a clear repeat in extreme values of the process), hidden period (a period that is not very clear in the time series but can be depicted by detail sub-series at different scales), dependence (correlation between two sub-series at two distinct scales) and jump or shift (e.g., a shift in the flow discharge time series due to construction of a dam) can be easily diagnosed through wavelet components.

4. Self-organizing map (SOM)

The SOM is an effective tool for the visualization of high-dimensional data. It implements an orderly mapping of a high-dimensional distribution onto a regular low-dimensional grid. Therefore, it is able to convert complex, nonlinear statistical relationships between high-dimensional data items into simple geometric relationships on a low-dimensional display while preserving the topology structure of the data (Kohonen, 1997). SOM reduces dimensions by producing a map of usually 1 or 2 dimensions that plots the similarities of the data by grouping similar data items together. Thus, SOMs accomplish two things: they reduce dimensions and display similarities. The SOM network generally consists of two layers, an input layer and a Kohonen layer. The input layer is fully connected to the Kohonen layer, which in most common applications is two-dimensional. A two-level SOM neural network is a promising approach to capture a preliminary overview of intricate data sets. It augments the conventional SOM network with an additional one-dimensional Kohonen layer in which each neuron is connected to neurons in the previous Kohonen layer (Hsu and Li, 2010). The schematic view of the two-level SOM network is shown in Fig. 2.

The SOM is trained iteratively, and initially the weights are randomly assigned. When the n-dimensional input vector x is sent through the network, the distance between the w weight neurons of SOM and the inputs is computed. The most common criterion to compute the distance is Euclidean distance (Kohonen, 1997):

$$\|x - w\| = \sqrt{\sum_{i=1}^{n} (x_i - w_i)^2} \tag{12}$$

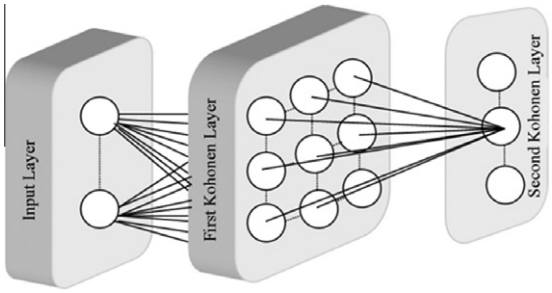

Fig. 2. Schematic view of the two-level SOM neural network.

The weight with the closest match to the presented input pattern is the winner neuron or the best matching unit (BMU). The BMU and its neighboring neurons are allowed to learn by changing the weights at each training iteration t, to further reduce the distance between the weights and the input vector (Kohonen, 1997):

$$w(t+1) = w(t) + \alpha(t) h_{lm}(x - w(t)) \tag{13}$$

where α is the learning rate, ranging in [0 1], l and m are the positions of the winning neuron and its neighboring output nodes, and h_{lm} is the neighborhood function. The most commonly used neighborhood function is the Gaussian function (Kohonen, 1997):

$$h_{lm} = \exp\left(-\frac{\|l - m\|^2}{2\sigma(t)^2}\right) \tag{14}$$

where h_{lm} is the neighborhood function of the best matching neuron l at iteration t; and $l - m$ is the distance between neurons l and m on the map grid; and σ is the width of the topological neighborhood. The training steps are repeated until convergence. After the SOM network is constructed, the homogeneous regions, or clusters, are defined on the map.

5. Study region and data

5.1. Study region

The Gilgel Abay watershed in Ethiopia is located in the northwestern highlands of Ethiopia (36°48′E–37°24′E and 10°56′N–11°23′N), and has an area of 1656 km^2 (Fig. 3). The topography of the watershed is complex, with elevations ranging from 1880 to 3530 m above sea level. Within the watershed there are rugged mountainous areas and a plateau with gentle slopes; overall more than 40% of the land has slopes greater than 10%. The climate is semi-humid, with a mean annual rainfall of 1300 mm and a mean annual temperature of 17–20 °C. The dry season, which occurs from October to May, lasts longer than the wet season. However, more than 70% of the precipitation falls during the summer monsoon season, and droughts are recurrent. The climate varies within the watershed, with the highlands having a temperate climate, and the areas at lower elevation having a tropical climate. Soil textural classes include clay (33.3%), clay loam (33.7%), and silt loam (33%). The land is used for rain-fed agriculture (74%), grassland (15%), and forest and shrubs (11%). Due to the recurrent droughts and dependence on rainfall for agriculture, the Gilgel Abay watershed requires effective management of its water resources. To enable this, a reliable method to translate rainfall values to runoff discharges is needed.

5.2. Rain gauges and runoff data

The watershed is equipped with one stream gauge at its outlet, and there are four rain gauges in its vicinity (Fig. 3), which record daily runoff and rainfall values, respectively. However, this

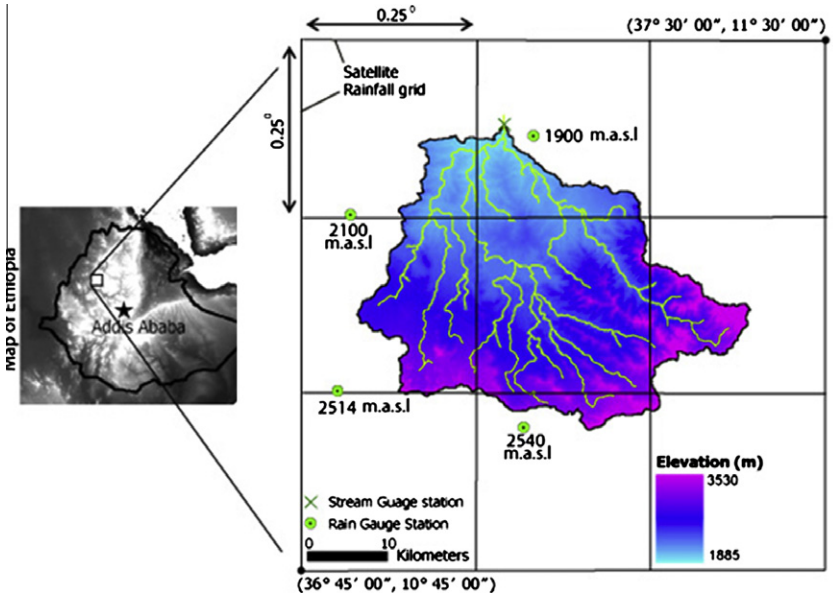

Fig. 3. Study area map.

network of ground based hydrological observations is sparsely distributed, and none of the rain gauge stations are located within the watershed. Therefore, accurate data that reflect changes in precipitation with topography is not available from rain gauges on the ground.

5.3. High resolution satellite rainfall products

Remote sensing, for example using satellite data, may provide better and/or additional coverage of precipitation data than ground-based rain gauges. The CMORPH and TMPA 3B42RT high-resolution satellite rainfall data were used in this study, as they offer the best performance for this region (Bitew and Gebremichael, 2011). Both products use microwave data as the primary source of rainfall information, and infrared data as the secondary source. The CMORPH (Climate Prediction Center Morphing) method obtains rainfall estimates from microwave data, but uses geostationary infrared data to construct the storm patterns between successive microwave overpasses (Joyce et al., 2004). The TMPA (Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis) method estimates rainfall from microwave data whenever they are available, and estimates rainfall from infrared data when microwave data are unavailable. The parameters of the infrared data are calibrated based on coinciding microwave-infrared overpasses (Huffman et al., 2007). In this study, the daily rainfall and runoff data from the gauges, and satellite for 2003–2007 were used for modeling; the first 3.5 years of data was used for training and the rest for verification purposes. Totally, 1827 data points were used, 1370 and 456 data points for training and verification purposes, respectively. In this way, it was tried to have extreme values of the observed data in the training period to help the models to learn the pattern of the process more reliably. The statistics for both rainfall data sources in daily time scale and for runoff data from the stream gauge at the outlet of the watershed are shown in Table 1. The rainfall time series of zone 2 from the satellite data and from the Gundi gauge are shown in Fig. 4 as examples.

Table 1
Statistics of rainfall and runoff data.

