Applied Soft Computing 68 (2018) 494–506

Contents lists available at ScienceDirect

Applied Soft Computing

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

Performance analysis of unorganized machines in streamflow forecasting of Brazilian plants

Hugo Siqueira [a,*], Levy Boccato [b], Ivette Luna [c], Romis Attux [b], Christiano Lyra [d]

[a] Department of Electric Engineering, Federal University of Technology – Paraná, Monteiro Lobato Av., Km4, s/n, Ponta Grossa, Paraná, 84016-210, Brazil
[b] Department of Computer Engineering and Industrial Automation, School of Electrical and Computer Engineering, University of Campinas, Albert Einstein Av., 400, Campinas, São Paulo, 13083-852, Brazil
[c] Institute of Economics, University of Campinas, Pitagoras St, 353, Campinas, São Paulo, 13083-857, Brazil
[d] Department of Systems and Energy, School of Electrical and Computer Engineering, University of Campinas, Albert Einstein Av., 400, Campinas, São Paulo, 13083-852, Brazil

ARTICLE INFO

Article history:
Received 2 October 2017
Received in revised form 3 February 2018
Accepted 3 April 2018
Available online 12 April 2018

Keywords:
Monthly seasonal streamflow series forecasting
Unorganized machines
Extreme learning machines
Echo state networks
Variable selection

ABSTRACT

This work performs an extensive investigation about the application of unorganized machines – extreme learning machines and echo state networks – to predict monthly seasonal streamflow series, associated to three important Brazilian hydroelectric plants, for many forecasting horizons. The aforementioned models are neural network architectures which present efficient and simple training processes. Moreover, the selection of the best inputs of each model is carried out by the wrapper method, using three different evaluation criteria, and three filters, *viz.*, those based on the partial autocorrelation function, the mutual information and the normalization of maximum relevance and minimum common redundancy method. This study also establishes a comparison between the unorganized machines and two classical models: the partial autoregressive model and the multilayer perceptron. The computational results demonstrate that the unorganized machines, especially the echo state networks, represent efficient alternatives to solve the task.

1. Introduction

Time series prediction [1] is an important task which can be solved by using neural networks. The possibility of having anticipated knowledge about some phenomena of interest constitutes a strategic advantage that can be decisive for decision-making processes in several areas, such as economics, meteorology and operations research [2–5]. A practical instance of this problem is related to monthly seasonal streamflow series, which has attracted the attention of scientific community due to: (*i*) its influence in many aspects related to energetic planning, such as pricing strategies and the expected availability of energy for both short and long-term horizons [6], especially when energy generation strongly depends on hydroelectric plants, as occurs in Brazil, and (*ii*) the challenging features of streamflow series in view of their non-stationary character, which is a consequence of a seasonal component that reflects the different rainfall periods/volumes in each season.

The forecasting process associated with monthly streamflow series generally involves five steps, as shown in Fig. 1.

The first step corresponds to the data acquisition, which is performed with the aid of sensors placed along the rivers that carry the water to the generating turbines of the hydroelectric plants. Hence, a degree of uncertainty is inevitably present in the obtained measurements.

The available data form a non-stationary discrete-time series that is pre-processed in the second step, which promotes a statistical treatment named deseasonalization, which aims at removing the seasonal component of the series in order to obtain an approximately stationary transformed series with zero mean and unit variance [1,7]. Unfortunately, this procedure may create a certain incompatibility, as a set of excellent estimates obtained in the deseasonalized domain does not necessarily lead to equally accurate predictions in the original series domain, after the seasonal component is reinserted [6].

The third step copes with the questions as to how many and which streamflow samples should be offered to the predictor as input variables. The techniques devoted to this task, known as vari-

* Corresponding author.
E-mail addresses: hugosiqueira@utfpr.edu.br (H. Siqueira), lboccato@dca.fee.unicamp.br (L. Boccato), ivette@eco.unicamp.br (I. Luna), attux@dca.fee.unicamp.br (R. Attux), chrlyra@densis.fee.unicamp.br (C. Lyra).

https://doi.org/10.1016/j.asoc.2018.04.007
1568-4946/

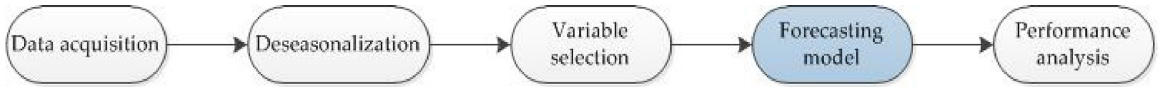

Fig. 1. Block diagram of the main steps involved in the streamflow series forecasting.

able selection, such as those based on wrappers and filters [8,9], usually take into account statistical properties concerning the candidate variables, such as correlation and mutual information, or even use the prediction error associated with the model considering different input sets as the criterion to select the variables that should be employed. Metrics that penalize a high number of inputs in the set can also be applied, as the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) [10,11].

