The cost of imperfect knowledge: how epistemic uncertainties influence flood hazard assessments

Mariano Balbi[1] and David Lallemant[2]

[1]University of Buenos Aires
[2]Earth Observatory of Singapore

July 9, 2023

Abstract

Classical approaches to flood hazard are obtained by the concatenation of a recurrence model for the events (i.e. an extreme river discharge) and an inundation model that propagates the discharge into a flood extent. The traditional approach, however, uses 'best-fit' models that do not include uncertainty from incomplete knowledge or limited data availability. The inclusion of these, so called epistemic uncertainties, can significantly impact flood hazard estimates and the corresponding decision-making process. We propose a simulation approach to robustly account for uncertainty in model's parameters, while developing a useful probabilistic output of flood hazard for further risk assessments. A Peaks-Over-Threshold Bayesian analysis is performed for future events simulation, and a pseudo-likelihood probabilistic approach for the calibration of the inundation model is used to compute uncertain water depths. The annual probability averaged over all possible models' parameters is used to develop hazard maps that account for epistemic uncertainties. Results are compared to traditional hazard maps, showing that not including epistemic uncertainties can underestimate the hazard and lead to non-conservative designs, and that this trend increases with return period. Results also show that the influence of the uncertainty in the future occurrence of discharge events is predominant over the inundation simulator uncertainties for the case study.

The cost of imperfect knowledge: how epistemic uncertainties influence flood hazard assessments

M. Balbi[1*], and D. C. B. Lallemant[2]

[1]Laboratorio de Materiales y Estructuras, School of Engineering, Universidad de Buenos Aires, Buenos Aires, Argentina
[2]Earth Observatory of Singapore, Nanyang Technological University, Singapore

Key Points:

- Flood hazard assessments involve sophisticated probability and physics-based models that require the specification of many parameters
- We propose a Bayesian methodology to include uncertainty in model's parameters into robust hazard estimates and useful hazard maps for risk-based decision-making
- The inclusion of uncertainty in parameters can significantly affect hazard estimates and its omission can lead to non-conservative planning and design

*Current address, Av. Las Heras 2214, Buenos Aires, Argentina

Corresponding author: Mariano Balbi, mabalbi@fi.uba.ar

Abstract

Classical approaches to flood hazard are obtained by the concatenation of a recurrence model for the events (i.e. an extreme river discharge) and an inundation model that propagates the discharge into a flood extent. The traditional approach, however, uses 'best-fit' models that do not include uncertainty from incomplete knowledge or limited data availability. The inclusion of these, so called epistemic uncertainties, can significantly impact flood hazard estimates and the corresponding decision-making process. We propose a simulation approach to robustly account for uncertainty in model's parameters, while developing a useful probabilistic output of flood hazard for further risk assessments. A Peaks-Over-Threshold Bayesian analysis is performed for future events simulation, and a pseudo-likelihood probabilistic approach for the calibration of the inundation model is used to compute uncertain water depths. The annual probability averaged over all possible models' parameters is used to develop hazard maps that account for epistemic uncertainties. Results are compared to traditional hazard maps, showing that not including epistemic uncertainties can underestimate the hazard and lead to non-conservative designs, and that this trend increases with return period. Results also show that the influence of the uncertainty in the future occurrence of discharge events is predominant over the inundation simulator uncertainties for the case study.

Plain Language Summary

Estimating the annual probability of some flood-depth level is a key input for risk analysis and engineering design. This is typically calculated via sophisticated probability and physics-based models that require many parameters. However, the classical approach uses a fixed set of 'best parameters' for this and do not include the degree of uncertainty, even when such uncertainties may be very high. This work proposes a method to estimate the annual probability of flood-depth including the uncertainty in the parameters used to compute it. More importantly, it shows that not including this uncertainty might severely underestimate the hazard and consequently lead to unsafe designs.

1 Introduction

As a key component of a comprehensive risk analysis, the hazard model is required to characterize the future occurrence of potentially damaging events. This 'potential' for damage is numerically quantified through an intensity measure (IM) metric that, in the case of flood hazard, usually is the water level, velocity and/or duration at any given point of interest (Pregnolato et al., 2015). It, ultimately, has the purpose of providing valuable input information for vulnerability (e.g. damage) models that allow decision-makers to help mitigate the impact of natural hazards.

In this context, flood hazard is typically defined as the probability of exceedance of an IM level at any point of interest during a given period of time. This usually comes in the form of 'hazard curves' that relate IM levels with an annual exceedance probability (AEP) or a mean time of recurrence, also known as return period (RP). In practice, the outcome is best conveyed through flood maps for different RPs to reflect the spatial distribution of hazard estimates.

Since observations of IMs usually scarce for most locations, a purely statistical description of their probability distribution is not possible. The typical approach, then, is to compute IMs as a result of the convolution of two distinct models: a 'recurrence model' that describes the probability of occurrence of extreme events, such as an extreme rainfall, river discharge, or sea-level rise; and a 'source-to-site propagation model' (or just propagation model for brevity) that represents how the triggering event is translated into a spatial (and temporal) distribution of IM levels (i.e. spatially distributed flood extent and depth). The former is inherently probabilistic and typically modelled via standard stochastic time process such as the Poisson process. The propagation model, on the other hand, is typically modelled through deterministic physics-based models such as a hydrologic and/or a hydraulic inundation model.

The classical method involves computing the event's magnitude, such as peak discharge in riverine flooding, for different RPs and use it as input of the inundation model to develop flood maps for the different recurrences. This is done using appropriately calibrated recurrence curves and inundation simulators based on available data and expert's knowledge. This is the standard approach for most practical applications in the industry, due to its conceptual simplicity and ease of implementation. It does, however, assume that the models used are perfectly 'true'. The only probabilistic nature of the approach comes from the inherent uncertainty in the future occurrence of extreme events; also termed 'aleatory uncertainty' (J. Hall & Solomatine, 2008).

A broad range of researchers during the last decades have brought attention to the importance of including other, more subjective, sources of uncertainty into risk analysis in general, and flood hazard modelling in particular (J. Hall & Solomatine, 2008; Beven, 2014; Merz & Thieken, 2005). Subjective uncertainties, also termed here 'epistemic', can arise from the data we use to constrain our models, our lack of knowledge regarding the true physical processes involved, or our limited analytic and computational capabilities for providing results (J. Hall & Solomatine, 2008). Their inclusion may lead to impactful modifications in the decision making process, at the cost of a significant increment in analytical and computational complexity.

Many researchers have dealt with the inclusion of epistemic uncertainties in flood hazard models, specifically riverine floods, in the last decades. Some works have focused mainly on dealing with uncertain representations on the recurrence of the input discharge events. That is, defining a distribution for the uncertain discharge for a given return period. This includes accounting for statistical fitting errors due to limited-length data and distribution family (Apel et al., 2008; G. T. Aronica et al., 2012; Neal et al., 2013; Romanowicz & Kiczko, 2016; Stephens & Bledsoe, 2020), secondary input variables as flood volume (Candela & Aronica, 2017), or more general hydrograph shape uncertainties through hydrological modelling (Grimaldi et al., 2013; Falter et al., 2015; Ahmadisharaf et al., 2018; Zahmatkesh et al., 2021). Others have focused on including uncertainty in the inundation model through its most sensitive parameters such as roughness coefficients (Di Baldassarre et al., 2010; Kalyanapu et al., 2012; G. T. Aronica et al., 2012; Kiczko et al., 2013; Romanowicz & Kiczko, 2016; Bharath & Elshorbagy, 2018), Digital Elevation Maps (DEM) (Apel et al., 2008), or cross-section geometrical properties (Stephens & Bledsoe, 2020). Furthermore, many of these have included both the epistemic uncertainties in the discharges recurrence as well as in the inundation model (Apel et al., 2008; Di Baldassarre et al., 2010; Kalyanapu et al., 2012; G. T. Aronica et al., 2012; Kiczko et al., 2013; Romanowicz & Kiczko, 2016; Bharath & Elshorbagy, 2018; Stephens & Bledsoe, 2020; Zahmatkesh et al., 2021).

The typical outcome from most of these approaches is in the form of 'probability of flood' maps for different return periods. That is, for a specific return period, different discharges and/or inundation model parameters are randomly sampled and used to obtain an ensemble of flood maps from which the probability of flooding is computed empirically (Di Baldassarre et al., 2010; Domeneghetti et al., 2013; Neal et al., 2013; Kiczko et al., 2013; Bharath & Elshorbagy, 2018; Stephens & Bledsoe, 2020; Zahmatkesh et al., 2021). A flood risk analysis requires estimating potential damages from the hazard outcomes, and this type of input is not very helpful since most flood damage models use as input the water depth above ground level (Pregnolato et al., 2015). For this reason, instead of translating uncertain flow discharges for given RP into a probability map (broadly known as 'event-based' approach), some researchers have aimed to develop recurrence curves for water depths at the points of interest (G. T. Aronica et al., 2012; Nuswantoro et al., 2016; Romanowicz & Kiczko, 2016).

The literature review indicates a lack of hazard methodologies that can (1) include epistemic uncertainties in both the recurrence and inundation models, (2) provide useful output for further risk assessments, while (3) also being probabilistically consistent and computationally tractable. This work explores a simulation methodology of flood scenarios using a Bayesian approach of extreme value theory and the Generalized Likelihood Uncertainty Estimation (GLUE) framework to account for epistemic uncertainty in the parameters of the recurrence model and the inundation simulator respectively. We propose to use the probability of exceedance averaged over the distribution of all possible parameters as a point measure of flood hazard due to its improved statistical properties as discussed in Fawcett and Green (2018), and its lower computational demand compared to obtaining full credible intervals. This framework allows the development of flood hazard curves, as well as flood hazard maps by computing this estimate of the recurrence at every point of interest.

Section 2 describes the mathematical model used to compute hazard estimates, a framework to include epistemic uncertainties through model's parameters posterior distributions, and a simulation procedure for its numerical evaluation. A small case study of riverine flooding is described in Sect. 3 and used as a working example to test this methodology. In Sect. 4, the resulting hazard curves and maps are compared to the traditional approach where no epistemic uncertainties are included. Insights in the hypothesis, results and implications of the model are analyzed in the discussions of Sect. 5, while a summary of main takeaways and potential future lines of research are drawn in the final section.

2 Methodology

2.1 The hazard model

As discussed in the introduction, flood hazard can be quantified as the annual probability of exceedance of a given IM level y, at any location of interest. Mathematically, this probability is calculated by a stochastic time process model. The most used one, due to its simplicity and well-known mathematical properties, is the Homogenous Poisson Process (HPP) for which events occur discretely with independent exponentially distributed inter-arrival times with a mean rate λ_0 that is constant over time. Under these simplifying assumptions, the probability of exceedance over a timespan T can be computed as per Eq. 1.

$$p_T(y) = 1 - \exp\left(-\lambda_0 T p\left(Y \geq y\right)\right) \tag{1}$$

Where Y is the random IM for any given event, T is the timespan of interest, and $p\left(Y \geq y\right)$ is the probability of exceedance of level y for any given event (this probability is constant over time in the HPP).

