Geophysical Research Abstracts, Vol. 7, 08651, 2005
SRef-ID: 1607-7962/gra/EGU05-A-08651

River flow forecasting using neural networks and wavelet analysis

B. Cannas (**??**), A. Fanni (**??**), **G. Sias** (**??**), S. Tronci (**??**), M. K. Zedda (**??**)

(**??**) Electric and Electronic Engineering Department, University of Cagliari, Italy, (**??**) Electric Engineering Department, University of Padova, Italy, (**??**) Chemical Engineering and Materials Department, University of Cagliari, Italy, (cannas@diee.unica.it / Fax: +390706755900)

Abstract

The paper deals with the evaluation of surface water resources for water management problems. A neural network has been trained to predict the hydrologic behavior of the runoff for the Tirso basin, located in Sardinia (Italy), at the S. Chiara section, by using the monthly time unit. In particular, due to high data non-stationarity and seasonal irregularity, typical of a Mediterranean weather regime, the role of data preprocessing through continuous and discrete wavelet transforms has been investigated.

Keywords: water management, runoff forecasting, neural networks, wavelet transform

1. Introduction

Monthly river flow forecast is a fundamental step for water resource system planning and management problems, since storage-yield sequences are frequently related to monthly periods.

Recently, artificial neural networks have been widely accepted as a potential useful way of modelling hydrologic processes, and have been applied to a range of different areas including rainfall-runoff, water quality, sedimentation and rainfall forecasting (Abrahart et al., 2004), (Cannas et al. 2004), (Baratti et al., 2003).

In this paper, we present a neural network technique for one month ahead forecasting of the runoff at the S. Chiara section in the Tirso basin located in Sardinia (Italy). Basic data for modelling are runoff time series with a monthly time step. The implementation of different neural network models to forecast runoff in a Sardinian basin was proposed in (Cannas et al. 2004), (Baratti et al., 2003). The results showed that most of the neural network models could be useful in constructing a tool to support the planning and management of water resources. The measures of efficiency obtained with the different models, although significantly greater than those obtained with traditional autoregressive models, were still only around 40%. A sizeable increase was obtained when the input data were manually partitioned into low, medium and high flows before training with three individual neural networks, indicating that this pre-processing technique warrants further investigation (Cannas et al. 2004). In fact, in general, and in Sardinian basins in particular, rainfall and runoff time series present high non-linearity and non-stationarity, and neural network models may not be able to cope with these two different aspects if no pre-processing of the input and/or output data is performed.

In this study wavelet transforms and neural networks have been applied to predict the hydrologic behavior of the runoff for the Tirso basin, located in Sardinia (Italy), at the S. Chiara section, by using the monthly time unit. *Wavelet* analysis is employed to pre-process the data to be inputted to a traditional Multi Layer Perceptron (MLP) neural network.

The wavelet decomposition of non-stationary time series into different scales provides an interpretation of the series structure and extracts the significant information about its history, using few coefficients. For these reasons, this technique is largely applied to times series analysis of non stationary signals (Nason and Von Sachs, 1999).

1 2. Wavelet analysis

The wavelet transform of a signal is capable of providing time and frequency information simultaneously, hence providing a time-frequency representation of the signal.

To do this, the data series is broken down by the transformation into its “wavelets”, that are a scaled and shifted version of the mother wavelet (Nason and Von Sachs, 1999).

The Continuous Wavelet Transform (CWT) of a signal $x(t)$ is defined as follows:

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

where s is the scale parameter, τ is the translation parameter and the ‘*’ denote the complex conjugate. Here, the concept of frequency is replaced by that of scale, determined by the factor s.

$\psi \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \cdot \psi^* \left[\frac{(t-\tau)}{a}\right] dt \psi$(t) is the transforming function and it is called *mother wavelet*. The term wavelet means small wave. The smallness refers to the condition that the function is of finite length. The wave refers to the condition that it is oscillatory. The term mother implies that the functions used in the transformation process are derived from one main function, the mother wavelet.

