Journal of Hydrology 561 (2018) 146–159

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Monthly streamflow forecasting based on hidden Markov model and Gaussian Mixture Regression

Yongqi Liu[a], Lei Ye[b], Hui Qin[a,*], Xiaofeng Hong[c], Jiajun Ye[a], Xingli Yin[a]

[a] College of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[b] School of Hydraulic Engineering, Dalian University of Technology, Dalian, China
[c] Changjiang River Scientific Research Institute, Wuhan, Hubei, China

ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Anne-Catherine Favre, Associate Editor

Keywords:
Gaussian Mixture Regression
Hidden Markov model
Kernelized K-medoids clustering
Probabilistic streamflow forecasting

ABSTRACT

Reliable streamflow forecasts can be highly valuable for water resources planning and management. In this study, we combined a hidden Markov model (HMM) and Gaussian Mixture Regression (GMR) for probabilistic monthly streamflow forecasting. The HMM is initialized using a kernelized K-medoids clustering method, and the Baum–Welch algorithm is then executed to learn the model parameters. GMR derives a conditional probability distribution for the predictand given covariate information, including the antecedent flow at a local station and two surrounding stations. The performance of HMM–GMR was verified based on the mean square error and continuous ranked probability score skill scores. The reliability of the forecasts was assessed by examining the uniformity of the probability integral transform values. The results show that HMM–GMR obtained reasonably high skill scores and the uncertainty spread was appropriate. Different HMM states were assumed to be different climate conditions, which would lead to different types of observed values. We demonstrated that the HMM–GMR approach can handle multimodal and heteroscedastic data.

1. Introduction

Streamflow forecasting plays a critical role in water resources planning and management (Chiew et al., 2003; Zhao and Zhao, 2014). Reliable and skillful streamflow forecasts can help water resource managers to make better decisions, as well as promoting the sustainable development of the local economy. The existing approaches for streamflow forecasting are classified into three main categories: physical based models, conceptual methods, and empirical models (Bourdin et al., 2012; Devia et al., 2015). The early empirical models were usually linear, such as autoregressive, autoregressive moving average and Linear Regression (Castellano-Méndez et al., 2004; Haltiner and Salas, 1988; Salas et al., 1985; Valipour et al., 2013; Wu et al., 2009). However, these models have a limited ability to handle the non-stationary and non-linear relations massively involved in hydrological processes (Zhang et al., 2015). In addition to the conventional linear statistical techniques, a wide range of machine learning methods have been developed and used for hydrological forecasting. The two most commonly used machine learning methods applied in the hydrologic community are the artificial neural network (ANN) (McCulloch and Pitts, 1943; Kohonen, 1988; Chiang et al., 2004; Cigizoglu, 2005; Mutlu et al., 2008) and support vector machine (SVM) (Vapnik, 1995; Collobert and Bengio, 2001; Dibike et al., 2001; Wu et al., 2009).

The reliable quantification of forecast uncertainty is also very important for water resource management. SVM and ANN are both point forecast algorithms and they cannot provide information about the intrinsic level of forecast uncertainty. One approach for quantifying this uncertainty involves obtaining estimates of upper and lower bound prediction intervals, which indicate the range within which the observed data are likely to occur with some probabilistic level of confidence (such as 90%) (Ye et al., 2014; Ye et al., 2016). The Bayesian technique is another approach for a direct quantification of prediction uncertainty, where it estimates the posterior distribution based on the prior probability and likelihood (Murphy, 2012). For example, the Bayesian joint probability method has been employed for monthly and seasonal streamflow forecasting (Wang et al., 2009; Wang and Robertson, 2011; Zhao et al., 2016), where this method assumes that the data are multivariate Gaussian and it mainly focuses on learning the parameters of an enhanced Box–Cox transform using Monte Carlo Markov chain sampling. However, the Bayesian joint probability method does not consider the potential climate states in the model. Different climate states will influence the rainfall runoff process so it is important and meaningful to consider the potential states and their variations during streamflow forecasting.

* Corresponding author at: School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China.
E-mail address: hqin@hust.edu.cn (H. Qin).

https://doi.org/10.1016/j.jhydrol.2018.03.057
Received 16 November 2017; Received in revised form 9 March 2018; Accepted 21 March 2018
Available online 26 March 2018
0022-1694/ © 2018 Elsevier B.V. All rights reserved.

Hidden Markov models (HMMs) (also known as Markov switching models or dependent mixture models) comprise a discrete-time, discrete-state Markov chain with hidden states plus an observation model, which can describe the potential states in a catchment. Previously, it has been reported that HMM is a robust method for simulating hydrologic time series. The hidden states of hydrologic time series are typically described as climate regimes. Akintug and Rasmussen (2005) investigated the properties of HMM for generating annual time series. A HMM was also developed to generate runoff scenarios and rainfall data respectively (Gelati et al., 2010; Sansom and Thomson, 2010). A HMM combined with climate indices was proposed for multidecadal streamflow simulation (Bracken et al., 2014). However, extensions of HMM are still needed for probabilistic monthly streamflow forecasting. Gaussian Mixture Regression (GMR) is an alternative method to obtain a predictive distribution from a joint distribution for a HMM, although it is generally used for mixture models. In contrast to other regression methods such as SVM and ANN, GMR does not model the regression directly, but instead it models a joint probability density function of the data and then derives the regression function from the joint density model (Carreau et al., 2009; Calinon et al., 2010; Lee et al., 2016). This is an advantage for streamflow forecasting since density estimation can handle different sources of missing data, non-concurrent data and data with many occurrences of zero flows (Wang et al., 2009; Zhao et al., 2016). Compared HMM with an unconditional mixture model (Calinon et al., 2007), it can be interpreted as an extension of a mixture model where the choice of mixture component for each observation is not selected independently but instead it depends on the choice of component for the previous observation.

