

## Discrete wavelet transform coupled with ANN for daily discharge forecasting into Três Marias reservoir

CELSO A. G. SANTOS<sup>1</sup>, PAULA K. M. M. FREIRE<sup>1</sup>, GUSTAVO B. L. SILVA<sup>1</sup> & RICARDE M. SILVA<sup>2</sup>

<sup>1</sup> Department of Civil and Environmental Engineering, Federal University of Paraíba, 58051-900 João Pessoa – PB, Brazil

[celso@ct.ufpb.br](mailto:celso@ct.ufpb.br)

<sup>2</sup> Department of Geosciences, Federal University of Paraíba, 58051-900 João Pessoa – PB, Brazil

**Abstract** This paper proposes the use of discrete wavelet transform (DWT) to remove the high-frequency components (details) of an original signal, because the noises generally present in time series (e.g. streamflow records) may influence the prediction quality. Cleaner signals could then be used as inputs to an artificial neural network (ANN) in order to improve the model performance of daily discharge forecasting. Wavelet analysis provides useful decompositions of original time series in high and low frequency components. The present application uses the Coiflet wavelets to decompose hydrological data, as there have been few reports in the literature. Finally, the proposed technique is tested using the inflow records to the Três Marias reservoir in São Francisco River basin, Brazil. This transformed signal is used as input for an ANN model to forecast inflows seven days ahead, and the error *RMSE* decreased by more than 50% (i.e. from 454.2828 to 200.0483).

**Key words** wavelet; ANN; forecasting

### INTRODUCTION

Wavelet analysis is an advanced tool in signal processing that has attracted much attention since its theoretical development by Grossmann and Morlet (1984). Thus, the main purpose of this study is to combine the use of the discrete wavelet transform (DWT) technique to filter the daily data of a 81-year-long streamflow time series to be used as input data of an ANN. In the continuous form, the wavelet transform has been well reported in several studies, even to detect signal periodicity (Santos *et al.* 2003, 2009), to identify precipitation zones (Santos and Silva 2013) or for synthetic generation of daily streamflow (Wang *et al.* 2011), but its discrete form is very suitable to filter signal data (Mallat 1989).

However, ANNs have shown good forecast time series, as used by Dawson and Wilby (1999), Birikundavyi *et al.* (2002) and Besaw *et al.* (2010), or even to use in daily reservoir operation (Farias *et al.* 2011). However, the noise present in the original time series could affect the final results, thus several attempts to combine DWT and ANN have been reported, e.g. Anctil and Tape (2004), Adamowski (2007, 2008a,b), Kisi (2008), Adamowski and Sun (2010). Most of them use the Daubechies mother-wavelet to filter the signal and probably the chosen mother-wavelet may interfere with the results. For example, Santos and Silva (2013), also using the Daubechies mother-wavelet, developed an ANN-Wavelet hybrid model for daily streamflow forecasting into Sobradinho reservoir, using combinations of DWT approximations as well. However, in this paper the Coiflet mother-wavelets is used because they are discrete wavelets designed to have scaling functions with vanishing moments, and they are near symmetric.

Finally, the proposed technique is applied to an 81-year-long time series of daily streamflow records into Três Marias reservoir, Brazil, which is part of the hydroelectric power generation system of Brazil. The paper presents the DWT, the ANN, the selected data and the results of the prediction using the filtered time series as input to the ANN.

### MATERIAL AND METHODS

#### Study area description

The Alto São Francisco basin, the upper part of São Francisco River basin, has an area of 50 870 km<sup>2</sup> and meteorological characteristics of the Minas Gerais State central area climate. The Alto São Francisco basin drains the main water course of the Minas Gerais State, and is a strategic area for

management of the Brazilian water resources. The data used in this work correspond to the re-naturalised daily inflows of Três Marias reservoir, São Francisco River, Brazil (Fig. 1). The data were extracted from the National System Operator (ONS – Operador Nacional do Sistema) website at [www.ons.org.br](http://www.ons.org.br), the institution responsible for the operation of the interlinked electric system in Brazil. The series had 29 585 daily data records, referring to the period of 1 January 1931 to 31 December 2011.

