Environmental Modelling & Software 104 (2018) 27–39

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Error correction-based forecasting of reservoir water levels: Improving accuracy over multiple lead times

Xiaojing Zhang [a, b], Pan Liu [a, b, *], Yan Zhao [a, b], Chao Deng [a, b], Zejun Li [a, b], Mengsi Xiong [a, b]

[a] *State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China*
[b] *Hubei Provincial Collaborative Innovation Center for Water Resources Security, Wuhan 430072, China*

A R T I C L E I N F O

Article history:
Received 24 March 2017
Received in revised form
9 February 2018
Accepted 24 February 2018

Keywords:
Flood forecasting
Lead time
Objective function
Xinanjiang model

A B S T R A C T

Providing reliable reservoir water level forecasts is a challenge because of the accumulative errors in hydrological and reservoir routing models. We present a novel forecasting model that addresses these issues. The model consists of a hydrological model to simulate inflow, a reservoir routing model to simulate water levels, and an autoregressive model for error correction. The parameters for the hydrological model were calibrated with the objective of forecasting water levels over multiple lead times, while a back-fitting algorithm was used to recalibrate the parameters sequentially for the hydrological and autoregressive models. The results show that: (1) the forecasting performance of effective lead times can be enhanced by minimizing the difference between the forecasted and observed water levels for multiple lead times; (2) the most recent errors method is better than the one-step-ahead recursive prediction method; and (3) the back-fitting algorithm is superior to the joint inference method.

Software availability

Name of software: HIRROF&BF-1 Developer: Xiaojing Zhang & Pan Liu
Contact address: State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430,072, China
Telephone, fax and email numbers: liupan@whu.edu.cn
Year first available: 2017
Hardware required: Personal Computer
Software required: MS Windows
Availability and cost:
free available at https://pan.baidu.com/s/1o7ZgvfO:
Program language: FORTRAN
Program size: 95.8 MB (downloaded zip file)

1. Introduction

Reliable reservoir inflow and water level forecasts contribute to efficient reservoir operations, which play a key role in flood control, hydropower generation, and water supply (Gragne et al., 2015; Liu et al., 2015a). However, temporal and spatial variations in climate as well as complex physical processes mean that forecasting reservoir water levels remains a complicated and challenging task (Chang and Tsai, 2016).

Data-driven and physically -based models are two basic approaches for forecasting reservoir water levels. The data-driven approaches include statistical and artificial intelligence (AI) models. Statistical methods include autoregressive (AR), autoregressive moving average, and autoregressive integrated moving average (ARIMA) models (Wang et al., 2015). These data-driven approaches offer easy implementation, but ignore the nonlinearity of hydrological series (Das et al., 2016; Sun et al., 2016). Although AI models, such as artificial neural networks (ANNs) and fuzzy inference (Gholami et al., 2015; Seo et al., 2015; Taormina and Chau, 2015), can solve the nonlinear problem, they do not represent the hydrological process (Chen et al., 2015b) but only depend on training data, which decreases their credibility. As a result, physically-based models, especially conceptual models, are widely used because they are both easy to understand and effective (Fang et al., 2017).

To forecast reservoir water levels in the context of physically -based modeling, the inflow should be simulated as the input to the reservoir routing model. However, the observed reservoir inflow used in hydrological models is generally inaccurate because the discharge from the tributaries of the reservoir is hard to measure

* Corresponding author. State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China.
E-mail address: liupan@whu.edu.cn (P. Liu).

https://doi.org/10.1016/j.envsoft.2018.02.017
1364-8152/

(Deng et al., 2015a). To address this issue, observed water levels can be employed as the fitting target of a forecasting model that integrates hydrological and reservoir routing (IHRR) models (Deng et al., 2015b). The advantage of IHRR models is that the hydrological parameters can be calibrated with more reliable raw data, thus improving the forecast accuracy.

To provide essential information for reservoir real-time operation, it is very important that accurate and reliable water level forecasts over multiple lead times are obtained. However, IHRR has a limited ability to meet these requirements because: (1) its applicability in forecasting over multiple lead times has not been validated; (2) inherent forecast errors are not corrected in real-time.

To improve lead times for forecasting, a direct approach is to forecast future rainfall, such as typhoon characteristics or remotely sensed variables (Chang and Tsai, 2016; Jhong et al., 2016; Lin et al., 2010). However, in areas without reliable rainfall forecasting, an alternative approach is to adopt the forecast streamflow from multiple lead times as the objective function, assuming the future rain is zero for the purpose of reducing the forecast errors. This improvement is feasible because the hydrological parameters can be adjusted to enhance the efficiency of longer lead times in the calibration procedure.

