Journal of Hydrology 394 (2010) 458–470

Contents lists available at ScienceDirect

Journal of Hydrology

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

Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach

Mukesh K. Tiwari, Chandranath Chatterjee *

Agricultural and Food Engineering Department, Indian Institute of Technology, Kharagpur, West Bengal 721 302, India

ARTICLE INFO

Article history:
Received 24 March 2010
Received in revised form 20 September 2010
Accepted 2 October 2010

This manuscript was handled by Dr. A. Bardossy, Editor-in-Chief, with the assistance of Vazken Andréassian, Associate Editor

Keywords:
Flood forecasting
Wavelet
Bootstrap
ANNs

SUMMARY

A hybrid wavelet–bootstrap–ANN (WBANN) model is developed in this study to explore the potential of wavelet and bootstrapping techniques for developing an accurate and reliable ANN model for hourly flood forecasting. The wavelet technique is used to decompose the times series data into different components which capture useful information on various resolution levels. Five years hourly water level data for monsoon season from five gauging stations in Mahanadi River basin, India are used in this study. The observed water level time series of a particular gauging station is decomposed to sub-series by discrete wavelet transformation and then appropriate sub-series are added up to develop new time series. The bootstrap resampling method is used to generate different realizations of the newly constructed datasets using discrete wavelet transformation to create a set of bootstrap samples that are finally used as input to develop WBANN model. Performance of WBANN model is also compared with three different ANN models: traditional ANNs, wavelet based ANNs (WANNs), bootstrap based ANNs (BANNs). The results showed that the hybrid models WBANN and BANN produced better results than the traditional ANN and WANN models. WBANN model simulated the peak water level better than ANN, WANN and BANN models, and in general, the overall performance of WBANN model is accurate and reliable as compared to the other three models. This study reveals that whereas wavelet decomposition improves the performance of ANN models, bootstrap resampling technique produces more consistent and stable solutions. WBANN model is also used to assess the predictive uncertainty in forms of confidence intervals (CI) to assess the predictive uncertainty for 1–10 h lead time forecasts. Results obtained indicate that WBANN forecasting model with confidence intervals can improve their reliability for flood forecasting.

1. Introduction

Hourly flood forecasting is desirable with sufficient lead time for taking appropriate flood prevention measures, evacuation plan and rehabilitation actions. A wide variety of rainfall–runoff models have been developed and applied for flood forecasting either based on mechanistic approach or on a systems theoretic approach. Spatially distributed modeling is a typical example of the mechanistic approach to construct a model that explicitly accounts for as much of the small-scale physics and the natural heterogeneity as computationally possible (Loague and VanderKwaak, 2004). The approach has been criticized for resulting in models that are overly complex, leading to problems of over parameterization and equifinality (Beven, 2006), which may manifest itself in large prediction uncertainty (Uhlenbrook et al., 1999), whereas system theoretic approach gives more emphasis on system operation and not the nature of the system or the physical laws governing its operation. Hydrological ANN models are simplification of more complex system where the natural processes are simulated with mathematical equations and the corresponding parameters are derived from observations and experience leading to uncertainty (Srivastav et al., 2007; Han et al., 2007a).

System theoretic approach in the form of artificial neural networks (ANNs) have gained great attentions by the researchers in last few decades for river flow forecasting. The ability of ANN in mapping complex nonlinear input–output relationship has increased the number of applications in rainfall–runoff modeling and river discharge forecasting (Jain and Srinivasulu, 2004; Altunkaynak, 2007; El-Shafie et al., 2007). There are several type of ANNs but the major advantage of feed forward backpropagation artificial neural network (FFBP ANN) is that it is less complex than other ANNs such as radial basis function (RBF) and support vector machine (SVM) and has similar nonlinear input–output mapping capability (Sudheer and Jain, 2003; Coulibaly and Evora, 2007; Bray and Han, 2004; Han et al., 2007b). Another type of ANN, generalized regression neural networks (GRNNs) has some advantages of being less sensitive to initial weights and do not produce negative values compared to the FFBP ANN (Cigizoglu, 2005a,b;

* Corresponding author. Tel.: +91 3222 283158.
E-mail addresses: mukesh_k_tiwari@yahoo.com (M.K. Tiwari), cchatterjee@agfe.iitkgp.ernet.in (C. Chatterjee).

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.10.001

Cigizoglu and Alp, 2006; Kisi and Cigizoglu, 2007; Sertel et al., 2008), but some literature show mixed performance of FFBP and GRNN (Toprak and Cigizoglu, 2008; Sertel et al., 2008; Tanrikulu, 2009; Cigizoglu, 2005a). Even though, fuzzy inference system has been successfully applied in river flow forecasting (Nayak et al., 2004; Mukerji et al., 2009) their application is rather limited in comparison to neural network models and there are several unresolved issues requiring further attention before more clear guidelines for the application of fuzzy inference systems can be given (Jacquin and Shamseldin, 2009). Hybrid GA-based ANN algorithm is found to avoid over-fitting and produces better accuracy in model performance but at the expense of additional modeling parameters and longer computation time (Wu and Chau, 2006). FFBP ANNs are known to have several dozens of successful applications in river basin management and related problems (Solomatine and Ostfeld, 2008). The FFBP ANNs are still widely applied and is a very popular tool compared to other data driven techniques in river basin management. Therefore in this study we have used the FFBP ANN (later on referred as ANN) model for hourly water level forecasting. Substantial literature on ANN have been reported in ASCE (2000a,b). Despite the good performance of ANN models, the outcome is highly dependent on the training data arrangement and there are undesirable uncertainties involved (Han et al., 2007a; Srivastav et al., 2007). The reliability of the model estimated discharge is affected by three major sources of uncertainties (Bates and Townley, 1988): data uncertainty (quality and representativeness of data), model structure uncertainty (ability of the model to describe the catchment's response), and parameter uncertainty (adequate values of model parameters). Han et al. (2007a) studied the uncertainties involved in real time forecasting using an ANN model. He concluded that for long term predictions, the ANN showed superior performance but that was only probabilistic depending on how the calibration and test events were arranged. Srivastav et al. (2007) proposed a method of uncertainty analysis for ANN hydrological models and showed that the ANN predictions contain a significant amount of uncertainty. In order to overcome the limitations inherent in the conventional treatment of uncertainty in ANN model predictions, recent trend has been to combine the outputs of several member bootstrap ANN models to reduce the uncertainty involved by controlling the generalization of final predictive model and to produce more reliable and consistent predictions (Cannon and Whitfield, 2002; Jeong and Kim, 2005).

