Journal of Hydrology 503 (2013) 222–232

Contents lists available at ScienceDirect

Journal of Hydrology

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

A data driven approach using Takagi–Sugeno models for computationally efficient lumped floodplain modelling

Vincent Wolfs *, Patrick Willems

KU Leuven, Hydraulics Division, Kasteelpark Arenberg 40, BE-3001 Leuven, Belgium

ARTICLE INFO

Article history:
Received 22 January 2013
Received in revised form 9 August 2013
Accepted 14 August 2013
Available online 27 August 2013
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Sheng Yue, Associate Editor

Keywords:
Adaptive neuro fuzzy inference systems
Computationally efficient
Floodplains
Floodplain hydraulics
Lumped modelling
Takagi–Sugeno

SUMMARY

Many applications in support of water management decisions require hydrodynamic models with limited calculation time, including real time control of river flooding, uncertainty and sensitivity analyses by Monte-Carlo simulations, and long term simulations in support of the statistical analysis of the model simulation results (e.g. flood frequency analysis). Several computationally efficient hydrodynamic models exist, but little attention is given to the modelling of floodplains. This paper presents a methodology that can emulate output from a full hydrodynamic model by predicting one or several levels in a floodplain, together with the flow rate between river and floodplain. The overtopping of the embankment is modelled as an overflow at a weir. Adaptive neuro fuzzy inference systems (ANFIS) are exploited to cope with the varying factors affecting the flow. Different input sets and identification methods are considered in model construction. Because of the dual use of simplified physically based equations and data-driven techniques, the ANFIS consist of very few rules with a low number of input variables. A second calculation scheme can be followed for exceptionally large floods. The obtained nominal emulation model was tested for four floodplains along the river Dender in Belgium. Results show that the obtained models are accurate with low computational cost.

1. Introduction

In support of water management decisions, hydrodynamic models are employed for applications such as uncertainty analyses, ensemble precipitation forecasts, real time control of hydraulic structures and long term simulations. These applications require hydrodynamic models with very little computation time. One- or two-dimensional distributed hydrodynamic models are often computationally too expensive for these purposes. Consequently, models with strong simplifications of the flow physics are necessary. Several models with conceptualization of the routing were examined in previous studies (e.g. Chen and Yu, 2007; Porporato and Ridolfi, 2001; Romanowicz et al., 2008; Shrestha et al., 2005), but little attention was given to the modelling of floodplains. The incorporation of floodplains is a vital element though, as floodplains exhibit a significant influence on the flow routing. Moreover, inundation depths and flood extents are often used in decision making. This information can for instance be linked to damage functions (e.g. Jonkman et al., 2008), which allow for control of the river with minimal damage.

The above mentioned applications in mind, it is often satisfactory to be able to predict the stage in the floodplain at one or more locations, together with the flow rate between river and floodplain. In order to obtain models with limited computation time, a conceptual modelling approach is examined in this paper, in which floodplains are represented by reservoirs. The presented modelling approach is built according to the data-based mechanistic (DBM) modelling philosophy (see e.g. Young, 1998, 2002). Hence, the model structure and parameter values should be chosen so that the model mimics the data obtained from running a detailed hydrodynamic model, while also allowing a physical interpretation of the underlying system. Due to the intricate topology and ditto flow properties, imposing a rigid functional structure to the pursued relationships is infeasible. This is also shown in the presented research. A solution should be searched in a very high dimensional parameter space in order to obtain adequate results.

To overcome this issue, this research uses soft computing techniques to emulate the results from running a full hydrodynamic model. These data-driven modelling techniques, such as artificial neural networks (ANN), fuzzy logic and support vector machines, are flexible, tolerant to imprecise data and can handle non-linearities (Zadeh, 1999). ANN techniques are widely used for prediction and forecasting in water resources and environmental engineering (e.g. Bowden et al., 2002, 2005; Chau, 2006; Chen and Adams, 2006; Kisi and Cigizoglu, 2007). When combined with fuzzy logic,

* Corresponding author. Tel.: +32 16321655; fax: +32 16 321989.
E-mail addresses: Vincent.Wolfs@bwk.kuleuven.be (V. Wolfs), Patrick.Willems@bwk.kuleuven.be (P. Willems).

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.08.020

the individual strengths of each approach can be exploited in a synergistic manner for the construction of powerful intelligent systems. In the last years, the integration of neural networks and fuzzy logic has given birth to neuro-fuzzy systems. Neuro-fuzzy systems have the potential to capture the advantages of both fuzzy logic and ANN in a single framework. Neuro-fuzzy systems eliminate the basic problem in fuzzy system design (obtaining a set of fuzzy IF-THEN rules) by effectively using the learning capability of an ANN for automatic fuzzy IF-THEN rule generation and parameter optimization. Many successful applications of adaptive neuro fuzzy inference systems (ANFIS) in water engineering have been reported (e.g. Chang and Chen, 2001; Jacquin and Shamseldin, 2006; Kisi et al., 2012; Lohani et al., 2006, 2007, 2012a, 2012b; Pramanik and Panda, 2009; Vernieuwe et al., 2005, 2006).

Given a river water level, the nominal emulating model provides an estimate of the flow rate between river and floodplain, and of the water level in the floodplain at one or more locations. In this paper, a modelling approach consisting of two calculation schemes is developed. In one of both methods, ANFIS play a crucial role. Different construction methods and input sets for ANFIS are compared by modelling a floodplain along the Dender River in Belgium. The best method and input set was then used for three other sites showing that this modelling approach delivers models suitable for water management practices.

