Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: https://www.tandfonline.com/loi/thsj20

Monthly streamflow forecasting using neuro-wavelet techniques and input analysis

André Gustavo da Silva Melo Honorato, Gustavo Barbosa Lima da Silva & Celso Augusto Guimarães Santos

To cite this article: André Gustavo da Silva Melo Honorato, Gustavo Barbosa Lima da Silva & Celso Augusto Guimarães Santos (2018) Monthly streamflow forecasting using neuro-wavelet techniques and input analysis, Hydrological Sciences Journal, 63:15-16, 2060-2075, DOI: 10.1080/02626667.2018.1552788

To link to this article: https://doi.org/10.1080/02626667.2018.1552788

Published online: 09 Jan 2019.

HYDROLOGICAL SCIENCES JOURNAL
2018, VOL. 63, NOS. 15–16, 2060–2075
https://doi.org/10.1080/02626667.2018.1552788

Monthly streamflow forecasting using neuro-wavelet techniques and input analysis

André Gustavo da Silva Melo Honorato[a], Gustavo Barbosa Lima da Silva[b] and Celso Augusto Guimarães Santos[b]

[a]Technische Universität Dresden, Dresden, Germany; [b]Department of Civil and Environmental Engineering, Federal University of Paraíba, João Pessoa, Brazil

ABSTRACT

Combinations of low-frequency components (also known as approximations) resulting from the wavelet decomposition are tested as inputs to an artificial neural network (ANN) in a hybrid approach, and compared to classical ANN models for flow forecasting for 1, 3, 6 and 12 months ahead. In addition, the inputs are rewritten in terms of the flow, revealing what type of information was being provided to the network, in order to understand the effect of the approximations on the forecasting performance. The results show that the hybrid approach improved the accuracy of all tested models, especially for 1, 3 and 6 months ahead. The input analyses show that high-frequency components are more important for shorter forecast horizons, while for longer horizons, they may worsen the model accuracy.

ARTICLE HISTORY
Received 19 December 2017
Accepted 11 October 2018

EDITOR
R. Woods

ASSOCIATE EDITOR
A. Efstratiadis

KEYWORDS
monthly streamflow forecasting; artificial neural networks; wavelet transform; low frequency

Introduction

Precise streamflow prediction is essential for many activities relating to water resources management, such as flood and drought control, reservoir operation, water supply planning and hydro-electric power generation. While both short- and long-term forecasts are important, reservoir operations are usually planned based on monthly periods; thus, monthly streamflow forecasting plays a major role in water resources management.

Nevertheless, predicting streamflow is difficult because hydrological systems are complex and are subject to many inherent uncertainties. In addition, river flow is the result of interaction among different properties and nonlinear processes, which vary across different spatial and temporal scales, predominantly precipitation, soil characteristics, land cover, interception, evapotranspiration and water usage.

Due to its great relevance, many different approaches have been tested in hydrology for flow forecasting, including adaptations of the classical autoregressive (AR) and moving average (MA) models and combinations of the two (ARMA) (Ciğizoğlu 2003, Koutsoyannis *et al.* 2008), fuzzy rule-based systems, genetic programming and model trees (Londhe and Charhate 2010) and artificial neural networks (ANNs) with different algorithms, such as the backpropagation, real-time recurrent learning and Kalman filters (Muluye 2011).

Because of their ability to simulate highly nonlinear relationships, ANNs have become a very popular tool among hydrologists (ASCE Task Committee 2000a). ANNs are black-box models that relate inputs and outputs using transfer functions; despite not having any physical basis, they have gained attention due to their good predictive capacity (Kişi and Ciğizoğlu 2007).

In hydrology, ANNs are being used for multiple purposes: groundwater modelling (Coulibaly *et al.* 2001, Daliakopoulos *et al.* 2005, Nayak *et al.* 2006, Mohanty *et al.* 2010), water quality modelling (Singh *et al.* 2009, Gazzaz *et al.* 2012), evapotranspiration estimation (Kumar *et al.* 2002, Trajkovic *et al.* 2003, Zanetti *et al.* 2007), precipitation prediction (Partal and Ciğizoğlu 2009) and sediment transport (Nagy *et al.* 2002, Tayfur 2002, Ciğizoğlu 2004, Ciğizoğlu and Alp 2006, Alp and Ciğizoğlu 2007, Melesse *et al.* 2011).

One of the most common uses of ANNs in hydrology is for streamflow prediction at multiple time scales, particularly monthly (Kişi 2004, Ciğizoğlu 2005, Jeong and Kim 2005, Prada-Sarmiento and Obregón-Neira 2009, Machado *et al.* 2011, Oliveira *et al.* 2014) and daily (Rajurkar *et al.* 2002, Sudheer *et al.* 2002, Riad *et al.* 2004, Kumar *et al.* 2005, Bravo *et al.* 2008, Cruz *et al.* 2010, Gomes *et al.* 2010, Debastiani *et al.* 2016).

Regardless of the good results achieved in modelling hydrological time series through ANNs, in some cases

CONTACT André Gustavo da Silva Melo Honorato andremelohonorato@gmail.com

the underlying processes are characterized by high nonlinearity and nonstationarity, which makes it difficult to provide reliable forecasts (Wang and Ding 2003, Cannas *et al.* 2006, Nourani *et al.* 2009). Due to this, researchers began to combine the ANN technique with wavelet transforms to pre-process the input data before introducing them to the ANN. This integrated analysis is drawing attention for surpassing the single ANN approach in terms of prediction accuracy and good fitting (Wei *et al.* 2013, Yaseen *et al.* 2015). The wavelet transform (WT) is a mathematical procedure that allows multiresolution analysis in the time–frequency domain in order to identify nonstationary variances at different scales (Torrence and Compo 1998, Kişi 2009).

One of the first works that used a combination of WT and ANN to predict streamflow was reported by Wang and Ding (2003). It employed the discrete wavelet transform (DWT) to decompose three times the daily discharge time series for the Yangtze River (China) at Cuntan Station. They then used the decomposed signal as input for a three-layer ANN, training it to forecast streamflow for 1, 2, 3, 4 and 5 days ahead. They also developed a threshold autoregressive model (TAR) to compare the results and verified the good performance of the hybrid model, especially for longer forecast horizons.

Cannas *et al.* (2006) tested the continuous wavelet transform (CWT), DWT and data partitioning prior to training a three-layered ANN for predicting monthly streamflow for one month ahead in a Sardinian catchment, where the ANN approach did not succeed. They concluded that pre-processing data with the DWT, decomposing the signal in three levels, gave the best results for the examined wavelet–neural network models.

Kişi (2008) also utilized DWT to decompose a monthly streamflow series in five levels. In particular, he analysed the correlation between the high-frequency components (referred to as details) resulting from the wavelet decomposition process and future flow and used as input only the most significant details. He tested this methodology in two Turkish rivers and compared the results with autoregressive models, multilinear regressions and a traditional ANN model, concluding that using previous flow processed with DWT as input for ANN improves the accuracy of the forecasting.

Kişi (2009) tested the aforementioned methodology to simulate intermittent daily flows for 1 day ahead. Such data exhibit significant difficulties for models because of their zero values. As a result, while the ANN model predicted negative values, the hybrid model was able to cope with the zero values and considerably improved the performance when compared to the classical ANN.

Similarly to Kişi's work, Partal (2009) also investigated the correlation between the eight times decomposed time series and future monthly flow. In this context, he selected the most relevant details and summed them, reducing the number of connections in the ANN and, thus, the number of weights. Moreover, he tested different types of ANN and hybrid models for two stations in the Beyderesi River (Turkey), concluding that the wavelet–feedforward back-propagation ANN was a useful tool for monthly streamflow prediction.

Pramanik *et al.* (2010) conducted a similar study for daily flow data in the Brahmani River, in India. They used the DWT to decompose the time series into 10 time scales and analysed the correlation of the details and several future flows for 1, 2, 3, 4 and 5 days ahead. They then summed the most relevant details and used them as inputs to the ANN. The results showed that pre-processing of data with wavelet transform enhanced the accuracy of the models for all prediction horizons.

Adamowski and Sun (2010) developed different versions of daily streamflow forecasting models for 1 and 3 days ahead, applied at two nonperennial rivers in Cyprus. They decomposed the time series eight times, used all details and approximations as inputs for the models, and concluded that the hybrid model is a promising tool for nonperennial rivers in semi-arid regions.

Mehr *et al.* (2013) investigated the WT–ANN technique for monthly predictions in the Çoruh River in the Black Sea Region using a downstream–upstream approach, and they compared it with the classical ANN. The results showed that, while both models performed well, the WT–ANN performed slightly better.

Wei *et al.* (2013) developed a hybrid model to predict streamflow for 12 future months, using as inputs the past 12 months. They used the Daubechies wavelet of order 5 (db5) as the mother wavelet to decompose the time series five times (producing six sub-series) in the Weihe River, a tributary of the Yellow River, in China. The results for the hybrid approach were significantly better than the results for the classical ANN.

Makwana and Tiwari (2014) tested the effect of using different parts of a time series for training, different mother wavelets and different training algorithms within a hybrid model for daily streamflow forecasting. Their model achieved better fitting in extreme events (which are particularly difficult to reproduce) when compared to classical ANNs.

As shown, most of the above works focus on using the details, i.e. high-frequency components of the signal, obtained from the wavelet transform as input for the ANN. However, Santos and Silva (2014) tested the use of approximations (i.e. low-frequency components), or a sum of them, as inputs for a daily streamflow forecasting model with increasing time horizons (1, 3, 5 and 7 days ahead) for the São Francisco River basin in Brazil. They

concluded that the use of approximations as inputs significantly increased the accuracy of the model when compared to the typical ANN model, especially for 5- and 7-day horizons.