The bootstrap is a computational procedure that uses intensive resampling with replacement, in order to reduce uncertainty (Efron and Tibshirani, 1993). It is also the simplest approach since it does not require complex computations of derivatives and Hessian-matrix inversion involved in linear methods or the Monte Carlo solutions of the integrals involved in the Bayesian approach (Dybowski and Roberts, 2000). Bootstrap technique has a wide variety of applications ranging from estimating means, CIs, parameter uncertainties and network design techniques (Lall and Sharma, 1996; Sharma et al., 1997; Tasker and Dunne, 1997). Bootstrap technique based ANNs have successfully been introduced in hydrological modeling. Abrahart (2003) employed bootstrap technique to continuously sample the input space in the context of rainfall–runoff modeling and reported that it offered marginal improvement in terms of greater accuracies and better global generalizations. Jeong and Kim (2005) used ensemble neural network (ENN) using bootstrap technique to simulate monthly rainfall–runoff. They concluded that ENN is less sensitive to the input variable selection and the number of hidden nodes than the single neural network (SNN). Jia and Culver (2006) used the bootstrap technique to estimate the generalization errors of neural networks with different structures and to construct the CIs for synthetic flow prediction with a small data sample. Han et al. (2007a) studied the uncertainties involved in real time forecasting in using an ANN model. They proposed a method to understand the uncertainty in ANN hydrologic models with the heuristic that the distance between the input vector at prediction and all the training data provide a valuable indication on how well the prediction would be. They concluded that for long term predictions, the ANN showed superior performance but that was only probabilistic depending on how the calibration and test events were arranged. However their method did not quantify the uncertainty of the model parameters or the predictions. Srivastav et al. (2007) proposed a method of uncertainty analysis for ANN hydrological models which was based on bootstrap technique. They developed an ANN model for forecasting the river flow at 1-h lead time and the results revealed that the proposed method of uncertainty analysis is very efficient and can be applied to an ANN based hydrological model. Tiwari and Chatterjee (2010) applied bootstrap technique for hourly flood forecasting and showed that bootstrap technique is capable of quantifying uncertainty in hourly flood forecasting and ensemble predictions were found to be more stable and accurate.

ANN models have limitation to consider any physics of the hydrologic processes in a catchment (Aksoy et al., 2007; Koutsoyiannis, 2007). Wavelet analysis provides a time–frequency representation of a signal at many different periods in the time domain (Daubechies, 1990) and gives considerable information about the physical structure of the data. Recently, variations, periodicities and trends in time series have been analyzed using wavelet transformation (Xingang et al., 2003; Yueqing et al., 2004; Partal and Kucuk, 2006). Wavelet based ANN (WANN) models have been employed recently in hydrological modeling successfully. Wang and Ding (2003) proposed a wavelet network model with a combination of the wavelet transform and the ANN, and decomposed the original time series into periodic components by wavelet transform. Later, sub-time series were used as the inputs for ANNs, and the resulting model was applied to forecast the original time series. This approach was used for monthly groundwater level and daily discharge forecasting. Kim and Valdes (2003) presented a hybrid neural network model combined with dyadic wavelet transforms to improve the forecasting of regional droughts. These researchers used the neural network model in two phases. First, neural networks were employed to forecast the signals decomposed by wavelet transform at various resolution levels, and then the forecasted decomposed signals were reconstructed into the original time series. The researchers applied this model to the monthly and annual inflow and rainfall data and showed that the model significantly improved the neural network's forecasting performance. Anctil and Tape (2004) used a wavelet–neural network model for 1 day lead rainfall–runoff forecasting. The time series was decomposed by wavelet transform into three sub-series: short, intermediate, and long wavelet periods. Then, multiple-layer ANN forecasting models were trained for each wavelet-decomposed sub-series, and later forecasted decomposed signals were reconstructed into the original time series. Partal and Cigizoglu (2008) predicted the suspended sediment load in rivers by a combined wavelet–ANN method. Measured data were decomposed into wavelet components via discrete wavelet transform, and the new wavelet series, consisting of the sum of selected wavelet components, was used as input for the ANN model. The wavelet–ANN model provided a good fit to observed data for the testing period. Partal and Cigizoglu (2009) predicted the daily precipitation using meteorological data from Turkey using the wavelet–neural network method. The new approach in estimating the peak values showed a noticeably high positive effect on the performance evaluation criteria. Kisi (2009) used wavelet–ANN conjunction model for daily intermittent streamflow forecasting. The results revealed that the proposed hybrid model could significantly increase the forecast accuracy of single ANN in forecasting daily intermittent streamflows.