Journal of Hydrology 509 (2014) 25–41

Contents lists available at ScienceDirect

Journal of Hydrology

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

Improving real time flood forecasting using fuzzy inference system

Anil Kumar Lohani [a,*], N.K. Goel [b], K.K.S. Bhatia [c]

[a] National Institute of Hydrology, Jal Vigyan Bhawan, Roorkee 247667, India
[b] Department of Hydrology, Indian Institute of Technology, Roorkee 247667, India
[c] Poornima Group of Institutions, Jaipur, India

ARTICLE INFO

Article history:
Received 5 February 2013
Received in revised form 9 November 2013
Accepted 14 November 2013
Available online 22 November 2013
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Ezio Todini, Associate Editor

Keywords:
Fuzzy inference system
Artificial neural network
Subtractive clustering
Self Organizing Map
Flood forecasting
Lead period

SUMMARY

In order to improve the real time forecasting of foods, this paper proposes a modified Takagi Sugeno (T–S) fuzzy inference system termed as threshold subtractive clustering based Takagi Sugeno (TSC-T–S) fuzzy inference system by introducing the concept of rare and frequent hydrological situations in fuzzy modeling system. The proposed modified fuzzy inference systems provide an option of analyzing and computing cluster centers and membership functions for two different hydrological situations, i.e. low to medium flows (frequent events) as well as high to very high flows (rare events) generally encountered in real time flood forecasting. The methodology has been applied for flood forecasting using the hourly rainfall and river flow data of upper Narmada basin, Central India. The available rainfall–runoff data has been classified in frequent and rare events and suitable TSC-T–S fuzzy model structures have been suggested for better forecasting of river flows. The performance of the model during calibration and validation is evaluated by performance indices such as root mean square error (RMSE), model efficiency and coefficient of correlation (R). In flood forecasting, it is very important to know the performance of flow forecasting model in predicting higher magnitude flows. The above described performance criteria do not express the prediction ability of the model precisely from higher to low flow region. Therefore, a new model performance criterion termed as *peak percent threshold statistics (PPTS)* is proposed to evaluate the performance of a flood forecasting model. The developed model has been tested for different lead periods using hourly rainfall and discharge data. Further, the proposed fuzzy model results have been compared with artificial neural networks (ANN), ANN models for different classes identified by Self Organizing Map (SOM) and subtractive clustering based Takagi Sugeno fuzzy model (SC-T–S fuzzy model). It has been concluded from the study that the TSC-T–S fuzzy model provide reasonably accurate forecast with sufficient lead-time.

1. Introduction

Real time flood forecasting is used to provide timely warning to people residing in flood plains and can alleviate a lot of distress and flood damage. Flood forecasting also provide useful information to water management personnel for making optimal decisions related to flood control structures and reservoirs operation. Floods are natural phenomena and are inherently complex to model. Conventional methods of flood forecasting are based on either simple empirical black box which do not try to mimic the physical processes involved or use complex models which aim to recreate the physical processes and the concept about the behavior of a basin in complex mathematical expressions (Lohani et al., 2005a). In between these two there is a wide variety of models, e.g. deterministic and stochastic, lumped and distributed, event driven and continuous or their combinations (Nielsen and Hansen, 1973; Box and Jenkins, 1976; Lundberg, 1982; Yakowitz, 1985; Yapo et al., 1993; Chatterjee et al., 2001), which are the basis of conventional flood forecasting system. Existing flood forecasting models are highly data specific and complex and make various simplified assumptions (Hecht-Nielsen, 1991; Hykin, 1992). For a reliable forecast Singh (1989) has listed three basic criteria, i.e. accuracy, reliability, and timeliness. Timeliness of forecasting is extremely important and this can be achieved by simple and robust forecasting models.

Recently there has been a growing interest in soft computing techniques viz. artificial neural networks (ANNs) and fuzzy logic. ANNs are basically data driven approach and are considered as black box models (Bishop, 1994) in hydrological context. These models are capable of adopting the non-linear relationship (Hecht-Nielsen, 1991; Flood and Kartam, 1994) between rainfall and runoff as compared to conventional techniques, which assume a linear relationship between rainfall and runoff. ANNs have strong generalization ability, which means that once they have been properly trained,

* Corresponding author. Tel.: +91 1332249214; fax: +91 1332272123.
E-mail addresses: lohani@nih.ernet.in, akl_nih@yahoo.co.in, aklnih@gmail.com (A.K. Lohani), goelnfhy@iitr.ernet.in (N.K. Goel), kksbneeru@yahoo.co.in (K.K.S. Bhatia).

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.11.021

they are able to provide accurate results even for cases they have never experienced before (Imrie et al., 2000; Lohani et al., 2012). Previous studies have shown that ANNs are capable of reproducing unknown rainfall–runoff relationship adequately (ASCE, 2000a, 2000b). ANN is also a powerful tool in solving complex nonlinear river flow forecasting problems (Hsu et al., 1995, 2002; Thirumalaiah and Deo, 1998a,b; Atiya et al., 1999; Toth, 2009; Birikundavyi et al., 2002; Kar et al., 2010) and in particular when the time required to generate a forecast is very short. Sahoo and Ray (2006) demonstrated that the ANN can outperform rating curves for discharge forecasting. Suitability of some deterministic and statistical techniques along with an ANN to model an event based rainfall–runoff process have been investigated by Jain and Indurthy (2003). This investigation on ANN with varying architecture, training rules and error back propagation establishes the suitability of ANN in flow forecasting. A comprehensive review of the ANN application in prediction and forecasting of water resources variables can be found in the works by Maier and Dandy (2000).