We could not find any work merely using approximations for monthly flow forecasting. Therefore, the main purpose of this article is to develop a hybrid model based on ANN and DWT using only low-frequency components or linear combinations of them as inputs to forecast monthly streamflow for 1, 3, 6 and 12 months ahead. A classical ANN model was also developed to serve as a comparative framework for the hybrid model. Furthermore, we aimed to provide new insight to understand the influence of this approach on model performance. In this sense, we propose an analysis of the used inputs to reveal what type of information improves the accuracy of the forecasts. The methodology was tested on a challenging flow forecasting problem in Brazil, involving an important hydro-electric reservoir in the semi-arid region.

Materials and methods

Study area and data

In several countries, a significant portion of the electricity generated comes from hydro-electric plants. For instance, in Brazil these plants produce 64% of electrical energy (EPE (Empresa de Pesquisa Energética) 2016). This further highlights the significance of precisely predicting inflows to hydro-electric reservoirs for several months ahead.

The objective of this research is the prediction of monthly inflows to the Sobradinho Dam, one of the largest artificial lakes in the world, located on the São Francisco River, Northeast Brazil (Fig. 1). It covers an area of 4214 km^2, feeds a hydro-electric plant with a power capacity of 1050 MW, drains an area of 498 968 km^2 and regulates the water resources of the broader region. It is important to emphasize that, due its large catchment, past and future flows have a strong correlation, which strongly influences (i.e. favours) the model performance.

The São Francisco River runs for 2697 km and its basin has a total area of 638 883 km^2, representing 7.5% of the entire Brazilian territory. Regarding its climate, the mean temperature ranges from 18 to 27°C, it has high solar incidence and two well-defined seasons, one rainy and one dry, with a mean rainfall of 1036 mm/year.

The data used in this research are the re-naturalized mean monthly inflow to the Sobradinho Dam and were made available by the National System Operator (ONS, Operador Nacional do Sistema), which is responsible for

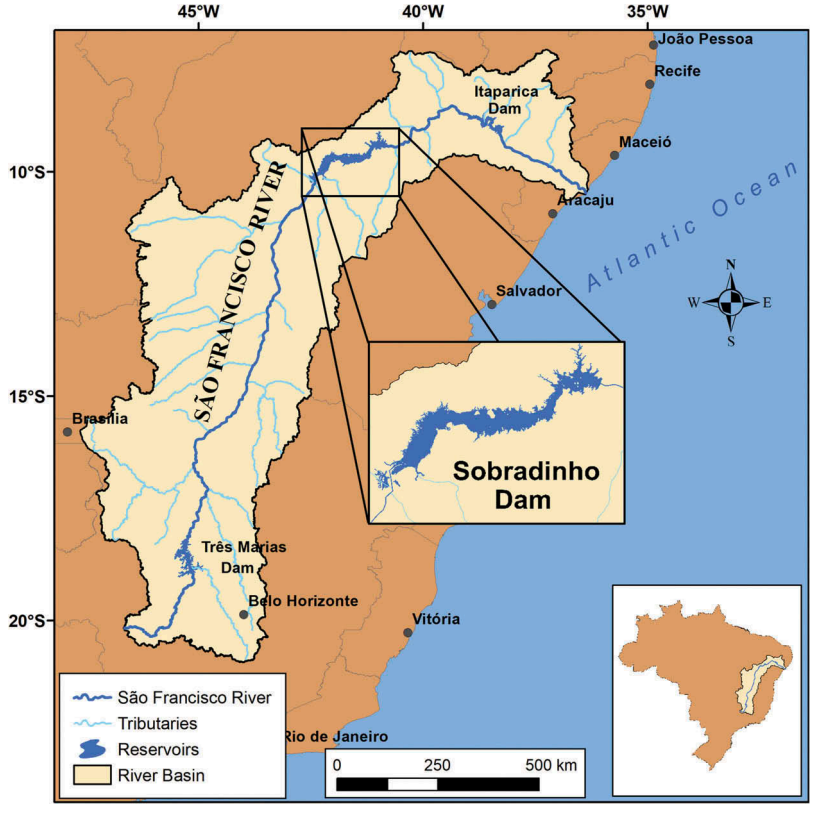

Figure 1. São Francisco River basin, Brazil, highlighting the Sobradinho Dam.

the operation of the interlinked electrical system in Brazil. A total of 940 values were available, from January 1931 to April 2009 (Fig. 2). The flow data exhibit high variability across hydrological years; for example, in May 2001 a flow of 845 $m^3 s^{-1}$ was observed, while in May 1945 a value of 8764 $m^3 s^{-1}$ was observed, which is more than 10 times higher. Table 1 shows the descriptive statitics for the analysed period.

Table 1. Descriptive statistics for the inflow time series (1931–2009).

Statistic	Value	Units
Arithmetic mean	2 666.30	$m^3 s^{-1}$
Mode	2 871.81	$m^3 s^{-1}$
Median	1 919.03	$m^3 s^{-1}$
Harmonic mean	1 707.63	$m^3 s^{-1}$
Geometric mean	2 107.40	$m^3 s^{-1}$
Range	15 170.04	$m^3 s^{-1}$
First quartile	1 187.96	$m^3 s^{-1}$
Third quartile	3 751.36	$m^3 s^{-1}$
Inter-quartile range	2 563.41	$m^3 s^{-1}$
Median absolute deviation	1 537.60	$m^3 s^{-1}$
Variance	3 837 145.80	$(m^3 s^{-1})^2$
Standard deviation	1 958.86	$m^3 s^{-1}$
Coefficient of variation	0.73	-
Skewness	1.65	-
Kurtosis	7.02	-

General description of methodology

The main objective of this work was to establish predictive models for future monthly flows (1, 3, 6 and 12 months ahead), based on ANNs and hybrid models. These were, in turn, based on ANN and wavelet transform using past flows as inputs. The results of both models for each time horizon were then compared.

Specifically, the flow time series were initially used to establish an ANN-based prediction model, in order to obtain a comparative framework for the performance of the hybrid model. The models that only use ANNs are symbolized as ANN_m, where m is the forecast horizon (1, 3, 6 and 12 months ahead). Furthermore, the wavelet transform was employed 10 times to decompose the flow time series into low-frequency (approximations) and high-frequency (details) components. Next, for the hybrid models, the approximations or a sum of them were used as input data for ANNs (the hybrid models are called wavelet–ANN_m, where m is the forecast horizon, i.e. 1, 3, 6 and 12 months). Finally, the performance of the models was evaluated through three statistical parameters and the inputs used for the hybrid models were analysed. A scheme of the general description of the methodology is shown in Figure 3.

Artificial neural networks

Artificial neural networks are empirical models composed of simple processing units distributed in parallel and connected to each other, which are inspired by the biological nervous system (Demuth and Beale 2002). Individually, artificial neurons have limited computing power, regardless of the activation function. However, by combining several elements in a network of neurons, arranged in layers, they become capable of solving high-complexity problems. Plenty of information about ANNs is available in the extant literature (ASCE Task Committee 2000a, 2000b, Demuth and Beale 2002, Haykin 2005); thus, the ANN's description here will be limited to its comprehension.

Multilayer perceptron (MLP) feedforward networks that have one or more hidden layers are the most common way of organizing ANNs in hydrology (ASCE Task Committee 2000a). This ANN structure has proved to be a powerful approach tool. Given that one hidden layer can theoretically approach any continuous function (Cybenko 1989), this structure was used in the following analyses.

The MLP (Fig. 4) is composed of an input layer that receives the data (x_1, x_2, …, x_n), which is progressively processed through the subsequent layers regulated by the weights (w_{ji} for the hidden layer and w_{kj} for the output) and the results (y_1, y_2, …, y_m) are calculated in the output layer. In general, the output y_k of a network with one hidden layer is given by:

$$y_k = f_0\left(\sum_{j=1}^{m} w_{kj} f_h\left(\sum_{i=1}^{n} w_{ji} x_i + b_j\right) + b_k\right) \quad (1)$$

where f_0 and f_h correspond to the activation functions, w_{ji} and w_{kj} are the weights, x_i is the input, and b is a bias.

Usually, the most common activation functions are the sigmoid functions, because they can be differentiated

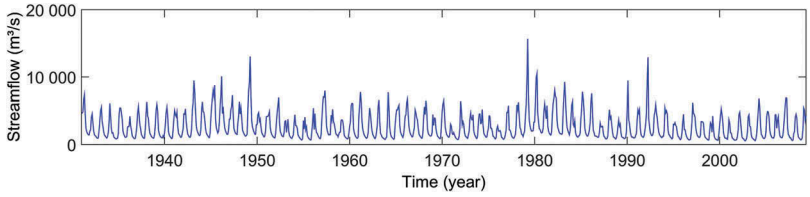

Figure 2. Time series of re-naturalized mean monthly inflows to the Sobradinho Dam.