Rainfall time series (mm)													
Zone	CMORPH satellite data set				3B42RT satellite data set				Rain gauge data sets				
	Max	Min	Mean	Standard deviation	Max	Min	Mean	Standard deviation	Station	Max	Min	Mean	Standard deviation
Zone 1	74.1	0	4.63	8.29	72.48	0	3.64	7.83					
Zone 2	69.69	0	4.14	7.89	96.09	0	4.05	8.96	Gundi (2540 m.a.s.l.)	98.9	0	6.74	12.26
Zone 3	61.53	0	3.55	6.90	86.1	0	3.33	7.53					
Zone 4	65.88	0	4.72	7.82	64.9	0	4.00	7.98	Kidmij (2514 m.a.s.l.)	96.1	0	6.07	10.9
Zone 5	58.98	0	4.06	6.82	82.71	0	3.92	8.03					
Zone 6	51.66	0	3.45	5.94	93.42	0	3.32	7.19	Dang (2100 m.a.s.l.)	61	0	4.28	8.13
Zone 7	69.81	0	4.54	7.05	64.83	0	4.16	7.75					
Zone 8	71.37	0	3.91	6.34	63.03	0	3.87	7.76	Wete (1900 m.a.s.l.)	74	0	4.46	8.53
Zone 9	57.81	0	3.45	5.83	80.13	0	3.66	7.45					
Runoff time series (m^3/s)						Max		Min		Mean		Standard deviation	
						400.2		0		54.5		69.9	

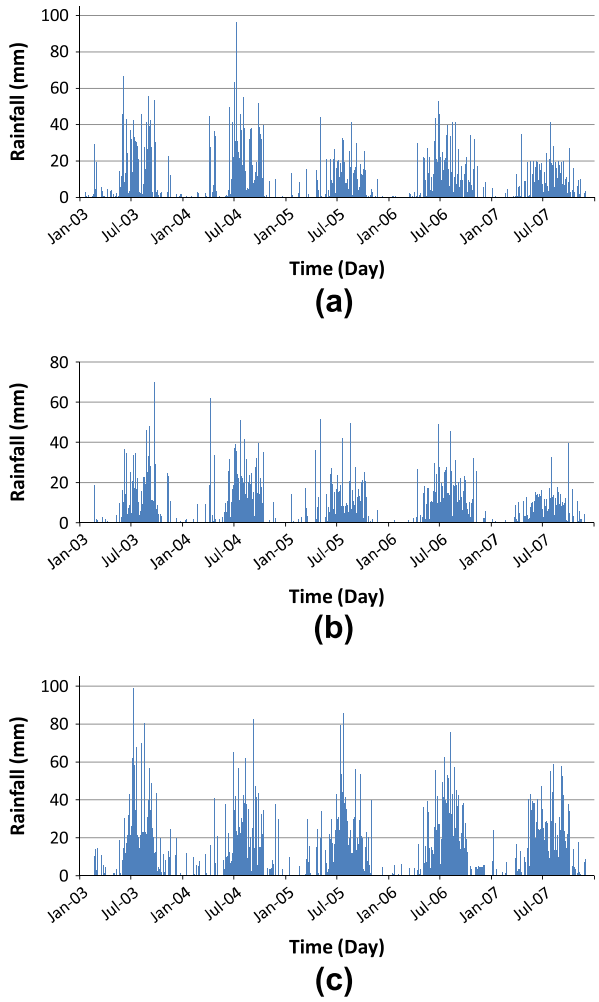

Fig. 4. Rainfall time series of (a) 3B42RT. (b) CMORPH satellite data. (c) Gundi rain gauge.

6. Methodology

To optimize the input layer and improve the efficiency of the ANN-based rainfall–runoff model of the watershed, the satellite and runoff data were spatially and temporally pre-processed, using a newly proposed method involving the SOM and wavelet transform (WT), respectively. For the SOM–FFNN model, the dominant input data selected by SOM-based spatial clustering were used in the FFNN model for single and multi-step-ahead rainfall–runoff modeling. For the proposed SOM–WT–FFNN method, wavelet analysis was also used for temporal pre-processing to improve the model's efficiency and its handling of the non-stationarity and seasonal effects of the process. Finally, a conventional ARIMAX model was developed for comparison with the two newly proposed FFNN based methods.

6.1. Data normalization

Input and output variables for ANN-based modeling are usually normalized by scaling between zero and one, to ensure that all variables receive equal attention during the training step of a model. The following simple linear mapping of the variables is the most common method for this purpose (Grimes et al., 2003):

$$r_i = \frac{b_i - b_{\min}}{b_{\max} - b_{\min}} \quad (15)$$

where b_i is the actual value and r_i is the respective normalized value. $b_{\min}$ and $b_{\max}$ are the minimum and maximum of the values, respectively. In this study, the normalized data were divided into training and verification sets. Three and a half years of data were used for model training, and the remaining 1.5 years of data were used for validation.

6.2. SOM clustering

The selection of the most relevant and appropriate inputs is an important step in ANN-based modeling of the rainfall–runoff process when various data sources, such as for precipitation, are available in the watershed. As satellite rainfall data are available over the nine zones of the watershed (Fig. 3), the zones that contribute highly to runoff generation and those that best represent the rainfall process in the watershed were identified. As the details are presented in the below lines, the SOM-based clustering was led to select zones 2, 5, 8 and zones 5, 8 as the representative zones using CMORPH and 3B42RT satellites data sets, respectively. Using only data from these zones allowed the dimensionality of the original set of inputs to be reduced, resulting in an appropriate and optimized FFNN for rainfall–runoff modeling of the watershed.

The conventional trial and error methodology for selecting the most dominant inputs from large datasets is a time consuming process, because many different combinations of the input variables need to be examined. The number of trials for a model with n input variables is $2^n - 1$.

Therefore, if the trial–error method were used in this study to determine the effective and dominant precipitation zones of the watershed, $2^9 - 1 = 511$ combinations of zones would need to be examined as the FFNN inputs for each satellite data set. Since the nine precipitation zones do not have an equal effect on runoff values or do not provide 'informative' input data, the use of only selected inputs into the FFNN simplifies the model structure and leads to better results.

Dominant precipitation zones that are representative of the watershed precipitation can be identified using a spatial clustering method, such as the two-step SOM clustering method. In the first step of this method, a two-dimensional SOM was used to classify the precipitation zones of each satellite data set into classes with similar precipitation patterns. This provided an overview of the homogeneous regions and the approximate number of clusters in relation to the watershed topography. Secondly, a one-dimensional SOM was applied to classify the rainfall zones with the specific class numbers determined in the first step. The number of neurons in the Kohonen layer was set to be equal to the number of clusters determined in the first step, so that each cluster was represented by a neuron. The Euclidean distance criterion (Bowden et al., 2005) was then utilized to select the centroid zone of each cluster, which is the best representation of the precipitation pattern of the cluster.

To apply the two-step SOM on the CMORPH satellite data, the size of the Kohonen layer used in the first step was determined as a 3-by-3 grid via a trial–error process. Since there is no theoretical principle for determining the optimum size of the Kohonen layer, the Kohonen layer should be large enough to ensure that a suitable number of clusters are formed from the training data (Cai et al., 1994). After creation of the 3-by-3 Kohonen layer, the relevant zones of each cluster on the map were identified according to the number of zones assigned to the neurons (Fig. 5a). A 1-dimensional SOM with the neurons organized in a 1-by-3 Kohonen layer was then formed as the second step of the clustering (Fig. 6). Zones 2 and 3 were classified in one cluster because of similar precipitation data. The other seven zones were classified individually into seven separate clusters. Fig. 5b shows the neighbor weight distances, where the dark hexagons represent

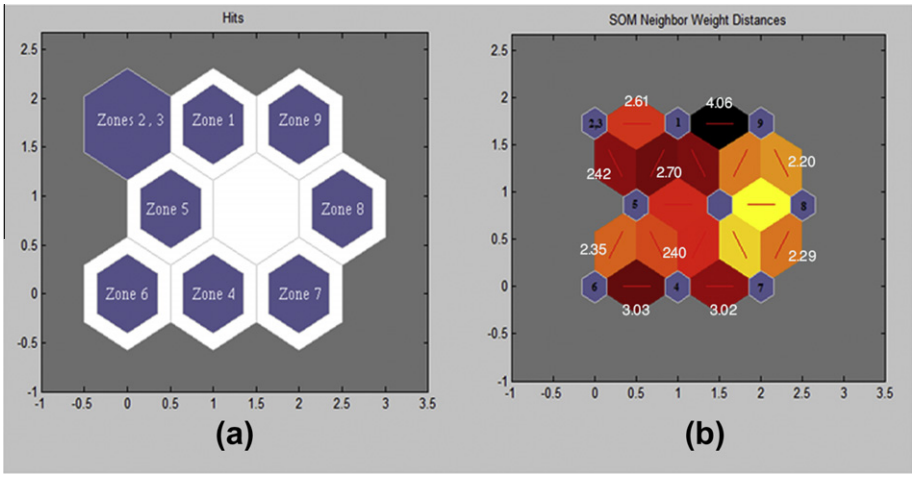

Fig. 5. The 2-Dimensional SOM clustering of CMORPH satellite data (a) SOM hits. (b). SOM neighbor weight distances plan.

