Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory

Guangtao Fu,[1] David Butler,[1] Soon-Thiam Khu,[1,2] and Si'Ao Sun[1]

[1] Uncertainty analysis is widely applied in water system modeling to quantify prediction uncertainty from models and data. Conventional methods typically handle various kinds of uncertainty using a single characterizing approach, be it probability theory or fuzzy set theory. However, using a single approach may not be appropriate, particularly when uncertainties are of different types. For example, in sewer flood estimation problems, random rainfall variables are used as model inputs and imprecise or subjective information is used to define model parameters. This paper presents a general framework for sewer flood estimation that enables simultaneous consideration of two types of uncertainty: randomness from rainfall data represented using imprecise probabilities and imprecision from model parameters represented by fuzzy numbers. These two types of uncertainties are combined using random set theory and then propagated through a hydrodynamic urban drainage model. Two propagation methods, i.e., discretization and Monte Carlo based methods, are presented and compared, with the latter shown to be much more computationally efficient and hence recommended for high-dimensional problems. The model output (flood depth) is generated in the form of lower and upper cumulative probabilities, which are best estimates given the various stochastic and epistemic uncertainties considered and which embrace the unknown true cumulative probability. The distance between the cumulative probabilities represents the extent of imprecise, incomplete, or conflicting information and can be reduced only when more knowledge is available. The proposed methodology has a more complete and thus more accurate representation of uncertainty in data and models and can effectively handle different uncertainty characterizations in a single, integrated framework for sewer flood estimation.

1. Introduction

[2] Sewer flooding caused by overloaded urban drainage systems is an unwelcome reality in many parts of the world, producing significant adverse economic, social, and environmental impacts. In the UK, for example, sewer flooding is considered to be the second most serious issue (after drinking water quality) facing water companies, with an estimated cost of 270 million GBP a year in England and Wales alone [*Parliamentary Office of Science and Technology* (*POST*), 2007]. Further, there is an increasing probability of sewer flooding due to the expansion of urban areas and the likely adverse impacts of global climate change [*Ryu*, 2008].

[3] Most sewer systems have been designed on the basis of simple deterministic methods, such as the rational method or time-area method [*Butler and Davies*, 2004; *Thorndahl and Willems*, 2008]. These methods commonly use a design storm with a typical return period from 1 to 10 years to determine the maximum (minor) system capacity. However, the return period of sewer flooding is certainly not equivalent to that of the design storm, as the system capacity is increased so that sewers can accommodate a considerable surcharge before surface flooding occurs [*Butler and Davies*, 2004]. For existing sewer systems, flood frequency estimation can be further complicated with the issues of pipe deterioration and network expansion to new developments. Thus, the estimation of sewer flood frequency statistics for an urban catchment is of great interest in practice, as it provides direct assessment of hydraulic performance of the sewer system and supports decision making for sewer flood risk management [*Schmitt et al.*, 2004; *Ryu*, 2008].

[4] Handling uncertainty is a major issue in modeling water systems, including sewer systems, given the complexity and extent of uncertainty sources involved. This uncertainty has received increasing attention in recent years [e.g., *Guo and Adams*, 1998; *Adams and Papa*, 2000; *Matott et al.*, 2009]. Uncertainty can be broadly classified as stochastic or epistemic. Stochastic uncertainty refers to the randomness observed in nature, which is normally irreducible due to the inherent variation of physical systems. Epistemic uncertainty arises from incomplete knowledge about a physical system,

---

[1]Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK.
[2]Now at School of Engineering and Physical Sciences, University of Surrey, Guildford, UK.

which can be reduced with improved understanding of the system. Various approaches to characterize uncertainty are available, such as probability distributions, fuzzy sets, and random sets. Selecting an appropriate characterizing approach is essentially subjective as, in general, it is difficult to recommend one over another.

[5] Classical probability theory (Bayesian methods) has most often been used to quantify uncertainties in, for example, rainfall [*Guo and Adams*, 1998; *Thorndahl and Willems*, 2008], model parameters [*Lei and Schilling*, 1994], model structures [*Freni et al.*, 2009], and system dimensions such as storage volume and runoff basins [*Korving et al.*, 2002]. However, it is increasingly recognized that the concept of uncertainty is too broad to be captured by probability measures alone [*Ross et al.*, 2009]. So, for example, the theory of fuzzy sets [*Zadeh*, 1965] has also been increasingly applied, albeit in an attempt to describe imprecision and vagueness arising from the modeling process. Numerous applications can be found in hydrologic and hydraulic engineering [e.g., *Revelli and Ridolfi*, 2002; *Jacquin and Shamseldin*, 2007], including applications in the decision-making context of urban water management [e.g., *Makropoulos et al.*, 2003].

[6] It is not uncommon to have to address different uncertainty types simultaneously in the modeling and decision making processes. For example, some uncertainties are represented by probability distributions when sufficient data is available, while others are better represented by fuzzy sets to capture linguistic expert knowledge (qualitative data). Effort has been made to accommodate both probabilistic and fuzzy uncertainties in a unified framework of uncertainty analysis. The straightforward way is to transform one type of uncertainty into another, for example, probability distributions can be transformed into fuzzy sets or vice versa with little difficulty [*Zhang et al.*, 2009]. *Guyonnet et al.* [2003] proposed a hybrid method by embedding the α-cut propagation method for fuzzy variables within each sample simulation of the Monte Carlo (MC) technique for random variables, and as a result a large number of fuzzy sets were obtained for output variables. However, these methods cannot effectively handle imprecise probabilities in which only probability bounds (rather than one precise probability) can be defined as a result of scarce, vague, or conflicting information [*Walley*, 1991].

[7] Random set theory [*Kendall*, 1974; *Matheron*, 1975; *Dubois and Prade*, 1991] has attracted increasing attention in recent years, as it can cope with varying levels of precision regarding information, and uncertainty can be represented directly using original uncertainty forms without further assumptions. It is a theory of set-valued stochastic processes, observations of which are intervals or sets rather than precise point values, and thus it can be viewed as a natural generalization of probability and statistics on random variables [*Nguyen*, 2006]. Most importantly, it can serve as a bridge between different uncertainty representations, thus allowing them to be handled simultaneously in a single modeling framework [*Hall*, 2003].

[8] The aim of this work is to present a new methodology for imprecise probabilistic evaluation of sewer flooding using random set theory. In this methodology, temporal uncertainty in rainfall data is considered (spatial distribution and measurement uncertainties are neglected) and rep-resented using imprecise probability distributions of rainfall depth and duration. Synthetic rainfall events of uniform shape are used both because they are simple and typically used in practice [*Butler and Davies*, 2004], and because they are generally assumed when there is a complete lack of evidence on the appropriateness of other shapes. Model parameter uncertainty is characterized by fuzzy numbers with assumed shapes only. The most commonly used discretization method is used initially to propagate the two different types of uncertainties, and a MC based method is then developed to improve computational efficiency. What results from the method are the lower and upper cumulative distribution functions (CDFs) for model outputs (flood depth), constructed using the propagated random set. This methodology can potentially handle different uncertainty characterizations simultaneously, and thus allows for a more complete, and arguably accurate representation of uncertainty in data and models in terms of the most appropriate form wherever they originally appear, i.e., without further assumptions that might reduce the information or lead to inappropriate conclusions.

2. Problem Statement

2.1. Conventional Method

[9] In this study, the hydraulic performance of an urban drainage system is measured by the maximum flood depth during a rainfall event, i.e., the maximum water level over the ground surface at a manhole $h = g(\mathbf{x})$, where g represents the urban drainage model and $\mathbf{x} = (x_1, \ldots, x_d)$ is a vector of d uncertain variables. The case of $g(\mathbf{x}) = 0$ is normally defined as the limit state function (or failure surface) in the field of system reliability analysis. When $h > 0$, flooding occurs and the urban drainage system is regarded as having failed. The probability that flood depth is less than or equal to a value h_f is defined as

$$P_f = P\{g(\mathbf{x}) \leq h_f\}. \tag{1}$$

[10] P_f is the so called failure probability when $h_f = 0$. Calculation of P_f is important to estimate flood risk in the process of sewer flood risk analysis and management. Conventionally, the variables $\mathbf{x}$ are regarded as random variables, and the uncertainties are characterized by the joint probability density function $f(\mathbf{x})$. Thus, equation (1) can be written as

$$P_f = \int_{g(\mathbf{x}) \leq h_f} f(\mathbf{x}) d\mathbf{x}. \tag{2}$$

[11] This precise probabilistic formulation is well established and various methods have been developed to estimate P_f. For instance, *Thorndahl and Willems* [2008] applied the first-order reliability method to estimate the failure probability of sewer system flooding, surcharge, and overflow, and this method was compared with the standard MC method.