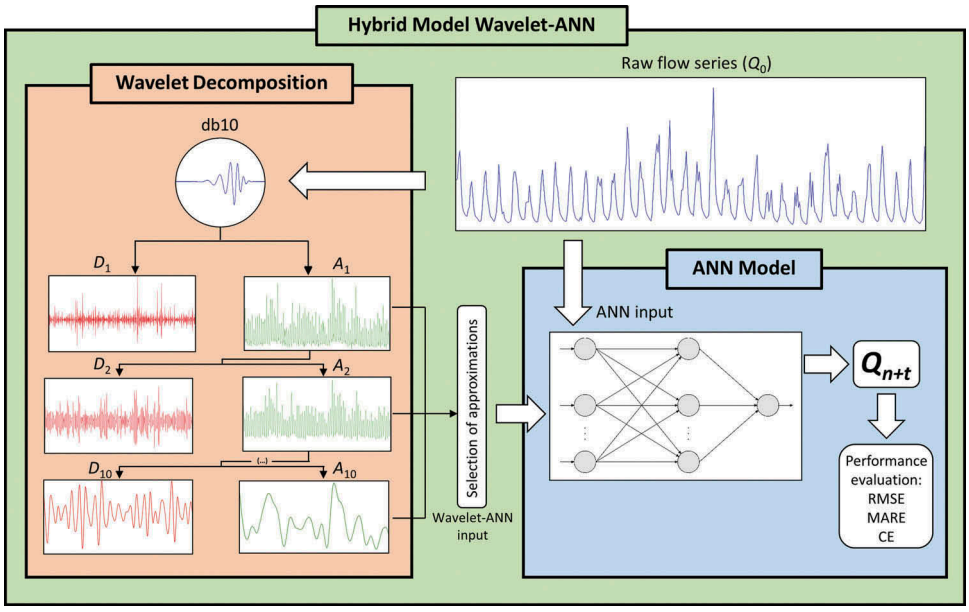

Figure 3. General scheme of the methodology, where *t* denotes the forecast horizon.

and due to their general behaviour (Haykin 2005). The definition of weights is performed through a network training process. During this training, a set of test data (inputs and associated outputs) is assigned to the network and the synaptic weights are adjusted iteratively to minimize the difference between the response provided by the network and the actual response. The learning level obtained at each iteration is generally quantified by means of typical statistical measures, such as the root mean square error (RMSE) (Haykin 2005). This procedure is repeated until the predictive capacity of the network is considered satisfactory.

Wavelet transform

The wavelet transform was developed to overcome the limitations of the windowed Fourier transform (WFT). This technique decomposes a one-dimensional series simultaneously in the time–frequency domain using short waves (mother wavelet) that preserve local, non-periodic, and multiscale phenomena, revealing information that the raw signal does not expose (Santos *et al.* 2013, 2019, Santos and Silva 2014). There are two main forms of wavelet transform, i.e. continuous (CWT) and discrete (DWT) (Wei *et al.* 2013). The DWT is used in this study.

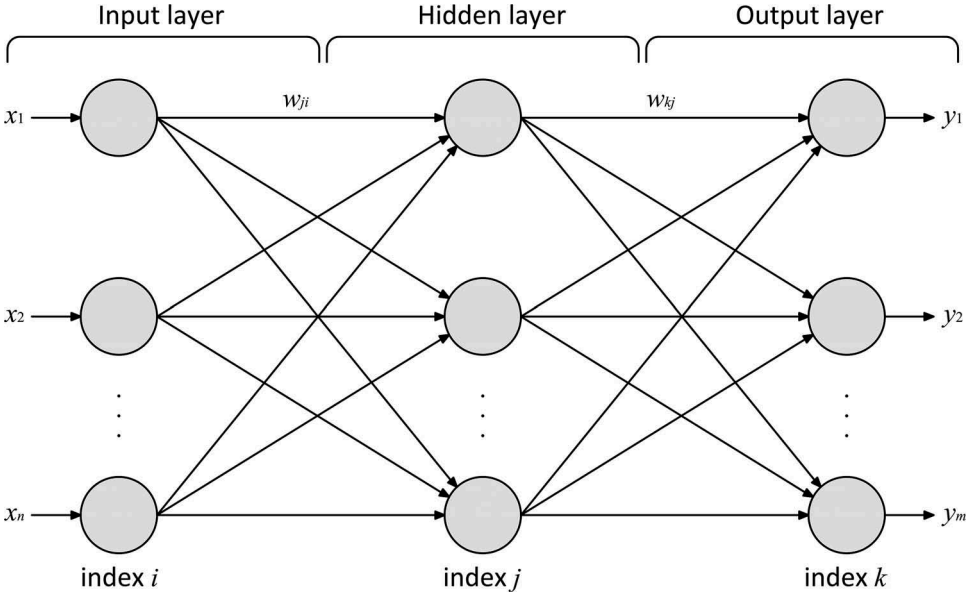

Figure 4. Multilayer perceptron (MLP) network with one hidden layer.

Mother wavelet

A mother wavelet (ψ_0) is a short duration wave that grows and decays in a limited period (Torrence and Compo 1998). This is crucial for the good performance of the transform; depending on the chosen wave, the method will filter specific information during the process (Holdefer and Severo 2015).

A wave to be considered as mother wavelet must satisfy two distinct properties. Its integral in time has to be zero:

$$\int_{-\infty}^{+\infty} \psi_0(t)\mathrm{d}t = 0 \tag{2}$$

and it must have unitary energy:

$$\int_{-\infty}^{+\infty} \left|\psi_0(t)\right|^2 \mathrm{d}t = 1 \tag{3}$$

Usually, for the DWT, the most commonly used mother wavelets are the Daubechies (Fig. 5), a family of orthogonal wavelets denoted by db$N$, where N is the indicator of their order. These waves do not have explicit expressions to characterize them (except for db1), and most of them are asymmetric (Santos and Silva 2014).

Discrete wavelet transform (DWT)

The DWT is generally used for time series decomposition and filtering, since it does not cause coefficient redundancies between scales and also because information about the position of certain events is not lost in the process (Daubechies 1992, Kişi 2009). The DWT is obtained from the discretization of the translation and expansion/compression parameters of the continuous form (Santos and Silva 2014). For its calculation, the simplest and most efficient method is that introduced by Mallat (1989), where the scale (a) and location (b) parameters are chosen based on powers of 2, called scales and dyadic positions, respectively, i.e.:

$$a = a_0^j,\ b = ka_0^j b_0 a_0 > 0 \tag{4}$$

where k is the translation factor and j the scaling factor. Then, the wavelet equation becomes:

$$\psi_{j,k}(t) = a_0^{-j/2}\psi_0\left(\frac{t - ka_0^j b_0}{a_0^j}\right) \tag{5}$$

which can be written as:

$$\psi_{j,k}(t) = a_0^{-j/2}\psi_0\left(a_0^{-j}t - kb_0\right) \tag{6}$$

Usually, $a_0 = 2$ and $b_0 = 1$ are chosen, thus the DWT becomes binary and the dyadic discrete wavelet transform is given by:

$$W_{j,k}(t) = 2^{-j/2}\sum_{j,k\in Z} f(t)\psi_0\left(2^{-j}t - k\right) \tag{7}$$

This simple algorithm causes DWT to operate as a bypass filter, quickly calculating the wavelet coefficients and thus decomposing the input signal into low- and high-frequency components (Misiti *et al.* 1996).

Wavelet as pass filter

The DWT operates as two functions that can be regarded as high-pass and low-pass filters (Fig. 6(a)). The original signal (S) passes through the low-pass and high-pass filters and the corresponding approximations (A) and details (D) are produced (Kişi 2008). The approximations produced by the low-pass filter correspond to the low-frequency signals of the time series, while the details are the high-frequency signals (Misiti *et al.* 1996). The decomposition may continue

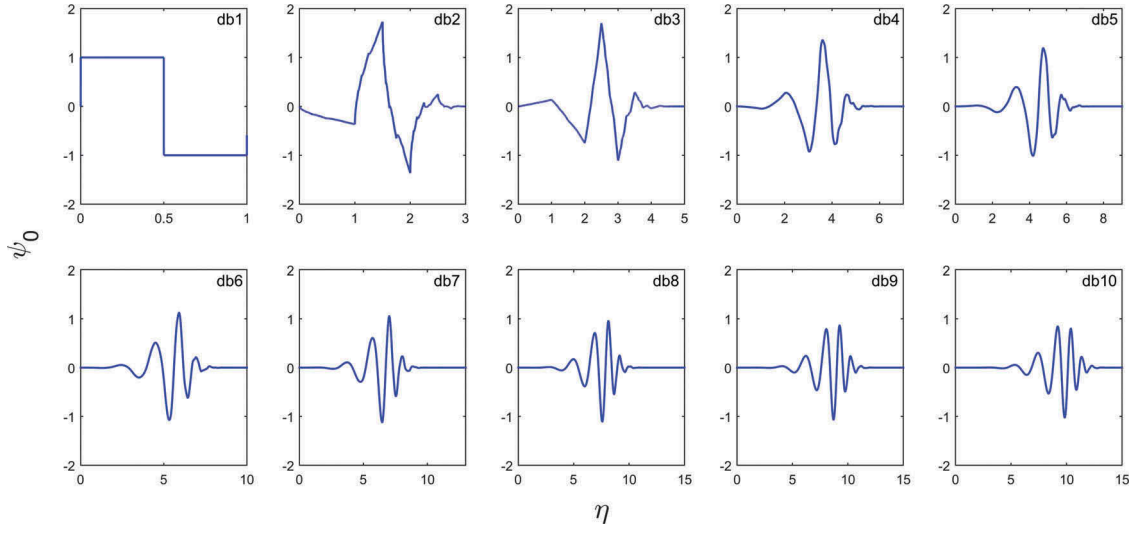

Figure 5. Family of Daubechies wavelets (order 1 to 10).

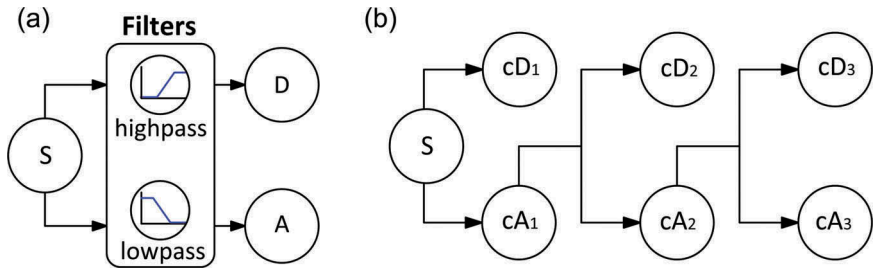

Figure 6. (a) The original signal (*S*) passes through two filters and is decomposed into complementary high-frequency (*D*) and low-frequency (*A*) signals. (b) Decomposition in three levels forming a wavelet decomposition tree.

in an iterative loop, with the approximations being decomposed at each step; thus the original signal is subdivided into several lower resolution components (Fig. 6(b)), a procedure called the wavelet decomposition tree (Misiti *et al.* 1996).

