Journal of Hydrology 532 (2016) 193–206

Contents lists available at ScienceDirect

Journal of Hydrology

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

# Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models

Yun Bai [a,*], Zhiqiang Chen [b], Jingjing Xie [c], Chuan Li [b,*]

[a] College of Architecture, Anhui Science and Technology University, Anhui 233100, China
[b] Research Base of Intelligent Manufacturing Service, Chongqing Technology and Business University, Chongqing 400067, China
[c] College of Resource and Environment, Anhui Science and Technology University, Anhui 233100, China

ARTICLE INFO

*Article history:*
Received 6 September 2015
Received in revised form 12 October 2015
Accepted 8 November 2015
Available online 27 November 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Purna Chandra Nayak, Associate Editor

*Keywords:*
Reservoir inflow
Forecast
Multiscale deep feature learning
Hybrid model
Deep belief network

SUMMARY

Inflow forecasting applies data supports for the operations and managements of reservoirs. A multiscale deep feature learning (MDFL) method with hybrid models is proposed in this paper to deal with the daily reservoir inflow forecasting. Ensemble empirical mode decomposition and Fourier spectrum are first employed to extract multiscale (trend, period and random) features, which are then represented by three deep belief networks (DBNs), respectively. The weights of each DBN are subsequently applied to initialize a neural network (D-NN). The outputs of the three-scale D-NNs are finally reconstructed using a sum-up strategy toward the forecasting results. A historical daily inflow series (from 1/1/2000 to 31/12/2012) of the Three Gorges reservoir, China, is investigated by the proposed MDFL with hybrid models. For comparison, four peer models are adopted for the same task. The results show that, the present model overwhelms all the peer models in terms of mean absolute percentage error (MAPE = 11.2896%), normalized root-mean-square error (NRMSE = 0.2292), determination coefficient criteria ($R^2$ = 0.8905), and peak percent threshold statistics (PPTS(5) = 10.0229%). The addressed method integrates the deep framework with multiscale and hybrid observations, and therefore being good at exploring sophisticated natures in the reservoir inflow forecasting.



## 1. Introduction

An accurate and reliable inflow forecast is a vital reference for making decisions of reservoir operation and management. Hence, it has attracted many researchers' attentions, and great deals of forecasting models have developed in the past decades.

Box–Jenkins models (Box and Jenkins, 1976), i.e., autoregressive model, autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA), have been widely used in the hydrologic time series forecasting (Valipour et al., 2012, 2013; Wang et al., 2015). Considering nonlinear and nonstationary characteristics of the real data of the reservoir inflow, researchers proposed other advanced methods, e.g., artificial neural network (ANN) (Lin and Wu, 2011; Kale et al., 2012; Taghi Sattari et al., 2012; Lohani et al., 2012; Abdellatif, 2015), support vector machine for regression (SVR) (Wang et al., 2010; Hwang et al., 2012; Hipni et al., 2013), Bayesian regression (Ticlavilca and McKee, 2011; Xu et al., 2014), fuzzy inference systems (Lin et al., 2012; Lohani et al., 2014), model tree (Maestre et al., 2013; Jothiprakash and Kote, 2014), spatial distribution-based model (Tsai et al., 2014; Cai et al., 2014), and hybrid model (Toro et al., 2013; Akrami et al., 2014; Li et al., 2014). These data-driven models have become appropriate alternatives to knowledge-driven models in hydrological forecasts. Their major advantage is that they only depend upon historical hydro-meteorological data without directly taking into account underlying physical processes and thus entailing much less input and parameter data. However, the success of data-driven models generally depends on data representation (Bengio et al., 2014; Längkvist et al., 2014). Therefore, feature extraction and learning of historical data are the key to ensure forecasting accuracy.