[12] There is seldom sufficient information to describe the joint probability function over the uncertain variables. In this study, the uncertain variables considered for sewer flood simulation include rainfall input variables and model

parameters. Stochastic rainfall is the driving force that has a significant effect on flood frequency. Continuous hydrograph simulation driven by long term historical rainfalls or event-based simulation from individual rainfall events are typically used in simulation-based flood analysis [*Ryu*, 2008]. The storm event analysis method, in which actual rainfall events are analyzed and fitted to appropriate CDFs, is used to characterize the stochastic nature of rainfall [*Adams and Papa*, 2000; *Thorndahl and Willems*, 2008]. Besides rainfall, uncertainty also arises from the way various physical processes are represented in the urban drainage model, including hydrologic processes that generate a runoff discharge at a particular location and hydraulic processes in the piped sewer network that result in surcharge water levels at that location. However, it is not normally possible to obtain distribution functions for model parameters due to insufficient data or cost constraints.

2.2. Random Set Theory and Its Use in Water Engineering

[13] Random set theory can be dated back to the work of *Kendall* [1974] and *Matheron* [1975] in the field of stochastic geometry. The theory can be regarded as set-valued random variables or multivalued mappings, and is equivalent to the Dempster–Shafer theory of evidence [*Dempster*, 1967; *Shafer*, 1976], in which probability masses are allocated to subsets rather than singletons of a given universe. It has proved to be a valuable theoretical framework for handling a range of different types of uncertainty information (e.g., intervals, probability distributions, fuzzy sets, and imprecise probability distributions), and the available information sources can be preserved in the original format in which they appear [*Hall*, 2003].

[14] Let X be a universal nonempty set containing all the possible values of a variable x, and P (X) the power set of X, i.e., the set of all the subsets of X. A random set can be defined as a pair $(\Im, m)$, where $\Im$ is the family of nonempty elements of P (X) and m is a mapping [*Dubois and Prade*, 1991]

$$m : \Im \rightarrow [0, 1],\qquad(3)$$

such that $m(\phi) = 0$ and

$$\sum_{A \in \Im} m(A) = 1,\qquad(4)$$

where $A \in$ P (X) and for which $m(A) > 0$ is called the focal element, and m is called the basic probability assignment. Each set A contains some possible values of the variable $x \in X$ and the value $m(A)$ expresses the probability that $x \in A$ but does not belong to any subsets of A. This does not exclude that some elements of A contribute to the probability of subset $B \in$ P (X) so that $A \cap B \neq \phi$. A random set is regarded as infinite provided that the cardinality of $\Im$, i.e., the number of elements in $\Im$, is infinite.

[15] A random set $(\Im, m)$ on X assigns a probability to all the subsets of X, while classical probability theory only considers the singleton subsets of X. Thus, random set theory can be seen as a generalization of probability theory to allow consideration of imprecision in the set definition of an event. This is designed to deal with the uncertainty

where the information is not sufficient enough to permit the probability assignment to single events. As a result of the imprecise nature of this formulation, it is impossible to calculate the precise probability of a subset $E \subset X$, i.e., P (E). Instead, the related imprecision of this probability can be bounded at the lower end by the belief function *Bel* [*Dempster*, 1967; *Shafer*, 1976]

$$Bel(E) = \sum_{A \subseteq E} m(A)\qquad(5)$$

and at the upper end by the plausibility function *Pl*

$$Pl(E) = \sum_{A \cap E \neq \phi} m(A) = 1 - Bel(\bar{E})\qquad(6)$$

where $\bar{E}$ is the complement of E. The belief $Bel(E)$ measures the minimum amount of evidence that fully supports $x \in E$, i.e., those cannot be removed out of E because the summation in equation (5) only involves A such that $A \subseteq E$. Similarly, the plausibility $Pl(E)$ measures the maximum amount of evidence that could be linked with the event E, i.e., those could be counted into E because the summation in equation (6) involves all A such that $A \cap E \neq \phi$.

[16] The application of random sets in modeling and decision making is relatively new in water and environmental engineering, and the first application was perhaps from *Caselton and Luo* [1992]. They described the theoretical aspects of the Dempster–Shafer theory with an intention to introduce this theory to decision makers and decision analysts as an alternative to the conventional Bayesian decision analysis. Its application was demonstrated with a water resource example to support decision making with scarce information in order to deal with rare design events and longer-term perspectives.

[17] *Hall* [2003] reviewed various mathematical methods for uncertainty analysis in hydroinformatic processes, including probability theory, fuzzy sets, and the Dempster–Shafer theory of evidence. He argued that purely probabilistic treatment of uncertainty does not lend itself to some of the more subtle aspects of uncertainty handling in hydroinformatics. The Dempster–Shafer theory was demonstrated as the best developed of the generalized mathematical methods for uncertainty handling. Different uncertainty types represented by probability distributions, interval measurements, and fuzzy sets were propagated through a simple equation to provide cumulative distribution bounds on overtopping discharge at a smooth sloping seawall.

[18] *Rubio et al.* [2004] applied random set theory to the uncertainty analysis of a combined slope hydrology and stability model. The uncertainties in model parameters were expressed in two forms: probability distributions and intervals, and were propagated through the model in the random set framework. This method generated the bounds on the cumulative probability distributions of slope safety with respect to rainfall-induced landslides. Application of random sets to slope stability analysis was also investigated by *Schweiger and Peschl* [2005].

[19] *Hall et al.* [2007] analyzed uncertainties in global mean temperature predictions using random set theory. Fuzzy emission scenarios were constructed to represent the

underlying uncertainties in social-economic constructs, and the lower and upper CDFs for a critical model parameter, derived from a number of published probability distributions, were used to characterize model parameter uncertainties. Using random set theory, these two types of uncertainty representations were then propagated through a climate change model to provide the bounds on the uncertainties in predicted global mean temperature rises.

[20] The above applications demonstrate the need for and potential of random sets in handling different representations of uncertainty in water and environmental modeling. However, in these applications, the only uncertainty propagation method used is the discretization method [*Tonon*, 2004], which is computationally expensive and intractable for high-dimensional problems. *Ross et al.* [2009] avoided the direct use of this method by transforming the random set, constructed for a model parameter (hydraulic conductivity) on the basis of expert knowledge, into a fuzzy set, and then using the vertex method [*Ross*, 2004] to propagate the fuzzy set through a groundwater model. However, the vertex method is essentially a discretization-based method and the model output is generated in the form of possibility distributions, which might need to be transformed back into upper and lower bounds of probability distributions for interpretation. In this paper, a MC-based method is applied, for the first time, to propagate uncertainties for sewer flood evaluation and is compared with the discretization method.

2.3. Random Set Approach for Flood Estimation

[21] Using random set theory, the problem of estimation of flood depth probability is to find the bounds on P_f, instead of a precise value as calculated in equation (2). Incomplete knowledge about the uncertain variables $\mathbf{x} = (x_1, \ldots, x_d)$, including their dependency, can be expressed as a random relation, i.e., a random set $(\Im, m)$ on the Cartesian product $X1 \times \ldots \times Xd$. According to the random set extension principle by *Dubois and Prade* [1991], the information about flood depth can be represented as a random set $(\Re, \rho)$, which is the image of $(\Im, m)$ through the urban drainage model g,

$$\Re = \{R_j = g(B_i), B_i \in \Im\}, \text{ where } g(B_i) = \{g(\mathbf{x}), \mathbf{x} \in B_i\} \quad (7)$$

$$\rho(R_j) = \sum_{B_i : R_j = g(B_i)} m(B_i). \quad (8)$$

[22] When the marginal random sets are stochastically independent the mass assignment in the joint space can be obtained as the product of the masses of the marginal random sets

$$m(A_1 \times \ldots \times A_d) = \prod_{i=1}^{d} m(A_i), \ A_1 \times \ldots \times A_d \in \Im. \quad (9)$$

[23] On the basis of the random set $(\Re, \rho)$, the lower and upper bounds of CDFs for flood depth H can be reconstructed using the belief and plausibility functions. Assume first that the flood depth axis is partitioned into adjoining intervals $[h_1, h_2]$, $[h_2, h_3]$, $\ldots$, $[h_s, h_{s+1}]$ denoted as R_1, R_2, $\ldots$, R_s, respectively. According to the definitions in

equations (5) and (6), the lower and upper CDFs $\underline{F}(h)$ and $\overline{F}(h)$ at some point h defined on the domain $[h_1, h_{s+1}]$ can be obtained as follows [*Tonon*, 2004; *Hall et al.*, 2007]

$$\underline{F}(h) = Bel(\{H \leq h\}) = \sum_{h \geq \sup(R_j)} \rho(R_j) \quad (10)$$

$$\overline{F}(h) = Pl(\{H \leq h\}) = \sum_{h \geq \inf(R_j)} \rho(R_j). \quad (11)$$

[24] These derived bounds represent the best possible knowledge about flood depth, given all kinds of uncertainties in variables, and should embrace the unknown true CDF of flood depth. The spread of the bounds represents the extent of imprecision and incompleteness in uncertainty representations and can only be reduced when more knowledge is available.