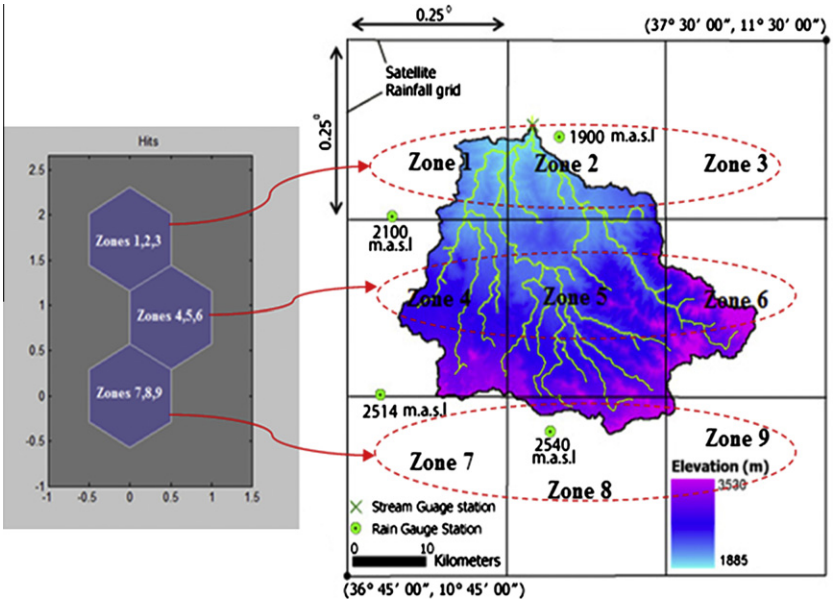

Fig. 6. The 1-dimensional SOM clustering of CMORPH satellite data.

the neurons. The colors in the regions indicate the distances between neurons, with the darker colors representing larger distances, and the lighter colors representing smaller distances.

6.3. Proposed SOM–FFNN model

For each type of satellite data, the rainfall values of representative zones were used in the FFNN to model the rainfall–runoff process in the watershed. This type of ANN model accompanied by a back-propagation training algorithm is widely used in hydrologic modeling (ASCE, 2000). The proposed FFNN models comprised both single- and multi-step ahead predictions.

Various combinations of the selected zones were examined using the FFNN model (Tables 2 and 3 for CMORPH and 3B42RT satellite data, respectively). Each FFNN was trained by examining 3–20 hidden neurons in a single hidden layer using the Levenberg–Marquardt training scheme (Hagan and Menhaj, 1994) and up to 200 training epochs, with 10^{-4} as the goal performance. The training was terminated at the point where the error in the validation data set began to rise to ensure that the network did not over fit the training data and then fail to generalize the un-seen test data set. No great improvement in model performance was found when the number of hidden neurons was increased above a threshold. At this stage, the model efficiency criteria for each satellite data set were used to determine the best model. Sensitivity analysis of the FFNN rainfall–runoff models with various input combinations allowed the best combination of the representative zones.

6.3.1. One-step-ahead forecasting using SOM–FFNN

In order to get appropriate 1-day-ahead forecasts of runoff, the input layer needs to include all relevant information on the target data. The size of the sliding window of the previous days of data for precipitation and runoff was identified by the trial–error procedure. Based on sensitivity analysis, the input layer was then optimized such that only the most important signals were included. The use of runoff values at current day together with one and two days lag (i.e. $Q(t)$, $Q(t-1)$, $Q(t-2)$) was determined to be optimal by sensitivity analysis over different sliding input windows of the previous days' runoff data (up to 7 past days were tested). This was therefore used in all models. The use of previous time steps up

Table 2
Results of single-step-ahead SOM–FFNN model for CMORPH satellite data.

Applied rainfall zones	Input variables[a]	Epoch	Network structure[b]	DC		RMSE[c]	
				Calibration	Verification	Calibration	Verification
Zone 2	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	190	4-10-1	0.88	0.77	0.059	0.064
Zone 5	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	180	4-9-1	0.87	0.66	0.061	0.080
Zone 8	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	150	4-12-1	0.88	0.63	0.058	0.082
Zones 2, 5	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	180	5-8-1	0.88	0.73	0.059	0.070
Zones 2, 8	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	150	5-12-1	0.91	0.63	0.051	0.083
Zones 5, 8	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	110	5-14-1	0.91	0.75	0.050	0.067
Zones 2, 5, 8	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	180	6-15-1	0.93	0.80	0.047	0.060

[a] Output is $Q(t+1)$ in all networks. The number of $I(t)$ refers to the number of applied rainfall zones.
[b] The result has been presented for the best structure.
[c] The RMSE in all tables is dimensionless.

Table 3
Results of single-step-ahead SOM–FFNN model for 3B42RT satellite data.

Applied rainfall zones	Input variables[a]	Epoch	Network structure[b]	DC		RMSE	
				Calibration	Verification	Calibration	Verification
Zone 5	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	170	4-7-1	0.87	0.60	0.060	0.200
Zone 8	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	180	4-8-1	0.88	0.66	0.059	0.079
Zones 5, 8	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	150	5-15-1	0.92	0.75	0.051	0.068

[a] Output is $Q(t+1)$ in all networks. The number of $I(t)$ refers to the number of applied rainfall zones.
[b] The result has been presented for the best structure.

to a week for rainfall data (e.g. $I(t)$, $I(t-1), \ldots, I(t-7)$) did not improve model performance. However, the current day's precipitation (i.e. I(t)) had a significant effect on runoff prediction. Therefore, the rainfall value for the current day was used in all FFNNs. The outcome suggested that daily rainfall cannot be considered as a Markovian process, meaning that the next event depends only on the current event and not on the sequence of events that preceded it. This is in agreement with the outcomes of previous studies (Nourani et al., 2009b), which found that the daily rainfall process does not present a strong Markov chain.

6.3.2. Multi-step-ahead forecasting using SOM–FFNN

In order to explore runoff predictions several time steps ahead, a multi-step-ahead FFNN approach was also applied to the selected precipitation zones for each type of satellite data obtained from the single-day-ahead SOM–FFNN model. Two methods can be used to carry out multi-step-ahead forecasts. The first method relies on the ability of the model to generate forecasts, where outputs of the model are selected directly from the original data set. The second method performs multi-step-ahead forecasting by using the ANN's own estimations as a source of information for further forecasts (Vos and Rientjes, 2005). Both of these approaches led to similar results in FFNN-based models up to 4-day-ahead forecasting in this study, thus, the first method was selected because it is simpler to use for two to 4-day-ahead (i.e. $Q(t+2)$, $Q(t+3)$, $Q(t+4)$) forecasts that were performed in our study. This is in agreement with the results of previous studies (e.g. Chang et al., 2007; Vos and Rientjes, 2005).

Precipitation data without any lag and runoff data with antecedents were used as the input neurons. Lag sequences for the runoff time series input were established based on the lead steps of predictions by sensitivity analysis. Details of the input values and predicted output values of daily data for each of the several-step-ahead models are provided in Table 4.

6.4. Proposed SOM–WT–FFNN model

In addition to spatial patterns, some temporal features may also exist in the rainfall–runoff process due to highly non-stationary fluctuations of the time series. The wavelet transform can be used

Table 4
Results of multi-step-ahead FFNN model.

Data set	Applied rainfall zones	Input variables	Output variable	Epoch	Network structure[a]	DC		RMSE	
						Calibration	Verification	Calibration	Verification
CMORPH	2, 5, 8	$I(t)$, $Q(t)$, $Q(t-2)$, $Q(t-4)$	$Q(t+2)$	80	6-9-1	0.87	0.75	0.061	0.072
		$I(t)$, $Q(t)$, $Q(t-3)$, $Q(t-6)$	$Q(t+3)$	70	6-10-1	0.85	0.67	0.067	0.078
		$I(t)$, $Q(t)$, $Q(t-4)$, $Q(t-8)$	$Q(t+4)$	180	6-10-1	0.85	0.51	0.067	0.093
3B42RT	5, 8	$I(t)$, $Q(t)$, $Q(t-2)$, $Q(t-4)$	$Q(t+2)$	120	5-9-1	0.86	0.73	0.065	0.070
		$I(t)$, $Q(t)$, $Q(t-3)$, $Q(t-6)$	$Q(t+3)$	90	5-12-1	0.85	0.61	0.066	0.084
		$I(t)$, $Q(t)$, $Q(t-4)$, $Q(t-8)$	$Q(t+4)$	150	5-15-1	0.84	0.58	0.065	0.087

[a] The result has been presented for the best structure.

as a temporal pre-processing technique for detection of temporal features and seasonalities in the rainfall–runoff process by decomposing the time series into multi-resolution sub-series. As data are processed at different temporal scales (levels), the large and small features of a signal can be separated. For this purpose, the observed runoff time series at the outlet was decomposed into one approximation sub-signal, $Q_a(t)$ and i detailed sub-signals $Q_{d1}(t), \ldots, Q_{di}(t)$ where i denotes to the decomposition level. In this manner, each sub-signal represents a different level of the seasonality relationship.