The hazard is then computed by setting $T = 1$ in Eq. 1 to obtain the annual probability of exceedance. For events with low recurrence λ_0 and low exceedance probability $p\left(Y \geq y\right)$, as is mostly the case in disaster risk analysis, the probability of Eq. 1 can be further simplified as in Eq. 2.

$$p_T(y) \approx \lambda_0 p\left(Y \geq y\right) \tag{2}$$

This is equivalent to the mean rate of exceedance of $Y \geq y$ or its multiplicative inverse, the mean time between occurrences $Tr\left(y\right)$, also known as 'return period' (see Eq. 3). In practical terms, hazard is measured by the annual probability of exceedance or by the mean rate of occurrence of IM level y, which are practically equivalent.

$$\lambda_0 p\left(Y \geq y\right) = \lambda\left(y\right) = \frac{1}{Tr\left(y\right)} \tag{3}$$

Since direct observations of water depths (or IMs in general) during flood events are very rare, probabilistic characterization of Y is usually done via mechanistic fluid dynamics models, here called 'inundation model' or simply 'simulator'. This model depends on a number of observable boundary conditions X that are considered to vary event to event, such as upstream river discharge, rainfall intensity or sea-level rise. It also depends on a set of unobserved calibration parameters β considered constant over events, such as the soil roughness parameters or the channel cross-section geometry (see Eq. 4).

$$Y = S\left(X, \beta\right) \tag{4}$$

In this context, an 'event' is characterized by a magnitude X which describes the impact potential of the phenomenon. For example, in riverine flooding, X can represent the river discharge flow and an event is triggered when it surpasses a given threshold. Analogously, in coastal flooding, X might stand for sea-level extreme rise, or in pluvial flooding where X stands for rainfall intensity. In more complex scenarios, X can be a vector representing multiple quantities, such as flow discharge and volume, rainfall intensity and duration, or a combination of flow discharge and sea-level rise. For the sake of simplicity, this work will focus on scalar X.

Since Y is a function of X, the probability $p\left(Y \geq y\right)$ in Eq. 2 can be computed by conditioning on the probability distribution of the event's magnitude X as given by,

$$p\left(Y \geq y\right) = \int_x \mathbf{1}\left\{S\left(x, \beta\right) \geq y\right\} p\left(x|\theta\right) dx \tag{5}$$

Where $\mathbf{1}\left\{\text{cond}\right\}$ is an indicator function that returns 1 when cond is true and 0 otherwise and θ is a vector of parameters that describe the probability distribution of X.

This expression is useful as long as it is easier to define the probability distribution of events magnitudes $p\left(x|\theta\right)$ than the distribution of IM levels $p\left(Y \geq y\right)$ from data or expert knowledge. As mentioned before, this is the typical case in flood hazard, where we usually have relatively robust historical measurements of river flow discharges or rainfall intensity, but very few of water depths at points of interest in the floodplain.

Introducing Eq. 5 into Eq. 2, we obtain the full expression for the flood hazard,

$$\lambda\left(y\right) = \int \underbrace{\mathbf{1}\left\{S\left(x, \beta\right) \geq y\right\}}_{\substack{\text{Inundation}\\ \text{model}}} \underbrace{\lambda_0 p\left(x|\theta\right)}_{\substack{\text{Events}\\ \text{recurrence}\\ \text{model}}} dx \tag{6}$$

An illustrative scheme of a realization of the described time process is shown in Figure 1. The varying sizes of the blue bubbles reflect the magnitude X_i of the events, while the black bars reflect the IM level (i.e. water depth) for each. According to the HPP model, the time T_i between events follows an exponential distribution with mean rate λ_i, while the time between IM exceedances $T_i\left(y\right)$ (the black bars that cross the dotted red line) follows an exponential distribution with mean rate $\lambda\left(y\right)$ as described in Eq. 6.

It is important to highlight that the model summarized here is replicated when analyzing the hazard for other types of natural phenomena. In the case of seismic hazard, X is the moment magnitude and spatial epicenter location of the earthquake and their probability distribution is typically given by the Gutenberg-Richter law and the source-to-site propagation model is defined by the Ground Motion Prediction Equations (GMPEs) (Baker et al., 2021). For typical hurricane winds hazard, the recurrence model describes the likelihood of the hurricane's central pressure and track, while the propagation model is described through a wind field model (Vickery et al., 2006).

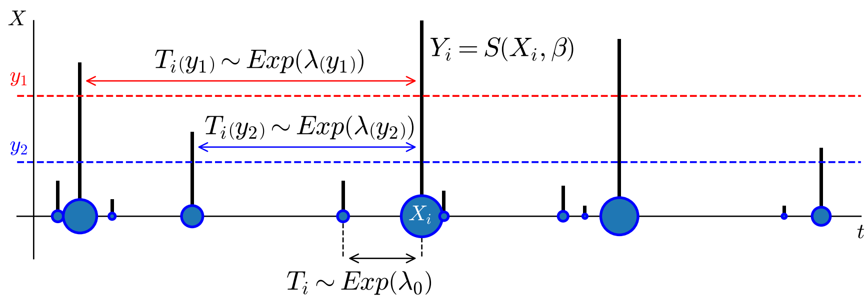

Figure 1. Schematic illustration of a realization of the HPP model, with two arbitrary IM levels y_1 and y_2 to define the hazard

2.2 Including epistemic uncertainties

The hazard problem is, as described by Eq. 6, tightly related to predicting an uncertain event in the future. Thus, it is strictly an uncertainty quantification task. At its core, that expression is a mathematical representation of what is known as 'aleatory uncertainty', here characterized by the exponentially distributed inter-arrival times with mean rate λ_0 and the probability distribution of the event's magnitude $p\,(x|\theta)$.

Aleatory uncertainty is considered an inherent component of the physical process and it does not depend on the amount of knowledge and information the modeller has. However, there are other sources of uncertainty around the estimation of the hazard that are related to our incomplete knowledge about the physical process and data available to characterize it. These are commonly known as 'epistemic uncertainties' (Spiegelhalter & Riesch, 2011).

As described by Spiegelhalter and Riesch (2011), epistemic uncertainty stems mainly from (1) limited information to properly characterize the models and variables involved and (2) limited knowledge to properly describe the true physical processes through the selected models. This more operational description of epistemic uncertainties allows for a more rigorous way of including them in the mathematical model.

Limited information appears in practice, as limited-length data, observation errors, or missing variables. It can, typically, be represented through uncertainty in the parameters that describe the models as data is not sufficient to perfectly identify them. Limited knowledge, on the other hand, is usually represented through simplifying assumptions as the ones in the HPP model, uncertainty in the distribution family chosen for $p\,(x|\theta)$, or the particular physics-based model chosen for S. It can be harder to represent this mathematically, although it has been done through model ensembles (i.e. considering and weighting many possible models) or statistical representations of model deficiencies (Kennedy & O'Hagan, 2001; Balbi & Lallemant, 2023).

Despite the epistemological differences between the two, it is not always clear which sources of uncertainty belong to each category, and it can vary depending on the context. In any case, the most important feature that differentiates aleatory and epistemic uncertainty is the fact that the former cannot be practically reduced since it is an inherent property of the system under analysis. The epistemic uncertainty, on the other hand, can be reduced by further collecting information and improving knowledge. This distinction is crucial when allocating resources for model improvement (Der Kiureghian & Ditlevsen, 2009; Merz & Thieken, 2005).

The inclusion of epistemic uncertainties greatly increases the complexity of the problem from an analytical and computational standpoint. In particular, this work focuses on the inclusion of epistemic uncertainty through model's parameters as a broad representation of limited data to define the models. In this context, hazard as calculated in Eq. 6 can be understood as being conditional to a given set of models and parameters as given by,

$$\lambda\left(y|\beta,\lambda_0,\theta\right)=\int\mathbf{1}\left\{S\left(x,\beta\right)\geq y\right\}\lambda_0 p\left(x|\theta\right)dx \tag{7}$$

One way of incorporating uncertainty regarding the values of the set of parameters $\{\beta,\lambda_0,\theta\}$ is provided by Bayesian decision theory. In this context, an appropriate estimate of the hazard should take into account the consequences of over or underpredicting its true value. Fawcett and Green (2018) discusses this when estimating return period levels for environmental extreme events, and they suggest the use of the predictive posterior return level as a point estimate that reliably incorporates epistemic uncertainty.

The posterior predictive estimate is obtained by averaging the conditional hazard of Eq. 7 over all possible values of the parameters $\{\beta,\lambda_0,\theta\}$ weighted by their posterior distribution $p\left(\Theta|\text{data}\right)$ as per Eq. 8.

$$\lambda\left(y|data\right)=\int_{\Theta}\lambda\left(y|\lambda_0,\beta,\theta\right)p\left(\lambda_0,\beta,\theta|\text{data}\right)d\lambda_0 d\beta d\theta \tag{8}$$

In the Bayesian framework, the posterior distribution of the parameters $p\left(\lambda_0,\beta,\theta|\text{data}\right)$, is the probability conditional on the available data and modeller's prior knowledge obtained by means of Bayes' Theorem (Gelman et al., 2013). The posterior distribution is proportional to the probability of observing the data given a set of parameters, also known as 'likelihood function', multiplied by the probability of a given set of parameters before incorporating the data, also known as 'prior distribution' (see Eq. 9). This can be colloquially described as the modeller's knowledge (i.e. prior distribution) times the information contained in the observations (i.e. the likelihood function).

$$p\left(\lambda_0,\beta,\theta|\text{data}\right)\propto p\left(\text{data}|\lambda_0,\beta,\theta\right)p\left(\lambda_0,\beta,\theta\right) \tag{9}$$

Equation 8 allows us to compute the hazard curve of water depths (or the IM chosen for the analysis) for any given point in space. Or inversely, compute the water depth for a given rate or return period. Hence, we can develop the T-years flood hazard map by marginally computing the y level for that return period from the expression.

In the following sub-sections we describe the methodology to estimate the posterior probability distributions of the parameters from Eq. 9, and the simulation methodology to estimate the hazard as per Eqs. 7 and 8.