Fuzzy rule based method, introduced by Zadeh (1965), is another soft computing technique recently received attention for modeling hydrological processes. It is a qualitative modeling scheme where the system behavior is described using natural language (Sugeno and Yasukawa, 1993). Dubois et al. (1998) state that the real power of fuzzy logic lies in its ability to combine modeling (constructing a function that accurately mimics the given data) and abstracting (articulating knowledge from the data). See and Openshaw (1999, 2000) indicated that the fuzzy logic can be used with a combination of soft computing technique to create sophisticated river level monitoring and forecasting system. Hundecha et al. (2001) and Lohani et al. (2009, 2011) have demonstrated the applicability of fuzzy logic approach in rainfall–runoff modeling. Rule based fuzzy logic modeling techniques for forecasting water supply was investigated by Mahabir et al. (2003). A number of studies demonstrated that the fuzzy rule-based models for deriving stage–discharge–sediment relationships and sediment concentration forecasts produce much better results than the conventional rating curve models (Kisi, 2004, 2005; Kisi et al., 2006; Lohani et al., 2007a).

Luchetta and Manetti (2003) have developed a fuzzy logic based approach to the forecasting of hydrological levels, particularly suitable to cope with extreme situations, by setting different rules for trivial and rare situations. Neurofuzzy technique based on the combination of backpropogation and least square error methods for the parameter optimization is applied in short term flood forecasting by Nayak et al. (2005b) and pointed out that the number of parameters grows exponentially with the number of membership functions resulting in large training time. Lohani et al. (2012) compared the performance of the ANFIS with back propagation algorithm based ANN and AR models for hydrological time series modeling. Kisi et al. (2012) applied ANN and ANFIS to forecast daily lake-level variations. Ren et al. (2010) presented a new classified real-time flood forecasting framework by integrating a fuzzy clustering model and neural network with a conceptual hydrological model. Takagi–Sugeno (T–S) fuzzy technique has been applied to rainfall–runoff modeling and flood forecasting by various researchers (Xiong et al., 2001; Vernieuwe et al., 2005; Jacquin and Shamseldin, 2006; Lohani et al., 2011; Kar et al., 2012b). The T–S fuzzy structure identification is obtained directly by fuzzy clustering approach (Chiu, 1994). Fuzzy clustering also plays an important role in finding out homogeneous region in regional flood frequency analysis (Kar et al., 2012a).

The above discussion reveals that the core of the Takagi–Sugeno fuzzy structure identification method is in the clustering and the projection. A limitation of the *Subtractive Clustering based Takagi Sugeno* (SC-T–S) fuzzy model is that if any data point falls away from the cluster or outside the clusters the model performance may not be satisfactory (Nayak et al.; 2005a). Particularly in non-structural flood management a slight improvement in the accuracy of the real time flood forecasts has many direct advantages. The input data vectors which are used to train and build the short term flood forecasting model do not have all the same importance. In such cases the time series of river flow values contains both low to medium (frequent events) as well as high to very high flows (rare events). In order to improve the real time forecasting of floods, this paper proposes a *Threshold Subtractive Clustering based Takagi Sugeno* (TSC-T–S) fuzzy inference system. The proposed TSC-T–S fuzzy model is applied for the forecasting of hourly river flow of river Narmada, India by evaluating the fuzzy membership functions for frequent and rare events. In the proposed method the input–output data space is classified into frequent and rare events to preserve generalization capability of the *Subtractive Clustering based Takagi Sugeno* (SC-T–S) fuzzy inference system with improved forecasting. The results of the proposed TSC-T–S fuzzy model are evaluated with the forecast from ANN, SOM and SC-T–S (or interchangeably used as TS or T–S fuzzy model) fuzzy models at different lead periods.

2. Fuzzy models

Unlike classical logic which requires a deep understanding of a system, exact mathematical equations, and precise numeric values, fuzzy logic incorporates an alternative way of thinking, which allows modeling of complex systems using a higher level of abstraction originating particularly from our knowledge and experience. Fuzzy logic allows expressing this knowledge in a subjective way which are mapped into exact numeric ranges. In ordinary (non fuzzy) set theory, elements either fully belong to a set or are fully excluded from it. The membership $\mu_A(x)$ of $x\psi$ of a classical set A, as a subset of the universe x, is defined by:

$$\mu_A(x) = \begin{Bmatrix} 1, \text{iff } x \in A \\ 0, \text{iff } x \notin A \end{Bmatrix} \tag{1}$$

This means that an element $x\psi$ is either a member of set $A(\mu_A(x) = 1)$ or not $\mu_A(x) = 0$. This strict classification is useful in the mathematics and other sciences.

The general linguistic fuzzy model of multi-input single-output system is interpreted by rules with multi-antecedent and single-consequent variables such as the following:

Rule Ri: if x_1 is A_{i1} and if x_2 is A_{i2} and ... and if x_n is A_{in}

$$\text{THEN } y \text{ is } B_i, \quad i = 1, 2, \ldots, k \tag{2}$$

where x_1, $x_2, \ldots, x_n$ are input variables and y is the output, A_{ij} $(i = 1, \ldots, k, j = 1, \ldots, n)$ and B_i $(i = 1, \ldots, k)$ are fuzzy sets.

The corresponding fuzzy rules, antecedent and consequent membership functions can be generated by (i) knowledge of human experts; and (ii) suitable identification technique. If no priori knowledge exists for a given system, fuzzy clustering technique can be useful.

3. Fuzzy structure identification

Data driven fuzzy identification is an effective tool for the approximation of uncertain non-linear systems (Hellendoorn and Driankov, 1997). Fuzzy models and their respective characteristics are developed using clusters derived from the measured input and output data. A number of clustering algorithms have been reviewed:

(1) K-means or C-means clustering (Krishnaiah and Kanal, 1982).
(2) Fuzzy C-means (FCM) clustering method (Bezdek, 1981; Bezdek and Pal, 1992).