The wavelet coefficient $CWT_x^{\Psi}(\tau, s)$ is large when the signal $x(t)$and the wavelet $\Psi^*\left(\frac{t-\tau}{s}\right)$ are similar; thus, the time series after the wavelet decomposition allows one to have a look at the signal frequency at different scales. ψThe CWT calculation requires a significant amount of computation time and resources. Conversely, the Discrete Wavelet Transform (DWT) allows one to reduce the computation time and it is considerably simpler to implement than CWT. High pass and low pass filters of different cutoff frequencies are used to separate the signal at different scales. The time series is decomposed into one containing its trend (the approximation) and one containing the high frequencies and the fast events (the detail). The scale is changed by upsampling and downsampling operations.

DWT coefficients are usually sampled from the CWT on a dyadic grid in the space-scale plane, i.e., $s_0 = 2$ and $\tau_0 = 1$, yielding $s = 2^j$, and $\tau = k \cdot 2^j$.

The filtering procedure is repeated every time some portion of the signal corresponding to some frequencies is removed, obtaining the approximation and one or more details, depending on the chosen decomposition level.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left[\frac{(t-b)}{a}\right] \psi x(t) = \sum_{k=-\infty}^{\infty} c_k \phi_{J_0,k}(t) + \sum_{k=-\infty}^{\infty} \sum_{J=J_0}^{\infty} d_{J,k} \psi_{J,k}(t) \phi CWT_x^{\Psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} x(t) \Psi^*\left(\frac{t-\tau}{s}\right) dt \Psi(t) \tau$$

3. Case study

Data used in this paper are from the Tirso basin, located in Sardinia, at the S. Chiara section. Tirso basin is of particular interest because of its geographic configuration and water resource management as a dam was built in the S. Chiara section in 1924, providing water resources for central Sardinia. The basin area is 2,082.01 km^2 and is characterized by the availability of detailed data from several rainfall gauges. Recently, a new "Cantoniera Tirso" dam was built a few kilometers down the river, creating a reservoir with a storage volume of 780 Mm^3, one of the largest in Europe.

The data used for the hydrological model are limited to monthly recorded numerical time series associated with the runoff at the considered station. In previous works (Baratti et al., 2003) it has been verified that monthly averaged data of temperature at gauge stations and rainfall data were not strictly correlated with the monthly runoff behavior, hence these data are not considered here in the development of the model.

4. Performance indexes

The following measures of evaluation have been used to compare the performance of the different models, where N is the number of observations, O_i are the actual data and P_i are the predicted values:

Coefficient of Efficiency (Nash and Sutcliffe, 1970):

$$R = 1 - \frac{\sum\limits_{i=1}^{N} (O_i - P_i)^2}{\sum\limits_{i=1}^{N} \left(O_i - \overline{O}\right)^2} \tag{2}$$

The seasonal Coefficient of Efficiency following the definition in Lorrai and Sechi (Lorrai and Sechi, 1995):

$$R_d = \frac{\left(\sum\limits_{d=1}^{D} E_d\right) - E}{\sum\limits_{d=1}^{D} E_d} \tag{3}$$

where $E_d = \sum\limits_{i=d}^{D} (P_i - \bar{O}_i)$ and d=1 to Dmonths.

Root mean squared error:

$$RMSE = N^{-1} \sqrt{\sum_{i=1}^{N} (O_i - P_i)^2} \tag{4}$$

Mean absolute error:

$$MAE = N^{-1} \sum_{i=1}^{N} |O_i - P_i| \tag{5}$$

Mean higher order error function (M4E):

$M4E = \frac{\sum\limits_{i=1}^{N} (O_i - P_i)^4}{N}$ (6)σ

The measures of evaluation were calculated for each model. Table 1 shows the values reported in literature (Cannas et al., 2004) feeding the network with unpreprocessed data and when the input data were manually partitioned into low, medium and high flows and then used as input to three individual MLPs.

5. Data preprocessing and neural networks

The reconstruction of the hydrological system was accomplished using traditional feedforward, MLP networks. Cross validation was used as stop criterion. For this reason the data set was split into three parts: the first 40 years (480 monthly values) are used as training set, the second 9 years (108 monthly values) are used for cross validation while the last 20 years (240 monthly values) as test set.

The input dimension and the number of hidden nodes for every input combination were determined with a heuristic procedure, i.e., trying different combinations of input and hidden node numbers for reasonably small networks and keeping the topology which gives the best result in terms of root mean square error.