The error correction method can mitigate the forecast uncertainty problems (Van Steenbergen et al., 2012; Yan et al., 2012). Owing to the difficulties in clarifying the errors from different sources and their interactions, it is preferable to consider the overall error (Deng et al., 2015a, 2015b), i.e., directly correcting the difference between the forecast and observations. The existing literature includes many studies on overall error correction. For example, Xiong and O'Connor (2002) compared four error forecasting models, including an AR, autoregressive-threshold, a fuzzy autoregressive-threshold, and an ANN model; they demonstrated that the AR model is efficient despite its simplicity. Similarly, Goswami et al. (2005) evaluated eight error updating methods, while Liu et al. (2016) compared three error correction techniques. These methods are a type of "post-processor." A feasible alternative method is "joint inference", in which the parameters of the hydrological and error-correction models are calibrated simultaneously. For example, Li et al. (2016) applied a restricted AR model to normalized errors and jointly calibrated all the parameters using shuffled complex evolution. Although the joint inference method can theoretically find the optimal global parameters, it incurs a heavy computational burden. As the back-fitting algorithm is a simple iterative procedure for fitting a generalized additive model (Sorokina et al., 2007), the back-fitting algorithm only focuses on either the parameter set of the hydrological model or the autoregression model during each recalibration. In contrast, the joint inference method simultaneously calibrates all the parameters. In cases where the number of parameters is very large, the back-fitting algorithm will significantly reduce the heavy computational burden.

This study integrates the IHRR with the objective function of minimizing the difference between the simulated and observed water levels over multiple lead times and error correction methods, and develops a real-time forecasting method for reservoir water levels. Our aim is to improve the accuracy of multiple-step-ahead forecasting. Compared with the IHRR, the method has two improvements: (1) the forecast streamflow for multiple lead times is derived simultaneously to provide longer flood warnings; (2) the back-fitting algorithm is used for multiple-step-ahead error correction to reduce the heavy computational burden.

This paper is organized as follows. In section 2, the details of the methodologies are described, and the performance evaluation criteria are provided. A case study focused on the Shuibuya reservoir is presented in section 3, followed by the results and discussion in sections 4 and 5. Finally, the conclusions are drawn in section 6.

2. Methodology

As shown in Fig. 1, the proposed method has two steps:

(1) The Xinanjiang model and reservoir routing model are integrated to simulate the water levels over multiple lead times, with the objective function of minimizing the difference between the simulated and observed water levels.
(2) Back-fitting-based error correction is used to estimate the errors over multiple lead times and recalibrate the parameters to find the overall global optimal parameters for the hydrological and error-correction models.

2.1. Simulating water levels over multiple lead times

2.1.1. Integrated hydrological and reservoir routing models

The model for simulating the water levels integrates the Xinanjiang and reservoir-routing models (Deng et al., 2015b). The Xinanjiang model is used as the rainfall-runoff model because it is the most popular conceptual rainfall-runoff model in China (Wu et al., 2017; Zhao, 1992). It has been widely used for streamflow simulation (Jayawardena and Zhou, 2000; Yao et al., 2014), e.g., in the China National Flood Forecasting System (WMO, 2011). A schematic overview of the model is presented in Fig. 2. The model inputs are precipitation and evaporation, and the simulated streamflow is calculated using four main modules: (1) evapotranspiration; (2) runoff generation; (3) runoff separation; and (4) flow concentration. The 15 parameters in the Xinanjiang model are defined in Table 1.

When evaporation, seepage, and other water losses are ignored, the reservoir-routing model based on the water balance equation (Zhang et al., 2016, 2017) can be written as

$$V_{t+1} = V_t + \left(\bar{I}_{t+1} - \bar{O}_{t+1}\right)\Delta t \tag{1}$$

where Δt is the time interval; V_t, V_{t+1} are the initial and final reservoir storage volumes at time $t+1$, respectively; and $\bar{I}_t$, $\bar{O}_t$ are the inflow and release of the reservoir during time period $t+1$, respectively.

The inflow of the reservoir is obtained by the Xinanjiang model, and the release (which contains flows for electricity generation, irrigation water supply, and ecological generation) is predetermined by reservoir operating rules (Liu et al., 2014). The simulated water levels are then derived via equation (1) and the reservoir stage storage curve (Liu et al., 2015b; Pan et al., 2016).

2.1.2. Objective function of the IHRR

The objective of the conventional water level simulating model is to minimize the sum of squares of the difference between the simulated and observed water levels, with the aim of making the simulated series match the observed one:

$$\min F_{con} = \sum_{t=1}^{N}\left(Z_t - \widehat{Z}_t\right)^2 \tag{2}$$

where Z_t and $\widehat{Z}_t$ are the observed and simulated water levels at time t, and N is the length of the data series. However, the forecast water levels for longer lead times are not considered.

A modified objective function is developed by adopting the