[25] Applying the extension principle for uncertainty propagation generally requires calculation of the image of the focal set B_i through the model g, which usually is a global optimization problem. The focal element $R_j = g(B_i)$ can be obtained by solving the following optimization problem,

$$R_j = [\min_{\mathbf{x} \in B_i} g(\mathbf{x}), \max_{\mathbf{x} \in B_i} g(\mathbf{x})]. \quad (12)$$

[26] Generally, the discretization method can be used to derive equation (12) [*Tonon*, 2004]. The idea is to use the interval-based vertex method [*Moore*, 1966]. As each focal element B_i is a d-dimensional box with 2^d vertices, if model g is a continuous function with no extreme points in the box or on its edges, it is only necessary to evaluate the model at the vertices to find the global minima and maxima. This means that model g has to be evaluated 2^d times for each focal element R_j. The number of model evaluations can be further reduced if g is monotonic to some or all variables. In the case of one or more extreme points existing in the interior of B_i or on its edges, the true global minima and maxima can be approached by increasing the number of each marginal random set, i.e., the fineness of discretization. *Tonon* [2004] suggested that doing so is computationally more efficient than invoking an optimization method.

[27] It should be noted that, in the discretization method, the number of model evaluations increases exponentially with the number of variables. Thus, this method becomes inefficient for high-dimensional problems. To improve computational efficiency, a MC-based method is proposed to propagate uncertainties for sewer flood evaluation.

3. Case Study

[28] A combined sewer network in the UK is used to demonstrate the methodology developed in this paper, as shown in Figure 1. The total catchment area is about 200 hectares, serving a population of 4000. The sewer system consists of 265 nodes, 265 pipes, 2 outfalls, and 1 weir, and has a total conduit length of 22,482 meters. The pipe gradients vary from 0.0001 to 0.0439. Flows are diverted downstream via the two outfalls: one is connected to a wastewater treatment works (Outfall 1) and the other to a
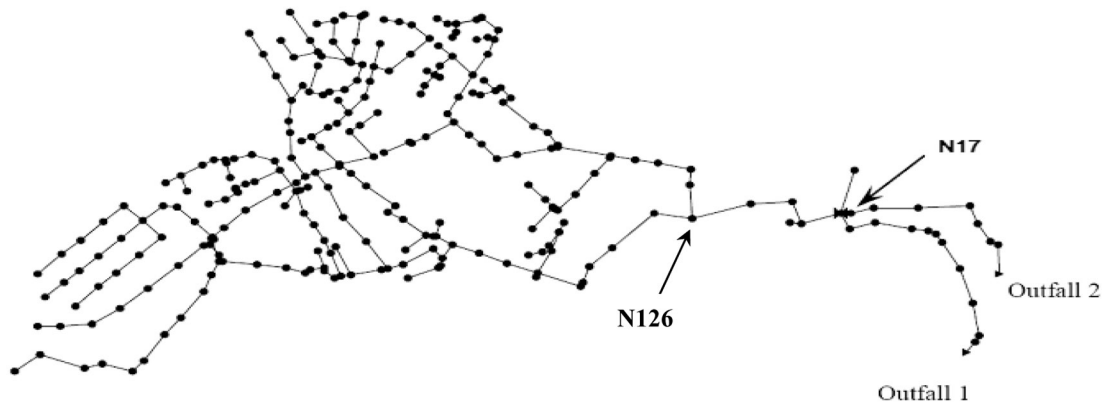
Figure 1. Layout of the case study network.

combined sewer overflow (Outfall 2), and both flows are eventually discharged into a river.

[29] The storm water management model (SWMM), developed by the U.S. Environmental Protection Agency (available at http://www.epa.gov/ednnrmrl/models/swmm/), was used for hydrologic simulation of rainfall runoff in the urban catchment and for hydrodynamic simulation of in-sewer transport through the urban drainage system. The sewer system model was originally set up and calibrated for flood evaluation in work by *Fullerton* [2004], as sewer systems historically have experienced significant flooding problems.

4. Uncertainty in Rainfall

4.1. Data Analysis

[30] The estimation of flood probability in an urban drainage system must take into consideration the stochastic nature of rainfall events that occur over the catchment. Duration and depth (equivalently, average intensity over the rainfall time interval) are the most important characteristics of a rainfall event. Thus, these two variables were used to analyze statistical characteristics of the actual rainfall events in the urban catchment, from 10 years of 5 min rainfall data from one rain gauge station.

[31] To analyze the rainfall data, the concept of interevent time definition (IETD) was adopted to separate individual rainfall events; that is, the time interval between two consecutive events should be no less than a predetermined IETD. In this case study, IETD was set to 30 min. This is greater than the maximum concentration time of the catchment, so the runoff response from an individual event is not affected by any other. A total of 767 events were identified, with an average of 77 rainfall events per year. Figure 2 shows the histogram of the events. There is a high frequency of low rainfall depth; about 50% of the rainfall events have a depth less than 5 mm, while the range of recorded values is up to 49 mm. Similarly, most events last a short period, although about 10% have a duration over 800 min.

[32] Independence between rainfall depth and duration is assumed for simplicity. This is reasonable in this case because the correlation coefficient (r) between these two variables is only 0.09 for the rainfall events considered, and

the scatterplot for rainfall depth and duration suggests that the two variables have no statistical relationship. Synthetic rainfall events are generated by applying a rectangular pulse, with duration as the width and average rainfall intensity as the height. If dependence between rainfall depth and duration was present it could be expressed as a random relationship, i.e., a random set $(\Im, m)$, and then handled using equations (7) and (8) for uncertainty propagation (refer to *Dubois and Prade* [1991] for more details). The rainfall, generated on the basis of statistics from one rain gauge station, is assumed to be uniformly distributed over the catchment as it is relatively small.

4.2. Imprecise Probabilistic Representation and *p*-Box Discretization

[33] In flood analysis, theoretical probability distributions are usually assumed for rainfall characteristics, and their parameters are calibrated using historical rainfall events. Choice of an appropriate distribution is arbitrary and dependent on expert knowledge, although it can have a significant effect on derived distributions of flood depth. However, in many situations, particularly when data are sparse, there can be more than one distribution that fits the data well and cannot be rejected on the basis of goodness-of-fit tests. Although Bayesian methods can be used to select or combine different probability distributions by using prior knowledge [*Korving et al.*, 2002], the principal difficulty is that the statistical parameters have to be represented by exactly specified, classical probability distributions, no matter how weak the information or prior knowledge [*Walley*, 1991; *Caselton and Luo*, 1992]. In this study, use of imprecise probabilities provides an alternative approach to describe the rainfall uncertainties in flood analysis.

[34] A number of CDFs were used to fit the rainfall data, and three goodness-of-fit tests, i.e., Kolmogorov–Smirnov (K–S), Anderson–Darling (A–D) and Chi-Square (χ^2) tests, were used to judge whether one specific distribution should be rejected. With the test at the 5% significance level, a number of distribution functions (Figure 2) were selected for rainfall depth or duration, including Frechet, Gamma, Generalized Pareto, Generalized Extreme Value (GEV), Inverse Gaussian, Log-Logistic, and Pearson Type 6. These CDFs are provided in Appendix A. It can be seen from Tables 1 and 2 that the statistics of the chosen CDFs
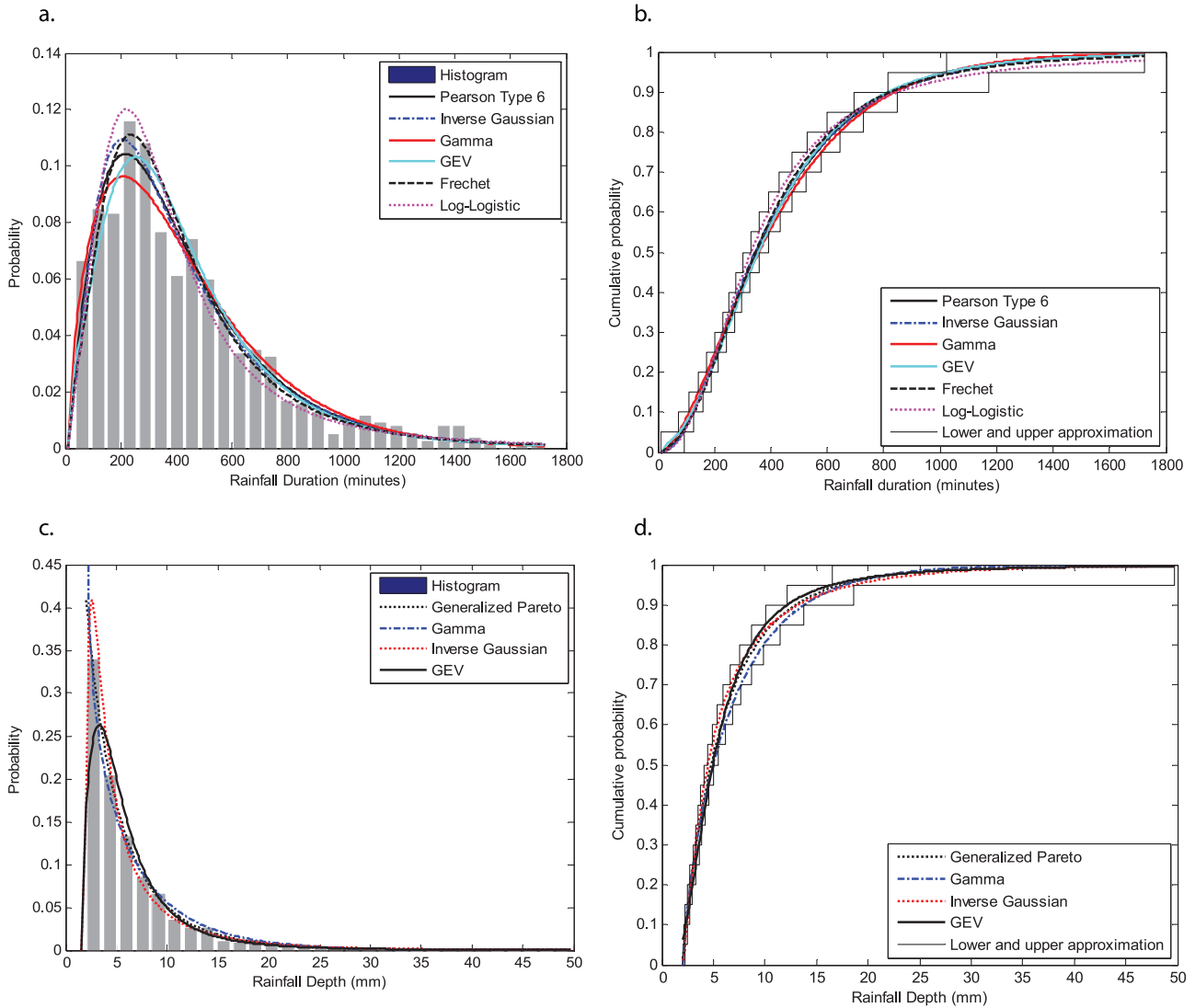
Figure 2.   Probability distributions for rainfall duration and depth. Gray bars show the histograms normalized by the total number of rainfall events. The stair-step curves show the 21-level discrete approximation for the lower and upper probability distributions.

