Journal of Hydrology 420–421 (2012) 360–372

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

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

Separating aleatory and epistemic uncertainties: Probabilistic sewer flooding evaluation using probability box

Siao Sun [a,*], Guangtao Fu [b], Slobodan Djordjević [b], Soon-Thiam Khu [c]

[a] LGCIE, INSA de Lyon, 34 avenue des Arts, 69621 Villeurbanne cedex, France
[b] College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, United Kingdom
[c] Civil Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, United Kingdom

ARTICLE INFO

Article history:
Received 27 July 2011
Received in revised form 23 November 2011
Accepted 14 December 2011
Available online 22 December 2011
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Vazken Andréassian, Associate Editor

Keywords:
Aleatory and epistemic uncertainty
Fuzzy sets
Probability box
Sewer flooding
Uncertainty analysis
Random sets

SUMMARY

Uncertainty is generally present in flood evaluation and can be divided into aleatory and epistemic categories. It is not uncommon that flood evaluation has to consider both aleatory and epistemic uncertainties when a simulation model is used. This paper presents a probability box methodology that enables various representations of uncertainty to be simultaneously propagated through a model while separation of aleatory and epistemic uncertainties is preserved in the model output. The proposed methodology is applied to a sewer flood evaluation problem, in which rainfall variables are characterized by probability boxes and two model parameters are respectively described by fuzzy sets and random sets. Consequently, the probabilistic flood evaluation is expressed by probability boxes. Simulation results demonstrate the critical importance of separating aleatory and epistemic uncertainties and of maintaining the uncertainty type (either aleatory or epistemic) in uncertainty propagation. It is suggested that the pooling of aleatory and epistemic uncertainties may lead to incoherent information in the model output.

1. Introduction

Determination of the magnitude of floods with a specified exceedance probability is required for many engineering works, such as design or rehabilitation of flood defences and for the development of flood risk management strategies. In general, the flood probability distribution can be determined from statistical analysis of historical flood data. However, a sufficiently long time series of measurements of flood is usually not available. This is particularly the case for sewer flood analysis in urban drainage systems (Thorndahl and Willems, 2008). In the case of flood data scarcity, the simulation method instead can be used to generate a series of flood data using rainfall data, which are more commonly available in many catchments (Balmforth et al., 2006; Rulli and Rosso, 2002).

Uncertainty is widely recognized in flood evaluation and there is an increasing concern in this area. There are two fundamentally different types of uncertainty: aleatory uncertainty and epistemic uncertainty (Apel et al., 2004; Ross et al., 2009). The former originates from variability in known (or observable) populations and therefore represents randomness in samples. It can be operationally defined as a feature of the population that conforms well to a probabilistic model (Hall, 2003). Epistemic uncertainty results from lack of knowledge of fundamental phenomena and is related to our ability to understand, measure, and describe the system under study. Aleatory uncertainty is a property of the system and epistemic uncertainty is a property of the analyst (Cullen and Frey, 1999). Aleatory uncertainty cannot be reduced due to its inherent nature while epistemic uncertainty can be reduced, for example, by obtaining more data or knowledge (Merz and Thieken, 2005). A good review of the two types of uncertainty was given in the special issue of the Journal of Reliability Engineering and Systems Safety (Helton and Burmaster, 1996). Moreover, ontological uncertainty is being recognized as a sort of uncertainty in addition to aleatory and epistemic uncertainty. It is "unknown unknowns" that is not perceived as being important because of lack of knowledge (Beven et al., 2011). This type of uncertainty is not discussed in this paper if it even cannot be perceived and it will become an epistemic uncertainty as soon as it is recognized as an issue.

It is usual that both aleatory and epistemic uncertainties are present when evaluating flood probabilities via modelling. In such cases, both uncertainties are propagated through a model. It is important to distinguish between aleatory and epistemic uncertainties as different types of uncertainty may trigger different re-

* Corresponding author. Tel.: +33 0472436250.
E-mail address: siao.sun@insa-lyon.fr (S. Sun).

0022-1694/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2011.12.027

sponses: a concrete action must be taken to circumvent the potentially dangerous effects of inherent variability, whereas the best decision for the presence of epistemic uncertainty is probably to try to reduce it by collecting more information (Dubois, 2010). Moreover, the superposition of aleatory and epistemic uncertainties may lead to erroneous inferences and the separation of aleatory and epistemic uncertainties gives a more differentiated picture of the complete uncertainty in the model output (Grum and Aalderink, 1999; Merz and Thieken, 2005).