Another important feature of wavelet decomposition is that the decomposed signals can be reconstructed from sums of approximations and details (Wei *et al.* 2013). For example, the raw signal (*S*) is the result of the sum of the first level approximation and details (A_1, D_1) (Equation 8). The first level approximation (A_1) can be recomposed through the sum of the second level approximation (A_2) with the second level details (D_2) (Equation (9)), and also the raw signal (*S*) can be reconstituted by the sum of the third level approximation (A_3) with details of third level (D_3), second level (D_2) and first level (D_1) (Equation (10)). Specifically:

$$S = A_1 + D_1 \tag{8}$$

$$A_1 = A_2 + D_2 \tag{9}$$

$$S = A_3 + D_3 + D_2 + D_1 \tag{10}$$

In this work, Matlab® was used to decompose the flow time series Q_0 in 10 levels using the db10 mother wavelet, the Daubechies wavelet of order 10, due to its more detailed format, which permits better representation of the studied series that have previously been applied in the São Francisco River catchment by Santos and Silva (2014). The results of this decomposition can be seen in Figure 7.

Model implementation

ANN model

Four different ANNs for monthly streamflow prediction were developed, which are symbolized ANN_m, where *m* is the number of months ahead of the prediction (in our analyses we considered 1, 3, 6 and 12 months). Assuming that the streamflow is the response of a dynamic system with a functional dependence between its future and past occurrences, past flow values were used as inputs to predict future flows.

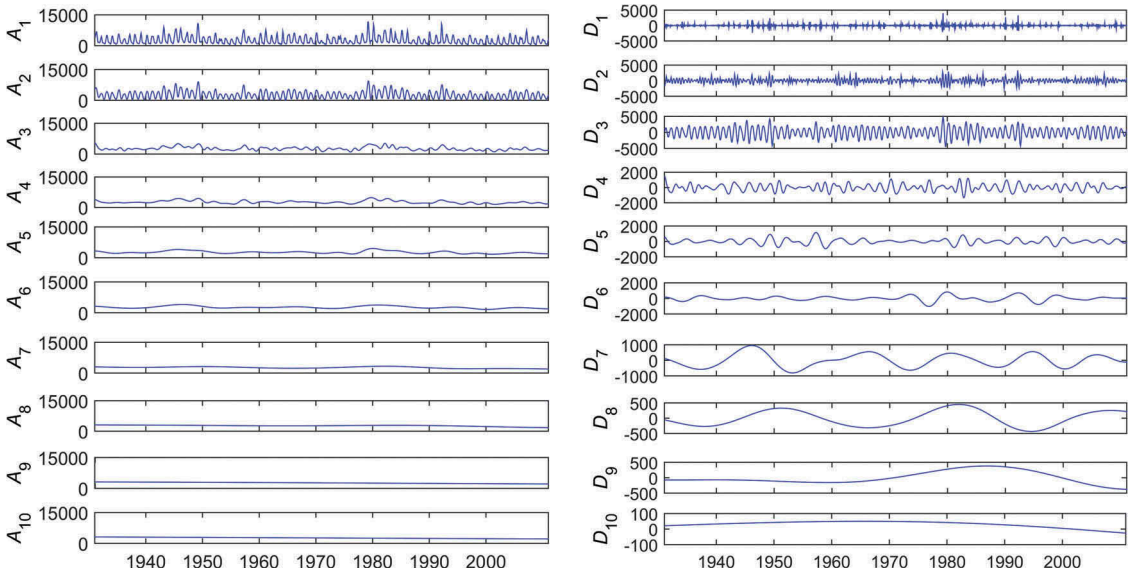

Figure 7. Approximations and details resulting from the decomposition procedure.

In this work, the number of neurons in the input and hidden layers was determined using a trial-and-error procedure. In particular, for the output layer (k) only one neuron was used, since the objective was to predict just one future flow value.

For the input layer, the process was started with three neurons (n = 3) and the values were tested in increments of 3 until reaching 12 neurons. Similarly, for the determination of the number of neurons in the hidden layer (m), the same process was employed, except for the fact that the increment was up to 15 instead of 12. Then, it was evaluated which architecture (n-m-k) had the best performance to be used in the predictive model.

Hybrid model wavelet–ANN

The methodology for establishing the neural networks for both models, ANN and wavelet–ANN, was the same, only their inputs differed. In particular, the classical ANN used the raw flows as inputs, while the wavelet–ANN model used the low-frequency components resulting from the wavelet decomposition of the original flow time series (Fig. 8).

The trial-and-error method was employed for the determination of the input (I) for the wavelet–ANN model. Different levels of decomposition were individually tested as input data for the network, starting with $I = A_1$ and continuing ($I = A_2$, $I = A_3$, …) until the approximations used as input no longer had the essential components to represent the behaviour of the time series and, therefore, the network could no longer generalize the results of the learning process.

In addition to testing the approximations individually, several combinations of sums between them (e.g. $I = A_1 + A_2$; $A_1 + A_2 + A_3$; $A_2 + A_3$) were also tested. Approximately 30 different inputs were tested for each prediction horizon (1, 3, 6 and 12 months) and the hybrid model that was finally established was the one ensuring the best predictive capacity (Table 2).

Standardization and data preparation

Before being used as input for the two models, i.e. ANN and wavelet–ANN, the data were standardized in the range [−1,1], assuming the output values of the activation functions used in this work are within that range. Other types of standardization may also be used, such as linear transformations or statistical standardizations (Shanker *et al.* 1996).

Training procedure

For both models, the time series were divided into two subsets, one for training (660 values) and the other for testing (280 values). The training subset was used in the learning procedure of the network, while the testing set was used to find whether the network could generalize the learning results and predict the flows for an independent period. The Levenberg-Marquardt algorithm was chosen for the training with μ = 0.001, where μ is the learning rate, a scalar that controls the adjustment of the weights of a network. Although this algorithm requires a higher computational and memory effort, it has been used successfully in small ANNs, generating fast and stable results and decreasing the frequency of entrapment in local minimums (Daliakopoulos *et al.* 2005). With respect to the stopping criteria for the training procedure, we assigned a maximum number of 600 iterations.

Two combinations of activation functions for each model were tested: two sigmoid tangents (TS–TS) and a sigmoid tangent followed by linear (TS–PL); the best

Table 2. Approximations used as inputs for hybrid models.

Model	Used approximation
Wavelet–ANN$_1$	$I = A_1 + A_2$
Wavelet–ANN$_3$	$I = A_2 + A_3$
Wavelet–ANN$_6$	$I = A_2 + A_3 + A_4$
Wavelet–ANN$_{12}$	$I = A_2 + A_3 + A_4$

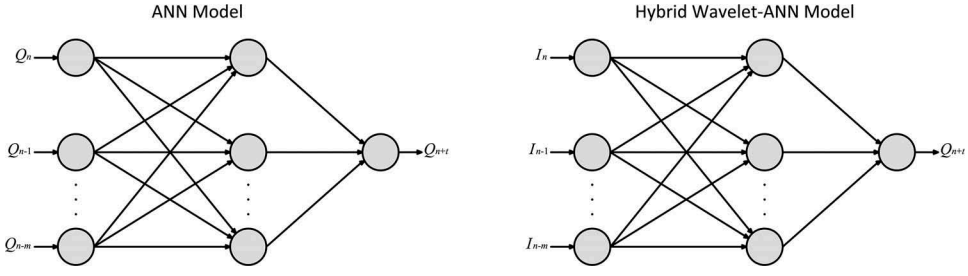

Figure 8. Differences between ANN and wavelet–ANN inputs, where t represents the lead time of the forecast (1, 3, 6 and 12 months).

performers were chosen. After its training, the network was applied to provide flow forecasts for the period related to the testing subset, and its performance was evaluated.

Performance evaluation

In addition to graphical analysis, three statistical metrics were used to verify the similarity of the data predicted by the models and those actually observed, thus allowing their quantitative comparison. The chosen performance metrics were the root mean square error (RMSE), the mean absolute relative error (MARE), and the coefficient of efficiency (CE), given by:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{11}$$

$$\mathrm{MARE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \hat{y}_i\right|}{y_i} \tag{12}$$

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)}{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)} \tag{13}$$

where y_i are the observed values, $\hat{y}_i$ are the values generated by the neural network, $\bar{y}_i$ is the mean of observed values, and n is the sample size.

The RMSE and MARE provide an idea of the absolute accuracy of the models, while the CE indicates how well the model is predicting values far from the mean estimator (Dawson *et al.* 2006). Therefore, the optimal models will be those with low RMSE and MARE and CE values close to unity.

The residual errors e_i were also calculated for each proposed model, expressed as the differences between the observed and predicted values, i.e.:

$$e_i = \hat{y}_i - y_i \tag{14}$$

By plotting the evolution of residual errors over time, it is possible to assess the model performance at each step.

Results

Monthly forecasts

Table 3 summarizes the statistical characteristics of the best models obtained. Figures 9–12 show the results of the forecasts obtained for 1, 3, 6 and

Table 3. Summary of all monthly models tested and their performances.