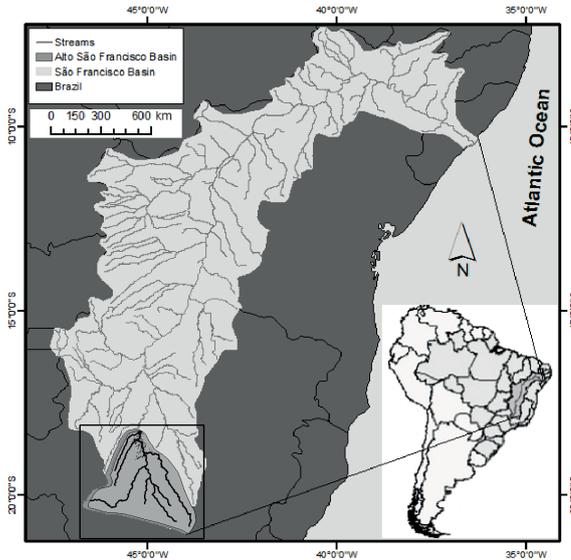


Fig. 1 Location of the São Francisco River basin and its sub-basin Alto do São Francisco, Brazil.

### Artificial neural networks

According to Haykin (1999), an Artificial Neural Network can be defined as a set of simple processing units working as a parallel distributed processor. These units, which are called neurons, are responsible for storing experimental knowledge for posterior disposal. The ANNs are inspired by the biological nervous system. As the brains, they learn through examples and have the acquired knowledge stored in the connection weights between neurons (Demuth and Beale 2005).

In this paper, a feed-forward network with one hidden layer was used. The data  $(x_1, x_2, \dots, x_n)$  are introduced in the input layer and the network progressively processes such data through subsequent layers, producing a result  $(y_1, y_2, \dots, y_k)$  in the output layer. The input neurons are linked to those in the intermediate layer through  $w_{ji}$  weights and the neurons in the intermediate layer are linked to those in the output layer through  $w_{ki}$  weights. The network maps out the relation between the input data and the output variables based on the nonlinear activation functions. The explicit correlation for the output values is given by:

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

in which  $w_{ji}$  is the weight connecting the  $i$ th neuron in the input layer and the  $j$ th neuron in the hidden layer,  $w_{kj}$  is the weight connecting the  $k$ th neuron in the output layer and the  $j$ th neuron in the hidden layer,  $b_j$  is the bias for the  $j$ th hidden neuron,  $b_k$  is the bias for the  $k$ th output neuron,  $f_h$  is the activation function of the nodes in the hidden layer, and  $f_o$  is the activation function of the nodes in the output layer.

The determination of the ANN weights is made through a procedure called calibration, in which several input–output examples are presented to the ANN and its weights are iteratively modified until ANN reaches an acceptable mapping capacity, which is defined by the user. In this approach, the Levenberg-Marquardt (LM) algorithm was used for training because it is considered one of the fastest methods for training ANNs. The efficiency of the models may be assessed using several

statistic parameters that describe the adhering degree among the data observed and predicted by the model (Abrahart *et al.* 2007, Dawson *et al.* 2007, Napolitano *et al.* 2011). In this paper, the root mean square error (*RMSE*) was used:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

in which  $y_i$  corresponds to the observed values,  $\hat{y}_i$  to the values produced by the network,  $\bar{y}$  is the mean of observed values, and  $n$  is the sample size. This error represents the variance of the error regardless of the size of the sample analysed. The best models are those which showed low *RMSE*.

### Discrete wavelet transform

A discrete wavelet transform is obtained from a continuous representation by discretizing dilation and translation parameters such that the resulting set of wavelets constitutes a frame. The dilation parameters are typically discretized by an exponential sampling with a fixed dilation step and the translation parameter by integer multiples of a dilation-dependent step (Daubechies 1992).

Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates a lot of data, but it would be possible to calculate only a subset of scales and positions. It turns out, rather remarkably, that if we choose scales and positions based on powers of two (so-called dyadic scales and positions) then the analysis will be much more efficient and just as accurate. Such an analysis is obtained from the DWT. An efficient way to implement this scheme using filters was developed by Mallat (1989). The Mallat algorithm is in fact a classical scheme known in the signal processing community as a two-channel sub-band coder. This very practical filtering algorithm yields a fast wavelet transform, i.e. a box into which a signal passes, and out of which wavelet coefficients quickly emerge.

The decomposition process can be iterated, with successive approximations being decomposed in turn, so that the original signal is broken down into many lower resolution components. This is called the wavelet decomposition tree. Thus, in the filtering process, at its most basic level, the original signal passes through two complementary filters (high- and low-pass) and emerges as two signals, which are called approximations and details. The approximations ( $A$ ) are the high-scale, low-frequency components of the signal, and the details ( $D$ ) are the low-scale, high-frequency components. It is also important to observe that the raw signal ( $Q$ ) is formed by those components, i.e.  $Q = A_1 + D_1$  or that  $A_1 = A_2 + D_2$ , and in the same way  $Q = A_3 + D_3 + D_2 + D_1$ .

There are many basic wavelets that can be used in this type of transformation. This paper uses the Coiflet wavelets (Fig. 2). Coiflets are discrete wavelets designed by Ingrid Daubechies, at the request of Ronald Coifman, to have scaling functions with vanishing moments. The wavelet is near symmetric, their wavelet functions have  $N/3$  vanishing moments and scaling functions  $N/3 - 1$ , in which  $N$  is the wavelet index, i.e. 5, in the present study.

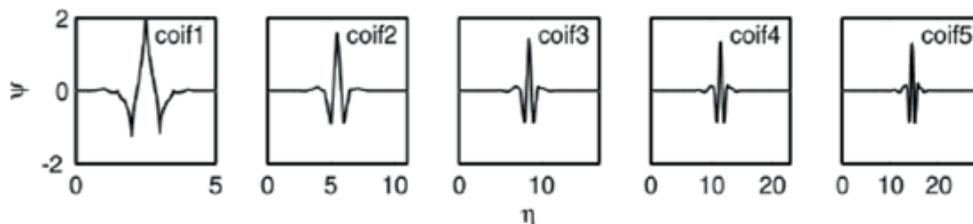


Fig. 2 Coiflet wavelets (from order  $N$  equal to 1 up to 5).

### ANN and Hybrid wavelet-ANN models

In this study, two models, called regular ANN and WA-ANN, were built to forecast streamflows seven days ahead. The regular ANN model used the original streamflow series observed in the current day and in the four previous days ( $Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$ ) as input, while the WA-ANN model used the subseries of approximations obtained by means of a multi-resolution analysis as

input data. Several input options were tested, and this was the best set. The decomposition was performed in 10 levels ( $A_1, \dots, A_{10}$ ), for each situation, the performance of the corresponding model was assessed. The tests revealed that the use of the approximations  $A_3$  yield the best results. Thus, the WA-ANN used as input values of  $A_3$  in the current day and in the four previous days ( $A_{3t}, A_{3t-1}, A_{3t-2}, A_{3t-3}, A_{3t-4}$ ). In both models the output layer had only one neuron, corresponding to the forecasted streamflow seven days ahead ( $Q_{t+7}$ ).

When an ANN is developed, the initial objective consists in finding an optimal architecture that permits the capture of the relationship between the input and output variables. There is not a general rule that indicates the number of neurons to be used in the input and hidden layers. Normally, these quantities are defined by means of a procedure of trial and error. In this study, the number of neurons in the hidden layer was set in 20 neurons. The tan-sigmoid function was chosen as the activation function for the hidden neurons ( $f_h$ ). For the output layer neuron, a linear activation function ( $f_o$ ) was used.

