International Journal of Forecasting 30 (2014) 464–476

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Modeling and forecasting of Brazilian reservoir inflows via dynamic linear models

L.M. Marangon Lima [a,*], E. Popova [a], P. Damien [b]

[a] Department of Operations Research and Industrial Engineering, The University of Texas at Austin, United States
[b] Department of Information, Risk and Operations Management, The University of Texas at Austin, United States

ARTICLE INFO

Keywords:
Reservoir inflow forecasting
Seasonal models
Bayesian updating
Climate predictors

ABSTRACT

This work focuses on developing a forecasting model for the water inflow at an hydroelectric plant's reservoir for operations planning. The planning horizon is 5 years in monthly steps. Due to the complex behavior of the monthly inflow time series we use a Bayesian dynamic linear model that incorporates seasonal and autoregressive components. We also use climate variables like monthly precipitation, El Niño and other indices as predictor variables when relevant. The Brazilian power system has 140 hydroelectric plants. Based on geographical considerations, these plants are collated by basin and classified into 15 groups that correspond to the major river basins, in order to reduce the dimension of the problem. The model is then tested for these 15 groups. Each group will have a different forecasting model that can best describe its unique seasonality and characteristics. The results show that the forecasting approach taken in this paper produces substantially better predictions than the current model adopted in Brazil (see Maceira & Damazio, 2006), leading to superior operations planning.

1. Introduction

The available methods for hydrological forecasting fall into two classes: conceptual (or physical) methods, which correspond to the rainfall-runoff hydrological models, and data-driven methods such as regression, time series, and artificial neural networks models. Interest lies in either forecasting river streamflow for water management and flood control or forecasting natural inflow to hydropower reservoir for operation and scheduling, with the latter being the focus of this work.

The rainfall-runoff conceptual model is a hydrological model that transforms rainfall (precipitation) into runoff (streamflow) based on physical and empirical equations. The components involved in the transformation process are evaporation, infiltration, interception, soil moisture, land use, and various meteorological conditions, including air temperature and solar radiation (Collisschonn, Allasia, da Silva, & Tucci, 2007; Moradkhani, Hsu, Gupta, & Sorooshian, 2004).

Since conceptual models rely on an accurate knowledge of the physical mechanisms of the underlying streamflow at a particular location, the data-driven techniques gained more popularity in the field of hydrology over the last decade (Wang, 2006). Data-driven models are defined on the basis of connections between state variables (input, internal and output), with little knowledge of the physical behavior of the system (Solomatine, 2002) being needed. Therefore, the forecasting procedure can easily be extended and applied to different locations and conditions. Examples of data-driven models are statistical models like time series models and artificial neural networks.

* Corresponding author. Tel.: +1 55 35 3621 3630.
E-mail address: luana_marangon@yahoo.com.br (L.M.M. Lima).

http://dx.doi.org/10.1016/j.ijforecast.2013.12.009
0169-2070/

The most popular univariate time series models applied to inflow forecasting are the autoregressive moving average (ARMA) models and their variants (Box & Jenkins, 1976). These are built on the assumption of stationarity, that is, the statistical properties of the process are not a function of time. (For a Bayesian perspective, see Marriott & Newbold, 1998, and West, 2013). Therefore, they are more commonly used for forecasting annual streamflows. Streamflow series with time scales of less than a year (e.g. monthly, quartely) usually exhibit seasonality because the hydrologic phenomena vary from one season to another. According to Hipel and McLeod (1994) three types of models can be applied to these series: the seasonal autoregressive integrated moving average (SARIMA), periodic ARMA (PARMA) and deseasonalized ARMA models. The deseasonalized and periodic models are used for describing data that possess stationarity within each season (e.g. Chen, 1997; Maceira & Damazio, 2006; Mondal & Wasimi, 2006; Yurekli, Kurunc, & Ozturk, 2005). The SARIMA family of models can be fitted to data where the level and perhaps other features change within each season across the years (e.g. Bender & Simonovic, 1994; Noakes, McLeod, & Hipel, 1985).