for the two rainfall variables are all below the critical values of the three tests, i.e., 0.049 for K−S, 16.919 for χ^2, and 2.502 for A−D, and that the calculated p-values exceed the significance level of 5%.

[35] The chosen distributions for rainfall depth or duration form a family of CDFs with different approximations to the unknown distribution. From these distributions, the lower and upper bounds of CDFs $\underline{F_X}$ and $\overline{F_X}$ can be derived

$$\underline{F_X}(x) = \inf_i F_i(x) \quad \forall x \in R \qquad (13)$$

$$\overline{F_X}(x) = \sup_i F_i(x) \quad \forall x \in R \qquad (14)$$

where $F_i(x)$ represents the ith CDF considered in the family. For each point x, there is a corresponding interval $\left[\underline{F_X}(x), \overline{F_X}(x)\right]$, the lower probability measures the evidence supported by the family and the upper probability reflects the lack of information against it. The interval provides a bracketing of some ill-defined CDFs and its spread represents the extent of incomplete knowledge about the true known distribution.

[36] Instead of the use of one single or linear combination of the chosen distributions, the lower and upper bounds of the family of distributions are used to represent the uncertainties in rainfall characteristics. The lower and upper probabilities can be transformed into a random set through the p-box discretization method [*Tonon*, 2004; *Hall et al.*, 2007; *Alvarez*, 2009]. An outer approximation is constructed by drawing $n+1$ horizontal lines with cumulative probabilities p_0, p_1, ..., p_n, where $p_i = i/n$ ($i = 0, \ldots, n$), and thus dividing the distributions into n boxes. Within each box, the interval $A_i = \left[\underline{x_i}, \overline{x_i}\right]$ ($i = 1, \ldots, n$) is defined by p_{i-1} and p_i such that $\overline{F_X}(\underline{x_i}) = p_{i-1}$ and $\underline{F_X}(\overline{x_i}) = p_i$. These intervals can be regarded as focal elements of a random set (κ, m)

$$\kappa = \left\{A_i = \left[\underline{x_i}, \overline{x_i}\right]\right\}, i = 1, \ldots, n \qquad (15)$$

Table 1. Test Statistics of the Chosen CDFs for Rainfall Duration[a]

Distribution	Distribution Parameter	K–S		χ^2		A–D
		D	p-Value	D	p-Value	D
Frechet	$\xi = 4.693, \sigma = 893.45, \mu = -620.33$	0.032	0.393	12.342	0.195	1.057
Gamma	$\alpha = 1.946, \beta = 220.18, \gamma = 0$	0.026	0.683	10.15	0.339	0.754
Pearson Type 6	$\alpha_1 = 2.463, \alpha_2 = 12.487, \beta = 1995.1$	0.021	0.892	9.408	0.401	0.397
GEV	$\xi = 0.137, \sigma = 202.48, \mu = 279.91$	0.032	0.407	7.73	0.562	1.033
Inverse Gaussian	$\lambda = 1216.7, \mu = 501.58, \gamma = -73.191$	0.024	0.746	13.388	0.146	0.541
Log-Logistic	$\alpha = 2.319, \beta = 329.25$	0.048	0.05	16.38	0.053	2.450

[a]CDFs are provided in Appendix A

with a probability mass assignment

$$m(A_i) = 1/n.\qquad(16)$$

[37] The random set approximates the lower and upper cumulative probabilities by touching them exactly at p_0, $p_1, \ldots, p_n$ levels. A 21-level approximation to the cumulative probabilities of rainfall duration and depth is shown in Figure 2. It should be noted that the accuracy of the p-box discretization method is dependent on the number of probability levels (n), thus a high number within the available computing resources should be used to improve the accuracy in propagated random sets.

4.3. Propagating Rainfall Uncertainty

[38] Three cases with an increasing approximation level were used to approximate the lower and upper cumulative distribution bounds of rainfall depth and duration. In each case, 21, 51, or 101 levels were used for both rainfall depth and duration, consequently a random set with 20, 50 or 100 focal elements, respectively, was generated for each variable. The joint random set was constructed using the Cartesian product and was propagated through the SWMM model in the first instance with the default model parameters (Manning's roughness coefficient = 0.013 and catchment runoff coefficient = 0.85). The vertex method was used to calculate the image of each joint focal element, i.e., one model simulation is run for each of the four vertices in the two-dimensional box, formed by every interval of rainfall depth and of rainfall duration. The propagated random set for flood depth is then used to construct the lower and upper CDFs using equations (10) and (11); results for the critical node N126 are shown in Figure 3.

[39] It can be seen that the lower and upper CDFs of flood depth with 21 levels has the widest gap, which is reduced when a higher-level approximation is used. When more levels are used to approximate the lower and upper CDFs of rainfall variables, the p-boxes derived can provide a more precise representation, i.e., less information contained in the

CDFs is lost in the transformation process. However, an increase from 51 and 101 levels provides almost the same results; this implies that result quality cannot be improved by further increasing the degree of discretization, i.e., the hard boundary for the lower and upper CDFs has been approached using the discretization method.

5. Uncertainty in Model Parameters

5.1. Fuzzy Representation of Model Parameter Uncertainty

[40] Two model parameters were considered: Manning roughness and runoff coefficients. For sewer systems after a certain period of service, the roughness coefficients of the pipes are difficult to estimate because of complex pipe aging processes [*Revelli and Ridolfi*, 2002]. In the rainfall runoff modeling process, the runoff coefficient is related to catchment and rainfall event characteristics. The uncertainties from these parameters arise from the difficulty in estimating the parameter values particularly when data are limited, thus in this study they are more adequately represented by fuzzy sets, rather than probability distributions. In a fuzzy set S defined on the universe X, the membership of an element x to the set is no longer binary, but is characterized by a membership function,

$$\mu_S : X \to [0, 1]\qquad(17)$$

which assigns each element $x \in X$ a degree $\mu_S(x)$ measuring how much the element belongs to the fuzzy set S. The set $\{x \in X \mid \mu_s(x) \geq 0\}$ is called the support of the fuzzy set, and the set $\{x \in X \mid \mu_s(x) = 1\}$ is called the kernel of the fuzzy set. The membership function can take many different shapes, and a trapezoidal shape for the two model parameters is assumed in the first place and then compared with triangular and narrower-support trapezoidal shapes in terms of their impacts on the probabilities of flood depth (Figure 4).