The model performance during the validation was evaluated by using the  $RMSE$ , as described in equation (2), and the correlation index ( $R^2$ ):

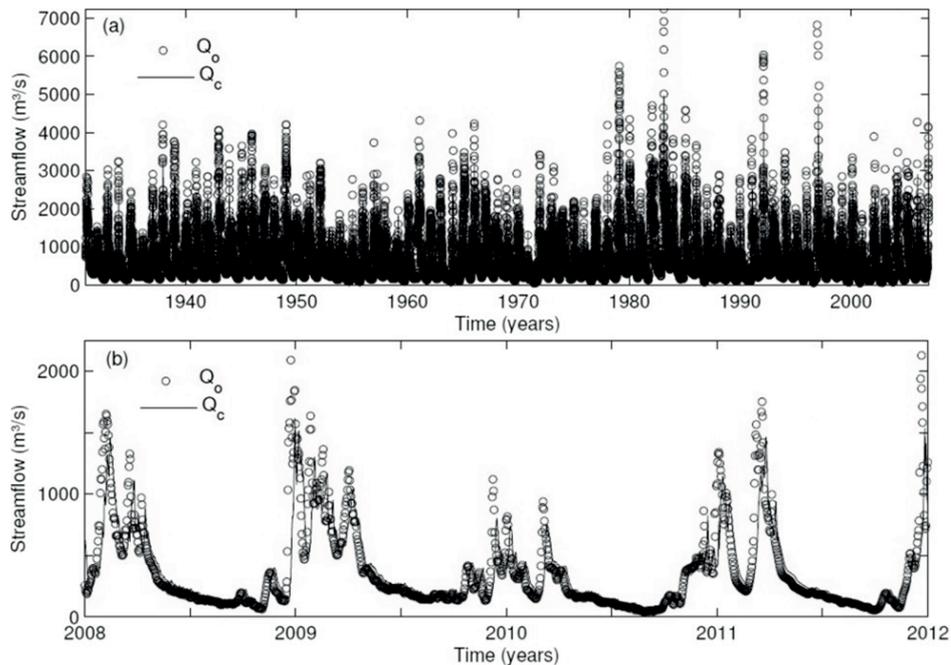
$$R^2 = \frac{\sum ((Q_o - \bar{Q})(Q_c - \bar{Q}))^2}{\sum (Q_o - \bar{Q})^2 \sum (Q_c - \bar{Q})^2} \quad (3)$$

where  $Q_o$  is the observed streamflow,  $\bar{Q}$  is the mean streamflow, and  $Q_c$  is the calculated streamflow.

## RESULTS

In order to forecast the inflows seven days ahead of Trés Marias reservoir, the daily streamflow data of 81 years was used, which correspond to the re-naturalised daily inflows. Then, the effect of the land use changes is minimized. Thus, the first 77 years (1 January 1931 to 31 December 2007) data were used to calibrate the model, and the remaining 4 years (1 January 2008 to 31 December 2011) were used to validate the calibrated ANN.

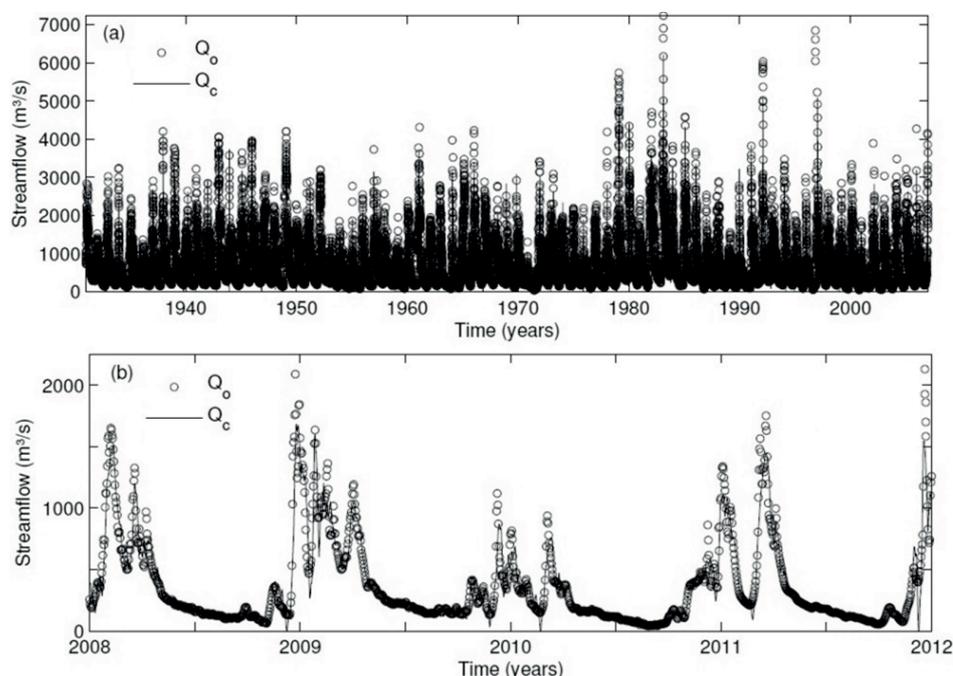
Figure 3 shows the forecast hydrograph for seven days ahead using the raw signal during the (a) calibration and (b) validation processes, and the results show a  $RMSE$  equal to 3.0024 and



