Journal of Hydrology (2008) **362**, 291–307

available at www.sciencedirect.com

ScienceDirect

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

Jordan recurrent neural network versus IHACRES in modelling daily streamflows

Elena Carla Carcano [a,*], Paolo Bartolini [a,1], Marco Muselli [b,2], Luigi Piroddi [c,3]

[a] *Dipartimento di Ingegneria Ambientale, Facoltà di Ingegneria Via Montallegro 1, 16145 Genova, Italy*
[b] *Istituto di Elettronica ed Ingegneria dell'Informazione e delle Telecomunicazioni, Consiglio Nazionale delle Ricerche di Genova — Via De Marini 6, 16149 Genova, Italy*
[c] *Dipartimento di Elettronica e Informazione, Politecnico di Milano — Via Ponzio 3415, 20133 Milano, Italy*

Received 18 April 2007; received in revised form 12 August 2008; accepted 30 August 2008

KEYWORDS
Feedforward neural network;
Recurrent neural network;
Tapped delayed line;
Memory effect;
Temporal dependence

Summary A study of possible scenarios for modelling streamflow data from daily time series, using artificial neural networks (ANNs), is presented. Particular emphasis is devoted to the reconstruction of drought periods where water resource management and control are most critical. This paper considers two connectionist models: a feedforward multilayer perceptron (MLP) and a Jordan recurrent neural network (JNN), comparing network performance on real world data from two small catchments (192 and 69 km^2 in size) with irregular and torrential regimes. Several network configurations are tested to ensure a good combination of input features (rainfall and previous streamflow data) that capture the variability of the physical processes at work. Tapped delayed line (TDL) and memory effect techniques are introduced to recognize and reproduce temporal dependence. Results show a poor agreement when using TDL only, but a remarkable improvement can be obtained with JNN and its memory effect procedures, which are able to reproduce the system memory over a catchment in a more effective way. Furthermore, the IHACRES conceptual model, which relies on both rainfall and temperature input data, is introduced for comparative study. The results suggest that when good input data is unavailable, metric models perform better than conceptual ones and, in general, it is difficult to justify substantial conceptualization of complex processes.

* Corresponding author. Tel.: +39 010 3532472; fax: +39 010 3532546.
E-mail addresses: carcano@diam.unige.it (E.C. Carcano), bart@diam.unige.it (P. Bartolini), marco.muselli@ieiit.cnr.it (M. Muselli), piroddi@elet.polimi.it (L. Piroddi).
[1] Tel.: +39 010 3532472; fax: +39 010 3532546.
[2] Tel.: +39 010 6475213; fax: +39 010 6475200.
[3] Tel.: +39 02 23993556; fax: +39 02 23993412.

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2008.08.026

Introduction

Modelling of the rainfall—streamflow (R—R) transformation at any time scale has been a primary concern of hydrological research for several decades and has resulted in plenty of models proposed in literature. Following Beck (1991), these models can be divided into three categories: metric (empirical), conceptual and physics-based. Metric models are deeply observation-based: they pursue the system response by extrapolating information from available data. These models are founded on the mathematical link between input and output series (e.g. rainfall and streamflow data) considering the catchment as a lumped unit, with no exploration of the spatial inhomogeneities of the basin. Besides ANNs, examples of this type of model include classical ARMA, initially developed by Box and Jenkins (1976) and all its extensions, and transfer function models (Hipel and McLeod, 1994).

The second category of models, on the other hand, describes all the relevant components of hydrological processes through simplified conceptualisations. A further step towards complexity is represented in physics-based models, as they use a theoretical equation for each process considered, e.g. the Saint Venant equation for simulation of flow WHERE. Examples include: IHDM (Beven et al., 1987), SWATCH (Morel-Seytoux and Al Hassoun, 1989). Despite their ambition to use spatially-distributed parameters which reflect the heterogeneity of the catchment, they are of limited practicality in most contexts use due to their complexity and data availability requirements. Herein ANNs and IHACRES are introduced for a comparative study.

