Journal of Hydrology 514 (2014) 358–377

Contents lists available at ScienceDirect

Journal of Hydrology

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

Review Paper

Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review

Vahid Nourani [a,*], Aida Hosseini Baghanam [a], Jan Adamowski [b], Ozgur Kisi [c]

[a] Dept. of Water Resources Eng., Faculty of Civil Eng., Univ. of Tabriz, Iran
[b] Dept. of Bioresource Eng., McGill Univ., Quebec, Canada
[c] Eng. Faculty, Canik Basari University, Samsun, Turkey

ARTICLE INFO

Article history:
Received 21 August 2013
Received in revised form 16 February 2014
Accepted 24 March 2014
Available online 1 April 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Fi-John Chang, Associate Editor

Keywords:
Hydro-climatology
Black box model
Artificial Intelligence
Wavelet transform
Hybrid model

SUMMARY

Accurate and reliable water resources planning and management to ensure sustainable use of watershed resources cannot be achieved without precise and reliable models. Notwithstanding the highly stochastic nature of hydrological processes, the development of models capable of describing such complex phenomena is a growing area of research. Providing insight into the modeling of complex phenomena through a thorough overview of the literature, current research, and expanding research horizons can enhance the potential for accurate and well designed models.

The last couple of decades have seen remarkable progress in the ability to develop accurate hydrologic models. Among various conceptual and black box models developed over this period, hybrid wavelet and Artificial Intelligence (AI)-based models have been amongst the most promising in simulating hydrologic processes. The present review focuses on defining hybrid modeling, the advantages of such combined models, as well as the history and potential future of their application in hydrology to predict important processes of the hydrologic cycle. Over the years, the use of wavelet–AI models in hydrology has steadily increased and attracted interest given the robustness and accuracy of the approach. This is attributable to the usefulness of wavelet transforms in multi-resolution analysis, de-noising, and edge effect detection over a signal, as well as the strong capability of AI methods in optimization and prediction of processes. Several ideas for future areas of research are also presented in this paper.

Contents

Abbreviations: ACO, Ant Colony Optimization; AI, Artificial Intelligence; ANFIS, Adaptive Neuro-Fuzzy Inference System; ANN, Artificial Neural Network; AR, Auto Regressive; ARIMA, Auto Regressive Integrated Moving Average; ARIMAX, ARIMA with exogenous input; CWT, Continues Wavelet Transform; DWT, Discrete Wavelet Transform; dbn, Daubechies order *n* wavelet; GA, Genetic Algorithm; GEP, Gene-Expression Programming; GP, Genetic Programming; GWL, Groundwater Level; LR, Linear Regression; MA, Moving Average; MAE, Mean Absolute Error; MARS, Multivariate Adaptive Regression Spline; MLP, Multi-Layer Perceptron; MLR, Multiple Linear Regression; NF, Neuro Fuzzy; NN, Neural Network; PSO, Particle Swarm Optimization; RMSE, Root Mean Square Error; SOM, Self-Organizing Map; SPI, Standard Precipitation Index; SRC, Sediment Rating Curve; SSA, Singular Spectrum Analysis; SSC, Suspended Sediment Concentration; SSL, Suspended Sediment Load; SST, Sea Surface Temperature; SVM, Support Vector Machine; SVR, Support Vector Regression; WANFIS, Wavelet–ANFIS; WANN, Wavelet–artificial neural network; WBANN, Wavelet–Bootstrapping ANN; WGRNN, Wavelet–Generalized Regression NN; WMF, Wavelet Modeling Framework; WMRA, Wavelet Multi-Resolution Analysis; WNF, Wavelet–Neuro Fuzzy; WR, Wavelet–Regression; WVC, Wavelet–Volterra.

* Corresponding author. Tel.: +98 914 403 0332; fax: +98 411 334 4287.

E-mail addresses: vnourani@yahoo.com, vnourani@umn.edu (V. Nourani), hosseinibaghanam@gmail.com (A. Hosseini Baghanam), jan.adamowski@mcgill.ca (J. Adamowski), okisi@basari.edu.tr (O. Kisi).

http://dx.doi.org/10.1016/j.jhydrol.2014.03.057
0022-1694/

1. Introduction

Characterized by high complexity, dynamism and non-stationarity, hydrological and hydro-climatologic forecasting has always presented a challenge to hydrologists who recognize its essential role in environmental and water resources management as well as in water-related disaster mitigation. Recent years have seen a significant rise in the number of scientific approaches applied to hydrologic modeling and forecasting, including the particularly popular 'data-based' or 'data-driven' approaches. Such modeling approaches involve mathematical equations drawn not from the physical process in the watershed but from an analysis of concurrent input and output time series (Solomatine and Ostfeld, 2008). Such models can be defined on the basis of connections between the system state variables (input, internal and output variables) with only a limited number of assumptions being made regarding the physical behavior of the system. Typical examples of data-driven models are rating curves, the unit hydrograph method and various statistical models (Linear Regression; LR, multi-linear, Auto Regressive Integrated Moving Average; ARIMA) and methods of machine learning. The conventional black box time series models such as ARIMA, ARIMA with exogenous input (ARIMAX) and Multiple Linear Regression (MLR) are linear models and assume stationarity of the dataset. Such models are unable to handle non-stationarity and non-linearity involved in hydrological processes. As a result, many researchers have focused on developing models that are able to model non-linear and non-stationary processes.