**Fig. 3** Forecasting seven days ahead using the raw data, in which  $Q_o$  and  $Q_c$  are the observed and calculated inflows (Regular ANN model), respectively.

454.2828, respectively. A small lag of 7 days between the observed and predicted values can be seen in Fig. 3(b); this phenomenon has already been reported by other authors (e.g. Wu *et al.* 2009) on forecasts using ANNs. They also reported that other authors suggested that an effective solution to the forecasting lag effect is to obtain new model inputs over the original discharge data.

The conventional discrete wavelet analysis of signals was performed on a flow time series using the multilevel 1-D wavelet decomposition method. This produced the wavelet transform of the input data at all dyadic scales. Rather than relying on an upsampling procedure, the DWT relies more on downsampling, which is excellent for de-noising. Thus, the time series of streamflow was transformed using the Coiflet (coif5) to decompose the raw signal into approximations and details, and then divided into two groups for: (a) calibration and (b) validation as well. Finally, the approximations were used as input to perform the forecasts for seven days ahead, as shown in Fig. 4 using  $A_3$ . The performance indices for the forecasting during the calibration and validation processes using the raw and transformed data are shown in Table 1. It is clear observed that all indexes improved when using the filtered signal and that the lag effect is eliminated.



**Fig. 4** Forecasting seven days ahead using the approximation  $A_3$ , in which  $Q_o$  and  $Q_c$  are the observed and calculated inflows (WA-ANN model), respectively.

**Table 1** Statistical indexes during the calibration and validation using the raw and transformed data.

Input data	Calibration		Validation	
	$RMSE$	$R^2$	$RMSE$	$R^2$
Raw	3.0024	0.8182	454.2828	0.7757
Transformed	1.5836	0.9602	200.0483	0.9603

## CONCLUSIONS

The discrete wavelet transform was used to remove the high-frequency components of the original signal (details) of the inflows of Três Marias reservoir, São Francisco River, Brazil. Such a procedure eliminated the noises present in the records to be used in an ANN model to forecast the inflows seven days ahead. The results showed that the forecasting could be improved with such a procedure. Using the raw data, the ANN presented for the validation  $RMSE = 454.2828$  and  $R^2 = 0.7757$ , but with the filtering process using the Coiflet the results were  $RMSE = 200.0483$  and  $R^2 = 0.9603$ . Thus, the DWT was shown to be an efficient tool to eliminate noises present in streamflow

records, and the choice of the approximation  $A_3$  shows that only the elimination of  $D_1$ ,  $D_2$  and  $D_3$  details was enough to improve the forecasting. The error  $RMSE$  decreased by almost 55%, while the coefficient  $R^2$  has an increase above 25%, when compared to the ANN using the raw signal.

**Acknowledgement** The financial support provided by CNPq (National Council for Scientific and Technological Development, Brazil) and the data from ONS (National System Operator, Brazil) are gratefully acknowledged.