2.2.1 A Bayesian recurrence model

In the framework described above, the recurrence model aims to quantify the mean rate of occurrence λ_0 of events, and characterize the distribution of the magnitude X and its parameters θ for any given event. These can be obtained from the statistical analysis of time series of past events. Poisson Point Process theory, as a generalization of the well-known Peaks-Over-Threshold (POT), methodology provides a robust framework for this (Bezak et al., 2014). Extreme events are individualized from historical records of daily discharges, by selecting an appropriate minimum threshold u and time separation to ensure independence. This results in a dataset of observed independent times between events $\hat{T}$ and a dataset of observed event's magnitudes $\hat{x}$. A Generalized Pareto Distribution (GPD) is then used as the probability model for the exceedances above the threshold threshold $x-u$, and the inter-arrival time between events is assumed to be exponentially distributed. This is,

$$\begin{aligned} T &\sim Exp\,(t|\lambda_0) & (10)\\ X-u &\sim \mathcal{GPD}\,(x|\xi,\sigma) & (11)\end{aligned}$$

Where λ_0 is the mean rate of arrival and ξ and σ are the shape and scale parameters that define the probability distribution, and u an appropriately defined threshold.

Hence, two probability models are required to describe the occurrence of events. An exponential distribution model of parameter λ_0 for the time between arrivals, and a GPD of parameters $\theta = \{\xi, \sigma\}$ for the magnitude of each event. Bayesian statistics provide an ideal framework to compute uncertainty in model's parameters that are consistent with the modeller's prior knowledge and proposed model (Bousquet & Bernardara, 2021).

For the dataset of observed interarrival times $\hat{T}$, the likelihood function is simply a product of n exponential densities, where n is the number of observations. The posterior distribution of the mean rate λ_0 can be obtained by assuming a weakly informative $Gamma\,(1/2, 1/2)$ prior distribution. A weakly informative prior distribution is relatively flat in the entire range of plausible values for the parameter. For this choice of prior, the posterior distribution has a closed form solution as given in Eq. 12 since it is a conjugate pair for the Exponential likelihood (Gelman et al., 2013).

$$p\left(\lambda_0|\hat{T}\right) = Gamma\,(n+1/2, t+1/2) \tag{12}$$

Where n is the number of events and $t = \sum_{\forall i} \hat{T}_i$ is the total number of years in the series.

The most likely value of the rate of occurrence given the observations is given by the mode of the posterior probability of Eq. 12, also known as the Maximum A-Posteriori (MAP) estimate. The MAP estimate and the mean value of the posterior distribution for λ_0 are given by Eqs. 13 and 14 respectively.

$$\begin{aligned} \lambda_0^* &= (n-1/2)\,/\,(t+1/2) & (13)\\ \overline{\lambda_0} &= n/t & (14)\end{aligned}$$

For the dataset of observed event's magnitudes $\hat{x}$, the likelihood function is given by a product of n GPD densities. There is no conjugate model for this likelihood, but a non-informative prior can built for the shape and scale parameters following Castellanos and Cabras (2007) (Eq. 15).

$$p\,(\xi,\sigma) \propto \sigma^{-1}\,(1+\xi)^{-1}\,(1+2\xi)^{-1/2} \tag{15}$$

Valid for $\xi > -0.5$ and $\sigma > 0$.

Then, the un-normalized expression for the posterior can be obtained by Bayes theorem as per Eq. 16, and the predictive posterior distribution for X can be subsequently computed as per Eq. 17. In both cases, there is no analytic solution, and samples from the distribution can be done via standard Markov Chain-Monte Carlo (MCMC) methods (Gelman et al., 2013). These can also be used to compute posterior mean and MAP estimates for ξ^*, σ^*.

$$p\,(\xi,\sigma|\hat{\mathbf{x}}) \propto p\,(\xi,\sigma) \prod_{i=1}^{n} (1+\hat{\mathbf{x}}_i\xi/\sigma)^{-(1+\xi)/\xi} \tag{16}$$

$$p\,(x-u|\hat{\mathbf{x}}) = \int \mathcal{GPD}\,(x-u|\xi,\sigma)\,p\,(\xi,\sigma|\hat{\mathbf{x}})\,d\xi d\sigma \tag{17}$$

2.2.2 Probabilistic inundation model

The inundation model is, in this context, a computational solver of some simplified version of the fluid dynamics equations that depend on variable inputs X and calibration parameters β. Epistemic uncertainties might come from lack of sufficient information to calibrate the parameters, observation errors, mechanistic simplifying assumptions, and numerical simplification of the equations solver (Kennedy & O'Hagan, 2001). We assume here, due to simplicity, that for a given simulator $S(X, \beta)$ these can be represented by uncertainty in the model's calibration parameters β. More complex procedures to include model uncertainty can be used to include uncertainty in the model structure. For example, a formally probabilistic calibration procedure that includes model structural uncertainty as an additive Gaussian Process is discussed in Balbi and Lallemant (2023).

Parameter's distributions can be obtained using nominal probability models from expert's knowledge (Kalyanapu et al., 2012; Stephens & Bledsoe, 2020) or statistically calibrated ones generally via the GLUE methodology (Di Baldassarre et al., 2010; G. T. Aronica et al., 2012; Kiczko et al., 2013; Romanowicz & Kiczko, 2016; Zahmatkesh et al., 2021). In this work, epistemic uncertainty will be represented by probability distributions in the roughness parameters only, for the floodplain and for the channel, considering all other inputs as constant regarding the calibration procedure. These distributions will be obtained by means of the GLUE framework, where all possible sets of parameters are assigned a normalized score (i.e. pseudo-likelihood) from an appropriately selected scoring rule. In the case of flood extent binary observations (as in the case study developed in this work), it is typical to use the F-score (as per Eq. 18), a variant of the classical Jaccard Index (G. T. Aronica et al., 2012; Papaioannou et al., 2017).

$$F(\boldsymbol{\beta}) = \frac{A - B}{A + B + C} \tag{18}$$

Where A is the number of correctly predicted pixels, B the number of over-predicted pixels (predicted flooded, observed non-flooded), and C is the number of under-predicted pixels (predicted non-flooded, observed flooded).

The details of the calibration procedure using the GLUE framework can be found in Balbi and Lallemant (2023), but they can be summarized in four steps:

1. Sample a large set of β from their prior distribution
2. Compute the F-score for the sampled β
3. Reject all 'non-behavioral' models using some thresholding criteria: $F < f^*$
4. Standardize the resulting F-scores so that they are all positive and integrate to 1

The model fit for each value of β is, then, a measure of its uncertainty, or pseudo-posterior probability (as it is not strictly obtained from a probabilistic likelihood). In addition, the MAP value for parameter β is the one that yields the best fit (i.e. largest F-score).

2.3 Numerical implementation

Computing the integral from Eq. 8 requires numerical procedures since no analytic solution exist for the posterior distributions of the parameters just described. Since many parameters are involved in its computation, a summary of variables and symbols can be found in Table 1.

Equation 8 can be slightly simplified, however, by noting that $\lambda(y|\lambda_0, \beta, \theta)$ is linear on λ_0 and its posterior distribution (Eq. 12) is independent from the rest of the pa-

Table 1. Summary of variables and symbols

Variable	Description
X	Flood event's magnitude (i.e. river peak discharge)
Y	Flood IM (i.e. flood depth) at a given point in space
λ_0	Mean rate of occurrence of events $X \geq 0$
$\lambda(y)$	Mean rate of occurrence of events $Y \geq y$
β	Inundation simulator calibration parameters
$S(X, \beta)$	Inundation simulator
ξ	$\mathcal{GPD}$ shape parameter
σ	$\mathcal{GPD}$ scale parameter
u	Threshold value for X POT model
$\overline{(.)}$	Mean value of parameter
$(.)^*$	MAP value of parameter (mode of its posterior distribution)

rameters as given by Eq. 19.

$$\lambda(y|data) = \overline{\lambda_0} \int_{\beta,\xi,\sigma} p(Y \geq y|\beta,\xi,\sigma)\, p(\beta,\xi,\sigma|\text{data})\, d\beta d\xi d\sigma \tag{19}$$

Where $\overline{\lambda_0}$ is the mean of the posterior distribution of λ_0 given by Eq. 14.

Standard Monte-Carlo (MC) integration techniques can be employed to compute such integral. Conceptually the task is straightforward: we need to sample from the posterior distribution of Y, by sampling first from the posterior of X. This can be done through the following steps:

1. Sample N values from the posterior distribution of β (as per Sect. 2.2.2) and $\{\xi, \sigma\}$ (from Eq. 16)
2. For each sample of $\{\xi_i, \sigma_i\}$, sample x_i from Eq. 17
3. For each sample x_i and β_i, compute water depth at all points of interest from the simulator $y_i = S(x_i, \beta_i)$
4. Estimate the mean rate of exceedance of y as:

$$\lambda(y|\text{data}) \approx \frac{\overline{\lambda_0}}{N} \sum_i \mathbf{1}\{S(x_i, \beta_i) \geq y\}$$

The number of samples N required for the simulation depends on the percentile of the curve (i.e. return period) we are trying to estimate and the precision desired. For example, to estimate the 100 years return level y_{100} we need to estimate an exceedance probability $p(Y \geq y_{100}) = (100\lambda_0)^{-1}$. According to the standard theory of empirical estimates of probabilities, based on the Central Limit Theorem (CLT), we can obtain an approximate minimum number of simulations for a 95% confidence interval as per,

$$N > \frac{1.96}{\varepsilon^2} \sqrt{\frac{1-p}{p}} \tag{20}$$

Where p is the actual probability being estimated (not exactly known) and ε is the width of the relative interval (Melchers & Beck, 2018).

On the other hand, the computation of the hazard in the classical approach (as given by Eq. 5), where no epistemic uncertainties are considered, is much simpler:

1. Obtain the return levels of the GPD distribution for each return period Tr of interest , and for fixed parameters ξ^* and σ^* as given by (see Appendix A for de-

tails),

$$x_{Tr} = u + \begin{cases} \frac{\sigma^*}{\xi^*}\left\{(\lambda_0^* \cdot Tr)^{\xi^*} - 1\right\}, & \text{if } \xi^* \neq 0 \\ \sigma^* \ln\left(\lambda_0^* \cdot Tr\right), & \text{if } \xi^* = 0 \end{cases}$$

2. For each Tr, compute water depth at all points of interest for fixed parameters β^* from the simulator $y_{Tr} = S\left(x_{Tr}, \beta^*\right)$

That is, we compute x for different return periods of interest (also known as return levels), and then evaluate the inundation model at each. It is important to note, that this two-step approach can be followed but for an entire ensemble of posterior realizations of parameters $\{\beta_i, \xi_i, \sigma_i\}$, to obtain an estimate of the predictive posterior estimate including epistemic uncertainties. This method requires many times more the number of calls to the inundation simulator S relative to the four-step procedure described above, but has the advantage of obtaining credible intervals for the estimate.