Table 2. Test Statistics of the Chosen CDFs for Rainfall Depth[a]

Distribution	Distribution Parameter	K–S		χ^2		A–D
		D	p-Value	D	p-Value	D
Gamma	$\alpha = 0.773, \beta = 6.235, \gamma = 2.0$	0.043	0.112	15.369	0.081	2.392
Generalized Pareto	$\xi = 0.190, \sigma = 3.823, \mu = 1.941$	0.020	0.910	5.393	0.80	0.496
GEV	$\xi = 0.360, \sigma = 2.358, \mu = 4.012$	0.042	0.057	15.806	0.071	4.533
Inverse Gaussian	$\lambda = 3.148, \mu = 5.07, \gamma = 1.591$	0.045	0.090	15.746	0.072	2.183

[a]CDFs are provided in Appendix A
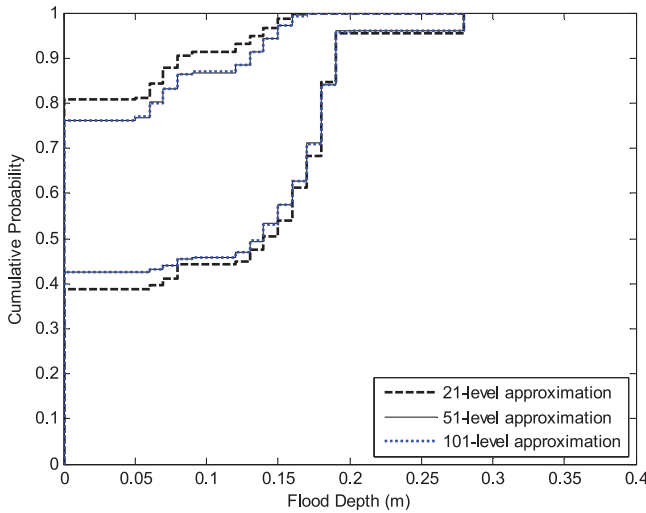
Figure 3. Lower and upper cumulative probabilities of flood depth at node N126. These curves were derived using the discretization method, considering only imprecise probabilistic rainfall uncertainties.

5.2. The α-Cut Discretization

[41] A fuzzy set can be represented as a nested set of intervals through the α-cut method [*Dubois and Prade*, 1991]. The α-cut of a fuzzy number S is defined as the set containing all the values x with membership degree no less than $\alpha \in [0, 1]$. For $n+1$ α-cut levels $0 = \alpha_0 < \cdots < \alpha_n = 1$, we have the a sequence of nested sets $S_0 \supset \cdots \supset S_n$, where $S_i = x \,|\, \mu_s(x) \geq \alpha_i\}$, $i = 0, \ldots, n$. These nested sets can be regarded as the focal elements in a random set (κ, m). For the set S_i, the probability mass can be calculated as

$$m(S_i) = \begin{cases} \alpha_i - \alpha_{i+1}, i = 1, \ldots, n \\ 0, i = 0 \end{cases} \qquad (18)$$

[42] In the case of the fuzzy set support S_0, the probability mass is thought to be completely unknown. In this way, the fuzzy set is transformed into a random set (κ, m) with n focal elements $S_i \in \kappa$ $(i = 1, \ldots, n)$. A 6-level discretization for the trapezoidal fuzzy numbers is shown in Figure 4.

5.3. Combining Model Parameter and Rainfall Uncertainties

[43] The imprecise probabilistic rainfall uncertainties and fuzzy model parameter uncertainties (using trapezoidal fuzzy numbers) can be combined in a joint random set using the discretization method. The 51-level approximation for rainfall depth and duration is used in this case, as it was proven to be accurate in the previous case of rainfall uncertainty. Two approximations, 6- and 11-level, are considered for the fuzzy model parameters, leading to a total of 62,500 (50 × 50 × 5 × 5) and 250,000 (50 × 50 × 10 × 10) focal elements in the joint random set, respectively. No correlation between any two variables is considered in constructing the joint random set. Similarly to the rainfall variable case, the image of all focal elements can be calculated using the convex method. However, the number of model simulations for each focal element can be reduced from 16

(2^4) to 8, as monotonicity can be observed for the two model parameters considered.

[44] Figure 3 is now redrawn as Figure 5 to show that the lower and upper cumulative probabilities from the high-level approximation bracket those from the low-level approximation. The gap is only reduced slightly when the higher-level approximation is used. The results from the 11-level approximation of model parameters are sufficiently accurate when compared with those from the MC method as discussed below.

6. Monte Carlo Simulations

[45] The Monte Carlo sampling method can be used to sample the fuzzy sets of model parameters and the lower and upper CDFs of rainfall variables. For imprecise probabilistic rainfall variables, each cumulative probabilistic level corresponds to a unique interval in the lower and upper CDFs and vice versa. Similarly, for fuzzy model parameters, each α-cut level of a fuzzy set relates to a unique α-cut subset. If the level is drawn from a uniform distribution on (0 1), then the MC method can be used to approximate equations (10) and (11) [*Alvarez*, 2006].

[46] In order to reduce the number of samples required and to provide a more efficient sampling, the Latin Hypercube Sampling (LHS) technique [*McKay et al.*, 1979] is used to generate a joint random set for uncertain variables. Assuming that n sample points $\mathbf{u}_i = \left(u_i^1, \ldots, u_i^d \right) \in (0, 1]^d$ $(i = 1, \ldots, n)$ are generated from the independent uniform distributions on (0 1] using the LHS technique, where d represents the number of variables and $d = 4$ in the case of considering both rainfall variables and model parameters. In the situations where the correlation between different variables has to be considered, the concept of copula can be used to generate the samples [*Alvarez*, 2006]. Each element of $\mathbf{u}_i$, i.e., u_i^j $(j = 1, \ldots, d)$, is then used to derive the corresponding focal element A_i^j for variable x_j. For the rainfall variables, the corresponding marginal focal element is obtained by choosing the lower and upper values at the probability level u_i^j, i.e.,

$$A_i^j = [\overline{F_X}^{-1}(u_i^j), \underline{F_X}^{-1}(u_i^j)] \qquad (19)$$

[47] For model parameters, the marginal focal element is obtained by deriving the interval at the u_i^j-level cut of the fuzzy number S, i.e., $A_i^j = \{x_j | u_S(x_j) \; u_i^j\}$. The joint focal element can be obtained as $A_i = A_i^1 \times \ldots \times A_i^d$. In this way, each sample point $\mathbf{u}_i$ $(i = 1, \ldots, n)$ has one corresponding joint focal element A_i. Thus a finite random set (F_n, m) is generated using the LHS sampling where $F_n = \{A_1, A_2, \ldots, A_n\}$ and $m(A_i) = 1/n$ implying that an equal weight is assigned to the sampled elements as they are generated randomly. The joint random set (F_n, m) contains all the information from all uncertainty sources of different types. The image of all focal elements in the joint random set can be calculated using the convex method, as in the discretization method. The lower and upper cumulative probabilities (belief and plausibility measures) can be constructed for flood depth on the basis of the propagated random set. The derived belief and plausibility measures using the above MC method can theoretically converge to their true values when $n \rightarrow \infty$ [*Alvarez*, 2006].
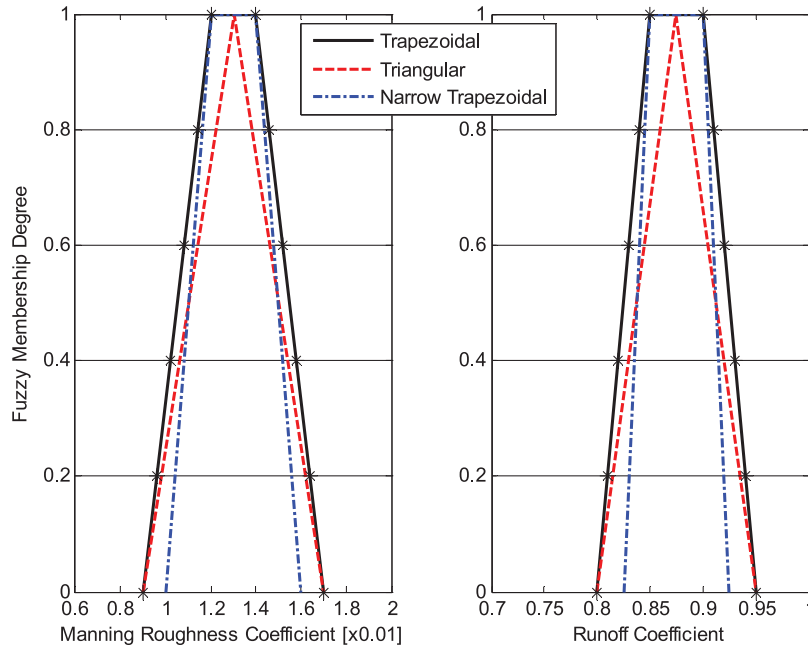
Figure 4. Fuzzy constructs for model parameters and their 6-level α-cut discretization.