Runoff series is decomposed using continuous and discrete wavelet transforms and the obtained coefficients are given in input to one neural network or to a system of several networks to predict the runoff one month ahead.

A sliding window was advanced one element at a time through the runoff time series and the obtained wavelet coefficients are givenas inputs to a neural network. Thus, the sliding window amplitude represents the network memory.

We trained different neural networks to predict either the unprocessed runoff or the wavelet coefficients one step ahead. In the second case we trained an additional neural network to reconstruct runoff values from the predicted wavelet coefficients.

5.1 Continuous wavelet transform

A sliding window was advanced one element at a time through the runoff time series and the obtained wavelet coefficients are givenas inputs to the neural network to predict either the unprocessed runoff or the wavelet coefficients one step ahead.

In the second case we trained a neural network to reconstruct runoff values from wavelet coefficients.

Wavelet decomposition was made on runoff time series. We tested different scales s, from 1 up to 10, and different sliding window amplitudes.

In this context, dealing with a very irregular signal shape, we opted for an irregular wavelet, the Daubechies wavelet of order 4, DB4, (Daubechies, 1992).

Test case 1

The neural network has been trained using as input the CWT coefficients and using as outputs the same coefficients one month ahead.

A second neural network, feed with the predicted coefficients reconstructs the runoff values.

The predicted coefficients before going trough the MLP, were normalized between –1 and 1.

We obtained the best results using only the first scale coefficients. This means that high frequencies make up part of the process and do not represent just noise. The sliding window amplitude was of 8 months.

Table 2 shows the performance indexes for the test set.

Test case 2

The neural network has been trained using as input the CWT coefficients and using as outputs the corresponding runoff one month ahead. The sliding window amplitude was of 13 months.

We obtained best results using only the first scale coefficients.

Table 3 shows the performance indexes for the test set.

As can be noted, both models present better performance with respect to the case of unpreprocessed inputs, but perform worst than the system of networks working on partitioned data (see Table 1).

Moreover, results obtained reconstructing runoff from wavelet predicted coefficients through a neural network are only slightly better with respect to the case of direct runoff prediction from wavelet coefficients. The efficiency increase is not so important to justify the higher computational effort, due to the training of an additional network.

5.2 ***Discrete Wavelet Transform***

The runoff time series is decomposed into the approximation and detail coefficients for different decomposition levels, l, from 1 up to 4.Then, it is normalized between –1 and 1.

Test case 1

The neural network has been trained using as input the approximation coefficients at level l and using as outputs the same coefficients one month ahead.

A second neural network, feed with the predicted coefficients reconstructs the runoff values.

Best results have been obtained using as input the approximation coefficients at level $l = 4$ and a sliding window amplitude of 8 months.

Table 4 shows the performance indexes.

Test case 2

In this case, the runoff prediction is the result of the combination of several neural predictors:

a neural network has been trained using as input the approximation coefficients at level l and using as outputs the same coefficients one month ahead; l neural networks have been trained for the prediction of the l detail coefficients.

Another neural network, feed with the coefficients predicted by the previous mentioned networks, reconstructs the runoff values.

Best results have been obtained for l = 3 and a sliding window amplitude of 8 months.

Table 5 shows the performance indexes for the test set.

Test case 3

In this case, the neural network has been trained using as input the approximation coefficients at level $l$and using as outputs the runoff one month ahead.

Best results have been obtained for l = 3 and a sliding window amplitude of 32 months.

Table 6 shows the performance indexes for the test set.

It is worth noting that we obtained the best results with the discrete wavelet transformation using the approximation coefficients at level $l = 3$in input and runoff values in output. Furthermore, in this case only one neural network is necessary to obtain the runoff forecasting, resulting in a small computational effort. This result evidences the promising rule of the discrete wavelet transform in the neural network modeling when faster dynamics are important in the correct understanding of the process, but are embedded in noise.

6. Conclusions

We trained a neural network to predict the hydrologic behavior of the runoff for the Tirso basin, located in Sardinia (Italy), at the S. Chiara section, by using the monthly time unit. We preprocessed neural network inputs and outputs through continuous and discrete wavelet transforms, to take into account non-stationarity and seasonal irregularity of runoff time series.