Models	Activation function	Months ahead	Architecture			Input variables	Inputs	Output	Training			Testing		
			n	m	k				RMSE ($m^3\ s^{-1}$)	CE (-)	MARE (-)	RMSE ($m^3\ s^{-1}$)	CE (-)	MARE (-)
ANN_1	TS–PL	1	6	6	1	Q	$Q_t, Q_{t-1}, \ldots, Q_{t-5}$	Q_{t+1}	873.3011	0.8144	0.2102	1 107.0449	0.5859	0.2911
Wavelet–ANN_1	TS–TS	1	9	9	1	$I = A_1 + A_2$	$I_t, I_{t-1}, \ldots, I_{t-8}$	Q_{t+1}	304.1540	0.9775	0.0860	401.7218	0.9454	0.1236
ANN_3	TS–TS	3	6	6	1	Q	$Q_t, Q_{t-1}, \ldots, Q_{t-5}$	Q_{t+3}	1 301.5595	0.5881	0.3431	1 297.0995	0.4303	0.5223
Wavelet–ANN_3	TS–PL	3	9	6	1	$I = A_2 + A_3$	$I_t, I_{t-1}, \ldots, I_{t-8}$	Q_{t+3}	575.8948	0.9195	0.1536	601.7661	0.8771	0.2033
ANN_6	TS–TS	6	9	6	1	Q	$Q_t, Q_{t-1}, \ldots, Q_{t-8}$	Q_{t+6}	1 146.9642	0.6802	0.2979	1 433.5464	0.3027	0.4443
Wavelet–ANN_6	TS–TS	6	9	6	1	$I = A_2 + A_3 + A_4$	$I_t, I_{t-1}, \ldots, I_{t-8}$	Q_{t+6}	769.4499	0.8561	0.2243	894.5346	0.7285	0.2686
ANN_{12}	TS–PL	12	6	9	1	Q	$Q_t, Q_{t-1}, \ldots, Q_{t-5}$	Q_{t+12}	1 173.9132	0.6648	0.3172	1 437.0581	0.3048	0.4917
Wavelet–ANN_{12}	TS–TS	12	6	9	1	$I = A_2 + A_3 + A_4$	$I_t, I_{t-1}, \ldots, I_{t-5}$	Q_{t+12}	958.9718	0.7763	0.2598	1 229.3590	0.4912	0.3955

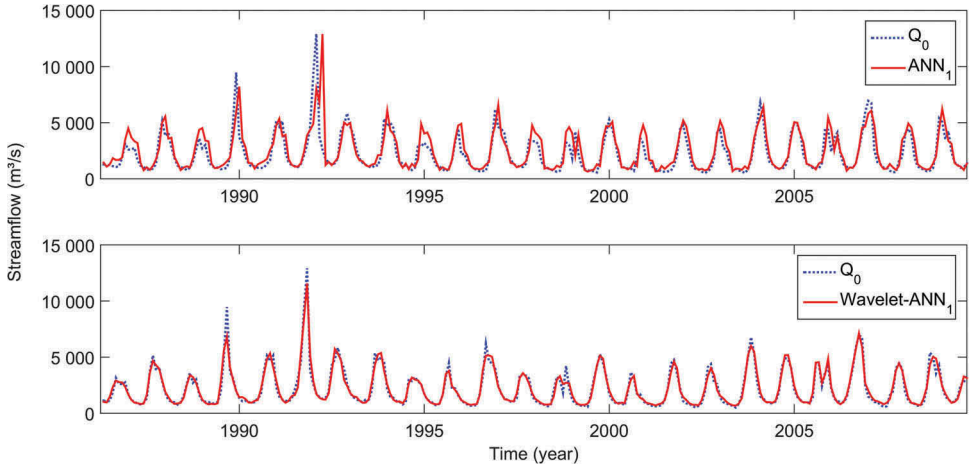

Figure 9. Forecasting for 1 month ahead generated by the ANN_1 and wavelet–ANN_1 models compared to the observed flows (Q_0).

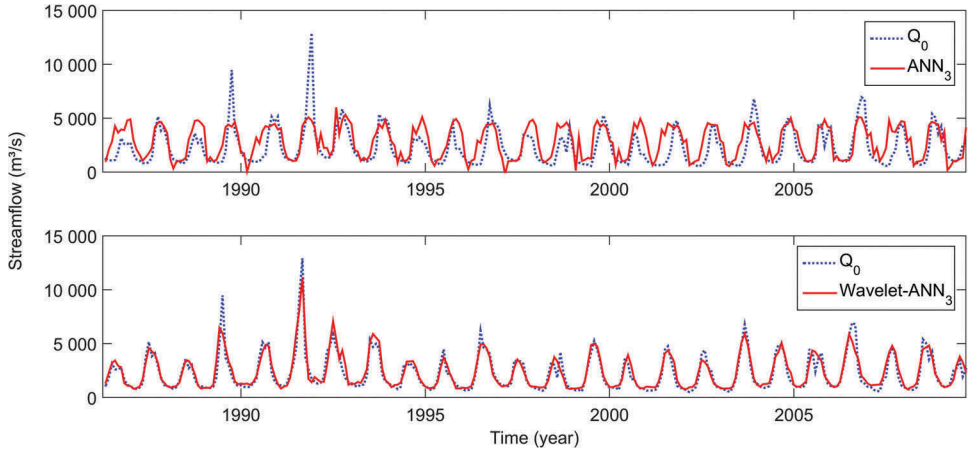

Figure 10. Predictions for 3 months ahead generated by the ANN_3 and wavelet–ANN_3 models compared to the observed flows (Q_0).

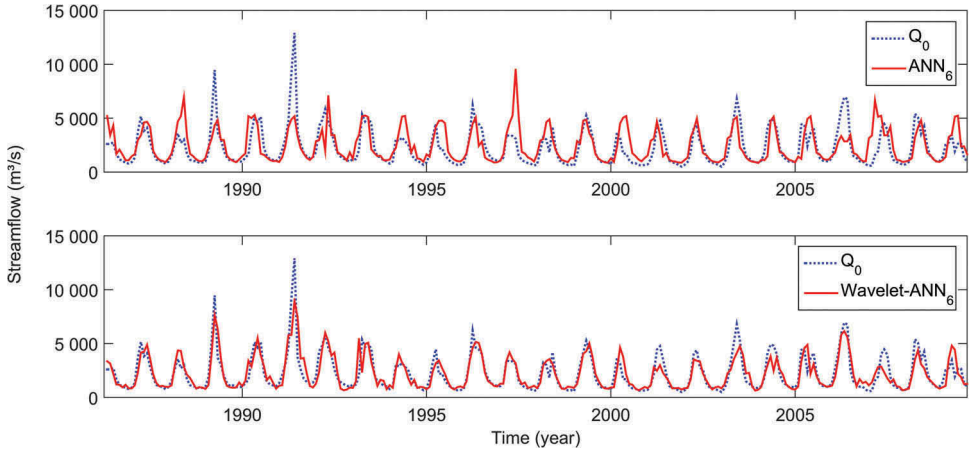

Figure 11. Predictions for 6 months ahead generated by the ANN_6 and wavelet–ANN_6 models compared to the observed flows (Q_0).

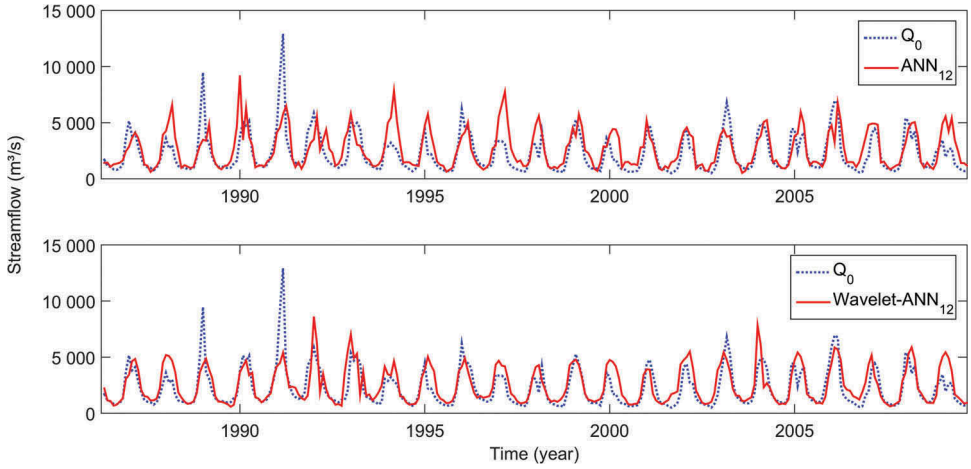

Figure 12. Predictions for 12 months ahead generated by the ANN_{12} and wavelet–ANN_{12} models compared to the observed flows (Q_0).

12 months, which are contrasted with the observed flows.

In contrast to the ANN_1 model, the hybrid prediction model for 1 month ahead, using as input the sum $I = A_1 + A_2$, obtained very good results: the RMSE was almost three times smaller than that of the classical model (401.7 $m^3\ s^{-1}$), its CE was close to 1 (0.945), and the MARE was close to zero (0.124).

Figure 9 shows that the wavelet–ANN_1 model was able to represent the general behaviour of the flow series with good precision, even predicting the highest peaks. In contrast, the ANN_1 model was not able to predict the peaks accurately, which are systematically delayed, and this bad behaviour is reflected in the high value of RMSE

Figure 10 clearly shows that the predictions by the classical model do not agree with the observed values. On the other hand, the predictions obtained with the hybrid model are noticeably better. This can be verified from the RMSE, CE and MARE values for the ANN_3 and wavelet–ANN_3 models.