For this reason, much of the actual effort puts into the processes of feature extraction, data pre-processing and transformations. Okkan and Ali Serbes (2013) applied wavelet transform technique to remove nonlinear dynamic noise of raw data, and established three hybrid wavelet-based box models for monthly inflow forecasts. Maheswaran and Khosa (2013) also used wavelet technique to improve non-linear model forecasting performances in Krishna River. Budu (2014) employed two kinds of data pre-processing methods, i.e., wavelet transform and moving average, to compare wavelet-based ANN and multiple linear regression models for the

* Corresponding authors. Tel.: +86 (23) 6276 8469.
*E-mail addresses:* baiy@ahstu.edu.cn (Y. Bai), chuanli@21cn.com (C. Li).

http://dx.doi.org/10.1016/j.jhydrol.2015.11.011
0022-1694/

daily inflow forecasting of Malaprabha reservoir, Belgaum, India. Nourani et al. (2014) summarized applications of hybrid wavelet technique and ANN models in various hydrology objects, e.g., precipitation modeling, flow forecasting, rainfall-runoff modeling, and sediment modeling, in the last couple of decades. Bai et al. (2015a, 2015b)proposed an additive model, combining feature term extraction (trend, period, and stochastic components) and separate forecasts using different models, for forecasting the monthly reservoir inflows of the Three Gorges reservoir, China. Wang et al. (2015) presented an improved ARIMA model based on data empirical decomposition for forecasting annual runoff time series of Biuliuhe reservoir, Dahuofang reservoir and Mopanshan reservoir, China. The successes for forecasting of these models depend on the better characterizations of raw data, because variations of time series in one scale maybe a noise component, but maybe a structural component in other smaller scales.

On the other hand, feature learning is also a significant unit in inflow forecasting. Recently, most popular pattern learning, e.g., ANN, SVM, are belong to "shallow" learning category, in which the instinct information represent insufficiently. In this regard, deep learning techniques, which can gain better forecasting performance owing to its "deeper" representations, are introduced to overcome these drawbacks effectively (Li et al., 2015b). Deep belief networks (DBNs), one of the deep learning methods, are proposed by Hinton et al. (2006), which is a probabilistic generative model. In deep structures, there are many hidden layers for digging the latent changes by layer-wise learning, realizing "deep" features representation. It has already been applied successfully in some fields, e.g., audio classification (Mohamed et al., 2012; Wu and Zhang, 2013), fault diagnosis (Li et al., 2015b), character recognition (Xie et al., 2014; Sazal et al., 2014), drought forecasts (Chen et al., 2012), and traffic prediction (Guo et al., 2014). However, few literatures were reported for the reservoir inflow forecasting.

Considering the above multiscale feature representations and feature learning ratios, this work addresses a multiscale deep feature learning (MDFL) method with hybrid models. Our philosophy is to integrate a deep framework with multiscale and hybrid observations for understanding sophisticated natures of the reservoir inflow series. To this end, ensemble empirical mode decomposition (EEMD) and Fourier spectrum (Li et al., 2015a) are employed to extract multiscale features of the inflow series, which are then represented by three deep belief networks (DBNs), respectively. Each DBN is fused into a neural network (NN) by initializing the weights, building a D-NN regression model. The outputs of the three-scale D-NNs are reconstructed for generating the forecasts. The proposed MDFL method is evaluated by the real daily time series of the Three Gorges reservoir, China, and is compared with the state-of-the-art models.

The rest of the paper is structured as follows. The modeling approaches are described in Section 2. The application to the Three Gorges reservoir and the forecasting performance criteria are introduced in Section 3. Section 4 illustrates the results and discussion. Conclusions are given in Section 5.