3 Case study

The proposed methodology described in Sect. 2 is applied here in a real-world case study, with the purpose of analyzing the influence of the inclusion of epistemic uncertainties in the recurrence model of river discharges (as per Sect. 2.2.1) and in the inundation model (as per Sect. 2.2.2).

3.1 Models and data

The case study is based on a short reach on the upper river Thames in Oxfordshire, England, just downstream from a gauged weir at Buscot (Fig. 2). The river at this reach has an estimated bankfull discharge of 40 m/s^3 and drains a catchment of approximately 1000 km^2. The topography was obtained from stereophotogrammetry at a 50 m scale with a vertical accuracy of ±25 cm, obtained from large-scale UK Environment Agency maps and surveys. This reach has also been study previously in G. Aronica et al. (2002), J. W. Hall et al. (2011) and Balbi and Lallemant (2023)

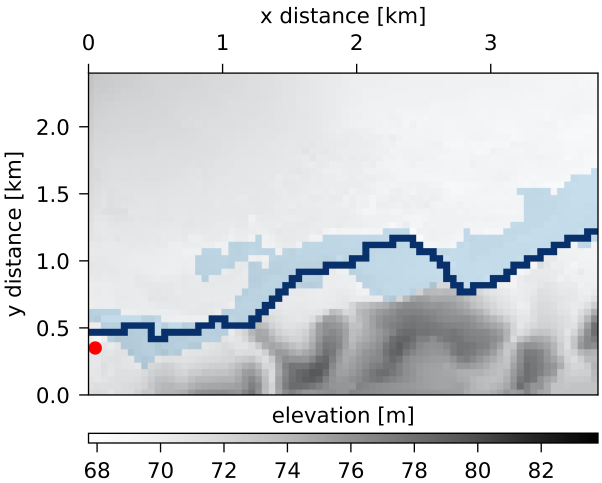

Figure 2. Floodplain topography at Buscot, SAR imagery of 1992 flood event (light blue), channel layout (dark blue) and gauge station location (red dot).

The events are characterized by the river discharge flow only. To develop the recurrence model for events, a publicly available daily discharge data series at Buscot weir was obtained from the UK National River Flow Archive (see Fig. 3). The series spans from 19 years from 1980 to 1998 with some minor gaps that are not expected to affect the extreme statistics analysis to perform.

On the other hand, for the calibration of the inundation model, a satellite observation of the flood extent of 1-in-5 year event occurred during December 1992 was used (see Fig. 2). The satellite SAR(synthetic aperture radar) image of the flood was captured 20 hours after the flood peak when discharge was at a level of $73m^3/s$ (G. Aronica et al., 2002). The resolution of the image is $50m$.

The computational inundation model used is the raster-based Lisflood-fp model (Neal et al., 2012). Lisflood-fp couples a 2D water flow model for the floodplain and a 1D solver for the channel flow dynamics. Its numerical structure makes it computationally efficient and suitable for the many simulations needed for probabilistic flood risk analysis and model calibration.

A simplified rectangular cross-section is used for the channel with a constant width of 20m for the entire reach and a varying height of around 2m. The observed event is defined by the boundary condition of a fixed input discharge of $x = 73m/s^3$ at the geographic location of the gauging station shown in Fig. 2, and by an assumed downstream boundary condition of a fixed water level of approximately 90cm above the channel bed height. The short length of the reach and the broadness of the hydrograph imply that a steady-state hydraulic model is sufficiently accurate for the calibration (G. Aronica et al., 2002).

The model's parameters used for calibration are the Manning's roughness parameters for the channel r_{ch} and for the floodplain r_{fp}, both considered spatially uniform in the domain of analysis. That is, $\beta = \{r_{ch}, r_{fp}\}$. For the calibration method described in Sect. 2.2.2, the inundation model was ran for a fixed observed discharge of $73m^3/s$ and for a uniform prior for both parameters in the range $0.01 - 0.15$.

3.2 Computational implementation

The statistical models and simulation method described in Sect. 2.3 were implemented in Python 3.X language (Van Rossum & Drake, 2009), using a 10-core Intel i9-10700k processor computer. Each evaluation of the inundation model $S(X, \beta)$ takes approximately 4s. For the calibration of the inundation model 19,600 evaluations of the simulator were needed to cover the entire grid of β values, and around 100,000 evaluations were needed for the hazard simulation procedure from Sect. 2.3.

4 Results

4.1 Discharge recurrence model

To define the events and their magnitudes, a threshold of $u = 12m^3/s$ and minimum distance between clusters of 7 days (i.e. there has to be 7 days of values below the threshold for two events to be considered as separate events) were selected aiming to satisfy the conditions required by POT standard theory (Bousquet & Bernardara, 2021):

- The minimum threshold for which the modified scale and shape parameters of the fitted GPD of the exceedances remain constant for higher thresholds.
- The resultant threshold exceedances (cluster's peaks) should form an independent sample.

A total of 73 clusters were identified in $18.8yrs$ of data as shown in Fig. 3. That is, on average, 3.9 events per year, and it shows the relative advantage of this type of analysis versus the standard annual maximum approach for which there would only be 18 data points. The posterior distribution of the mean rate λ_0, as given by Eq. 12, has a mode (i.e. MAP) $\lambda_0^* = 3.8yrs^{-1}$ and a mean value $\overline{\lambda_0} = 3.9yrs^{-1}$. Statistical graphical tests, shown in Fig. 4, showed that the resulting series of extreme discharges can be considered independent and that the time between events has a good fit to the exponential model as assumed in the HPP model.

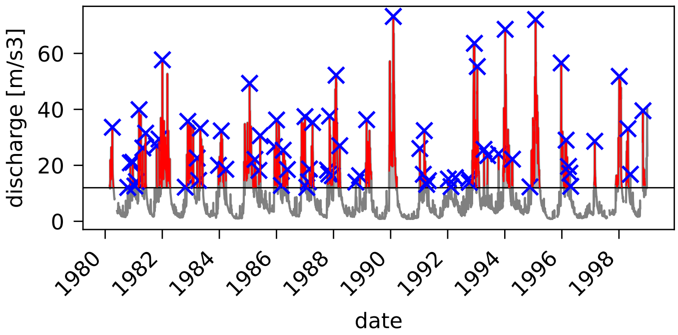

Figure 3. Daily discharge for Buscot and identification of flood events by clustering with a 12 m/s^3 threshold and 7 days of minimum return. Blue cross indicates event's peak discharge.

Samples of the posterior distribution of the GPD parameters (Eq. 16) were drawn by a standard MCMC algorithm of 4 chains of 15,000 samples each. Discarding the first half of sample from each as burn-in stage, convergence of the chains was assessed by verifying that Gelman-Rubin R-scores remains below 1.01 (Gelman et al., 2013). Goodness-of-fit tests showed a good agreement of the exceedances with the GPD. The shape parameter ξ is centered around -0.05 while the scale parameter σ is centered around 16.5, both with a relatively small skewness (see Fig. 5). The MAP values for the parameters practically coincide with these values.

For each posterior sample of ξ and σ, the probability distribution of the discharges follows a GPD. With the ensemble of distributions for each sample, we computed the mean curve (i.e. the predictive posterior distribution of X) and the 90% confidence posterior intervals. These are shown as return period curves in Fig. 6 as computed by,

$$Tr\,(x) = 1/\lambda_0^* p\,(X \geq x) \tag{21}$$

A 'deterministic' hazard curve was also computed using the MAP values for the GPD parameters, following the classical approach. It can be seen that epistemic uncertainties have the effect of increasing the discharges for a given return period, and that this effect increases with increasing return period. This is intuitive, as larger return periods are more uncertain with limited-length data. Similar results have been obtained before (Merz & Thieken, 2005; Romanowicz & Kiczko, 2016; Fawcett & Green, 2018).

4.2 Inundation model

The statistical calibration of the inundation model was done by using a uniform grid for both parameters in the range $(0.01, 0.15)$ with a step of 0.01, and a threshold

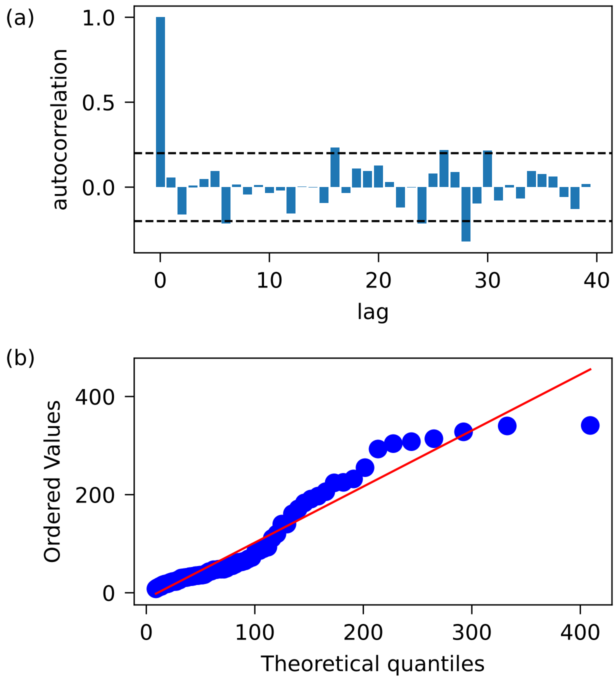

Figure 4. (a) Autocorrelation plot of discharge series; (b) Probability plot of interarrival times compared to the exponential distribution.

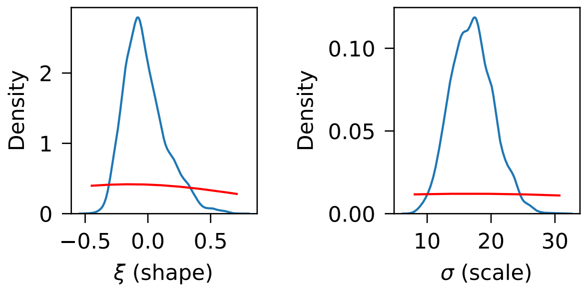

Figure 5. Posterior (blue) and prior (red) distributions of the parameters of the frequency model (blue)

of 0.5 to filter out non-behavioral models. This resulted in a total of 543 accepted simulations out of 19,600. The bivariate pseudo-posterior distribution for the roughness parameters is shown in Fig. 7. The set of parameters that yields the maximum F-score (i.e. MAP parameters) were $r_{ch} = 0.029, r_{fp} = 0.045$ giving $F = 0.54$.