Analysing the predictions for 6 months ahead, the hybrid model can give a general idea of the behaviour of the series, but the 6-month flow correlation is too small to be learned, even by the hybrid model (Fig. 11). Nevertheless, the performance of the hybrid model was superior to that of the classical one, with RMSE equal to 849.5 $m^3\ s^{-1}$ and CE up to 0.878, compared to 1433.5 $m^3\ s^{-1}$ and 0.303, respectively.

The performance of the two models was very similar for the 12-month prediction (Fig. 12). For the 12-month lead time, none of the two modelling approaches could predict the observed flows with reasonable accuracy. However, despite not being able to foresee the flows for 6 and 12 months ahead with full accuracy, the hybrid model performed better overall than the classical model, according to the three statistical parameters for all forecast horizons.

Regarding the residuals, Figure 13 shows that the hybrid model performed well for predictions with 1 and 3 months ahead. It is also possible to note that, while the hybrid model performed slightly better than the ANN model for 6-month forecasts and that for the annual forecasts, both models had similar performances.

In percentile terms, the highest error of ANN_1 is overestimated by more than 300%, while wavelet–ANN_1 deviated from the corresponding observed value by only 27.9%. For 3 and 6 months ahead the hybrid model also performed better, with the ANN models missing the observed value by 62.8 and 59.9%, respectively, and the hybrid only 36.9 and 30.7%, respectively.

Finally, regarding the annual (i.e. 12-month ahead) forecasts, both models obtained errors above 55%, in relation to the observed values.

Input analysis

The tests revealed that A_2 contains significant information for the learning procedure, as by omitting it the network could not predict future flows accurately. This can also be shown by comparing A_2 and A_3 (Fig. 7), thus indicating significant loss of information.

In order to evaluate the influence of the approximations on the model performance, it is possible to rewrite the inputs of the hybrid models as functions of flow Q_0. For example, if the combination $I = A_3 + A_4$ is used as input, the approximation A_3 is nothing other than the sum of A_4 and D_4:

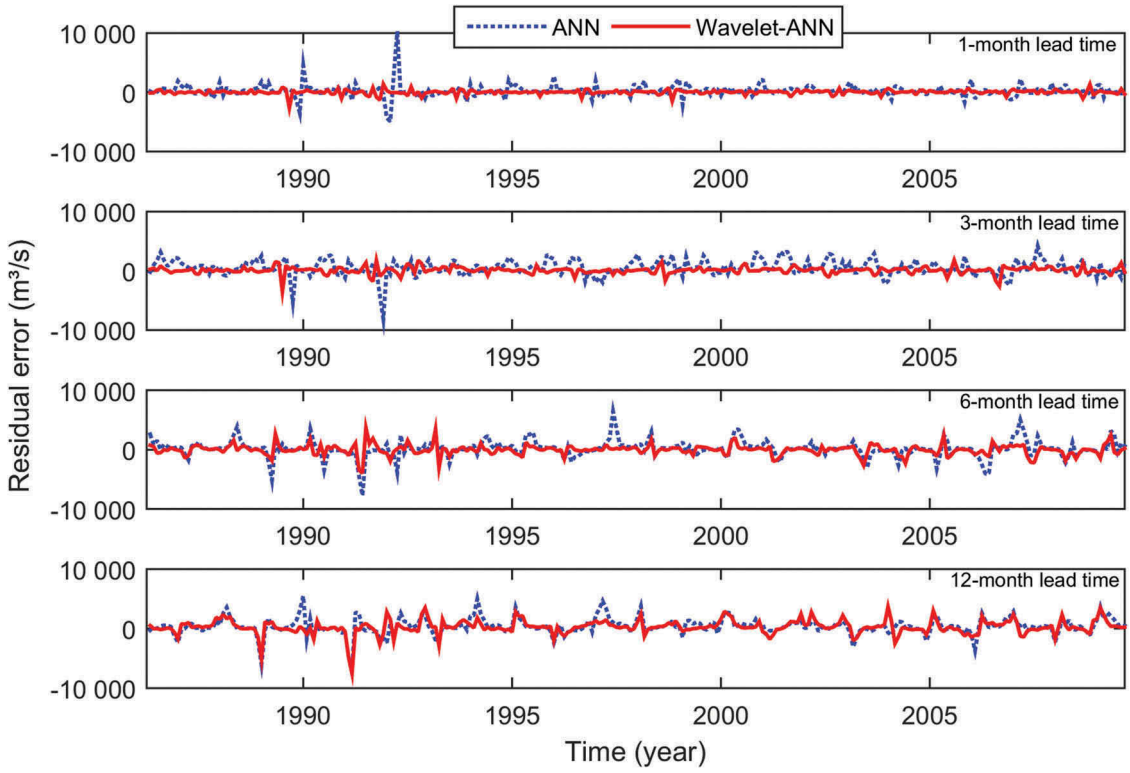

Figure 13. Residual errors for predictions with ANN and wavelet–ANN models.

$$A_3 = A_4 + D_4 \tag{15}$$

Thus, the input $I = A_3 + A_4$ can be rewritten as:

$$A_3 + A_4 = 2A_3 - D_4 \tag{16}$$

Following the previous rationale, we can substitute the approximations and then divide by 2, so that the sum is written in terms of flow, i.e.:

$$\frac{A_3 + A_4}{2} = Q_0 - D_1 - D_2 - D_3 - \frac{D_4}{2} \tag{17}$$

Thus, using the sum of the approximations A_3 and A_4 as model input is the same as using the raw series minus the details D_1, D_2, D_3 and half of D_4.

In Table 4 we visualize the inputs of hybrid models rewritten in terms of Q_0. The rewritten inputs provide an idea of the type of information that the ANN is using to correlate I with future monthly flows. We remark that, by rewriting the variable I in terms of Q_0, we clearly observe that adding different approximations is equivalent to subtracting, totally or partially, different high-frequency components (details) from the raw signal. This improves our understanding of the performance of the hybrid models, tested as a function of the use of the combinations of the approximations as inputs. The first level of details (D_1) not only does not contribute to good forecasting, but actually makes it worse: the results are clearly improved when D_1 are extracted from the inputs, for all examined forecast horizons.

Table 4. Hybrid model inputs rewritten in terms of flow.

Model	Used approximation	Rewritten in terms of Q_0
Wavelet–ANN$_1$	$I = A_1 + A_2$	$Q_0 - D_1 - D_2/2$
Wavelet–ANN$_3$	$I = A_2 + A_3$	$Q_0 - D_1 - D_2 - D_3/2$
Wavelet–ANN$_6$	$I = A_2 + A_3 + A_4$	$Q_0 - D_1 - D_2 - 2D_3/3 - D_4/3$
Wavelet–ANN$_{12}$	$I = A_2 + A_3 + A_4$	$Q_0 - D_1 - D_2 - 2D_3/3 - D_4/3$

Another interesting aspect that can be seen from the rewritten inputs is that lower level details are needed for shorter forecast horizons, while for longer horizons it seems that they are not essential: for 1 month ahead, D_1 and half of D_2 were subtracted from the inputs; for 3 months ahead, D_2 was taken completely with half of D_3; for the 6-month ahead model, two-thirds of D_3 and also one-third of D_4 had to be taken from the inputs (see Table 4). This may happen because the first levels of detail contain information regarding short-term events, which certainly influence the streamflow. On the other hand, the higher-level details comprehend long-term events that have a smaller effect on the streamflow, thus they are only important for long-term forecasts.

Furthermore, in order to demonstrate the influence of each approximation on the predictions, 10

prediction models were established using each approximation (from A_1 to A_{10}) as inputs. The objective of this investigation was only to analyse the influence of the approximations in the predictions, thus the architecture (9–6–1) and the activation functions (TS–TS) were kept constant and the 3-month forecast horizon was assumed. Then, we calculated the statistical parameters and plotted the residual errors in order to evaluate the accuracy of each prediction. However, it is important to remember that, since the data processing procedure is nonlinear, simple direct conclusions should not be drawn. Nevertheless, this analysis can give hints about the success of the hybrid model.

Figure 14 shows the predictions for 3 months ahead using single approximations as inputs. The forecast improves when only A_2 is used instead of A_1; however, without the components that are present in A_2, the network is not able to learn the general behaviour of the streamflow, thus the latter cannot be forecasted accurately. This is revealed by the poor prediction obtained when using A_3 as sole input.

Table 5 shows the results for the hybrid models using only one approximation as input for predicting flows for 3 months ahead, while Figure 15 illustrates the residual errors for each model. This analysis also confirms that the best prediction model for the 3-month horizon was the one that used A_2 as input (RMSE = 626.9 $m^3 s^{-1}$, CE = 0.867 and MARE = 0.203), corresponding to the smallest residual error in Figure 15 (dashed red line). It is also possible to infer that A_2 (or D_3) has crucial information for the learning procedure of the network, since its removal worsened the predictions. As the details are removed, the flows can no longer be predicted using the hybrid model, which is evident by the increase of oscillations in Figure 15 and the bad values of all performance metrics.

Table 5. Statistical outcomes for 3-month ahead predictions using single approximations as inputs, 9–6–1 architecture and TS–TS activation functions. The best results are highlighted.

Used approximation	RMSE ($m^3 s^{-1}$)	CE (-)	MARE (-)
A_1	1 116.2166	0.577	0.3948
A_2	626.8974	0.8666	0.2032
A_3	1 557.1187	0.1768	0.8291
A_4	1 689.4307	0.031	0.8289
A_5	1 848.7474	−0.1604	0.8113
A_6	2 539.8769	−1.1901	1.1786
A_7	2 184.7238	−0.6204	1.4511
A_8	2 315.4646	−0.9059	1.0947
A_9	2 153.8505	−0.575	0.4423
A_{10}	2 377.7996	−0.9195	0.6084

Conclusions

In general, the results obtained from the proposed scheme are encouraging, indicating that the methodology tested can be an important alternative for monthly streamflow forecasting.