## 2. Methodologies

Generally, a reservoir inflow series $\mathbf{X} = [x(1), x(2), \ldots, x(N)]$ ($N$ is the length of the series) can be regarded as the multiple feature fusion time series, which consists of trend, periodicity, and random mutation. That is, the $\mathbf{X}$ can be represented by the combination of trend term $\mathbf{T} = [x_T(1), x_T(2), \ldots, x_T(N)]$, periodic term $\mathbf{P} = [x_P(1), x_P(2), \ldots, x_P(N)]$, and the stochastic term $\mathbf{S} = [x_S(1), x_S(2), \ldots, x_S(N)]$ (Bai et al., 2015b), i.e.,

$$\mathbf{X} = \mathbf{T} + \mathbf{P} + \mathbf{S}. \tag{1}$$

For the reservoir inflow, the $\mathbf{T}$ reflects the long-term change rules in several hydrologic years, the $\mathbf{P}$ shows the variations within a seasonal time, and the $\mathbf{S}$ indicates the random interferences in the entire historical time. Thus sophisticated natures of the reservoir inflow series $\mathbf{X}$ cannot be represented sufficiently using a single model, which will lead to poor forecasting performances. To strengthen the characterization and weaken the influences of the complex features on forecasting performances, a deep and hybrid feature representation within multiscale framework is proposed. The details of these approaches involved, multiscale feature extraction, deep belief learning and hybrid model forecasting, are introduced in following subsections, respectively.

### 2.1. EEMD and Fourier transform for multiscale feature extracting

The EEMD (Wu and Huang, 2009), a developed version of the empirical mode decomposition (EMD), is a self-adaptive method to decompose original data into the sum of sub-spaces time series called intrinsic mode functions (IMFs). Compared with other decomposition techniques, e.g., wavelet transform, the EEMD has some special characteristics, e.g., empirical, intuitive, direct, self-adaptive, and without preliminary knowledge on the decomposition levels and the basic functions (Li and Liang, 2011). Therefore, the EEMD is adopted as a multiscale decomposition tool for the daily reservoir inflow.

The EMD (Huang et al., 1998) is an effective method, based on local characteristics of the data, for extracting original data series generated in noisy nonlinear processes and has been successfully applied in many areas. However, one of the major drawbacks of the EMD is the frequent appearance of mode mixing (i.e., scale separation problem). To alleviate this drawback, the EEMD is realized by the following means: (1) add a white noise series into the targeted data series, which can facilitate the multiscale frequencies separations and reduce the occurrence of mode mixing; (2) extract features of the data with added white noise by performing EMD repeatedly (specially, with different white noise series each time); and (3) achieve the ensemble means of corresponding IMFs of the decompositions as the final result. The procedures of the EEMD above show that the core algorithm of the EEMD is still the EMD, with added white noise, which sufficiently reveals the instinct characteristics of the original data. Details of the EEMD are described in the literatures (Huang et al., 1998; Wu and Huang, 2004, 2009). Brief introductions are introduced in this paper.

In the EMD approach, the daily reservoir flow series $\mathbf{X}$ is decomposed into $n+1$ IMFs, i.e., $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_i, \ldots, \mathbf{C}_n, \mathbf{R}_{n+1}$ (for convenience, the residue is represented as the last IMF in this study, remarked as $\mathbf{R}_{n+1}$). The EEMD results can be expressed as

$$\mathbf{X} = \sum_{i=1}^{n} \mathbf{C}_i + \mathbf{R}_{n+1}. \tag{2}$$

Two parameters, ensemble members $M$ and amplitude of the added noise $\varepsilon$, affect the decomposing results of the EEMD, which is evaluated by the final standard deviation of error $e$ (the distinctions between the input time series and the corresponding IMFs),

$$e = \frac{\varepsilon}{\sqrt{M}}. \tag{3}$$

In general, the $M$ of a few hundred will gain a good result, and the remaining noise may cause only less than 1% of error (Wu and Huang, 2009). In addition, under an appropriate level of the $\varepsilon$, it has no obvious improvement for decomposing results by constantly increasing the $M$. Based on these findings above, meaningful IMFs can be well specified when $M = 100$, and $\varepsilon \in [0.01, 0.5]$ (Wu and Huang, 2009).