The data-driven methods of Artificial Intelligence (AI) have shown promise in modeling and forecasting non-linear hydrological processes and in handling large amounts of dynamicity and noise concealed in datasets. Such properties of AI-based models are well suited to hydrological modeling problems. Numerous AI tools or techniques have been used, including versions of search optimization, mathematical optimization, as well as logic-, classification-, statistical learning- and probability-based methods (Luger, 2005). In particular, three sub-sets of AI have been widely used in the hydro-climatologic and environmental fields:

(1) *Evolutionary computation:* A branch of optimization methods that includes swarm intelligence algorithms such as Ant Colony Optimization (ACO; Dorigo et al., 1996) or Particle Swarm Optimization (PSO; Kennedy and Eberhart, 1995) and evolutionary algorithms such as Genetic-Algorithms (GA; Goldberg, 2000), Gene-Expression Programming (GEP), and Genetic-Programming (GP; Koza, 1992).
(2) *Fuzzy logic:* Fuzzy systems (Zadeh, 1965) can be used for uncertain reasoning, which provide a logic perspective in AI techniques.
(3) *Classifiers and statistical learning methods:* These models employ statistical and machine-learning approaches. The most widely used classifiers are Neural Networks (NNs; Haykin, 1994), kernel methods such as the Support Vector Machine (SVM; Vapnik, 1995), *k*-nearest neighbor algorithms such as Self-Organizing Map (SOM; Kohonen, 1997), Gaussian mixture model, naive Bayes classifier, and decision tree. NNs, the predominant AI method, are used in hydrology via two approaches: (i) supervised, including acyclic or feed-forward NNs (where the signal passes in only one direction) and recurrent NNs (which allow feedback), and (ii) unsupervised (*e.g.*, SOM).

Among the broader applications of AI methods, GA, GP, Fuzzy, NNs, and SVM are widely used in different fields of hydrology. Since their emergence in hydrology, the efficient performance of AI techniques such as data-driven models has been reported over a wide range of hydrological processes (*e.g.*, precipitation, stream-flow, rainfall–runoff, sediment load, groundwater, drought, snowmelt, evapotranspiration, water quality, *etc.*). The number of researchers active in this area has increased significantly over the last decade, as has the number of publications. Several dozen successful applications for hydrological process modeling (*e.g.*, stream-flow, rainfall–runoff, sediment, groundwater, water quality) using ANN, Fuzzy, GP, GA, and SVM have been reported, with some examples listed in Table 1.

Despite the flexibility and usefulness of AI-based methods in modeling hydrological processes, they have some drawbacks with highly non-stationary responses, *i.e.*, which vary over a wide scale of frequencies, from hourly to multi-decadal. In such instances of 'seasonality', a lack of input/output data pre/post-processing, may not allow AI models to adequately handle non-stationary data. Here, hybrid models which combine data pre/post-processing schemes with AI techniques can play an important role.

Hybrid hydrological models may take advantage of black box (here AI-based) models and their ability to efficiently describe observed data in statistical terms, as well as other prior information, concealed in observed records. The hybrid models discussed here represent the joint application of AI-based methods with the wavelet transform to enhance overall model performance.

As an advance in signal processing, wavelet transforms can reliably obviate AI model shortcomings in dealing with non-stationary behavior of signals. A mathematical technique useful in numerical analysis and manipulation of multidimensional signal sets, wavelet analysis provides a time-scale representation of the process and of its relationships. Indeed, the main property of the wavelet transform is its ability to provide a time-scale localization of a process. The wavelet transform has attracted significant attention since its theoretical development in 1984 (Grossmann and Morlet, 1984). A number of recent hydrological studies have implemented wavelet analysis (*e.g.*, Adamowski and Sun, 2010; Kim and Valdes, 2003; Kisi, 2009a,b, 2010; Nourani et al., 2009a,b, 2011; Maheswaran and Khosa, 2012a; Partal and Kisi, 2007; Sang, 2012; Tiwari and Chatterjee, 2010; Zhou et al., 2008).

The Wavelet transform is applicable in extracting nontrivial and potentially useful information, or knowledge, from the large data sets available in experimental sciences (historical records, reanalysis, global climate model simulations, *etc.*). Providing explicit information in a readable form, it can be used to solve diagnostic, classification or forecasting problems. In a review of the applications of the wavelet transform in hydrologic time series modeling, Sang (2013a) highlighted the multifaceted information that can be drawn from such analysis: characterization and understanding of hydrologic series' multi-temporal scales, identification of seasonalities and trends, and data de-noising. Therefore, the ability of the wavelet transform to decompose non-stationary signals into