In this study, we considered a framework that combines HMM and GMR (HMM–GMR) for monthly streamflow forecasting. The main outcomes of this study are as follows. (1) We used a kernelized K-medoids clustering technique as a stable initialization strategy to avoid the model becoming trapped by poor local minima. (2) We extended GMR by recursively computing a likelihood through the HMM representation, thereby considering the predictors as well as the sequential information probabilistically encapsulated in the HMM. (3) We verified the effectiveness of HMM–GMR based on Yichang station in the upstream region of Yangtze River by using a monthly streamflow series, where we mainly focused on the forecast reliability and skill.

The remainder of this paper is organized as follows. In Section 2, we introduce the model formulation, the learning algorithm for HMM and GMR for probabilistic forecasting. In Section 3, we explain the forecast verification methods. In Section 4, we present an application of HMM–GMR to forecasting streamflows at three hydrological stations in the upstream region of the Yangtze River in Chain. In Sections 5 and 6, we discuss our results and give our conclusions, respectively.

2. Methods

2.1. HMM

A HMM comprises a discrete-time, discrete-state Markov chain with hidden states $z_t \in \{1,...,K\}$, plus an observation model $p(x_t \mid z_t)$. Where t stands for the time indices and $1 \leqslant t \leqslant T$. The probability of z_t depends on the state of the previous latent variable z_{t-1} via a conditional distribution $p(z_t|z_{t-1})$. The latent variables may be represented as K-dimensional binary variables with K-1 0's and a single 1 at position k indicating the state value, so this conditional distribution corresponds to a table of numbers, which we denote by **A** and its elements are known as transition probabilities:

$$\mathbf{A} = \begin{pmatrix} A_{11} & \dots & A_{1K} \\ \vdots & \ddots & \vdots \\ A_{K1} & \cdots & A_{KK} \end{pmatrix} \tag{1}$$

where $A_{ij} = p(z_t = j \mid z_{t-1} = i)$, and they satisfy $0 \leqslant A_{ij} \leqslant 1$ with $\sum_j A_{ij} = 1$.

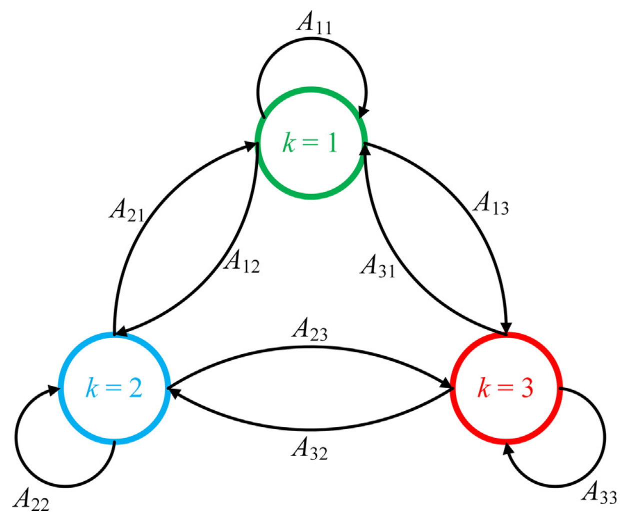

Fig. 1. Transition diagram of a Markov chain with three possible states.

The transition matrix is illustrated diagrammatically by drawing the states as nodes in a state transition diagram, as shown in Fig. 1 for the case where $K = 3$.

In this formulation, the observed sequence x_t depends on the current hidden state z_t. Thus, the conditional distributions of the observed variables are defined as $p(x_t|z_t = k,\phi_k)$, where ϕ is a set of parameters that govern the distribution. We consider d variables comprising both monthly streamflows to be forecast and their predictors such as climate and catchment indicators:

$$x^T = [x^1\, x^2 \cdots x^d] \tag{2}$$

In this study, the distribution of the observations of each state is represented as a multivariate Gaussian:

$$p(x_t|z_t = k,\phi_k) \hbar N\left(x_t|\mu_k, \sum_k\right). \tag{3}$$

where μ_k is the mean vector and $\sum_k$ is the covariance matrix, and one way of estimating the parameters is the maximum likelihood estimate (MLE), as follows.

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} x_t = \overline{x} \tag{4}$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} (x_t-\overline{x})(x_t-\overline{x})^T = \frac{1}{T}\left(\sum_{t=1}^{T} x_t x_t^T\right) - \overline{x}\,\overline{x}^T \tag{5}$$

The joint probability distribution over both the latent and observed variables is then given by:

$$p(\mathbf{X},\mathbf{Z}|\theta) = p(z_1|\pi)\left[\prod_{t=2}^{T} p(z_t|z_{t-1},\mathbf{A})\right]\prod_{t=1}^{T} p(x_t|z_t,\phi) \tag{6}$$

where $\mathbf{X} = \{x_1,...,x_T\},\mathbf{Z} = \{z_1,...,z_T\}$, and $\theta = \{\pi,\mathrm{A},\phi\}$ denotes the set of parameters of HMM. π_k is the initial probability of being in state k, A_{ij} is the transitional probability from state i to state j, and $\phi_k = \{\mu_k, \sum_k\}$,where μ_k and $\sum_k$ represent the center and the covariance matrix of the kth Gaussian distribution of the HMM, respectively.

2.2. Learning for the HMM

We now explain how to estimate the parameters $\theta = \{\pi,\mathbf{A},\phi\}$. The Baum–Welch algorithm (Baum et al., 1970), which is a variant of the expectation maximization (EM) algorithm, is used to learn the parameters. In the same manner as the EM algorithm, we must be careful

Download English Version:

https://daneshyari.com/en/article/8894728

Download Persian Version:

https://daneshyari.com/article/8894728

Daneshyari.com