Since runoff values are measured in discrete form, the dyadic discrete wavelet transform was used rather than a continuous wavelet. The Daubechies with four vanishing moments (db4, Fig. 8d), as well as Haar mother wavelets as a step function (Mallat, 1998), were used to decompose the runoff time series into sub-series at different levels, as their effects on rainfall–runoff modeling has already been investigated by one of the authors (Nourani et al., 2009a, 2011).

The decomposed sub-series accompanied by the satellite rainfall data of the representative zones were then used in the FFNN model in order to predict 1-day-ahead and multi-step ahead runoff values. For instance, in decomposition level 4 there is one approximation and 4 details, i.e. 2^1-day mode, 2^2-day mode, 2^3-day mode (which is a nearly weekly mode), and 2^4-day mode (which is a nearly 2-week mode) (Fig. 7).

6.4.1. One-step-ahead forecasting using SOM–WT–FFNN

The SOM–WT–FFNN method enabled the wavelet decomposed levels that had the most influence on the conversion of rainfall into runoff over the watershed to be determined and used for forecasting. In the training step, rainfall data and decomposed runoff values from the calibration dataset at distinct wavelet levels were entered into the FFNN as input neurons in order to forecast the runoff 1 day-ahead. The trained model was then validated using the verification dataset.

When multi-level wavelet decomposed sub-signals are entered into the model as input neurons, the weights applied to them by the FFNN differ between the decomposition levels. In this case, the inputs were the rainfall time series of representative zones (obtained via SOM) and the runoff sub-signals at different resolutions (obtained via WT), with each resolution demonstrating a specific seasonality feature of the process. In this way, the most important level of a signal receives a high weight compared to the other levels, which reflects the dominant seasonality of the rainfall–runoff process.

6.4.2. Multi-step-ahead forecasting using SOM-WT–FFNN

The multi-step-ahead forecasting approach (in this case two to four days ahead forecasting) was also investigated in order to develop and test the SOM–WT–FFNN approach in terms of its ability to provide accurate rainfall–runoff forecasts with sufficient lead times for watershed management and flood mitigation in the watershed. Detailed sub-signals of the runoff time series along with the precipitation data were used in the SOM–WT–FFNN model to forecast runoff 2–4-days-ahead.

6.5. Proposed ARIMAX model

ARIMAX is a classic time series-forecasting model, and therefore was used as a comparison model to evaluate the efficiencies of the proposed hybrid models. Although a linear model like ARIMAX is likely not able to model a complex nonlinear hydrological process, it is still a commonly used method in practice and as such is useful as a comparison model. The ARIMAX (p, q, d) model was checked using different values of p, d and q. Precipitation and antecedents of runoff data were used as exogenous inputs to predict future runoff as the output. As with the other models, the ARIMAX model was first calibrated using the training data set, and the calibrated model was then validated using the verification data set.

6.6. Evaluation of model precision

Three different criteria were used to measure the efficiency of the proposed forecasting methods; the root mean square error

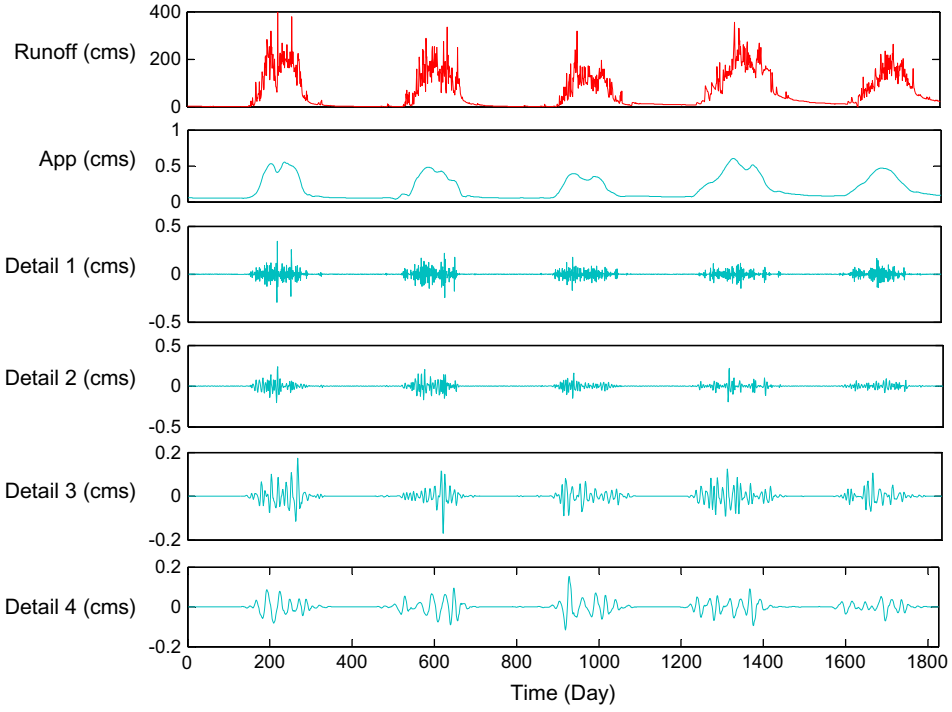

Fig. 7. Approximation and detail sub-signals of runoff time series.

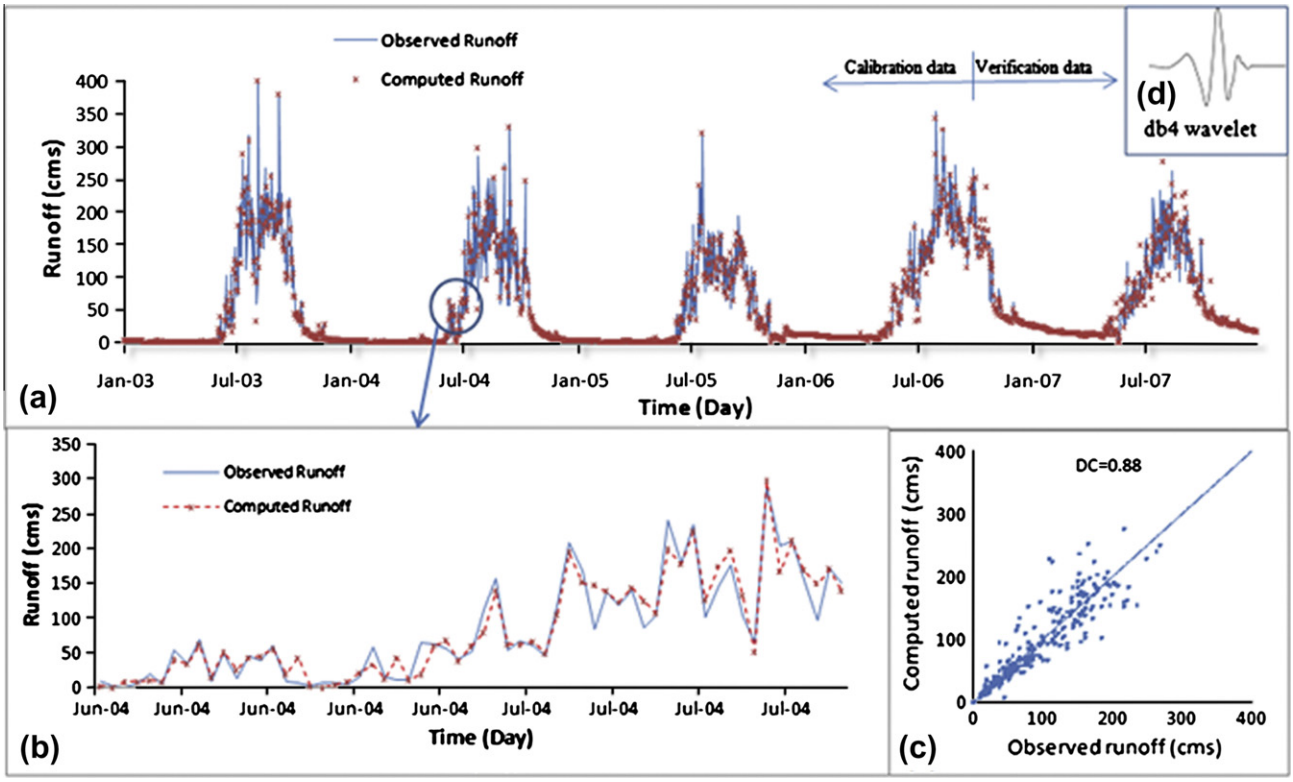

Fig. 8. The results of SOM–WT–FFNN model using 3B42RT data (a) Computed and observed runoff time series. (b) A detail. (c) Scatter plot of verification data. (d) db4 Mother wavelet.