4.3 Flood hazard

Close to 15,000 thousand posterior samples of water depth Y at all points in the region were obtained, following the simulation procedure described in Sect. 2.3. This allowed to empirically estimate the posterior exceedance recurrence $\lambda(Y \geq y|\text{data})$ for ev-

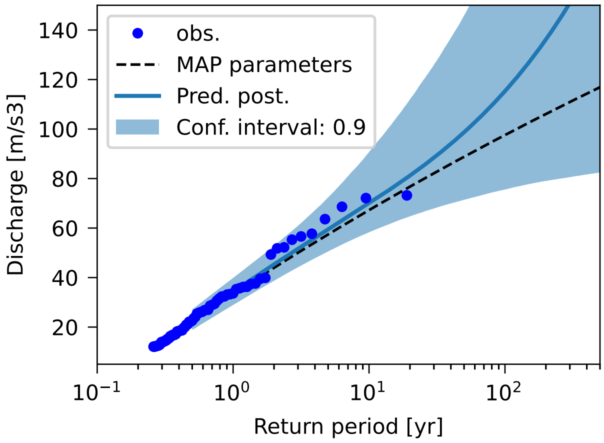

Figure 6. Hazard curves for river discharge

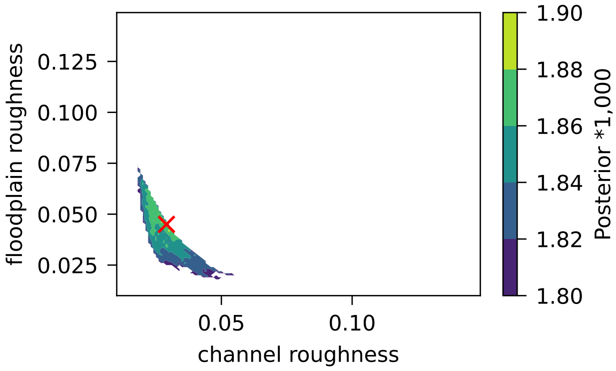

Figure 7. Bivariate posterior distribution for inundation model parameters

ery pixel and, consequently, the hazard curve. The number of simulations ran implies that exceedance probabilities as small as 10^{-4} can be estimated with a 10% interval according to Eq. 20. For a mean value $\overline{\lambda_0} = 3.8yrs^{-1}$, this is equivalent to RPs of up to $1,500yrs$.

Figure 8 shows the flood depth hazard curves for different points in the floodplain. The posterior predictive curves are compared with the classical approach that uses the deterministic discharge hazard curve (dotted black curve in Fig. 6) and the MAP parameters for the inundation simulator. In every case, it can be seen that the flood depth values of the posterior predictive curves increase faster than the classical approach as the return period grows, in a similar fashion observed in the discharge hazard model of Fig. 6. Furthermore, the two curves are very similar for lower return periods.

The 100yrs flood hazard map was developed by computing the 100yrs flood depth from the posterior predictive curves at each point (see Fig. 9). This map is compared in plot (a) of Fig. 10 with the traditional hazard map computed using the best inundation model with the deterministic estimate of the 100yrs discharge. The increasing flood depth for the posterior predictive map can be seen to be replicated for every pixel in the region of analysis with the exception of some isolated pixels right next to the channel.

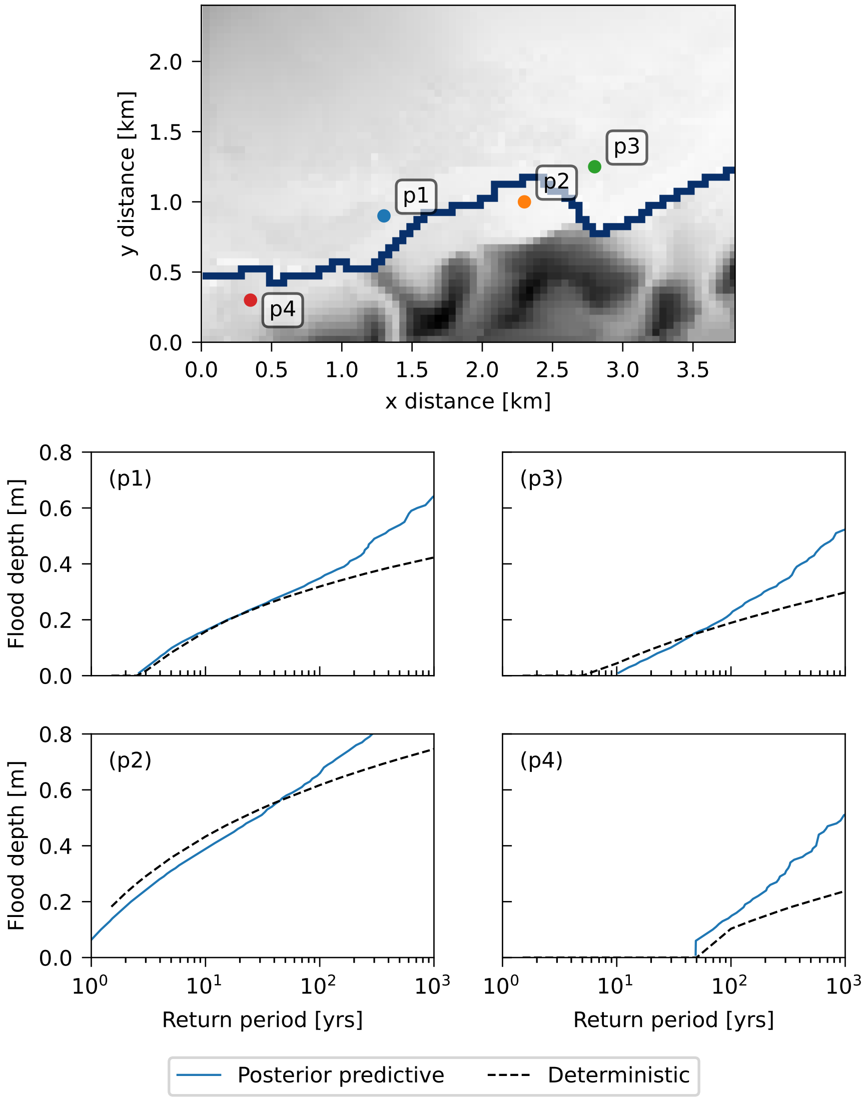

Figure 8. Hazard curves for flood depths at different locations in the floodplain

This effect, as in the discharge hazard curve, is exacerbated with increasing return periods as can be seen in plot (b) and (c) of the figure for the 250yrs and 500yrs maps comparison.

5 Discussion

5.1 On the influence of epistemic uncertainties in hazard estimates

From an engineering design perspective, the water height used to design a structure for a specified safety level (i.e. return period) will be larger when including epistemic uncertainties. Results show that this is true for every pixel in the region of analysis as can be seen from the maps in Fig. 10 and the curves in Fig. 11 for a specific pixel. That is, when including our lack of knowledge and information on the process, we need to be more conservative in design to ensure an appropriate level of reliability. Furthermore, this trend increases with the return period, and is practically negligible for more

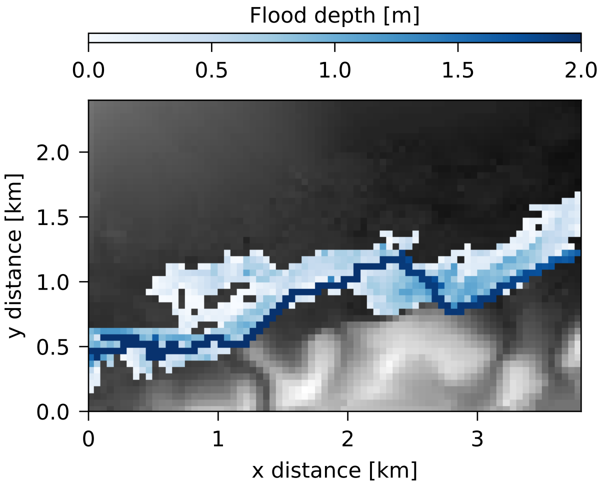

Figure 9. 100yrs flood hazard map using posterior predictive flood depth at each point

recurrent events. This result is somewhat intuitive as we usually have less knowledge on rare events, and similar conclusions have been obtained by other researchers in hydrological hazards (Merz & Thieken, 2005; Fawcett & Green, 2018).

To further understand and generalize this result, however, we need to understand the relative influence of epistemic uncertainties on the discharge recurrence model and on the inundation model. To do this, we obtained the hazard curves while including epistemic uncertainties one model at a time, as shown in Fig. 11. It can be seen that the effect of more conservative water depths for any given return period is entirely due to epistemic uncertainties in the recurrence model. This trend reflects the heavier tails of the posterior distribution of the discharges (as also seen in Fig. 6) that mainly represents uncertainty due to the limited-length observed time series used to build the model. That is, using an 18-years data record, there are practically no observations of much higher return periods which is reflected in the larger uncertainty.

Uncertainty on the inundation model parameters, on the other hand, seems to have a relatively constant decreasing effect over return periods. That is, it gives lower (i.e. less conservative) water depths for a given recurrence relative to the classical approach. This is related to the shape of the posterior distribution of the parameters, but also on the non-linear nature of the S transformation. Thus, the influence of epistemic uncertainty in the inundation model parameters is related to the type of observations used for calibration (i.e. binary flood extent observations in this case), the statistical procedure used (i.e. F-score pseudo-likelihood), and also the non-linearity of the inundation simulator itself. Given that water levels are relatively constrained by the topography, it is not expected that the inundation simulator presents a radically high non-linearity. Thus, the highly skewed shape of the roughness parameters posterior distribution (Fig. 7), obtained from the GLUE method, with respect to the MAP values might be the main driver responsible for the underestimation in flood hazard. However, further analysis is required in order to deeply understand how this influence varies in different settings (i.e. different observations, different calibration methods and inundation models), and understand if this effect can be magnified in some contexts.