## REFERENCES

- Abrahart, R. J., Heppenstall, A. J. and See, L. M. (2007) Timing error correction procedures applied to neural network rainfall-runoff modelling. *Hydrol. Sci. J.* 52(3), 414–431.
- 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 watersheds. *J. Hydrol.* 390, 85–91.
- Adamowski, J. (2007) Development of a short-term river flood forecasting method based on wavelet analysis. *Polish Academy Sci. Monograph*, Warsaw, 124.
- 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. Processes* 22, 4877–4891.
- Ancil, F. and Tape, D. G. (2004) An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. *J. Environ. Engng and Sci.*, 3(s1), s121–s128(8).
- Besaw, L. E. *et al.* (2010) Advances in ungauged streamflow prediction using artificial neural networks. *J. Hydrol.* 386, 27–37.
- Birikundavyi, S., *et al.* (2002) Performance of Neural Networks in Daily Streamflow Forecasting. *J. Hydrol. Engng* 7(5), 392–398.
- Daubechies, I. (1992) *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM.
- Dawson, C. W. and Wilby, R. L. (1999) A comparison of artificial neural networks used for river forecasting. *Hydrol. and Earth System Sci.* 3(4), 529–540.
- Dawson, C. W., Abrahart, R. J. & See, L. M. (2007) HydroTest: a web-based toolbox of statistical measures for the standardised assessment of hydrological forecasts. *Environ. Modell. Softw.* 27, 1034–1052.
- Demuth, H. and Beale, M. (2005) *Neural network toolbox: for use with Matlab*. The MathWorks, Inc., Natick, USA.
- Farias, C. A. S., Santos, C. A. G. and Celeste, A. B. (2011) Daily reservoir operating rules by implicit stochastic optimization and artificial neural networks in a semi-arid land of Brazil. In: *Risk in Water Resources Management* (ed. by Gunter Bloeschl *et al.*) IAHS Publ. 347, 191–197. IAHS Press, Wallingford, UK.
- Grossman, A. and Morlet, J. (1984) Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.* 15, 723–736.
- Haykin, S. (1999) *Neural Networks: A Comprehensive Foundation*. 2 edn. Prentice Hall, Upper Saddle River, USA.
- Kisi, Ö. (2007) Streamflow forecasting using different artificial neural network algorithms. *J. Hydrol. Engng* 12(5), 532–539.
- Kisi, Ö. (2008) Stream flow forecasting using neuro-wavelet technique. *Hydrol. Processes* 22(20), 4142–4152.
- Mallat, S. (1989) A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions Pattern Analysis and Machine Intelligence* 11(7), 674–693.
- Napolitano, G., Serinaldi, F. and See, L. (2011) Impact of EMD decomposition and random initialisation of weights in ANN hindcasting of daily stream flow series: An empirical examination. *J. Hydrol.* 406(3–4), 199–214.
- Santos, C. A. G. and Silva, G. B. L. (2013) Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. *Hydrol. Sci. J.* 59(2), 321–324. doi: 10.1080/02626667.2013.800944.
- Santos, C. A. G., Galvão, C. O. and Trigo, R. M. (2003) Rainfall data analysis using wavelet transform. In: *Hydrology in the Mediterranean and Semiarid Regions* (ed. by Eric Servat *et al.*) IAHS Publ. 278, 195–201. IAHS Press, Wallingford, UK.
- Santos, C. A. G., Morais, B. S. and Silva, G. B. L. (2009) Drought forecast using artificial neural network for three hydrological zones in San Francisco river basin. IAHS Publ. 333, 302–312. IAHS Press, Wallingford, UK.
- Wang, W., Hu, S. and Li, Y. (2011) Wavelet transform method for synthetic generation of daily streamflow. *Water Resour. Management* 25(1), 41–57.
- Wu, C. L., Chau, K. W. and Li, Y.S. (2009) Methods to improve neural network performance in daily flows prediction. *J. Hydrol.* 372, 80–93.