(RMSE), the determination coefficient (DC) and the ratio of absolute error of peak flow ($RAE_p(\%)$). The RMSE and DC are used to demonstrate discrepancies between forecasts and observations. The RAE_p measures the closeness of observed and estimated runoff peak values. They are calculated as (Nourani and Kalantari, 2010):

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(b_i - \hat{b}_i)^2}{n}} \tag{16}$$

$$\mathrm{DC} = 1 - \frac{\sum_{i=1}^{n}(b_i - \hat{b}_i)^2}{\sum_{i=1}^{n}(b_i - \bar{b}_i)^2} \tag{17}$$

$$\mathrm{RAE}_p(\%) = \frac{100}{M}\sum_{m=1}^{M}\frac{|b_m - \hat{b}_m|}{I_m} \tag{18}$$

In Eqs. (16) and (17), n is the data number, b_i and $\hat{b}_i$ are the observed data and the calculated values, respectively, and $\bar{b}_i$ is the averaged value of the observed data. In Eq. (18), M, b_m and $\hat{b}_m$ are the number of peak values, observed and predicted peak values, respectively. High values for DC and small values for RMSE indicate that the model is highly efficient (for the best model, the values of DC and RMSE would be one and zero, respectively). Legates and McCabe (1999) indicated that a hydrological model can be sufficiently evaluated by DC and RMSE, but RAEp is also used in this study due to the importance of peak and extreme values in hydrologic processes.

7. Results and discussion

The results of the proposed single and multi-step-ahead FFNN-based models using satellite data were compared with the models using rain gauge data. The FFNN results were also compared with those of the conventional ARIMAX model.

7.1. Results of SOM clustering

The first step of the SOM clustering method for the CMORPH dataset resulted in the zones being clustered into three distinct groups. Zone 1 is in the same cluster as zones 2 and 3 due to the small distance between their neurons. Zones 4, 5 and 6 are clustered together, while zones 7, 8 and 9 are located in the third cluster according to the weight distances. Three clusters were again identified during the second step of the method for the CMORPH dataset. The first cluster contained zones 1, 2 and 3, which were located at approximately the same altitude near the outlet of the watershed. Zones 4, 5 and 6 that cover the central area of the watershed were classified in the second cluster. Zones 7, 8 and 9, covering the mountainous area, were situated in the third cluster. Therefore, the second step of the clustering method reconfirmed the classification of the zones that was obtained from the first step.

The clustering results of the CMORPH satellite data are consistent with the physical characteristics of the watershed, meaning that the precipitation zones are spatially divided according to the topography of the watershed. Therefore, by sampling one representative zone from each cluster, it was possible to remove highly correlated, redundant zones from that cluster. In this way, three zones were selected from the original nine. Zones 2, 5 and 8 were selected from clusters 1, 2 and 3, respectively, by means of the Euclidean distance criterion to represent the rainfall pattern over the whole watershed using CMORPH satellite rainfall data.

Similarly, the most important zones were extracted from the 3B42RT satellite data using the proposed 2-step SOM method. A 2-by-2 Kohonen layer was obtained from trial–error, in which the first cluster contained zones 2, 3, 5, 6 and 9. Zones 7 and 8 were classified in the second cluster, and zones 1 and 4 were dissimilar to all others and hence were classified in two separate clusters. For the second step of SOM, a one-dimensional 1-by-2 Kohonen layer

was selected to capture more organized clusters. This revealed two clusters, the first containing only zone 8, with the second including the other eight zones. From the second cluster, zone 5 was identified as the central zone using the Euclidean distance criterion. Therefore, zones 5 and 8 were considered as representative of the rainfall pattern over the watershed using the 3B42RT satellite data.

Although used SOM in both steps are independent, processing of data by passing through the first SOM (Eq. (14)) leads the data to be stationary prior to input to the second SOM. This pre-processing can help the second SOM to be trained quickly. On the other hand, the result of the first step (2D SOM) could be used to determine the number of clusters (neurons) in the second SOM without any need to the trial–error process in the second step. Therefore, the results of steps could reconfirm the outcome.

This clustering technique, which identified the homogeneous zones of each satellite dataset, can be considered as an input screening procedure that encompasses some advantages. The examination of 511 input combinations for FFNN was decreased to seven cases for the CMORPH dataset, and to three cases for the 3B42RT satellite dataset. From the spatial pre-processing point of view, the selection of dominant zones led to suitable coverage of precipitation patterns over the watershed. Furthermore, the SOM process has the potential to increase the accuracy of the rainfall–runoff model, as only dominant or important inputs are used in the FFNN.

7.2. Results of the SOM–FFNN models

The results of the SOM–FFNN forecasting models indicated that almost all of the models produced acceptable outcomes, and confirmed the appropriate identification of the representative rainfall zones and the precipitation patterns over the watershed. The SOM–FFNNs with CMORPH rainfall data from zones 2, 5 and 8 and 3B42RT data from zones 5 and 8 with outlet runoff values (highlighted in Tables 2 and 3) yielded better performances in terms of DC and RMSE.

Among the SOM–FFNN models, the important structures were derived from the combinations of dominant precipitation zones (Tables 2 and 3). For the CMORPH satellite dataset the effective combination of precipitation data is from three zones, whereas for the 3B42RT satellite data it is from two zones. Although the inputs for the two data sources differ, the SOM–FFNN models perform almost equally in terms of the efficiency criteria (Tables 2 and 3). The presence of one redundant precipitation zone when using the CMORPH dataset leads to the model results being relatively similar to those from the 3B42RT dataset for which only two precipitation zones were used.

In comparing the suitability of two satellite datasets for the SOM–FFNN model, it is sensible to choose the dataset that performs equally well but requires fewer input neurons. Therefore, in this study the 1-day-ahead FFNN model using 3B42RT satellite data is preferable to the model that uses the CMORPH data. It is anticipated that future hydrological forecasting in the Gilgel Abay watershed using 3B42RT satellite data will provide effective, informative, and rapid forecasts with minimum time and cost.

Precipitation zones for the SOM–FFNN inputs were spatially selected in an attempt to produce a more efficient rainfall–runoff model. There are physical concepts that explain the selection of the zones in the center of the watershed (zones 2, 5 and 8) by the SOM approach, in addition to the non-linear patterns that relate rainfall to runoff. These zones cover a large area of the watershed and a wide altitudinal range, and therefore use of their precipitation data leads to the rainfall–runoff process being appropriately modeled (Fig. 3).

The Markovian property of the runoff was more perceptible by determination of current day along with one and two days lag of runoff data ($Q(t)$, $Q(t-1)$, $Q(t-2)$) to predict the 1 day ahead discharge ($Q(t+1)$). Hence, the combination of runoff antecedents and rainfall data produced an appropriate rainfall–runoff pattern. Moreover, transforming rainfall into runoff using 1 and 2-day lags for the discharge values as inputs is appropriate for the size of the watershed.

The results of the FFNNs that use individual precipitation zones in the input layer (Tables 2 and 3) indicated that including rainfall data from downstream zones 2 and 5 led to more efficient models than including data from upstream zone 8. This may be because zone 5 covers a large area of the watershed and because zone 2 is adjacent to its outlet. Thus, the amount of precipitation over such areas has the greatest influence on runoff generation.

Forecasting accuracy decreased when the forecasting lead time was increased (up to 4 days ahead) for the SOM–FFNN model (Table 4). Comparison of the results of single-step-ahead forecasting (Tables 2 and 3) with the results of multi-step-ahead models (Table 4) reveals that the SOM–FFNN model is capable of making one-step-ahead forecasts with reasonable accuracy, but that multi-step-ahead predictions are not as accurate. This may be related to the small size of the catchment, resulting in a lag-time that is less than a day. As a result of this, multi-step-ahead predictions with a lead time of longer than a day yield poor results by increasing the steps of forecasting for this watershed.

7.3. Results of SOM–WT–FFNN models

Determination of dominant sub-signals increases forecasting capability by employing the synchronic effects of non-linear correlation between the detailed sub-signals and the main time series. Hence, it is important to select the input components, such as wavelet sub-signals, for the FFNN model by taking account of the hydrologic characteristics of the process. Models with inputs including detail signals at level 3 obtained via decomposition using the db4 mother wavelet performed well in terms of the efficiency criteria (i.e. DC and RMSE) (Tables 5 and 6). In both modeling procedures relevant to the CMORPH and 3B42RT datasets, the appropriate structures of the SOM–WT–FFNNs were selected following the results of DC and RMSE in the calibration and verification phases. When the decomposition level was increased from three levels, the accuracy of the models decreased. This is because high decomposition levels introduce a large number of parameters with complex nonlinear relationships into the FFNN. The errors for each parameter magnify the total error of the network, and thus the FFNN efficiency diminishes. Although this effect is not perceptible when moving from levels 3 to 4 in the calibration phase, it is apparent in the results of the verification phase and the examination of higher levels in the sensitivity analysis. Level 4 was identified as the level at which accuracy starts to decrease (Tables 5 and 6). The selection of the decomposition level is directly related to the data length in which the periodicity and seasonal cycle of the hydrologic time series are apparent (Fig. 8a–c). According to Tables 5 and 6, for 1-day-ahead SOM–WT–FFNN, the best model in the validation phase had a DC of 0.93 and 0.91 for CMORPH and 3B42RT satellite data, respectively.