Section 2 first provides a detailed overview of the fuzzy logic techniques used in this study, where after in Section 3 our methodology is presented. Section 4 describes the case study, on which the methodology was tested. Results are shown in Section 5, followed by general conclusions on the research in Section 6.

2. Fuzzy logic

Fuzzy inference systems (FIS) are non-linear modelling approaches that map an input space to an output space using a set of fuzzy IF-THEN rules. The classical notion of binary membership in a set has been modified to include partial membership ranging from 0 to 1 (Zadeh, 1965), hereby creating fuzzy sets with imprecise boundaries. Therefore, uncertainties inherent in the system can be reckoned with.

The IF-THEN rule statements are used to formulate the conditional statements that comprise fuzzy logic, e.g. "*IF* antecedent proposition, *THEN* consequent proposition". Depending on the structure of the consequent, the following two types of fuzzy systems can be distinguished: linguistic (Mamdani Type) (Mamdani, 1977; Zadeh, 1973) and Takagi–Sugeno (TS) fuzzy models (Sugeno and Kang, 1988; Takagi and Sugeno, 1985). The first group has fuzzy sets in both the antecedent and consequent proposition, whereas TS models have crisp output functions of the input. Considering their simplicity and efficiency, first-order TS models are employed in this work.

2.1. Takagi–Sugeno (TS) fuzzy inference system

The process of formulating the mapping from a given input to an output using fuzzy logic is called fuzzy inference (Jang, 1993). Characteristics of input data are allocated to input membership functions, input membership functions to rules, rules to a set of output characteristics and output characteristics to a single-valued output (Jang et al., 2002).

In the first-order TS fuzzy systems, the rule consequent forms a first order polynomial of the input variables $\boldsymbol{x}$:

$$R_i : \text{IF } \boldsymbol{x} \text{ is } A_i \text{ THEN } y_i = \boldsymbol{a}_i^T \boldsymbol{x} + b_i, \quad i = 1, 2, \ldots, M \tag{1}$$

where $\boldsymbol{x} \in \mathcal{R}^p$ is a premise part, $y_i \in \mathcal{R}$ is the consequent of the ith rule and M is the number of rules. In the consequent, $\boldsymbol{a}_i$ is the parameter vector and b_i the scalar offset. The terms A_i in the antecedents of the rules represent fuzzy sets (Zadeh, 1965) which are used to partition the input space into overlapping regions. Each multivariate fuzzy set A_i of the ith rule is defined by the degree of fulfilment (DOF, μ_i) of the ith rule, which is in this study evaluated using a t-norm (Piegat, 2001; Zimmerman, 2001), such as the algebraic product:

$$\mu_i(\boldsymbol{x}) = \prod_{j=1}^{p} \mu_{ij}(x_j) \tag{2}$$

where x_j is the jth input variable in the p dimensional input data space, and μ_{ij} the membership degree of x_j to the fuzzy set describing the jth premise part of the ith rule. This evaluation is also known as the "AND fuzzy operator" with the "product method". Several types of membership functions of the fuzzy sets in the antecedents of the rules are examined later on in this study.

For the input $\boldsymbol{x}$ the total output y of the TS model is computed by aggregating the individual rules contributions:

$$y = \sum_{i=1}^{M} u_i(\boldsymbol{x}) y_i \tag{3}$$

where u_i is the normalized DOF of the antecedent clause of the ith rule:

$$u_i(\boldsymbol{x}) = \frac{\mu_i(\boldsymbol{x})}{\sum_{i'=1}^{M} \mu_{i'}(\boldsymbol{x})} \tag{4}$$

2.2. Generation of TS fuzzy model

TS fuzzy models can be built using several identification methods. The computation of the antecedent membership functions using Grid Partitioning (GP) is widely used. Alternatively, fuzzy clustering techniques can be employed, which divide the data space into fuzzy clusters. Such methods generate a FIS with the minimum number of rules required to distinguish the fuzzy qualities associated with each of the clusters. Both antecedent identification techniques are described and applied in this study. Next, the consequent parameters can be computed by solving a linear least square problem.

2.2.1. Identification of the antecedent membership functions

2.2.1.1. Grid partitioning. This method proposes independent partitions of each antecedent variable (Jang, 1993). Using prior knowledge and experience, the modeller can define the membership functions of all antecedent variables. However, the division into the different partitions is often impeded due to the lack of knowledge. Therefore, the domains of the antecedents are simply divided in homogeneously spaced and equally shaped membership functions. If input–output data is available, the location of these membership functions can be optimized. A major drawback is that the membership functions for every variable are constructed independently of each other. Therefore, relationships between the variables are ignored. Secondly, GP generates rules by enumerating all possible combinations of membership functions of all inputs. This leads to an exponentially increasing number of rules with increasing number of inputs.

2.2.1.2. Subtractive clustering. Clustering is a popular unsupervised pattern classification technique which partitions the input space into C regions based on similarity or dissimilarity metric (Jain and Dubes, 1988). Various methods have been described in literature, such as mountain clustering (Yager and Filev, 1994), subtractive clustering (Chiu, 1994) and fuzzy C-means clustering (Bezdek,