[48] In this study, the impact of the number of samples on the precision of simulation results has been investigated. Three different sample sizes, 1000, 2000, and 5000, were used to simulate the CDFs of flood depth. In each case, five runs were conducted, taking into account the randomness of the MC method. The variation of the derived lower and upper CDFs of flood depth between different runs is reduced when more samples are used. Particularly, the results from the five different runs are almost identical in the case of 5000 samples. The linear averages of different runs from 1000 samples have excellent agreement with those of 5000 samples, as shown in Figure 6. This confirms
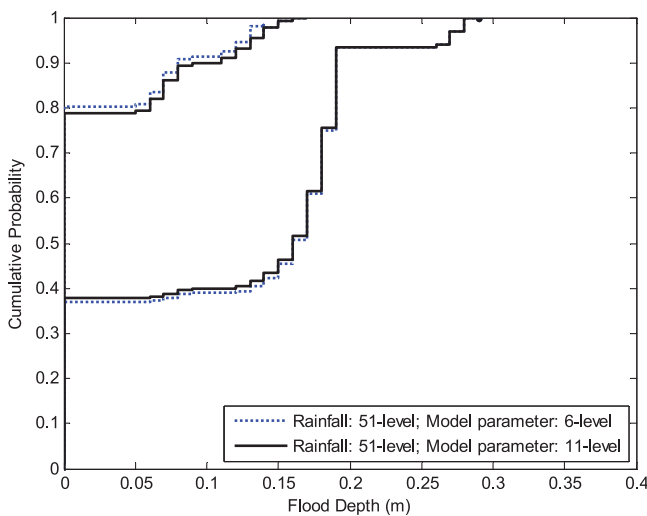
that the results derived with the MC method are unbiased [*Alvarez*, 2009]. A set of 5000 samples was considered to be sufficiently accurate in this case study, and thus was used in the following investigations.

7. Discussion of Results

7.1. Imprecise Probabilistic Representation of Rainfall

[49] Stochastic rainfall is conventionally characterized by probability distributions, and subjective knowledge expressed through Bayesian analysis. However, in many situations, prior information or beliefs may be very weak, incomplete, or even nonexistent, implying that it is inappropriate for uncertain statistical parameters to be represented using exactly specified probability distributions [*Caselton and Luo*, 1992]. Such potentially ambiguous beliefs can be captured in the form of imprecise probabilities, which are bounded by a pair of lower and upper probabilities. In the context of generalized game theory [*Walley*, 1991], the lower bound can be interpreted as the highest betting rate at which the decision maker is sure to buy a gamble, and the upper bound can be interpreted as the lowest betting rate at which the decision maker is sure to buy the opposite of the gamble (equivalent to selling the original gamble). The use of imprecise probabilities might be particularly appropriate for characterizing rainfall uncertainty in urban catchments where the rainfall data are either scarce or not suitable for sewer system modeling that requires small time step data.

[50] There are different ways to construct the imprecise probabilities of rainfall depth and duration. In this paper, a number of possible probability distributions were fitted to a set of historical rainfall data, and those that could not be rejected by all three statistical tests (K–S, χ^2, and A–D) were chosen to constrain the uncertainties in the rainfall variables. As a result, the family of chosen distributions consists of different probability functions. Alternatively, a family of distributions may be simple to obtain when one probability

Figure 5. Lower and upper cumulative probabilities of flood depth at node N126. These curves were derived using the discretization method, combining imprecise probabilistic rainfall and fuzzy model parameter uncertainties.
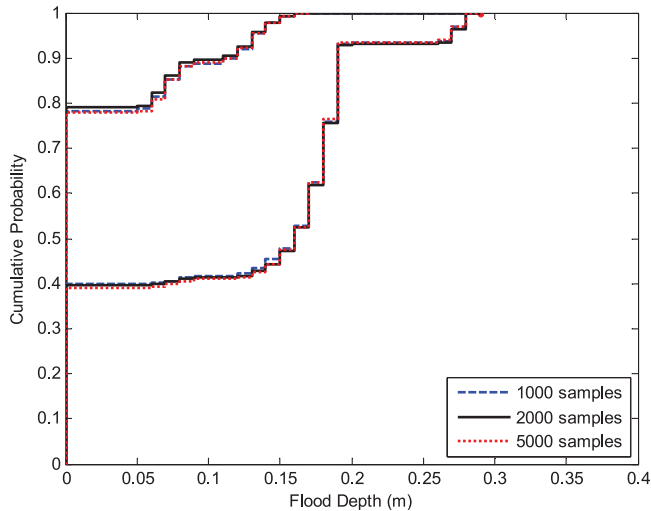
Figure 6. Lower and upper cumulative probabilities of flood depth at node N126 using the Monte Carlo method. The curves are the linear averages of five random seed runs.


Figure 7. Comparison between the conventional stochastic approach and the random set method of the cumulative probabilities of flood depth at node N126.

function is used but its parameters such as the mean value or the variance are poorly known, for instance, if the parameters are derived from subjective prior knowledge or fitted to lie in an interval. In this case, the lower and upper bounds of the family can be easily obtained.

[51] From a practical point of view, flood estimation aims to elicit the probability of flood depth, or calculate the corresponding return period. Considering the uncertainties in the modeling process, confidence intervals can be derived using probability or Bayesian approaches [e.g., *Korving et al.*, 2002]. When imprecise probabilities are used to characterize rainfall uncertainty, lower and upper probability bounds can be derived to describe the resultant flood depth. These bounds not only provide the interval that brackets the true probability of flood depth, but also they allow for an exclusion of those probabilities that would be incommensurate with the currently available information or experts' beliefs. Furthermore, using imprecise probabilities, the imprecisions are expressed explicitly to reflect the appropriate level of confidence ascribed to them. The magnitude of confidence can be indicated by the distance between the upper and lower probabilities (belief and plausibility measures). Thus, the narrower the distance, the more confidence a decision-maker can have in them.

[52] The CDFs of flood depth from an imprecise probabilistic representation of rainfall uncertainty are compared with those derived when both rainfall depth and duration are represented by classic, precise probability distributions. Several combinations of the probability distributions of rainfall duration and depth in Tables 1 and 2 are propagated using the standard MC method. Figure 7 shows the comparative results that consider rainfall variables only. As expected, the lower and upper probability bounds of flood depth bracket any of the combinations from the precise probabilistic cases. None of these combinations are close to the upper or lower bounds, indicating how unwise it is to rely solely on such estimates.

[53] According to the lower and upper CDFs of flood depth in Figure 7, for each rainfall event, the probability of
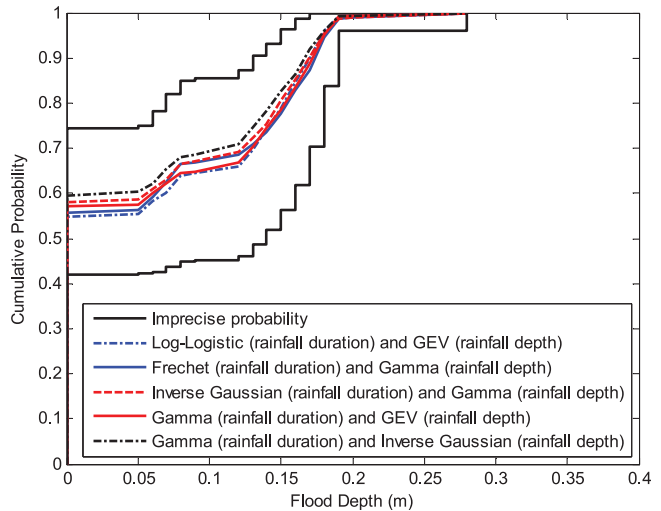
no flooding occurring (flood depth = 0) at node N126 lies in the range (0.42–0.74). In other words, when a rainfall event occurs, the probability of flooding at this node is from 0.26 to 0.58, which is equivalent to the probability of system failure. Because the average number of rainfall events is 77 per year, the number of flood events ranges from 20 to 45 in one year on average. This result is consistent with the number of floods (about 30 per year) obtained when the 10 year rainfall series is used for simulation. This high number of failures at this critical node is caused by the expansion of the network to the (left) upstream due to urban development. Similarly, the probability bounds for any specific flood depth can be derived; for example, the likelihood of flood depth greater than 0.15 m is confined to the range (0.06–0.43). Although these probability gaps are apparently rather large, they do represent the belief interval given the substantial uncertainties in rainfall, which include the epistemic uncertainty in choosing the distribution types for rainfall depth and duration as well as their combinations, and complete ignorance of the shape of synthetic rainfall events. These gaps can be reduced only when more data or knowledge are available. For example, Figure 7 illustrates how the probability gap is narrowed substantially when and if the distribution type for rainfall depth and duration and their combinations are known with certainty.

[54] In addition to imprecise probabilities and fuzzy sets, other types of uncertainty representations can also be combined in the integrated methodology using random sets [*Hall*, 2003; *Tonon*, 2004]. For example, precise probabilities or intervals can be used to characterize one variable where necessary, and then combined with either or both imprecise probabilities and fuzzy sets.