The db4 mother wavelet provided comparatively better outcomes than Haar wavelet in terms of efficiency criteria (Tables 5 and 6). The better performance of the db4 wavelet may be due to its form (Fig. 8d), which is similar to the runoff signal fluctuations. The sudden onset and cessation of rainfall over the watershed produced several spikes in the runoff time series. As the db4 wavelet has similar signal properties, it was likely able to capture the runoff signal features well, particularly the peak values, and therefore performed better (Fig. 8). In future studies, it is strongly recom-

Table 5
Results of single-step-ahead SOM–WT–FFNN model for CMORPH satellite data.

Decomposition level	Mother wavelet	Epoch	Network structure[a]	DC		RMSE	
				Calibration	Verification	Calibration	Verification
2	Haar	180	6-10-1	0.96	0.89	0.036	0.048
3	Haar	170	7-15-1	0.97	0.82	0.028	0.058
4	Haar	80	8-15-1	0.97	0.78	0.028	0.065
2	db4	130	6-10-1	0.96	0.88	0.036	0.047
3	db4	120	7-15-1	0.98	0.93	0.023	0.035
4	db4	160	8-15-1	0.98	0.80	0.023	0.061

Note: Input variables are $I(t)$ of zones 2, 5, 8 and $Q_a(t)$, $Q_{di}(t)$; $Q(t+1)$ is output.
[a] The result has been presented for the best structure.

Table 6
Results of single-step-ahead SOM–WT–FFNN model for 3B42RT satellite data.

Decomposition level	Mother wavelet	Epoch	Network structure[a]	DC		RMSE	
				Calibration	Verification	Calibration	Verification
2	Haar	140	5-10-1	0.94	0.89	0.039	0.045
3	Haar	110	6-13-1	0.96	0.88	0.033	0.046
4	Haar	200	7-17-1	0.98	0.80	0.027	0.060
2	db4	150	5-7-1	0.95	0.90	0.038	0.042
3	db4	80	6-15-1	0.97	0.91	0.030	0.041
4	db4	170	7-16-1	0.98	0.88	0.027	0.047

Note: Input variables are $I(t)$ of zones 5, 8 and $Q_a(t)$, $Q_{di}(t)$; $Q(t+1)$ is output.
[a] The result has been presented for the best structure.

Table 7
Results of multi-step-ahead SOM–WT–FFNN model.

Data set	Input variables[a]	Output variables	Epoch	Network structure[b]	DC		RMSE	
					Calibration	Verification	Calibration	Verification
CMORPH	$I(t)$, $Q_i(t)$	$Q(t+2)$	180	7-10-1	0.94	0.81	0.042	0.060
	$I(t)$, $Q_i(t)$	$Q(t+3)$	190	7-9-1	0.93	0.86	0.045	0.050
	$I(t)$, $Q_i(t)$	$Q(t+4)$	170	7-9-1	0.91	0.79	0.055	0.059
3B42RT	$I(t)$, $Q_i(t)$	$Q(t+2)$	140	6-11-1	0.94	0.86	0.043	0.051
	$I(t)$, $Q_i(t)$	$Q(t+3)$	200	6-8-1	0.92	0.88	0.049	0.046
	$I(t)$, $Q_i(t)$	$Q(t+4)$	100	6-12-1	0.91	0.79	0.058	0.061

[a] $I(t)$, refers to zones 2, 5, 8 in CMORPH data set, and zones 5, 8 in 3B42RT data set. $Q_i(t)$ refers to decomposed runoff at level 3: $Q_a(t)$, $Q_{d1}(t)$, $Q_{d2}(t)$, $Q_{d3}(t)$.
[b] The result has been presented for the best structure.

mended that the mother wavelet be selected by taking into account the nature and formation of the hydrologic signal, as was done in this study.

As was the case for the multi-step-ahead predictions in the SOM–FFNN, the accuracy of the SOM–WT–FFNN model also decreased compared to the single-step-ahead model, but the accuracy did not decrease as the number of prediction steps increased (Table 7). The 2-day-ahead model produced the best results compared to the 1- and 3-day-ahead forecasts (using wavelet decomposition level three). This might be because the effect of the decomposed levels up to three levels (2^3-day mode) coincides with the 2-day sliding window at the 2-day lead-time forecast.

Although the results of the multi-step-ahead forecasts for the SOM–WT–FFNN model were weak in comparison to single-step-ahead forecasts, the evaluation criteria for the multi-step-ahead SOM–WT–FFNN model were better than the multi-step-ahead SOM–FFNN model. Overall, the 1-day-ahead forecasts of both the SOM–FFNN and SOM–WT–FFNN rainfall–runoff models provided reasonably accurate forecasting results.

7.4. Results of ARIMAX model

Among the various single-step-ahead ARIMAX forecasting models that were developed using datasets of satellites and rain gauges, the models with the best forecasting performance are shown in Table 8. The lower accuracy of the conventional ARIMAX forecasting model compared to the FFNN-based models can be attributed to the linearity of ARIMAX and its shortcomings in modeling nonlinear rainfall–runoff process. For example, the results for the CMORPH satellite data during the verification stage revealed that there is an approximately 50% decrease in the DC for the ARIMAX model in comparison with the best single-step-ahead FFNN-based model.

Table 8
Results of ARIMAX model with lead-time of 1-day.

Rainfall data source	ARIMAX structure	DC		RMSE	
		Calibration	Verification	Calibration	Verification
CMORPH	satellite	(1, 0, 1)	0.85	0.47	0.0061
0.0085					
3B42RT	satellite	(1, 2, 1)	0.83	0.77	0.0060
0.0056					
Gundi gauge	(2, 1, 2)	0.85	0.63	0.0061	0.0012

8. Discussion

In order to provide a comprehensive comparison of the various rainfall–runoff models in the watershed, the results of the single-

Table 9
Results and structures of daily SOM–FFNN model using rain gauge data.

Rain gauge	Input variables[a]	Epoch	Network structure	DC		RMSE	
				Calibration	Verification	Calibration	Verification
Dang	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	150	3-9-1	0.88	0.74	0.061	0.069
Gundi	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	150	3-10-1	0.89	0.82	0.058	0.059
Kidmij	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	120	3-10-1	0.88	0.78	0.061	0.064
Wete	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	140	3-10-1	0.87	0.82	0.062	0.058
Gundi-Wete	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	60	4-10-1	0.89	0.71	0.059	0.074
Dang-Wete	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	140	4-9-1	0.88	0.75	0.059	0.069
All stations	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	140	6-10-1	0.90	0.79	0.055	0.062
Average	$I(t)$, $Q(t)$, $Q(t-1)$, $Q(t-2)$	70	3-9-1	0.88	0.82	0.061	0.058

Note: $I(t)$ at gauge(s) $Q(t+1)$ is output.
[a] The result has been presented for the best structure.

Table 10
Comparison of the results for models with the lead-time of 1-day.

Rainfall data set	Model	Structure	DC		RAE_p (%)	
			Calibration	Verification	Calibration	Verification
CMORPH satellite	SOM-FFNN	6-15-1	0.93	0.80	12.73	28.92
	SOM–WT–FNN	7-15-1	0.98	0.93	6.46	18.33
	ARIMAX	(1, 0, 1)	0.85	0.47	23.04	41.74
3B42RT satellite	SOM–FFNN	5-15-1	0.92	0.75	15.87	28.71
	SOM–WT–FFNN	6-15-1	0.97	0.91	7.56	21.06
	ARIMAX	(1, 2, 1)	0.83	0.77	24.92	26.57
Gundi gauge	SOM-FFNN	4-10-1	0.89	0.82	18.84	26.02
	SOM–WT–FFNN	5-7-1	0.95	0.90	9.31	17.48
	ARIMAX	(2, 1, 2)	0.85	0.63	24.16	32.15

step-ahead FFNN models obtained using the data of rainfall gauges are gathered in Table 9, where the highlighted row shows the dominant gauge that was selected for comparison purpose with the satellite data.