As expected from the performance of hybrid models in previous works for monthly forecasting (Partal 2009, Mehr *et al.* 2013), the hybrid approach using only approximations as inputs outperformed the classical ANN model for all forecasting horizons tested for the Sobradinho Dam, São Francisco River basin in Brazil. Specifically, for 1 and 3 months ahead, the hybrid model considerably reduced the oscillations of the forecasts, fixed the lags, and correctly predicted the streamflow peaks. Even for 6 months in advance, the hybrid model was able to represent the broad behaviour of future flows, while the classical model could not ensure any correlation among future and past flows. For annual predictions, no model could establish a correlation between past and future flows, although the hybrid approach performed slightly better.

This article also brings a new way to analyse the combination of approximations used as inputs. When the

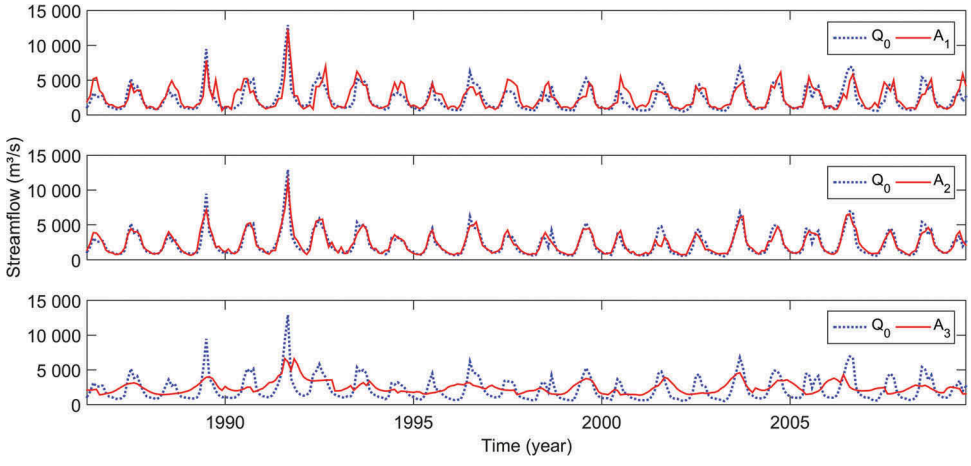

Figure 14. Predictions for 3 months ahead generated by the hybrid model using single approximations.

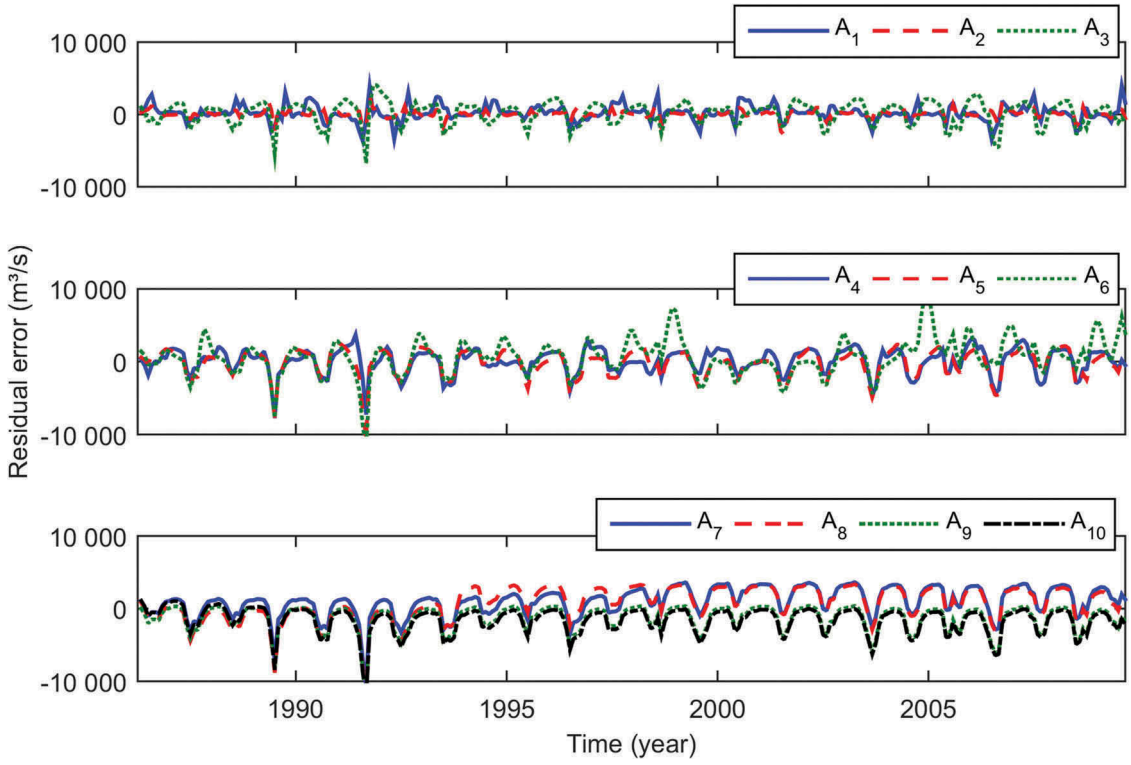

Figure 15. Residual errors for each monthly hybrid model that used approximations as sole inputs.

inputs are rewritten in terms of raw flows, the information that is provided to the network is exposed, which can give us a clue about the better performance of the hybrid models. The first levels of details carry unnecessary information for the learning process of the ANN and actually worsen the forecasting, especially for longer forecasting horizons. Additionally, the input analysis revealed which level of detail carries essential information for the forecasting.

However, it is important to consider that the São Francisco River has a large catchment, where the correlation between past and future flows is high, and this may be the reason why the approximations were successful. Thus, this approach could be tested in different catchments with different sizes and physiographic characteristics and across different climates, to extract stronger conclusions.

Although limited to one application, the results shown here can inspire the development of new research on this theme. For instance, future research may focus on studying the effect of different mother wavelets on forecasting, comparing this hybrid model with different wavelet–ANN approaches (e.g. the methods published by Kişi 2009, Partal 2009 and Pramanik *et al.* 2010), in order to assess which performs better, and also compare this approach with adaptive neuro-fuzzy inference system (ANFIS) and wavelet–ANFIS models (e.g. Moosavi *et al.* 2013).

Acknowledgements

The authors thank the ONS for the re-naturalized mean monthly streamflow data provided and the reviewers and editor for the relevant contributions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by National Council for Scientific and Technological Development, Brazil (CNPq) [grant numbers 471342/2012-2 and 304213/2017-9], and was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES) [finance code 001].

References

Adamowski, J. and Sun, K., 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid water-

sheds. *Journal of Hydrology*, 390 (1–2), 85–91. doi:10.1016/j.jhydrol.2010.06.033

Alp, M. and Ciğizoğlu, H.K., 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling & Software*, 22 (1), 2–13. doi:10.1016/j.envsoft.2005.09.009

ASCE Task Committee, 2000a. Artificial neural networks in hydrology. I: preliminary concepts. *Journal of Hydrologic Engineering*, 5 (2), 115–123. doi:10.1061/(ASCE)1084-0699(2000)5:2(115)

ASCE Task Committee, 2000b. Artificial neural networks in hydrology. II: hydrologic applications. *Journal of Hydrologic Engineering*, 5 (2), 124–132. doi:10.1061/(ASCE)1084-0699(2000)5:2(124)

Bravo, J.M., *et al.*, 2008. Previsões de Curto Prazo de Vazão Afluente ao Reservatório de Furnas Utilizando Redes Neurais Artificiais. *Revista Brasileira de Recursos Hídricos*, 13 (2), 77–88. doi:10.21168/rbrh.v13n2.p77-88

Cannas, B., *et al.*, 2006. Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Physics and Chemistry of the Earth*, 31 (18), 1164–1171. doi:10.1016/j.pce.2006.03.020

Ciğizoğlu, H.K., 2003. Incorporation of ARMA models into flow forecasting by artificial neural networks. *Environmetrics*, 14 (4), 417–427. doi:10.1002/(ISSN)1099-095X

Ciğizoğlu, H.K., 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Advances in Water Resources*, 27 (2), 185–195. doi:10.1016/j.advwatres.2003.10.003

Ciğizoğlu, H.K., 2005. Generalized regression neural network in monthly flow forecasting. *Civil, Engineering and Environmental Systems*, 22 (2), 71–84. doi:10.1080/10286600500126256

Ciğizoğlu, H.K. and Alp, M., 2006. Generalized regression neural network in modelling river sediment yield. *Advances in Engineering Software*, 37 (2), 63–68. doi:10.1016/j.advengsoft.2005.05.002

Coulibaly, P., *et al.*, 2001. Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, 37 (4), 885–896. doi:10.1029/2000WR900368

Cruz, M.F.M., Rodrigues, L.D., and Versiani, B.R., 2010. Previsão de Vazões com a Metodologia DPFT e com Redes Neurais Artificiais. *Revista Brasileira de Recursos Hídricos*, 15 (1), 121–132. doi:10.21168/rbrh.v15n1.p121-132

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2 (4), 303–314. doi:10.1007/BF02551274

Daliakopoulos, I.N., Coulibaly, P., and Tsanis, I.K., 2005. Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309 (1–4), 229–240. doi:10.1016/j.jhydrol.2004.12.001

Daubechies, I., 1992. *Ten lectures on wavelets*. Philadelphia, PA: SIAM.