7.2. Model Parameter Uncertainties

[55] The impact of model parameter uncertainties on the lower and upper bounds of CDFs for flood depth can be established by comparing the probability gap in uncertainties for rainfall only with those of the rainfall and model parameters. Figure 8 shows the results obtained from the
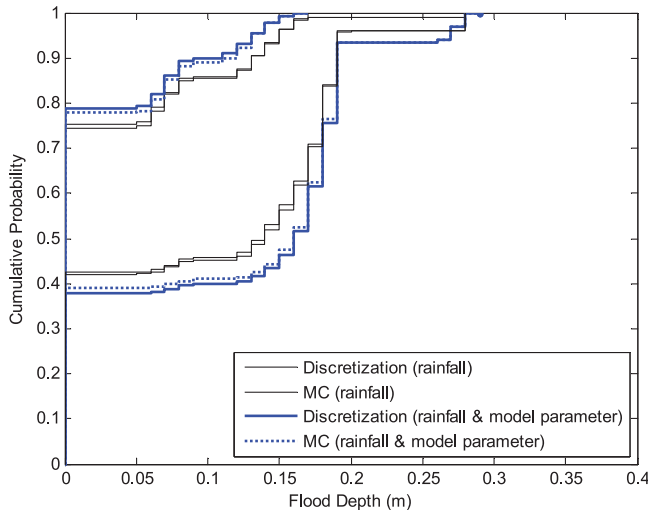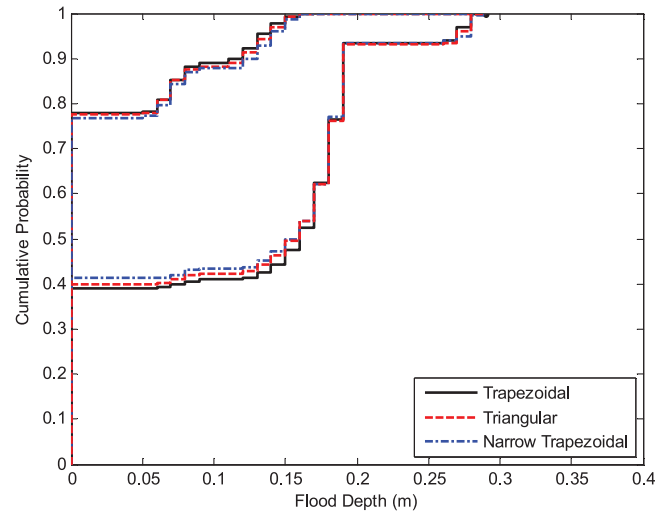
Figure 8. Comparison of the discretization and MC methods regarding the cumulative probabilities at Node 126.

Figure 9. Influence of different constructs of fuzzy model parameters on the cumulative probabilities at Node 126.

two propagation methods: the discretization method with a 51-level approximation for rainfall depth and duration and an 11-level approximation for the two model parameters and the MC method using a sample size of 5000 (averaged over five random runs). The gap between the two pairs of solid curves (or dotted curves) shows the influence of model parameter uncertainties, which is rather small compared with that of rainfall uncertainties. Recall that the gap between the lower and upper CDFs indicates the magnitude of (reducible) epistemic uncertainty, thus the gap reduction realized by considering additional information indicates the value of such information. To effectively reduce the overall uncertainty in flood depth estimation, effort should be made toward reducing rainfall prediction uncertainty.

[56] The impact of different fuzzy constructs has also been investigated. Compared with the original trapezoidal fuzzy numbers, two different constructs for runoff and roughness coefficients have also been simulated: a triangular fuzzy number with the same support and a narrower-support trapezoidal fuzzy number (Figure 4). Figure 9 shows the linear average CDFs of flood depth at node N126 from 5 runs using the MC method with 5000 samples. The variation in fuzzy numbers has very small impact on the probability gaps; however, the reduction in the fuzzy number kernel has less impact than the reduction in the fuzzy number support. This implies that, to reduce fuzzy model parameter uncertainties, it is more valuable to seek expert knowledge constraining the maximum range with the lowest possibility (the support) than the minimum range with the highest possibility (the kernel).

7.3. Comparison of the Discretization and MC Methods

[57] For imprecise probabilities, there exists a clear difference between the discretization and MC method in constructing the random set. In the discrete method, the interval extremes of each focal element in equation (15) are derived using different probability levels, while in the MC method the same level is used, in equation (19). When the degree of discretization is low, the information contained in the lower and upper CDFs can be lost in the transformation process through p-boxes. Thus, a danger exists that the discretization method

can severely underestimate the lower bound and overestimate the upper bound [*Tonon*, 2004], although it ensures that the p-boxes completely envelop the lower and upper CDFs.

[58] Figure 8 shows a comparison of the results from the two methods. In both of the two cases, i.e., considering rainfall uncertainty only or together with model parameter uncertainty, the two methods generated very similar lower and upper CFDs. However, the underlying computational requirements are different: in the case of rainfall and model parameter uncertainty, the discretization method used 250,000 ($50 \times 50 \times 10 \times 10$) focal elements in the joint random set, the MC method used only 5000 focal elements.

[59] The discretization method requires $(2N)^d$ model evaluations for uncertainty propagation, where N is the number of focal elements for each of d variables, as there are N^d focal elements and 2^d evaluations for each focal element using the convex method (8 model evaluations, instead of $2^4 = 16$, are necessary for consideration of rainfall and model parameter uncertainties because of the monotonicity of the model parameters). The number of model evaluations increase dramatically when the number of discretizations N or the dimension d increases. As shown in Figure 3, the quality of results can be poor when the discretization is low. Although results quality can be improved by increasing the degree of discretization for uncertain variables, the computation required makes it impractical for high-dimensional problems. However, accuracy of the MC method mainly depends on the number of samples and is insensitive to the number of variables [*Alvarez*, 2009]. Thus, the MC method is recommended for high-dimensional problems.

8. Conclusions

[60] Uncertainty in sewer flooding modeling, when it is estimated at all, is typically handled using probability theory. However, it has been argued that the type uncertainty is too broad to be captured by probability measures alone. Given the presence of imprecise data, vague expert knowledge, and incomplete understanding of the system, the challenge of sewer flood analysis is to assess what system failures can be ruled out as unlikely, rather than simply ruled in as probable or

possible. In this paper, a random set-based framework for sewer flood analysis has been presented, in which two different uncertainty characterizations, i.e., imprecise probabilities from rainfall and fuzziness from model parameters, are handled simultaneously. This provides a more complete and therefore accurate means of capturing uncertainties in data and models by applying the most appropriate uncertainty characterization to each uncertainty source, rather than assuming one for all sources regardless of their underlying nature. This new methodology is promising, in that it provides a single mathematical framework to handle fuzzy sets and (imprecise) probabilities in the uncertainty analysis process. The following conclusions are presented on the basis of this study:

[61] 1. Imprecise probabilities are more appropriate to describe stochastic uncertainty in rainfall when more than one probability distribution fits well based on the available data. Excluding any possible probability distributions or imposing a false probability distribution may lead to inappropriate conclusions being drawn from the flood analysis process.

[62] 2. Random set theory provides a general framework to accommodate different uncertainty characterizations, and can be applied to handle stochastic rainfall uncertainty and fuzzy model parameter uncertainty in a single modeling framework as demonstrated in the case study. The benefit is that the stochastic and epistemic uncertainties can be represented in the most appropriate form wherever they originally appear.

[63] 3. Discretization and MC methods are presented and compared for uncertainty propagation within the random set-based flood analysis framework. Results show that the MC method is more computationally efficient at deriving the lower and upper probability bounds of flood depth. From 5000 samples, the MC method can provide an accurate estimate of the lower and upper CDFs of flood depth. The discretization method is only practical for low-dimensional problems, as the computation required increases exponentially with problem dimension.

[64] 4. Lower and upper probabilities can be derived for flood depths using belief and plausibility functions as a result of imprecise representation of information, which embrace the true unknown probabilities. The distance between the lower and upper probabilities reflects the imprecision, incompleteness, or conflict in data and models, and so provides an indication of the magnitude of confidence for use by the decision maker.

[65] 5. Rainfall (epistemic) uncertainty contributes to a significant portion of the probability gaps in flood depth estimation in the sewer system studied in this paper, compared with model parameter uncertainty, and this indicates where effort should be made to reduce the overall uncertainty of flood estimation. Furthermore, comparative results show that constraining the support (the maximum range with the lowest possibility) of fuzzy model parameters is more valuable in reducing the overall uncertainty than constraining the kernel (the minimum range with the highest possibility).

[66] The methodology presented can be extended to estimate flood risk, by including lower and upper bounds for the possible cost caused by sewer flooding, and thus potentially provides a useful tool for sewer flood management. Further, the framework can be readily developed to include other representations of uncertainty, for example, intervals or classic probabilities.

Appendix A: Cumulative Distribution Functions

A1.   Frechet Distribution

[67]

$$F(x) = \exp\left\{-[(x - \mu)/\sigma]^{-\xi}\right\}$$

where $\xi > 0$, $\sigma > 0$, and μ are shape, scale, and location parameters, respectively

A2.   Gamma Distribution

[68]

$$F(x) = \Gamma_{(x-\gamma)/\beta}(\alpha)/\Gamma(\alpha)$$

where $\alpha > 0$, $\beta > 0$, and γ are shape, scale, and location parameters, respectively. Γ is the Gamma function,

$$\Gamma(\alpha) = \int\limits_0^\infty t^{\alpha-1}e^{-t}dt,$$

for $(\alpha > 0)$ and Γ_z is the incomplete Gamma function,

$$\Gamma_z(\alpha) = \int\limits_0^z t^{\alpha-1}e^{-t}dt$$

for $(\alpha > 0)$.