The results of the best hybrid models for single-step-ahead forecasting (Table 10) indicate that use of SOM as a spatial pre-processor provides some advantages, due to the selection of dominant precipitation zones over the watershed, which reduces the trial–error process. The SOM–FFNN results in terms of the efficiency criteria are not as strong compared to the results of the SOM–WT–FFNN model that uses both spatial and temporal pre-processing methods. Therefore, the SOM–WT–FFNN model, which makes use of both SOM and wavelet techniques to capture the space and/or time variations involved in the process, is an appropriate and promising rainfall–runoff forecasting method.

In the SOM–FFNN model, which uses only spatially pre-processed data, only the short-term autoregressive features of the process ($Q(t)$, $Q(t-1)$, $Q(t-2)$) were considered, without the long term seasonalities being taken into account. This weakness was addressed by incorporating the wavelet concept into the SOM–WT–FFNN model. In this case, RAE_p values decreased by 30–50% because the extreme values of previous seasonal periodicity were used in the model training step (Table 10). The simultaneous application of spatial clustering and temporal pre-processing allows both the spatial distribution of the rainfall data and the multi-scale seasonality signature of the runoff to be taken into account. The comparison of results obtained by the satellite-based and gauge-based modeling confirms that the satellite data offers appropriate accuracy in the simulation of the rainfall–runoff process in the watershed. Thus, in watersheds where ground-based hydro-meteorological observations are sparse (such as in some developing countries like Ethiopia), satellite data provide relatively reliable data for rainfall–runoff modeling. For this purpose, estimated precipitation based on meteorological satellite radiances sensed by microwave (on-board low-orbiting satellites) and infrared (on-board geostationary satellites) sensors can be useful (Lábó, 2012).

In this study, it was found that single-step-ahead forecasts are accurate and multi-step-ahead predictions in spite of less accuracy than single-step-ahead approach can provide useful horizon of forecasts. Such results suggest that a weak link may exist between input and output data of multi-step-ahead models. However, the comparative evaluation of single and multi-step-ahead forecasts from the SOM–FFNN and SOM–WT–FFNN models demonstrated that use of the wavelet transform could produce forecasts that are more accurate. This could be because the wavelet transform extracts the periods and seasonalities that are important in rainfall–runoff modeling. The FFNN weighs the important periods of the signal, which may decrease the effects of redundant information that is not important in the conversion of rainfall to runoff. This can be interpreted as a noise reduction property of the wavelet transform.

Finally, the two FFNN-based (SOM–FFNN and SOM–WT–FFNN) and ARIMAX models were compared. The results indicate the superiority of FFNN-based models to ARIMAX type models due to the inherent ability of FFNNs to forecast nonlinear processes. The SOM–FFNN model, which simulates the process only based on spatial pre-processed data and uses the window input values of the previous time steps, performs much better than the ARIMAX model. This demonstrates the existence and importance of non-linearity and non-stationarity in the rainfall–runoff process.

9. Conclusion

Data pre-processing via SOM and wavelet transforms was shown to be useful in improving FFNN based rainfall–runoff forecasting models, and warrant further exploration. It should be noted that in general, and in the Gilgel Abay watershed in particular, rainfall and runoff time series are characterized by non-linearity, non-stationarity and seasonality. FFNN models may be unable to cope with these features without pre-processing of the input or/and output data.

In this study, the SOM, FFNN, and wavelet transform concepts were combined for the first time to develop hybrid black box models for multivariate daily and multi-step ahead rainfall–runoff forecasting using two satellite rainfall datasets, as well as runoff data. First, an un-supervised ANN technique (i.e. SOM) was used to determine the dominant zones of each satellite data set among all zones, which best represent the rainfall pattern over the watershed. The selection of dominant zones reduced time and labor in modeling, as the dimensionality of the original dataset was reduced as well as the number of trial–error process required to optimize the model. Furthermore, the positions of the selected zones were compatible with the topological characteristics of the watershed. Following spatial pre-processing, the FFNN rainfall–runoff model was constructed to find the non-linear relationship between the selected precipitation data and runoff. Multi-time-step forecasts were carried out in addition to single-step-ahead forecasts of runoff. The performance of the multi-step-ahead FFNN-based models decreased for steps greater than 1 day ahead.

In order to improve model efficiency and consider seasonality effects, the wavelet transform, which can capture the multi-scale features of a signal, was used to decompose the runoff time series into different sub-signals or levels. The sub-signals were then used as inputs to the FFNN model to predict the runoff discharge. The hybrid SOM–WT–FFNN model resulted in an improvement in rainfall–runoff modeling compared to the model that only incorporated spatial pre-processing (i.e. SOM–FFNN). Finally, the SOM–FFNN, SOM–WT–FFNN and conventional ARIMAX models were compared. Results indicate that FFNN based models are more suitable than the linear ARIMAX model, which cannot cope with the non-linear characteristics of the rainfall–runoff process.

Overall, the results of this study provide promising evidence for combining spatial and temporal data pre-processing methods, and more specifically the SOM and WT methods, to forecast runoff values using the FFNN method. In order to build on the current study, it is recommended in the future that the proposed SOM–WT–FFNN method is used to forecast runoff by adding other hydrological time series and variables, such as temperature and/or evapotranspiration, to the input layer of the model. Moreover, due to the uncertainty of the rainfall process and the ability of the Fuzzy concept to handle uncertainties, the conjunction of the ANN and fuzzy inference system (FIS) models as an adaptive neural-fuzzy inference system (ANFIS) model, could provide useful results. It would also be useful to apply the proposed methodology on other heterogeneous watersheds in order to investigate the overall effect of watershed climatic conditions on the performance of SOM–WT–FFNN model.

Acknowledgement

This study was partially funded by an NSERC Discovery Grant held by Jan Adamowski.

References

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

Adamowski, J., 2008b. River flow forecasting using wavelet and cross-wavelet transform models. Hydrol. Process. 22, 4877–4891.

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

Adamowski, J., Chan, H.F., Prasher, S.O., Ozga-Zielinski, B., Sliusarieva, 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 forecasting in Montreal, Canada. Water Resources Res. 48, W01528. http://dx.doi.org/10.1029/2010WR009945.

Addison, P.S., Murrary, K.B., Watson, J.N., 2001. Wavelet transform analysis of open channel wake flows. J. Eng. Mech. 127 (1), 58–70.

Akhtar, M.K., Corzo, G.A., Van Andel, S.J., Jonoski, A., 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges River basin. Hydrol. Earth Syst. Sci. 13, 1607–1618.

Anctil, F., Tape, G.D., 2004. An exploration of artificial neural network rainfall runoff forecasting combined with wavelet decomposition. J. Environ. Eng. Sci. 3, 121–128.

Antar, M.A., Elassiouti, I., Alam, M.N., 2006. Rainfall–runoff modeling using artificial neural networks technique: a Blue Nile catchment case study. Hydrol. Process. 20 (5), 1201–1216.

Artan, G., Gadain, H., Smith, J.L., Asante, K., Bandaragoda, C.J., Verdin, J.P., 2007. Adequacy of satellite derived rainfall data for stream flow modeling. Nat. Hazards. 43, 167–185.

ASCE task committee on application of Artificial Neural Networks in hydrology, 2000. Artificial neural networks in hydrology 2: hydrologic applications. J. Hydrol. Eng. 5 (2), 124–137.

Aussem, A., Campbell, J., Murtagh, F., 1998. Wavelet-based feature extraction and decomposition strategies for financial forecasting. J. Comput. Intell. Finance 6 (2), 5–12.

Bitew, M.M., Gebremichael, M., 2011. Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. Water Resources Res. 47, W06526. http://dx.doi.org/10.1029/2010WR009917.

Bowden, G.J., Dandy, G.C., Maier, H.R., 2005. Input determination for neural network models in water resources applications. Part 1 – Background and methodology. J. Hydrol. 301, 75–92.

Box, G.E.P., Jenkins, G., 1976. Time Series Analysis: Forecasting and Control, second ed. Holden-Day, San Francisco.

Cai, S., Toral, H., Qiu, J., Archer, J.S., 1994. Neural network based objective flow regime identification in air–water two phase flow. Can. J. Chem. Eng. 72, 440–445.

Chang, F.J., Chang, Y.M., Chang, L.C., 2007. Multi-step-ahead neural networks for flood forecasting. Hydrol. Sci. J. 52 (1), 114–130.

Cleaveland, M.K., Stahle, D.W., 1989. Tree ring analysis of surplus and deficit runoff in the White River, Arkansas. Water Resources Res. 25 (6), 1391–1401.

Dawson, C.W., Wilby, R., 1998. An artificial neural network approach to rainfall–runoff modeling. J. Hydrol. 43, 47–66.

Foufoula-Georgiou, E., Kumar, P., 1995. Wavelet in Geophysics, first ed. Academic Press, New York.

Furundzic, D., 1998. Application of neural networks for time series analysis: rainfall–runoff modelling. Signal Process. 64, 383–396.