IHACRES (Jakeman et al., 1990; Littlewood et al., 1997) is an example of a hybrid conceptual-metric model as it uses a conceptual module to estimate the effective rainfall and a transfer function module to convert effective rainfall into streamflow. ANNs are an example of black-box models.

In black-box models, unfortunately, no physical insight is possible and the structure of the model is generally chosen from a family that shows good flexibility and has been successfully employed in a similar situation elsewhere (Sjöberg et al., 1995). ANNs, in general, have been proved to provide useful solutions when applied to (1) complex systems that, otherwise, may be poorly reproduced, (2) problems tainted by noise, and (3) circumstances where input is incomplete or ambiguous by nature. ANNs are suited for modelling the R—R relationship due to their ability to synthesize a reliable model without needing any prior knowledge of the functional relationship between dependent and independent variables and to treat difficult issues such as the high nonlinearity involved in such a processes.

It has been proven (Cybenko, 1989) that ANNs are able to approximate with arbitrary accuracy (by increasing the number of neurones) any function with a finite number of discontinuities.

Although feedforward ANNs were introduced in 1986, mainly through the book of Rumelhart, Hinton and McClelland, their application in hydrological modelling began only in the middle of 1990s. Pioneer researches by Zhu and Fujita (1994) compared the performance of a feedforward ANN to fuzzy logic in predicting a 3 h lead runoff. Runoff dependence was expressed by using a window of previous rainfall inputs. Later, Campolo et al. (1999) made use of a feedforward network to predict the occurrence of flood events from distributed rainfall and hydrometer data on an hourly time scale.

Recurrent neural networks (RNNs) have also been used in hydrology and related fields. Connor et al. (1994) highlighted the advantages of recurrent over feedforward neural networks for forecasting time series that include moving average components. Anomala et al. (2000) found out that RNNs perform better when compared to other ANN architectures for predicting watershed streamflows. Recent work done by Kumar et al. (2004) compares the traditional feedforward approach to RNNs (trained with ordered partial derivatives), to forecast monthly river flows.

Problems in daily rainfall—streamflow modelling

Recent severe droughts in many European countries (and elsewhere) have had a significant impact on water supplies which, in turn, has serious economic and social consequences. Hence, the need to find better tools for management and design of water resources has become a more important issue than ever before. Despite considerable improvements introduced into the catchment modelling field, hydrologists are still limited by a number of factors in rainfall—streamflow modelling. Inadequate data availability and poor forecast/simulation capability of mostly used models are two of the most significant.

Regarding data availability, it is known that the transformation of inflow to watershed streamflow depends on a plethora of hydrological and climatic factors such as precipitation, evapotranspiration, temperature, soil moisture and snow water equivalence, and many more.

In a favourable scenario, where most of these variables have actually been measured over a reasonable period of time, conceptual and physics-based models produce different levels of understanding of the R—R process. Sophisticated conceptual models like TOPMODEL and MIKE SHE (Abbott et al., 1986) implement both climatic and catchment descriptive data (e.g., topography, vegetation, and soil) and represent the inner mechanism of the R—R transformation in a detailed way. Thus the level of physical process understanding that can be gleaned from them is quite high. On the other hand, less data-demanding models, like IHACRES, which require rainfall, streamflow and evapotranspiration only, produce only the more practical results and understanding. When a poor level of information available does not allow the conceptualization of complex phenomena, or in general, when data are collected only at a few specific sites in a catchment, empirical models become the best choice (Coulibaly et al., 2000). In many practical circumstances, where the main concern can only be to make accurate predictions with no insight on the internal structure of the process involved, the authors believe black-box models can provide suitable and accurate solutions.

The authors believe that, aside from data requirements, the choice between conceptual and empirical models depends on the temporal scale under consideration. Both conceptual and physics-based models perform well in continuous or short time scales (daily, sub-daily) where the