A3.   Generalized Extreme Value Distribution

[69]

$$F(x) = \exp\left\{-[1 + \xi(x - \mu)/\sigma]^{-1/\xi}\right\}$$

where $\xi \neq 0$, $\sigma > 0$, and μ are shape, scale, and location parameters, respectively.

A4.   Generalized Pareto Distribution

[70]

$$F(x) = 1 - [1 + \xi(x - \mu)/\sigma]^{-1/\xi}$$

where $\xi \neq 0$, $\sigma > 0$, and μ are shape, scale, and location parameters, respectively.

A5.   Inverse Gaussian Distribution

[71]

$$F(x) = \Phi\left\{[\lambda/(x - \gamma)]^{1/2}[(x - \gamma)/\mu - 1]\right\} \\ + \Phi\left\{-[\lambda/(x - \gamma)]^{1/2}[(x - \gamma)/\mu + 1]\right\}\exp(2\lambda/\mu),$$

where $\lambda > 0$, $\mu > 0$, and γ are parameters, and Φ is the Laplace Integral, i.e., the CDF of the standard normal distribution

$$\Phi(x) = (2\pi)^{-1/2} \int_0^x e^{-t^2/2} dt \,.$$

A6. Log-Logistic Distribution

[72]

$$F(x) = [1 + (\beta/x)^\alpha]^{-1}$$

where $\alpha > 0$ and $\beta > 0$ are parameters.

A7. Pearson Type 6 Distribution

[73]

$$F(x) = B_{x/(x+\beta)}(\alpha_1, \alpha_2)/B(\alpha_1, \alpha_2)$$

where $\alpha_1 > 0$ and $\alpha_2 > 0$ are shape parameters, and $\beta > 0$ is the scale parameter. B is the Beta function

$$B(\alpha_1, \alpha_2) = \int_0^1 t^{\alpha_1-1}(1-t)^{\alpha_2-1} dt$$

for $\alpha_1 > 0$ and $\alpha_2 > 0$,

$$B_z(\alpha_1, \alpha_2) = \int_0^z t^{\alpha_1-1}(1-t)^{\alpha_2-1} dt$$

for $\alpha_1 > 0$, $\alpha_2 > 0$, and $0 \le x \le 1$.

References

Adams, B. J., and F. Papa (2000), *Urban Stormwater Management Planning With Analytical Probabilistic Models*, 358 pp., John Wiley, Chichester, U.K.

Alvarez, D. A. (2006), On the calculation of the bounds of probability of events using infinite random sets, *Int. J. Approximate Reasoning*, *43*(3), 241–267.

Alvarez, D. A. (2009), A Monte Carlo-based method for the estimation of lower and upper probabilities of events using infinite random sets of indexable type, *Fuzzy Sets Syst.*, *160*, 384–401.

Butler, D., and J. W. Davies (2004), *Urban Drainage*, 2nd ed., Spon Press, London.

Caselton, W. F., and W. Luo (1992), Decision making with imprecise probabilities: Dempster-Shafer theory and application, *Water Resour. Res.* *28*(12), 3071–3083, doi:10.1029/92WR01818.

Dempster, A. P. (1967), Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.*, *38*, 325–339.

Dubois, D., and H. Prade (1991), Random sets and fuzzy interval analysis, *Fuzzy Sets Syst. 42*, 87–101.

Freni, G., G. Mannina, and G. Viviani (2009), Urban runoff modelling uncertainty: Comparison among Bayesian and pseudo-Bayesian methods, *Environ. Modell. Software*, *24*(9), 1100–1111.

Fullerton, J. N. (2004), A simulated modelling approach for storm water flow optimisation, Ph.D. thesis, Cent. for Water Syst., Univ. of Exeter, U. K.

Guo, Y., and B. J. Adams (1998), Hydrologic analysis of urban catchments with event-based probabilistic models 1. Runoff volume, *Water Resour. Res.*, *34*(12), 3421–3431, doi:10.1029/98WR02449.

Guyonnet, D., B. Bourgine, D. Dubois, H. Fargier, B. Côme, and J.-P. Chilès (2003), Hybrid approach for addressing uncertainty in risk assessments, *J. Environ. Eng.*, *129*(1), 68–78. doi:10.1061/(ASCE)0733-9372(2003)129:1(68).

Hall, J. (2003), Handling uncertainty in the hydroinformatic process, *J. Hydroinformatics*, *5*, 215–232.

Hall, J., G. Fu, and J. Lawry (2007), Imprecise probabilities of climate change: aggregation of fuzzy scenarios and model uncertainties, *Clim. Change*, *81*(3-4), 265–281.

Jacquin, A. P., and A. Y. Shamseldin (2007), Development of a possibilistic method for the evaluation of predictive uncertainty in rainfall-runoff modeling, *Water Resour. Res.*, *43*, W04425, doi:10.1029/2006WR005072.

Kendall, D. G. (1974), Foundations of a theory of random sets, in *Stochastic Geometry*, edited by E. F. Harding and D. G. Kendall, pp. 322–376, New York, Wiley.

Korving, H., F. Clemens, J. van Noortwijk, and P. van Gelder (2002), Bayesian estimation of return periods of CSO volumes for decision-making in sewer system management, in *Global Solutions for Urban Drainage, Proceedings of the Ninth International Conference on Urban Drainage*, [CD-ROM], edited by E. W. Strecker and W. C. Huber, Am. Soc. of Civ. Eng., Reston, Va.

Lei, J. H., and W. Schilling (1994), Parameter uncertainty propagation analysis for urban rainfall runoff modelling, *Water Sci. Technol.*, *29*(1-2), 145–154.

Makropoulos, C., D. Butler, and C. Maksimovic (2003), Fuzzy logic spatial decision support system for urban water management, *J. Water Resour. Plann. Manage. 129*(1), 69–77.

Matheron, G. (1975), *Random Sets and Integral Geometry*, Wiley, New York.

Matott, L. S., J. E. Babendreier, and S. T. Purucker (2009), Evaluating uncertainty in integrated environmental models: A review of concepts and tools, *Water Resour. Res.*, *45*, W06421, doi:10.1029/2008WR007301.

McKay, M. D., W. J. Conover, and R. J. Beckman (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, *211*, 239–245.

Moore, R. E. (1966), *Interval Analysis*, Prentice Hall, Englewood Cliffs, N. J.

Nguyen, H. T. (2006), *An Introduction to Random Sets*, Chapman and Hall, Boca Raton, Fla.

Parliamentary Office of Science and Technology (POST) (2007), POST note 289—Urban flooding, report, London, (available at www.parliament.uk/documents/upload/postpn289.pdf)

Revelli, R., and L. Ridolfi (2002), Fuzzy approach for analysis of pipe networks, *J. Hydraul. Eng.*, *128*(1), 93–101.

Ross, J. L., M. M. Ozbek, and G. F. Pinder (2009), Aleatoric and epistemic uncertainty in groundwater flow and transport simulation, *Water Resour. Res.*, *45*, W00B15, doi:10.1029/2007WR006799.

Ross, T. (2004), *Fuzzy Logic With Engineering Applications*, 654 pp., John Wiley, Chichester, U. K.

Rubio, E., J. W. Hall, and M. G. Anderson (2004), Uncertainty analysis in a slope hydrology and stability model using probabilistic and imprecise information, *Comput. Geotech. 31*, 529–536.

Ryu, J. (2008), *Decision Support for Sewer Flood Risk Management*, Ph.D. thesis, 323 pp., Dep. of Civil and Environ.. Eng.., Imperial College London, U. K.

Schmitt, T. G., M. Thomas, and N. Ettrich (2004), Analysis and modeling of flooding in urban drainage systems, *J. Hydrol.*, *299*(3-4), 300–311.

Schweiger, H. F., and G. M. Peschl (2005), Reliability analysis in geotechnics with the random set finite element method, *Comput. Geotech.*, *32*, 422–435.

Shafer, G. (1976), *A Mathematical Theory of Evidence*, 297 pp., Princeton Univ. Press, Princeton, N. J.

Thorndahl, S., and P. Willems (2008), Probabilistic modelling of overflow, surcharge and flooding in urban drainage using the first-order reliability method and parameterization of local rain series, *Water Res.*, *42*(1-2), 455–466.

Tonon, F. (2004), Using random set theory to propagate epistemic uncertainty through a mechanical system, *Reliab. Eng. Syst. Safety*, *85*, 169–181.

Walley, P. (1991), *Statistical Reasoning With Imprecise Probabilities*, Chapman and Hall, London.

Zadeh, L. (1965), Fuzzy sets, *Inf. Control*, *8*, 338–353.

Zhang, K., H. Li, and G. Achari (2009), Fuzzy-stochastic characterization of site uncertainty and variability in groundwater flow and contaminant transport through a heterogeneous aquifer, *J. Contam. Hydrol. 106* (1-2), 73–82.

D. Butler, G. Fu, and S. Sun, College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK. (g.fu@exeter.ac.uk)

S.-T. Khu, School of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, UK.