Dawson, C.W., *et al.*, 2006. Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, 319 (1), 391–409. doi:10.1016/j.jhydrol.2005.07.032

Debastiani, A.B., Silva, R.D., and Neto, S.L.R., 2016. Eficácia da arquitetura MLP em modo closed-loop para simulação de um Sistema Hidrológico. *Revista Brasileira de Recursos Hídricos*, 21 (4), 821–831. doi:10.1590/2318-0331.011615124

Demuth, H. and Beale, M., 2002. *Neural network toolbox – for use with MATLAB©*. Natick, MA: The MathWorks.

EPE (Empresa de Pesquisa Energética), 2016. *Brazilian energy balance – final report year 2015*. Brasília, DF: Ministry of Mines and Energy – MME.

Gazzaz, N.M., *et al.*, 2012. Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Marine Pollution Bulletin*, 64 (11), 2409–2420. doi:10.1016/j.marpolbul.2012.08.005

Gomes, L.F.C., Montenegro, S.M.G.L., and Valença, M.J.S., 2010. Modelo Baseado na Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco. *Revista Brasileira de Recursos Hídricos*, 15 (1), 5–15. doi:10.21168/rbrh.v15n1.p5-15

Haykin, S., 2005. *Neural networks: a comprehensive foundation*. New Delhi: Prentice-Hall.

Holdefer, A.E. and Severo, D.L., 2015. Análise por ondaletas sobre níveis de rios submetidos à influência de maré. *Revista Brasileira de Recursos Hídricos*, 20 (1), 192–201. doi:10.21168/rbrh.v20n1.p192-201

Jeong, D. and Kim, Y., 2005. Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrological Processes*, 19 (19), 3819–3835. doi:10.1002/(ISSN)1099-1085

Kişi, Ö., 2004. River flow modeling using artificial neural networks. *Journal of Hydrologic Engineering*, 9 (1), 60–63. doi:10.1061/(ASCE)1084-0699(2004)9:1(60)

Kişi, Ö., 2008. Stream flow forecasting using neuro-wavelet technique. *Hydrological Processes*, 22 (20), 4142–4152. doi:10.1002/hyp.v22:20

Kişi, Ö., 2009. Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *Journal of Hydrologic Engineering*, 14 (8), 773–782. doi:10.1061/(ASCE)HE.1943-5584.0000053

Kişi, Ö. and Ciğizoğlu, K., 2007. Comparison of different ANN techniques in river flow prediction. *Civil Engineering and Environmental Systems*, 5 (2), 211–231. doi:10.1080/10286600600888565

Koutsoyannis, D., Yao, H., and Georgakakos, A., 2008. Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods. *Hydrological Sciences Journal*, 53 (1), 142–164. doi:10.1623/hysj.53.1.142

Kumar, A.R.S., *et al.*, 2005. Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrological Processes*, 19 (6), 1277–1291. doi:10.1002/(ISSN)1099-1085

Kumar, M., *et al.*, 2002. Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128 (4), 224–233. doi:10.1061/(ASCE)0733-9437(2002)128:4(224)

Londhe, S. and Charhate, S., 2010. Comparison of data-driven modelling techniques for river flow forecasting. *Hydrological Sciences Journal*, 55 (7), 1163–1174. doi:10.1080/02626667.2010.512867

Machado, F., *et al.*, 2011. Monthly rainfall–runoff modelling using artificial neural networks. *Hydrological Sciences Journal*, 56 (3), 349–361. doi:10.1080/02626667.2011.559949

Makwana, J.J. and Tiwari, M.K., 2014. Intermittent streamflow forecasting and extreme event modelling using wavelet based artificial neural networks. *Water Resources Management*, 28 (13), 4857–4873. doi:10.1007/s11269-014-0781-1

Mallat, S., 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 11 (7), 674–693. doi:10.1109/34.192463

Mehr, A.D., *et al.*, 2013. Successive-station monthly streamflow prediction using neuro-wavelet technique. *Earth Science Informatics*, 7 (4), 217–229. doi:10.1007/s12145-013-0141-3

Melesse, A.M., *et al.*, 2011. Suspended sediment load prediction of river systems: an artificial neural network approach. *Agricultural Water Management*, 98 (5), 855–866. doi:10.1016/j.agwat.2010.12.012

Misiti, M., *et al.*, 1996. *Wavelet toolbox – for use with MATLAB©*. Natick, MA: The MathWorks.

Mohanty, S., *et al.*, 2010. Artificial Neural Network modeling for groundwater level forecasting in a river island of eastern India. *Water Resources Management*, 24 (9), 1845–1865. doi:10.1007/s11269-009-9527-x

Moosavi, V., *et al.*, 2013. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. *Water Resources Management*, 27 (5), 1301–1321. doi:10.1007/s11269-012-0239-2

Muluye, G.Y., 2011. Improving long-range hydrological forecasts with extended Kalman filters. *Hydrological Sciences Journal*, 56 (7), 1118–1128. doi:10.1080/02626667.2011.608068

Nagy, H.M., Watanabe, K.A.N.D., and Hirano, M., 2002. Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulic Engineering*, 128 (6), 588–595. doi:10.1061/(ASCE)0733-9429(2002)128:6(588)

Nayak, P.C., Rao, Y.R.S., and Sudheer, K.P., 2006. Groundwater level forecasting in a shallow aquifer using Artificial Neural Network approach. *Water Resources Management*, 20 (1), 77–90. doi:10.1007/s11269-006-4007-z

Nourani, V., Alami, M.T., and Aminfar, M.H., 2009. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence*, 22 (3), 466–472. doi:10.1016/j.engappai.2008.09.003

Oliveira, G.G., Pedrollo, O.C., and Castro, N.M.R., 2014. O Desempenho das Redes Neurais Artificiais (RNAs) para Simulação Hidrológica Mensal. *Revista Brasileira de Recursos Hídricos*, 19 (2), 251–265. doi:10.21168/rbrh.v19n2.p251-265

Partal, T., 2009. River flow forecasting using different artificial neural networks algorithms and wavelet transform. *Canadian Journal of Civil Engineering*, 36 (1), 26–38. doi:10.1139/L08-090

Partal, T. and Ciğizoğlu, H.K., 2009. Prediction of daily precipitation using wavelet-neural networks. *Hydrological Sciences Journal*, 54 (2), 234–246. doi:10.1623/hysj.54.2.234

Prada-Sarmiento, F. and Obregón-Neira, N., 2009. Forecasting of monthly streamflows based on artificial neural networks. *Journal of Hydrologic Engineering*, 14 (12), 1390–1395. doi:10.1061/(ASCE)1084-0699(2009)14:12(1390)

Pramanik, N., Panda, R.K., and Singh, A., 2010. Daily river flow forecasting using wavelet ANN hybrid models. *Journal of Hydroinformatics*, 13 (1), 49–63. doi:10.2166/hydro.2010.040

Rajurkar, M.P., Kothyari, U.C., and Chaube, U.C., 2002. Artificial neural networks for daily rainfall – runoff modelling. *Hydrological Sciences Journal*, 47 (6), 865–877. doi:10.1080/02626660209492996

Riad, S., *et al.*, 2004. Rainfall-runoff model using an artificial neural network approach. *Mathematical and Computer Modelling*, 40 (7–8), 839–846. doi:10.1016/j.mcm.2004.10.012

Santos, C.A.G., Freire, P.K.M.M., and Torrence, C., 2013. A transformada wavelet e sua aplicação na análise de séries hidrológicas. *Revista Brasileira de Recursos Hídricos*, 18 (3), 271–280. doi:10.21168/rbrh.v18n3.p271-280

Santos, C.A.G. and Silva, G.B.L., 2014. Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. *Hydrological Sciences Journal*, 59 (2), 312–324. doi:10.1080/02626667.2013.800944

Santos, C.A.G., *et al.*, 2019. Hybrid wavelet neural network approach for daily inflow forecasting using Tropical Rainfall Measuring Mission data. *Journal of Hydrologic Engineering*, 24 (2). doi:10.1061/(ASCE)HE.1943-9105584.0001725

Shanker, M., Hu, M.Y., and Hung, M.S., 1996. Effect of data standardization on neural network training. *Omega*, 24 (4), 385–397. doi:10.1016/0305-0483(96)00010-2

Singh, K.P., *et al.*, 2009. Artificial neural network modeling of the river water quality – a case study. *Ecological Modelling*, 220 (6), 888–895. doi:10.1016/j.ecolmodel.2009.01.004

Sudheer, K.P., Gosain, A.K., and Ramasastri, K.S., 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes*, 16 (6), 1325–1330. doi:10.1002/(ISSN)1099-1085

Tayfur, G., 2002. Artificial neural networks for sheet sediment transport. *Hydrological Sciences Journal*, 47 (6), 879–892. doi:10.1080/02626660209492997

Torrence, C. and Compo, G.P., 1998. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79 (1), 61–78. doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Trajkovic, S., Todorovic, B., and Stankovic, M., 2003. Forecasting of reference evapotranspiration by artificial neural networks. *Journal of Irrigation and Drainage Engineering*, 129 (6), 454–457. doi:10.1061/(ASCE)0733-9437(2003)129:6(454)

Wang, W. and Ding, J., 2003. Wavelet network model and its application to the prediction of hydrology. *Nature and Science*, 1 (1), 67–71.

Wei, S., *et al.*, 2013. A wavelet-neural network hybrid modelling approach for estimating and predicting river monthly flows. *Hydrological Sciences Journal*, 58 (2), 374–389. doi:10.1080/02626667.2012.754102

Yaseen, Z.M., *et al.*, 2015. Artificial intelligence based models for stream-flow forecasting: 2000–2015. *Journal of Hydrology*, 530, 829–844. doi:10.1016/j.jhydrol.2015.10.038

Zanetti, S.S., *et al.*, 2007. Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering*, 133 (2), 83–89. doi:10.1061/(ASCE)0733-9437(2007)133:2(83)