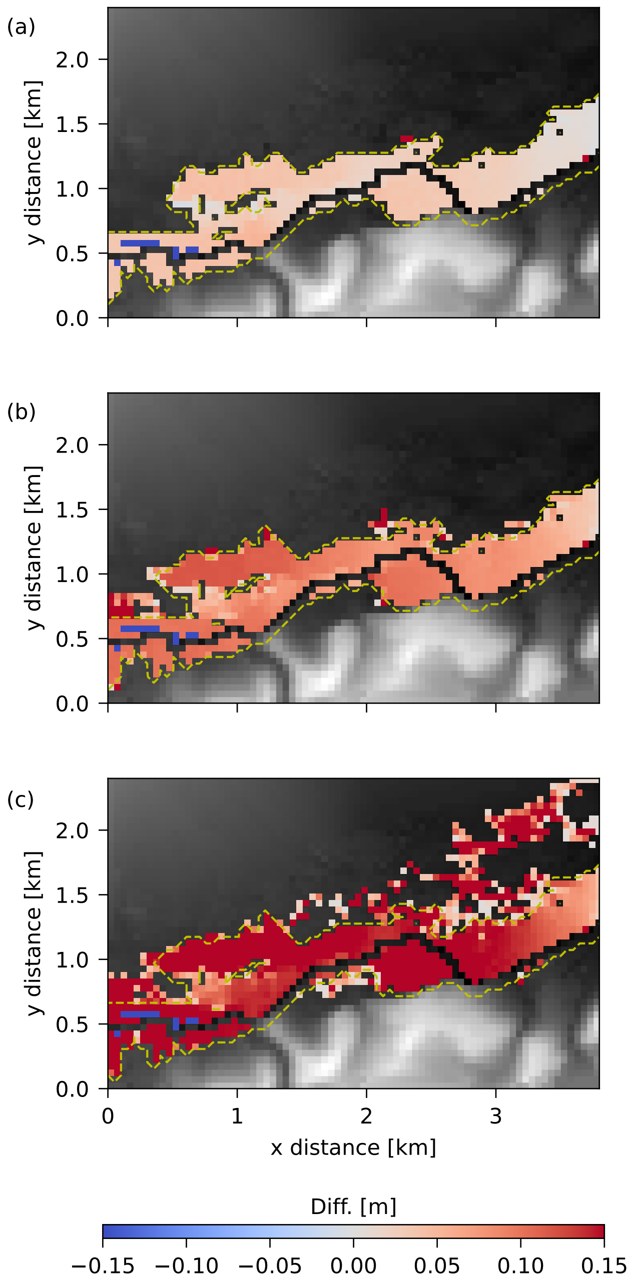

Figure 10. Difference maps between posterior predictive estimates and deterministic estimates for (a) 100yrs, (b) 250yrs, (c) 500yrs

5.2 On the usefulness of the output

From a risk analysis standpoint, we might be interested in computing some damage measure, or any higher-level decision metric, that reflects the impact of the flood in human communities. The framework is analogous to the one described here, but instead we are interested in the distribution of the decision metric Z over all the potential hazard events. We can straightforwardly compute this from the probability distribution of

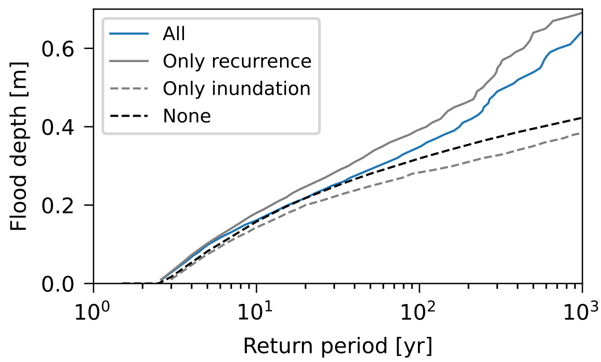

Figure 11. Hazard curves for water depth at location $x = 1.30km, y = 1.05km$

the IM, as in Eq. 22, since the vulnerability model is generally dependent on the y level.

$$\lambda(z) \approx \int \underbrace{p\left(Z \geq z|y\right)}_{\substack{\text{Vulnerability}\\ \text{model}}} \underbrace{\lambda_0 p\left(y\right)}_{\substack{\text{Hazard}\\ \text{model}}} dy \tag{22}$$

Equation 22 shows that we actually need the recurrence of water depth y (and eventually other IMs like duration, or velocity) at any site of interest in order to compute the risk. For most modern vulnerability models then, a probability of flood map for a given return period is not useful since it does not provide the required information. The hazard maps, as developed in this work, provide a reliable estimate of the recurrence $\lambda_0 p\left(y\right)$ while also accounting for epistemic uncertainties. Specifically the maps show the water depths for a given return period that can be transformed into an exceedance probability as per Eq. 2.

It is important to note that the maps reflect marginal probabilities and do not take into account spatial correlation in the flood process, as they are built by individually computing the hazard curves at each point. In other words, the resulting hazard maps (as in Fig. 9) do not show a real flood event. For this reason, the hazard maps are useful for site-specific hazard, and eventually risk analysis, but not for analyzing spatially-distributed assets. However, these maps were built from an ensemble of simulated flood maps as per the MC simulation approach described in Sect. 2.3. This ensemble is a direct output of the inundation simulator, and thus, can be used to estimate damage for spatially distributed exposure while propagating epistemic uncertainties in the hazard model.

5.3 On the separation of aleatory and epistemic uncertainties

The posterior predictive estimate for the hazard proposed in this work combines in one metric both the aleatory and epistemic uncertainty. While this is important for risk-based decision making, it does not directly help exploiting the main distinction between the two: that epistemic uncertainties can potentially be reduced by further collecting data or improving our models.

Qualitatively speaking, the departure of the posterior predictive curves from the deterministic hazard curve (Fig. 11) can be interpreted as a measure of the relevance of epistemic uncertainties on top of aleatory uncertainties. In this sense, it can be noted that for the case study developed, epistemic uncertainty on the future occurrence of discharge events seems to be more impactful than uncertainty on the inundation model parameters; at least for return periods of 100 years and above.

Developing uncertainty bounds for the hazard curve estimates is the typical way of showing the sensitivity of the hazard estimates on the model's parameters, and their overall relevance. This was done for the hazard curve of discharges in Fig. 6. Doing the same for the flood depths would require to compute a hazard curve for each posterior sample of parameters with the consequent added computational demand. An ensemble of 100 hazard curves was computed and is shown in Fig. 12 together with the posterior predictive (i.e. the mean curve of the ensemble) and the deterministic hazard curve.

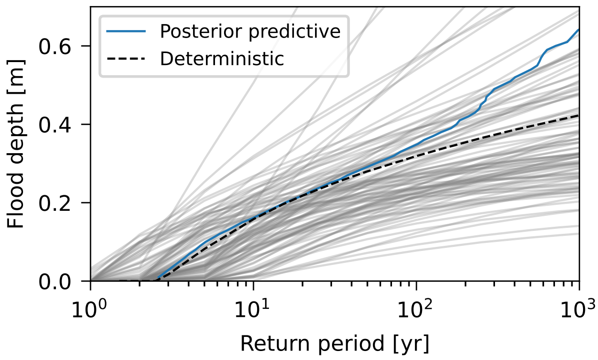

Figure 12. Ensemble of posterior samples of the hazard curves for water depth at location $x = 1.30km, y = 1.05km$

Figure 12 shows that uncertainty in the model's parameters can yield hazard curves that are vastly different from one another, resulting in large uncertainty bounds. This type of analysis can encourage modellers to obtain more data and/or refine the knowledge of the models used in order to reduce it.

5.4 On the modelling methodology

The inclusion of epistemic uncertainties in the computation of flood hazard result in a non-uniqueness of the parameters (and models eventually) used to compute the hazard curves. Thus, there is not a single discharge for a given return period, and there is not a single flood map for any given input discharge. As a result, there is no direct translation between the discharge for a given return period and its corresponding flood map. The water depth for a given return period, as per Eq. 19, has now contributions from all possible discharges which is a more accurate representation of knowledge (and uncertainty).

There are many methodologies to include this uncertainty in the modelling process, and we have chosen to rely on rigorous probabilistic modelling based on a Bayesian framework. The Bayesian methodology allows to consistently include modeller's knowledge and data from different sources into posterior estimates of probabilities. In particular, this work has limited the epistemic uncertainties to uncertainty over the parameters of some models appropriately chosen (i.e. the GPD model for the discharges and the Lisflood simulator for the inundation model), but further prior distributions can be set over different models without affecting the workflow of the method proposed.

Finally, other deeper sources of uncertainty cannot be discarded in risk analysis and include what Spiegelhalter and Riesch (2011) named 'indeterminacy' and 'ignorance'. The former are associated with known limitations in understanding and modelling ability, while the latter is associated with unknown limitations of understanding. Different approaches have been proposed to deal with epistemic uncertainties that are not fully quantifiable including these deeper sources of uncertainties; these were not treated in this work and the reader is referred to (Spiegelhalter & Riesch, 2011; Goldstein, 2011; Beven, 2014).

5.5 On the computational challenges

From a computational perspective, the inclusion of epistemic uncertainties via the Bayesian posterior predictive distribution of IMs means that we need to perform lots of simulations of the inundation model in order to compute the hazard curves. This contrasts with the traditional approach where we only need to run the inundation model once to obtain the 100yrs flood hazard (see Sect. 2.3). Furthermore, the number of simulations needed grow with the return period according to Eq. 20.

This can result in an unfeasible computational burden, particularly for very large return periods. Ongoing advances in computer technology, on the other hand, are producing faster computers that will make simulation approaches like the one proposed, easier to deal as time progresses. In this context, it is expected that this type of numerical analysis will become more common in future research.

In any case, it is also important the implementation of efficient simulation techniques in order to reduce the computational cost. There are mainly two families of techniques that can be implemented in order to reduce the computational time required to compute the desired probabilities: using a more efficient simulation algorithm that targets the desired return period faster, such as importance sampling (Zio, 2013); and making each run of the inundation model faster by using a statistical emulator (Jiang et al., 2020).

6 Conclusions

We propose and develop in this work, a methodology to compute flood hazard curves and maps including epistemic uncertainty on the model's parameters. The framework aims, not only to consistently include this uncertainty into robust hazard estimates, but also to produce useful output for risk analysis and engineering decision-making.

Rather than computing probabilities for an uncertain T-yrs peak discharge event, we propose to compute the flood depth distributions from all possible events within a given time-frame. The posterior distribution of the models' parameters were computed to include epistemic uncertainty, and the average recurrence-rate over these distributions (i.e. predictive posterior distribution) was used as a point estimate for the hazard. As a result, the flood hazard maps developed provide information related to the water-depths for different recurrence rates (or return periods) that can be readily used for further damage analysis including the forward propagation of epistemic uncertainties. The mathematical notation was kept intentionally generic to encourage its use in other natural hazards applications.

The results of the real-world case study developed, showed significant differences between the hazard estimates considering epistemic uncertainties and the classical ones without. In particular, the results show that not considering epistemic uncertainties might underestimate the water-depths hazard at any point in the region of analysis, resulting in less reliable decisions and designs. Furthermore, this tendency aggravates with larger return periods, that tend to be the main focus in many risk analysis. A deeper analysis shows that uncertainty in the prediction of future discharges as a product of a short-length observation series, is the main driver responsible for the underestimation in flood hazard. A similar pattern can be observed in the recurrence curves for the peak discharges in line with some results from the literature (Merz & Thieken, 2005; Romanowicz & Kiczko, 2016; Fawcett & Green, 2018).