The fourth step is of paramount importance, since it involves the definition of the model that shall be effectively used for producing the future estimates of the series. In this context, there are two crucial aspects regarding the forecasting model: (1) the selection of the prediction structure; and (2) the strategy used for adapting the characteristics of the chosen model to the specific task. The first choice is vital since it may limit the approximation capability of the model and, thus, the attainable performance. Additionally, it also affects the complexity of the corresponding adaptation process, which usually amounts to the task of adjusting the parameters of the model, such as coefficients in weighted combinations, in order to achieve a high similarity between the predicted values and the actual measured data.

Finally, we can point out as the fifth step the assessment of the prediction accuracy of the chosen models according to predefined error measures, such as the mean squared error (MSE) and the mean absolute error (MAE).

Although the design of an effective prediction strategy essentially relies on three choices – those of the deseasonalization method, the variable selection strategy and the forecasting model –, it is possible to affirm, in light of the previous exposition, that selecting the prediction structure constitutes the key element and, thus, a significant portion of the literature related to time series prediction (including streamflow series) has focused on studying the use of different forecasting models [12,13].

In this context, artificial neural networks (ANNs) constitute important tools to solve forecasting problems due to their capability of generating complex input-output mappings [12]. In fact, neural models, such as the multilayer perceptron (MLP) and the radial basis function (RBF) networks, that are universal approximators [14,15]. Recurrent neural networks (RNNs), which are endowed with feedback connections between the internal neurons, also present the universal approximation capability [16,63], but have the advantage of automatically creating an internal memory about the history of the involved signals, which can be particularly useful in scenarios where there is a temporal dependence between the information sources, as occurs in the case of time series prediction [17].

The effective use of neural networks requires the adaptation of their free parameters. This process usually is formulated according to a supervised learning approach that involves minimizing a cost function based on certain error metric with respect to the synaptic weights, which is commonly carried out with the aid of gradient-based algorithms [12,53].

Alternatively, two proposals, known as extreme learning machines (ELMs) [18,19] and echo state networks (ESNs) [20,21], simplify the training process by using intermediate layers whose parameters can be randomly set in advance, i.e., without taking into account any information about an error signal. Hence, only the linear combiner at the output layer is effectively adapted, which essentially requires the solution of a linear regression task [22–24]. Therefore, ELMs and ESNs can be seen as trade-off solutions that achieve a noticeable simplification with respect to the complexity of the training process, while the processing capability of the underlying structure is, to a certain degree, preserved. Due to these features, these models, which can be called unorganized machines, as suggested in Boccato et al. [25], represent interesting candidates to be employed in the streamflow series forecasting problem.

This perspective has been initially addressed in Sacchi et al. [26], where a performance study involving the original architecture proposed by Jaeger [24] has been performed. Then, Siqueira et al. [27–29,6] have extended this investigation in several important aspects by: (*i*) including ELMs and alternative ESN architectures, such as those proposed by Boccato et al. [21] and Butcher et al. [30] in the comparative study, (*ii*) considering scenarios beyond the classical one-step ahead prediction case, and (*iii*) evaluating the performance for several streamflow series with different hydrological behaviors.

Motivated by the encouraging results reported in Siqueira et al. [6], we intend to continue investigating the application of ELMs and ESNs to the streamflow series forecasting problem by considering three possibilities. The first perspective involves the use of different variable selection mechanisms. The previous efforts of Siqueira et al. [27–29,6] have selected the input past samples based on a simple correlation measure, more specifically, the partial autocorrelation function (PACF) [31,32]. In this work, by resorting to more sophisticated criteria to select the number and the set of past samples to be used, we aim at reaching a more parsimonious solution in terms of: (1) the computational cost associated with the variable selection stage, (2) the size of the input dataset, and (3) the accuracy attained by the prediction model.

The second perspective is associated with the methodology adopted for the predictor design: instead of constructing a single predictor for the whole series, as performed in Siqueira et al. [27–29,6], we shall use a specially-tailored predictor for each month – so that the prediction model actually consists of twelve predictors –, which are designed to generate one-step ahead estimates, similarly to the currently adopted solution in the electric sector of Brazil. The expectation behind this approach is that it may be possible to deal more adequately with the seasonal character of the series if we treat each month separately, since the series exclusively related to a specific month preserves, in a more extensive manner, its main characteristics along the years when compared with the entire streamflow series [33,48].

However, unlike the currently adopted solution, which is based on linear models, the prediction structures we shall consider are nonlinear and, in the case of ESNs, recurrent, which have an extended information processing capability and, thus, offer the potential of improving the prediction accuracy.

The third perspective refers to the strategy employed for multistep ahead forecasting [34]. In the previous works, the authors considered the direct approach, whereas, in this investigation, the recursive approach is adopted [35], which means that the designed predictors yield estimates for multiple steps ahead by making successive one-step predictions.

Our objective is to provide a comparative analysis of several prediction models to the monthly streamflow series forecasting problem, as well as to point out promising tendencies that contribute to reaching better results in this task. By analyzing the performance of the models considering the three aforementioned