Tests showed that the networks trained with pre-processed data present better performance with respect to networks trained with undecomposed noisy raw signals. In particular, we obtained best results preprocessing data through the discrete wavelet transformation and training one neural network using as input the approximation coefficients at level three and runoff values in output.

This results, and those reported in literature with a data partitioning technique, evidences the promising role of data clustering and discrete wavelet transform techniques combination, in water flow forecasting.

References

(Abrahart et al, 2004) Abrahart, R.J., Kneale, P.E., See, L., Neural Networks in Hydrology, A.A. Balkema, Rotterdam.

(Baratti et al., 2003) Baratti, R., Cannas, B., Fanni, A., Pintus, M., Sechi, G.M., Toreno, N., 2003, River flow forecast for reservoir management through neural networks, NeuroComputing, v. 55, p. 421-437.

(Cannas et al., 2004) Cannas, B., Montisci, A., Fanni, A., See, L., Sechi, G.M. Comparing artificial neural networks and support vector machines for modelling rainfall-runoff, In: Liong, Phoon & Babovic editors, World Scientific Publishing Company, Proceedings of the 6^{th} International Conference on Hydroinformatics.

(Daubechies, 1992) Daubechies, I., (1992), Ten Lectures on wavelets., CSBM – NSF Series Application Mathematics, 61, SIAM publication, Philadelphia, PA.

(Lorrai and Sechi, 1995) Lorrai, M., Sechi, G.M., 1995. Neural nets for modeling rainfall-runoff transformations, Water Resources Management, v. 9, p.299-313.

(Nash and Sutcliff, 1970) Nash, J.E., Sutcliffe, J.V., 1970, River flow forecasting through conceptual models, I: A discussion of principles, Journal of Hydrology, v. 10, p. 398-409.

(Nason and Von Sachs, 1999) Nason, G. P., and Von Sachs, R., 1999, Wavelets in time series analysis, Phil. Trans. Roy. Soc. A, v. 357, p. 2511 – 2526.

Tables

	R	Rd	RMSE $(Mm^3)^{1/2}$	MAE (Mm^3)	M4E $(Mm^3)^4 x10^{-6}$
MLP with not preprocessed Input	0.42	0.38	29	19	0.5
Three MLPs with Data Partitioning	0.57	0.48	10.09	8.7	0.02

TABLE 1: Performance indexes for the test data set: Literature results

CWT	R	Rd	RMSE $(Mm^3)^{1/2}$	MAE (Mm^3)	M4E $(Mm^3)^4 x10^{-6}$
First scale coefficients; Sliding window of 8 months	0.45	0.36	11.74	8.46	0.1

TABLE 2: CWT- Performance indexes for the test data set: Wavelet coefficients in input and in output; Runoff reconstruction through a neural network.

CWT	R	Rd	RMSE $(Mm^3)^{1/2}$	MAE (Mm^3)	M4E $(Mm^3)^4$x10^{-6}
First scale coefficients; Sliding window of 13 months	0.44	0.34	11.86	8.11	0.1

TABLE 3: CWT- Performance indexes for the test data set: Wavelet coefficients in input, runoff in output.

DWT	R	Rd	RMSE $(Mm^3)^{1/2}$	MAE (Mm^3)	M4E $(Mm^3)^4$x10^{-6}
Level 4 Sliding window of 8 months	0.38	0.27	12.52	8.56	0.15

TABLE 4: DWT- Performance indexes for the test data set: Wavelet approximation coefficients in input and in output; Runoff reconstruction through a neural network.

DWT	R	Rd	RMSE $(Mm^3)^{1/2}$	MAE (Mm^3)	M4E $(Mm^3)^4$x10^{-6}
Level 3 Sliding window of 8 months	0.40	0.29	12.31	8.05	0.18

TABLE 5: DWT- Performance indexes for the test data set: Combination of several neural predictors. Performance indexes for the test data set.

TABLE 6: DWT- Performance indexes for the test data set: Wavelet approximation coefficients in input, runoff in output.

DWT	R	Rd	RMSE $(Mm^3)^{1/2}$	MAE (Mm^3)	M4E $(Mm^3)^4$x10^{-6}
Level 3 Sliding window of 32 months	0.47	0.37	11.59	7.53	0.13