A more general class of regression models, the dynamic linear models (DLMs), have the capability to deal with nonstationarity within a season (West & Harrison, 1997). Krishnaswamy, Halpin, and Richter (2001) introduced a Bayesian dynamic linear regression model as a useful tool for studying the dynamics of hydrology in systems which are subject to high natural variability and land-use change. The model was applied to the Terraba River basin in the southern part of Costa Rica. Kumar and Maity (2008) apply a Bayesian dynamic model to the Devil's Lake basin, located in North Dakota, USA. They claim that the major strength of this type of model lies in its quantification of prediction uncertainty, particularly under different climate change scenarios.

Migon and Monteiro (1997) propose a dynamic nonlinear Bayesian model for the Fartura river basin in Brazil, where the complex system of equations that defines the physical processes is replaced by a simple one that tries to mimic the runoff's behavior given current and past precipitations. An extension of this model to Brazil's Grande river basin is presented by Ravines, Schmidt, Migon, and Renno (2008). Other applications of the Bayesian dynamic model in the field of hydrology include those of Berger and Rios-Insua (1998), Krishnaswamy, Lavine, Richter, and Korfmacher (2000) and Rios-Insua, Salewicz, Muller, and Bielza (1997).

This paper presents a Bayesian DLM for forecasting the water inflow at the Brazilian hydropower reservoirs. The idea is to initially group the reservoirs by basin, then develop a model for each basin based on its particular characteristics. The model will be then used as an input to a multi-stage stochastic optimization problem that solves the hydrothermal planning. The planning horizon is five years ahead, meaning that we are dealing with long-term forecasting. We also want to incorporate relevant climate variables as predictors.

The data are non-stationary. The monthly mean inflows for the basins located in the southern part of Brazil show a tendency to increase. However, most importantly, we observe non-stationarity in the seasonal pattern of the time series; for instance, a delay in the wet season for some of the basins. By 'a delay in the wet season' we mean that the window with the peak water inflow, which used to be December to February, is now from January to March. Therefore, we are also dealing with non-stationarity within the season, which means that the series cannot be reduced to a stationary process by differencing, so we need to work with a general dynamic model.

We want to model the process in its original scale, i.e., without performing any transformation of the data in order to achieve normality. As was noted above, the forecasts from the DLM are fed into a stochastic optimization algorithm, which requires the forecasting approach to assume a linear error structure in the time series regression. The DLM approach in this paper does precisely that, unlike the approaches of Migon and Monteiro (1997) and Ravines et al. (2008).

The remainder of the paper is organized as follows. Section 2 presents the Brazilian framework, with its major river basins and hydroelectric capacity. Section 3 describes the basin time series and correlation analysis, while Section 4 describes the climate variables that will be used as predictors in the model. Section 5 offers a brief description of the Periodic Autoregressive model which is currently used in Brazil. In Section 6 we describe the Bayesian model for basin inflows. Section 7 details the forecasting results and model performance criteria for each basin. Section 8 concludes the work.

2. Brazilian framework

Brazil has many rivers that form twelve major drainage basins, as shown in Fig. 1, of which only ten have hydropower plants. The Parana basin has the highest hydroelectric potential, around 54 gigawatts [GW], which represents more than 50% of the total capacity. It can be further subdivided into six minor basins, based on its major rivers: Paranaiba, Grande, Tiete, Paranapanema, Parana and Iguacu. Table 1 shows the total installed capacity for each basin, which is the sum of the generation capacities of each of the hydro plants within the basin: treating the Parana as 6 sub-basins leads to a total of 15 basins.

There are around 140 hydroelectric power plants currently in operation, and these plants operate in a cascade scheme. In order to determine how much electricity each one will produce in the future, one needs to know how much water will be available in the reservoirs. The available historical data are the natural inflow for each generator on a monthly basis, starting from January 1931, and measured in cubic meters per second [m^3/s]. The natural inflow is the average incoming water per unit of time at each generator's reservoir from affluent rivers, lakes and its own drainage area. Since the reservoirs operate in a cascade scheme, decisions taken at the upstream reservoirs will influence the inflow of the downstream reservoirs. The available data exclude the upstream reservoir operation by summing the natural inflow of the reservoir upstream in the cascade and the incremental inflow. Consider an example with two reservoirs, represented by the two triangles depicted in Fig. 2.