On the other hand, the influence of the uncertainty in the inundation model parameters (i.e. floodplain and channel roughness) proved to be less significant. Results show that this uncertainty seems to have a constant effect over return periods, towards the conservative side; that is, it yields lower hazard values than using the classical approach. This seemingly non-intuitive influence requires further studies to understand how it generalizes to different applications and contexts (i.e. different calibration approaches, different computational simulators, different available data), although a potential explanation is the highly skewed shape of the posterior distribution (see Fig. 7) relative to the MAP parameters used in the classical approach.

On the computational aspects, the proposed numerical methodology requires the simulation of thousands of inundation maps. This ensemble can be straightforwardly used in spatially distributed vulnerability models covering a continuous range of return periods. It comes at the expense, however, of a much larger computational burden than the classical approach where only a few runs for selected return periods are needed. Technological advances are expected to rapidly reduce the development-time of this simulation approach, particularly in applications that are practically feasible today. Additionally, this should also encourage researchers to find ways of optimizing the computation of the estimates by using more efficient sampling algorithms or cheaper emulators of the inundation model.

To summarize, we have shown that the inclusion of epistemic uncertainties can significantly modify the estimates of hazard estimates that are later used for risk assessments and damage mitigation decision-making. This is particularly exacerbated for rare and big events (i.e. longer return periods). A framework that can consistently include these in robust probabilistic outputs is, we believe, therefore a major advance for risk analysis.

Appendix A The GPD distribution

This appendix summarizes the probability functions for the Generalized Pareto Distribution (GPD), to avoid confusion in the definition and meaning of the parameters used.

$$p\left(x|\xi,\sigma\right) = \begin{cases} \left(1+\frac{\xi}{\sigma}y\right)^{-\frac{\xi+1}{\xi}}, & \text{if } \xi \neq 0 \\ \exp\left(-\frac{y}{\sigma}\right), & \text{if } \xi = 0 \end{cases} \tag{A1}$$

$$F\left(x|\xi,\sigma\right) = \begin{cases} 1-\left(1+\frac{\xi}{\sigma}y\right)^{-1/\xi}, & \text{if } \xi \neq 0 \\ 1-\exp\left(-\frac{y}{\sigma}\right), & \text{if } \xi = 0 \end{cases} \tag{A2}$$

Where the support of X is $x \geq 0$ when $\xi \geq 0$, and $0 \leq x \leq -\sigma/\xi$ when $\xi < 0$.

Return levels x_{Tr} for this distribution can be computed from,

$$Tr = \frac{1}{1-F\left(x_{Tr}|\xi,\sigma\right)} \tag{A3}$$

Open Research Section

All the data used as input in this work (DEM, daily discharge series, flood extent observation) is publicly available and the sources were mentioned in the manuscript. The open-source software Lisflood-fp was used as inundation simulator (http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/). All the simulations and figures were performed Python 3.X (https://www.python.org/) scripts developed by the authors. All the code used to develop the results and figures mentioned in this manuscript, as well as the input data and the Lisflood-fp binaries used, are publicly available on GitHub https://github.com/mbalbi/epistemic_flood_hazard.git.

Acknowledgments

Jim W. Hall, from Oxford University, is thanked for freely providing the satellite observation for the 1992 flood event at the reach under study. Paul Bates and the LISFLOOD-FP development team are thanked for their support during the simulator learning curve and for the free access to their software, including manuals and tutorial examples, which this case study was drawn from. The research was partially funded by the School of Engineering of the University of Buenos Aires, Argentina, through a Peruilh doctoral scholarship. This project is supported by the National Research Foundation, Prime Minister's Office, Singapore under the NRF-NRFF2018-06 award.

References

Ahmadisharaf, E., Kalyanapu, A. J., & Bates, P. D. (2018, September). A probabilistic framework for floodplain mapping using hydrological modeling and unsteady hydraulic modeling. *Hydrological Sciences Journal*, *63*(12), 1759–1775. doi: 10.1080/02626667.2018.1525615

Apel, H., Merz, B., & Thieken, A. H. (2008, June). Quantification of uncertainties in flood risk assessments. *International Journal of River Basin Management*, *6*(2), 149–162. doi: 10.1080/15715124.2008.9635344

Aronica, G., Bates, P. D., & Horritt, M. S. (2002, July). Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes*, *16*(10), 2001–2016. doi: 10.1002/hyp.398

Aronica, G. T., Candela, A., Fabio, P., & Santoro, M. (2012, January). Estimation of flood inundation probabilities using global hazard indexes based on hydrodynamic variables. *Physics and Chemistry of the Earth, Parts A/B/C*, *42–44*, 119–129. doi: 10.1016/j.pce.2011.04.001

Baker, J., Bradley, B., & Stafford, P. (2021). *Seismic Hazard and Risk Analysis* (1st ed.). Cambridge University Press. Retrieved 2022-06-23, from https://www.cambridge.org/core/product/identifier/9781108348157/type/book doi: 10.1017/9781108425056

Balbi, M., & Lallemant, D. C. B. (2023, March). Bayesian calibration of a flood simulator using binary flood extent observations. *Hydrology and Earth System Sciences*, *27*(5), 1089–1108. doi: 10.5194/hess-27-1089-2023

Beven, K. (2014, March). A Framework for Uncertainty Analysis. In *Applied Uncertainty Analysis for Flood Risk Management* (pp. 39–59). IMPERIAL COLLEGE PRESS. doi: 10.1142/9781848162716_0003

Bezak, N., Brilly, M., & Šraj, M. (2014, May). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, *59*(5), 959–977. doi: 10.1080/02626667.2013.831174

Bharath, R., & Elshorbagy, A. (2018, November). Flood mapping under uncertainty: A case study in the Canadian prairies. *Natural Hazards*, *94*(2), 537–560. doi: 10.1007/s11069-018-3401-1

Bousquet, N., & Bernardara, P. (Eds.). (2021). *Extreme Value Theory with Applications to Natural Hazards: From Statistical Theory to Industrial Practice.* Cham: Springer International Publishing. doi: 10.1007/978-3-030-74942-2

Candela, A., & Aronica, G. T. (2017, March). Probabilistic Flood Hazard Mapping Using Bivariate Analysis Based on Copulas. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *3*(1). doi: 10.1061/AJRUA6.0000883

Castellanos, M. E., & Cabras, S. (2007, February). A default Bayesian procedure for the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, *137*(2), 473–483. doi: 10.1016/j.jspi.2006.01.006

Der Kiureghian, A., & Ditlevsen, O. (2009, March). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105–112. doi: 10.1016/j.strusafe.2008.06.020

Di Baldassarre, G., Schumann, G., Bates, P. D., Freer, J. E., & Beven, K. J. (2010). Flood-plain mapping: A critical discussion of deterministic and probabilistic approaches. *Hydrological Sciences Journal*, *55*(3), 364–376. doi: 10.1080/02626661003683389

Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., & Brath, A. (2013, August). Probabilistic flood hazard mapping: Effects of uncertain boundary conditions. *Hydrology and Earth System Sciences*, *17*(8), 3127–3140. doi: 10.5194/hess-17-3127-2013

Falter, D., Schröter, K., Dung, N. V., Vorogushyn, S., Kreibich, H., Hundecha, Y., ... Merz, B. (2015, May). Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. *Journal of Hydrology*, *524*, 182–193. doi: 10.1016/j.jhydrol.2015.02.021

Fawcett, L., & Green, A. C. (2018, August). Bayesian posterior predictive return levels for environmental extremes. *Stochastic Environmental Research and Risk Assessment*, *32*(8), 2233–2252. doi: 10.1007/s00477-018-1561-x

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition.* CRC Press.

Goldstein, M. (2011, October). External Bayesian Analysis for Computer Simulators. In J. M. Bernardo et al. (Eds.), *Bayesian Statistics 9* (pp. 201–228). Oxford University Press. doi: 10.1093/acprof:oso/9780199694587.003.0007

Grimaldi, S., Petroselli, A., Arcangeletti, E., & Nardi, F. (2013, April). Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling. *Journal of Hydrology*, *487*, 39–47. doi: 10.1016/j.jhydrol.2013.02.023

Hall, J., & Solomatine, D. (2008, June). A framework for uncertainty analysis in flood risk management decisions. *International Journal of River Basin Management*, *6*(2), 85–98. doi: 10.1080/15715124.2008.9635339

Hall, J. W., Manning, L. J., & Hankin, R. K. (2011). Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research*, *47*(5), 5529. doi: 10.1029/2009WR008541

Jiang, P., Zhou, Q., & Shao, X. (2020). *Surrogate Model-Based Engineering Design and Optimization.* Singapore: Springer Singapore. doi: 10.1007/978-981-15-0731-1

Kalyanapu, A., Judi, D., McPherson, T., & Burian, S. (2012, March). Monte Carlo-based flood modelling framework for estimating probability weighted flood risk: Monte Carlo-based flood modelling framework. *Journal of Flood Risk Management*, *5*(1), 37–48. doi: 10.1111/j.1753-318X.2011.01123.x

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(3), 425–464. doi: 10.1111/1467-9868.00294

Kiczko, A., Romanowicz, R. J., Osuch, M., & Karamuz, E. (2013, December). Maximising the usefulness of flood risk assessment for the River Vistula in Warsaw. *Natural Hazards and Earth System Sciences*, *13*(12), 3443–3455. doi: 10.5194/nhess-13-3443-2013

Melchers, R. E., & Beck, A. T. (2018). *Structural reliability analysis and prediction* (Third edition ed.). Hoboken, NJ: Wiley.

Merz, B., & Thieken, A. H. (2005, July). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, *309*(1-4), 114–132. doi: 10.1016/j.jhydrol.2004.11.015

Neal, J., Keef, C., Bates, P., Beven, K., & Leedal, D. (2013, April). Probabilistic flood risk mapping including spatial dependence. *Hydrological Processes*, *27*(9), 1349–1363. doi: 10.1002/hyp.9572

Neal, J., Schumann, G., & Bates, P. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, *48*(11), 1–16. doi: 10.1029/2012WR012514

Nuswantoro, R., Diermanse, F., & Molkenthin, F. (2016). Probabilistic flood hazard maps for Jakarta derived from a stochastic rain-storm generator. *Journal of Flood Risk Management*, *9*(2), 105–124. doi: 10.1111/jfr3.12114

Papaioannou, G., Vasiliades, L., Loukas, A., & Aronica, G. T. (2017, April). Probabilistic flood inundation mapping at ungauged streams due to roughness coefficient uncertainty in hydraulic modelling. *Advances in Geosciences*, *44*, 23–34. doi: 10.5194/adgeo-44-23-2017

Pregnolato, M., Galasso, C., & Parisi, F. (2015, July). A compendium of existing vulnerability and fragility relationships for flood : Preliminary results. In *12th International Conference on Applications of Statistics and Probability in Civil Engineering.* Vancouver, Canada. doi: 10.14288/1.0076226

Romanowicz, R. J., & Kiczko, A. (2016, July). An event simulation approach to the assessment of flood level frequencies: Risk maps for the Warsaw reach of the River Vistula: Event Simulation Approach to Flood Risk Assessment. *Hydrological Processes*, *30*(14), 2451–2462. doi: 10.1002/hyp.10857

Spiegelhalter, D. J., & Riesch, H. (2011, December). Don't know, can't know: Embracing deeper uncertainties when analysing risks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1956), 4730–4750. doi: 10.1098/rsta.2011.0163

Stephens, T. A., & Bledsoe, B. P. (2020, March). Probabilistic mapping of flood hazards: Depicting uncertainty in streamflow, land use, and geomorphic adjustment. *Anthropocene*, *29*, 100231. doi: 10.1016/j.ancene.2019.100231

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual.* Scotts Valley, CA: CreateSpace.