Graumlich, L.J., 1987. Precipitation variation in the Pacific Northwest (1675–1975) as reconstructed from tree rings. Ann. Assoc. Am. Geogr. 77 (1), 19–29.

Grimes, D.I.F., Coppola, E., Verdecchia, M., Visconti, G., 2003. A neural network approach to real-time rainfall estimation for Africa using satellite data. J. Hydrometeorol. 4, 1119–1133.

Grossmann, A., Morlet, J., 1984. Decomposition of Hardy function into square integrable wavelets of constant shape. J. Math. Anal. 5, 723–736.

Hagan, M.T., Menhaj, M.B., 1994. Training feed-forward networks with the Marquardt algorithm. IEEE Trans. Neural. Netw. 5 (6), 989–993.

Hansen, J.V., Nelson, R.D., 1997. Neural networks and traditional time series methods: a synergistic combination in state economic forecasts. IEEE Trans. Neural. Netw. 8 (4), 863–873.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feed-forward networks are universal approximators. Neural Networks 2 (5), 359–366.

Hsu, K., Li, S., 2010. Clustering spatial–temporal precipitation data using wavelet transform and self-organizing map neural network. Adv. Water Resources 33, 190–200.

Hsu, K., Gupta, H.V., Sorooshian, S., 1995. Artificial neural network modeling of rainfall runoff process. Water Resources Res. 31, 2517–2530.

Hsu, K., Gupta, H.V., Gao, X., Sorooshian, S., Imam, B., 2002. Self-organizing linear output map (SOLO): an artificial neural network suitable for hydrologic modeling and analysis. Water Resources Res. 38 (12), 1–38.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM multi-satellite precipitation analysis (TMPA): quasi-global, multilayer, combined-sensor, precipitation estimates at fine scale. J. Hydrometeorol. 8, 38–55.

Jain, A., Sudheer, K.P., Srinivasulu, S., 2004. Identification of physical processes inherent in artificial neural network rainfall–runoff models. Hydrol. Process. 18, 571–581.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimation from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5, 487–503.

Kalteh, A.M., Hjorth, P., Berndtsson, R., 2008. Review of self-organizing map (SOM) in water resources: analysis, modeling, and application. Environ. Model. Softw. 23, 835–845.

Kohonen, T., 1997. Self-Organizing Maps. Springer-Verlag, Berlin, Heidelberg.

Kohonen, T., 1998. The self-organizing map. Neurocomputing 21, 1–6.

Labat, D., 2005. Recent advances in wavelet analyses: Part 1 – A review of concepts. J. Hydrol. 314, 275–288.

Labat, D., Ababou, R., Mangin, A., 2000. Rainfall–runoff relation for karstic spring. Part 2: Continuous wavelet and discrete orthogonal multi resolution analyses. J. Hydrol. 238, 149–178.

Lábó, E., 2012. Validation studies of precipitation estimates from different satellite sensors over Hungary – analysis of new satellite-derived rain rate products for hydrological purposes. J. Hydrol. http://dx.doi.org/10.1016/j.jhydrol.2012.08.031 (Published online)

Lallahem, S., Maina, J., 2003. A nonlinear rainfall–runoff model using neural network technique: example in fractured porous media. Math. Comput. Modell. 37, 1047–1061.
Legates, D.R., McCabe Jr., G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydro-climatic model validation. Water Resources Res. 35 (1), 233–241.
Lin, G.F., Wu, M.C., 2007. A SOM-based approach to estimate design hyetographs of un-gauged sites. J. Hydrol. 339, 216–226.
Lin, G.F., Wu, M.C., 2011. An RBF network with a two-step learning algorithm for developing a reservoir inflow forecasting model. J. Hydrol. 405, 439–450.
Liu, Y., Weisberg, R.H., 2011. A review of self-organizing map applications in meteorology and oceanography. In: Mwasiagi, J.I. (Ed.), Self-Organizing Map-Applications and Novel Algorithm Design. InTech, Rijeka, Croatia.
Mallat, S.G., 1998. A Wavelet Tour of Signal Processing, second ed. Academic Press, San Diego.
Mirbagheri, S.A., Nourani, V., Rajaee, T., Alikhani, A., 2010. Neuro-Fuzzy models employing wavelet analysis for suspended sediment concentration prediction in rivers. Hydrol. Sci. J. 55 (7), 1175–1189.
Nourani, V., Kalantari, O., 2010. Integrated artificial neural network for spatiotemporal modeling of rainfall–runoff–sediment processes. Environ. Eng. Sci. 27 (5), 411–422.
Nourani, V., Mano, A., 2007. Semi-distributed flood runoff model at the sub-continental scale for southwestern Iran. Hydrol. Process. 21, 3173–3180.
Nourani, V., Komasi, M., Mano, A., 2009a. A multivariate ANN–wavelet approach for rainfall–runoff modeling. Water Resources Manage. 23, 2877–2894.
Nourani, V., Alami, M.T., Aminfar, M.H., 2009b. A combined neural–wavelet model for prediction of Lighvanchai watershed precipitation. Eng. Appl. Artif. Intell. 16, 1–12.
Nourani, V., Kisi, Ö., Komasi, M., 2011. Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. J. Hydrol. 402, 41–59.
Nourani, V., Hosseini Baghanam, A., Daneshvar, F., Alami, M.T., 2012a. Classification of groundwater level data using SOM to develop ANN-based forecasting model. Int. J. Soft Comput. Eng. 2 (1), 464–469.
Nourani, V., Kalantari, O., Hosseini Baghanam, A., 2012b. Two semi-distributed ANN-based models for estimation of suspended sediment load. J. Hydrol. Eng.. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000587 (Published on-line).
Pankratz, A., 1991. Forecasting with Dynamic Regression Models. John Wiley & Sons, Inc., New York.
Partal, T., Cigizoglu, H.K., 2008. Estimation and forecasting of daily suspended sediment data using wavelet–neural networks. J. Hydrol. 358 (3–4), 317–331.
Partal, T., Kisi, O., 2007. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. J. Hydrol. 342, 199–212.
Petty, G.W., 1995. The status of satellite-based rainfall estimation over land. Remote Sens. Environ. 51, 125–137.
Pulido-Calvo, I., Portela, M.M., 2007. Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. J. Hydrol. 332, 1–15.
Sajikumara, N., Thandaveswara, B.S., 1999. A non-linear rainfall–runoff model using an artificial neural network. J. Hydrol. 216, 32–55.
Salas, J.D., Delleur, J.W., Yevjevich, V., Lane, W.L., 1980. Applied Modeling of Hydrological Time Series, first ed. Water Resources Publications, Littleton.
Sawunyama, T., Hunghes, DA., 2010. Using satellite-based rainfall data to support the implementation of environmental water requirements in South Africa. Water SA 136 (4), 379–385.
Senthil Kumar, A.R., Sudheer, K.P., Jain, S.K., Agarwal, P.K., 2004. Rainfall–runoff modeling using artificial neural network: comparison of networks types. Hydrol. Process. 19 (6), 1277–1291.
Shrestha, M.S., Artan, G.A., Bajracharya, S.R., Sharma, R.R., 2008. Using satellite-based rainfall estimates for stream flow modeling: Bagmati Basin. J. Flood Risk Manage. 1, 89–99.
Sudheer, K.P., Gosain, A.K., Ramasastri, K.S., 2000. A data-driven algorithm for constructing artificial neural network rainfall–runoff models. Hydrol. Process. 16 (6), 1325–1330.
Tankersley, C., Graham, W., Hatfield, K., 1993. Comparison of univariate and transfer function models of groundwater fluctuations. Water Resources Res. 29 (10), 3517–3533.
Tokar, A.S., Johnson, P.A., 1999. Rainfall–runoff modeling using artificial neural networks. J. Hydrol. Eng. 4, 232–239.
Toth, E., 2009. Classification of hydro-meteorological conditions and multiple artificial neural networks for stream flow forecasting. Hydrol. Earth Syst. Sci. 13, 1555–1566.
Vos, N.J., Rientjes, T.H.M., 2005. Constraints of artificial neural networks for rainfall–runoff modelling: trade-offs in hydrological state representation and model evaluation. Hydrol. Earth Syst. Sci. 9, 111–126.
Wang, W., Ding, S., 2003. Wavelet network model and its application to the predication of hydrology. Nat. Sci. 1 (1), 67–71.
Yonaba, H., Anctil, F., Fortin, V., 2010. Comparing sigmoid transfer functions for neural network multistep ahead stream flow forecasting. J. Hydrol. Eng. 15 (4), 275–283.
Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50, 159–175.
Zhang, Q., Wang, B.D., He, B., Peng, Y., Ren, M.L., 2011. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. Water Resources Manage. 25, 2683–2703.