Uncertainty is traditionally represented by probability theory. Probability is based on the additivity axiom, which implies that the relevant evidence is a complete and consistent description of a problem. However, under the circumstances with scarce data, incomplete information or possibly inconsistent knowledge, the additivity axiom is difficult to justify (Hall, 2003). Hence other mathematical methods such as fuzzy set, possibility, random set, and probability box were devised to address cases where the objective information is incomplete or subjective judgment is modelled. It is very possible that uncertainties in inputs or model parameters are characterized by different representations according to available information or evidence in modelling. Hence the question of propagating various representations of uncertainty through models arises naturally. Effort has been made to develop methodologies that can propagate various uncertainty representations simultaneously through a model. For example, a hybrid approach for combining probability distribution functions and fuzzy numbers was introduced for risk estimation (Guyonnet et al., 2003); a random set based method was employed to bridge the gap between probability and fuzzy set to model contaminant transport in groundwater flow (Ross et al., 2009); a methodology based on random set theory was developed for sewer flooding evaluation, which can deal with imprecise probabilities and fuzzy numbers together (Fu et al., 2011). All of these studies focused on the methodology of integrating uncertainties represented using different mathematical methods, however, not enough attention has been paid on the separation of aleatory and epistemic uncertainties, which is essential to assure an appropriate physical meaning of the result.

This paper illustrates that it is important to maintain the nature of uncertainty when propagating it, that is, the aleatory uncertainty in inputs or parameters should be maintained to be aleatory in the output and similarly for epistemic uncertainty. The pooling or mixture of different types of uncertainty can lead to incoherent output representation and thus inform inappropriate decision making. This paper aims to solve a sewer flooding evaluation problem where both aleatory and epistemic uncertainties are present and the uncertainty sources are represented by various mathematical methods. Probability box, which represents a range of probability distributions bounded by a lower and an upper bound (Williamson and Downs, 1990), serves as a bridge between different uncertainty representations. Monte Carlo simulation (MCS) is employed to propagate uncertainty under the framework of probability box. The method of characterizing uncertainty sources in sewer flooding evaluation is presented and the way to maintain the nature of uncertainty (aleatory or epistemic uncertainty) is also described.

2. Problem statement

2.1. Sewer flooding evaluation with uncertainty

Sewer flooding is mainly caused by hydraulic overloading of urban drainage systems. The damage potential of sewer flooding is especially high as it occurs in densely populated urban areas. It often leads to serious consequences with not only direct damage to properties and infrastructures, but also social disruptions. Furthermore, there will possibly be increasing number of sewer flooding events due to global climate change and urbanization (Brown and Damery, 2002; Plate, 2002). Therefore the evaluation of sewer flooding is of great importance to identify critical components that likely cause system failure and thus develop urban flood risk management and mitigation strategies.

As many factors in sewer systems such as precipitation and some system parameters are stochastic variables that vary with time, sewer flooding events also have stochastic characteristics. It is commonly accepted that probability theory is an ideal tool for the characterization of aleatory uncertainty (Hall, 2003). The inherent variation of sewer floods can be well characterized by a probability distribution. Consequently, the generally interesting variants, such as sewer flood volume or depth of specific return periods and failure probability of a system, can also be elicited from this probability distribution.

When probabilistic sewer flood is evaluated via modelling, uncertainties of inputs and model parameters are propagated through the model to predict the uncertainty in the model output of interest. If information is perfectly known (i.e., without epistemic uncertainty), the resultant evaluation will be a probability distribution revealing the stochastic variation of sewer flood events. However, due to the lack of knowledge, epistemic uncertainty may exist in inputs or parameters with or without aleatory uncertainty. For example, the probability distribution of precipitation cannot be exactly derived from the available historical data; the percentage of impervious area of the catchment and the sewer pipe parameters such as the roughness coefficient cannot be precisely detected. However, these epistemic uncertainties cannot be simply ignored in the modelling process. As a result, with both aleatory and epistemic uncertainty present in the flood simulation model, the resultant evaluation is in the form of imprecise probabilities (probability boxes) that bound the true but unknown probability distribution.

Arnbjerg-Nielsen and Harremoes (1996) found that, in urban drainage modelling, the uncertainty in the description of rainfall variables was generally the most important contributor to the overall uncertainty in the model output, followed by the uncertainty in the description of surface runoff. This paper focuses on the method for propagating uncertainties through a model. It considers uncertainties stemming from inputs and model parameters, e.g. rainfall, pipe roughness coefficient and catchment percentage imperviousness. The developed method can easily incorporate other sources of uncertainty in inputs and parameters if necessary. It is assumed that the model simulation can capture the reality well, thus the model uncertainty from model structure and uncertainty in other parameters except those mentioned above are not taken into account.

2.2. Case study

A real world combined sewer network shown in Fig. 1 is used to demonstrate the proposed method. This network was first investigated and calibrated by Fullerton (2004) and was also studied by Fu et al. (2011) for urban flood evaluation using a random set based method. The network historically had significant flooding problems. The overall catchment area is around 2 km^2 with a population of 4000. The combined sewer network consists of 265 manholes, 265 pipes and 2 outfalls. The total conduit length is 22.5 km. Water is conveyed to a treatment plant and then released to a local river (outfall 1). During rainfall events, the flow exceeding the capacity of the treatment plant is diverted to outfall 2 and discharged to the river directly.

The system performance simulation including the hydrologic simulation of rainfall–runoff process and the hydraulic simulation in sewer system was performed using the Storm Water Management Model (Rossman, 2008).