Vickery, P. J., Lin, J., Skerlj, P. F., Twisdale, L. A., & Huang, K. (2006, May). HAZUS-MH Hurricane Model Methodology. I: Hurricane Hazard, Terrain, and Wind Load Modeling. *Natural Hazards Review*, *7*(2), 82–93. doi: 10.1061/(ASCE)1527-6988(2006)7:2(82)

Zahmatkesh, Z., Han, S., & Coulibaly, P. (2021, January). Understanding Uncertainty in Probabilistic Floodplain Mapping in the Time of Climate Change. *Water*, *13*(9), 1248. doi: 10.3390/w13091248

Zio, E. (2013). *The Monte Carlo Simulation Method for System Reliability and Risk Analysis.* London: Springer London. doi: 10.1007/978-1-4471-4588-2

The cost of imperfect knowledge: how epistemic uncertainties influence flood hazard assessments

M. Balbi[1]*, and D. C. B. Lallemant[2]

[1]Laboratorio de Materiales y Estructuras, School of Engineering, Universidad de Buenos Aires, Buenos Aires, Argentina
[2]Earth Observatory of Singapore, Nanyang Technological University, Singapore

Key Points:

- Flood hazard assessments involve sophisticated probability and physics-based models that require the specification of many parameters
- We propose a Bayesian methodology to include uncertainty in model's parameters into robust hazard estimates and useful hazard maps for risk-based decision-making
- The inclusion of uncertainty in parameters can significantly affect hazard estimates and its omission can lead to non-conservative planning and design

*Current address, Av. Las Heras 2214, Buenos Aires, Argentina

Corresponding author: Mariano Balbi, `mabalbi@fi.uba.ar`

Abstract

Classical approaches to flood hazard are obtained by the concatenation of a recurrence model for the events (i.e. an extreme river discharge) and an inundation model that propagates the discharge into a flood extent. The traditional approach, however, uses 'best-fit' models that do not include uncertainty from incomplete knowledge or limited data availability. The inclusion of these, so called epistemic uncertainties, can significantly impact flood hazard estimates and the corresponding decision-making process. We propose a simulation approach to robustly account for uncertainty in model's parameters, while developing a useful probabilistic output of flood hazard for further risk assessments. A Peaks-Over-Threshold Bayesian analysis is performed for future events simulation, and a pseudo-likelihood probabilistic approach for the calibration of the inundation model is used to compute uncertain water depths. The annual probability averaged over all possible models' parameters is used to develop hazard maps that account for epistemic uncertainties. Results are compared to traditional hazard maps, showing that not including epistemic uncertainties can underestimate the hazard and lead to non-conservative designs, and that this trend increases with return period. Results also show that the influence of the uncertainty in the future occurrence of discharge events is predominant over the inundation simulator uncertainties for the case study.

Plain Language Summary

Estimating the annual probability of some flood-depth level is a key input for risk analysis and engineering design. This is typically calculated via sophisticated probability and physics-based models that require many parameters. However, the classical approach uses a fixed set of 'best parameters' for this and do not include the degree of uncertainty, even when such uncertainties may be very high. This work proposes a method to estimate the annual probability of flood-depth including the uncertainty in the parameters used to compute it. More importantly, it shows that not including this uncertainty might severely underestimate the hazard and consequently lead to unsafe designs.

1 Introduction

As a key component of a comprehensive risk analysis, the hazard model is required to characterize the future occurrence of potentially damaging events. This 'potential' for damage is numerically quantified through an intensity measure (IM) metric that, in the case of flood hazard, usually is the water level, velocity and/or duration at any given point of interest (Pregnolato et al., 2015). It, ultimately, has the purpose of providing valuable input information for vulnerability (e.g. damage) models that allow decision-makers to help mitigate the impact of natural hazards.

In this context, flood hazard is typically defined as the probability of exceedance of an IM level at any point of interest during a given period of time. This usually comes in the form of 'hazard curves' that relate IM levels with an annual exceedance probability (AEP) or a mean time of recurrence, also known as return period (RP). In practice, the outcome is best conveyed through flood maps for different RPs to reflect the spatial distribution of hazard estimates.

Since observations of IMs usually scarce for most locations, a purely statistical description of their probability distribution is not possible. The typical approach, then, is to compute IMs as a result of the convolution of two distinct models: a 'recurrence model' that describes the probability of occurrence of extreme events, such as an extreme rainfall, river discharge, or sea-level rise; and a 'source-to-site propagation model' (or just propagation model for brevity) that represents how the triggering event is translated into a spatial (and temporal) distribution of IM levels (i.e. spatially distributed flood extent and depth). The former is inherently probabilistic and typically modelled via standard stochastic time process such as the Poisson process. The propagation model, on the other hand, is typically modelled through deterministic physics-based models such as a hydrologic and/or a hydraulic inundation model.

The classical method involves computing the event's magnitude, such as peak discharge in riverine flooding, for different RPs and use it as input of the inundation model to develop flood maps for the different recurrences. This is done using appropriately calibrated recurrence curves and inundation simulators based on available data and expert's knowledge. This is the standard approach for most practical applications in the industry, due to its conceptual simplicity and ease of implementation. It does, however, assume that the models used are perfectly 'true'. The only probabilistic nature of the approach comes from the inherent uncertainty in the future occurrence of extreme events; also termed 'aleatory uncertainty' (J. Hall & Solomatine, 2008).

A broad range of researchers during the last decades have brought attention to the importance of including other, more subjective, sources of uncertainty into risk analysis in general, and flood hazard modelling in particular (J. Hall & Solomatine, 2008; Beven, 2014; Merz & Thieken, 2005). Subjective uncertainties, also termed here 'epistemic', can arise from the data we use to constrain our models, our lack of knowledge regarding the true physical processes involved, or our limited analytic and computational capabilities for providing results (J. Hall & Solomatine, 2008). Their inclusion may lead to impactful modifications in the decision making process, at the cost of a significant increment in analytical and computational complexity.

Many researchers have dealt with the inclusion of epistemic uncertainties in flood hazard models, specifically riverine floods, in the last decades. Some works have focused mainly on dealing with uncertain representations on the recurrence of the input discharge events. That is, defining a distribution for the uncertain discharge for a given return period. This includes accounting for statistical fitting errors due to limited-length data and distribution family (Apel et al., 2008; G. T. Aronica et al., 2012; Neal et al., 2013; Romanowicz & Kiczko, 2016; Stephens & Bledsoe, 2020), secondary input variables as flood volume (Candela & Aronica, 2017), or more general hydrograph shape uncertainties through hydrological modelling (Grimaldi et al., 2013; Falter et al., 2015; Ahmadisharaf et al., 2018; Zahmatkesh et al., 2021). Others have focused on including uncertainty in the inundation model through its most sensitive parameters such as roughness coefficients (Di Baldassarre et al., 2010; Kalyanapu et al., 2012; G. T. Aronica et al., 2012; Kiczko et al., 2013; Romanowicz & Kiczko, 2016; Bharath & Elshorbagy, 2018), Digital Elevation Maps (DEM) (Apel et al., 2008), or cross-section geometrical properties (Stephens & Bledsoe, 2020). Furthermore, many of these have included both the epistemic uncertainties in the discharges recurrence as well as in the inundation model (Apel et al., 2008; Di Baldassarre et al., 2010; Kalyanapu et al., 2012; G. T. Aronica et al., 2012; Kiczko et al., 2013; Romanowicz & Kiczko, 2016; Bharath & Elshorbagy, 2018; Stephens & Bledsoe, 2020; Zahmatkesh et al., 2021).

The typical outcome from most of these approaches is in the form of 'probability of flood' maps for different return periods. That is, for a specific return period, different discharges and/or inundation model parameters are randomly sampled and used to obtain an ensemble of flood maps from which the probability of flooding is computed empirically (Di Baldassarre et al., 2010; Domeneghetti et al., 2013; Neal et al., 2013; Kiczko et al., 2013; Bharath & Elshorbagy, 2018; Stephens & Bledsoe, 2020; Zahmatkesh et al., 2021). A flood risk analysis requires estimating potential damages from the hazard outcomes, and this type of input is not very helpful since most flood damage models use as input the water depth above ground level (Pregnolato et al., 2015). For this reason, instead of translating uncertain flow discharges for given RP into a probability map (broadly known as 'event-based' approach), some researchers have aimed to develop recurrence curves for water depths at the points of interest (G. T. Aronica et al., 2012; Nuswantoro et al., 2016; Romanowicz & Kiczko, 2016).

The literature review indicates a lack of hazard methodologies that can (1) include epistemic uncertainties in both the recurrence and inundation models, (2) provide useful output for further risk assessments, while (3) also being probabilistically consistent and computationally tractable. This work explores a simulation methodology of flood scenarios using a Bayesian approach of extreme value theory and the Generalized Likelihood Uncertainty Estimation (GLUE) framework to account for epistemic uncertainty in the parameters of the recurrence model and the inundation simulator respectively. We propose to use the probability of exceedance averaged over the distribution of all possible parameters as a point measure of flood hazard due to its improved statistical properties as discussed in Fawcett and Green (2018), and its lower computational demand compared to obtaining full credible intervals. This framework allows the development of flood hazard curves, as well as flood hazard maps by computing this estimate of the recurrence at every point of interest.

Section 2 describes the mathematical model used to compute hazard estimates, a framework to include epistemic uncertainties through model's parameters posterior distributions, and a simulation procedure for its numerical evaluation. A small case study of riverine flooding is described in Sect. 3 and used as a working example to test this methodology. In Sect. 4, the resulting hazard curves and maps are compared to the traditional approach where no epistemic uncertainties are included. Insights in the hypothesis, results and implications of the model are analyzed in the discussions of Sect. 5, while a summary of main takeaways and potential future lines of research are drawn in the final section.

2 Methodology

2.1 The hazard model

As discussed in the introduction, flood hazard can be quantified as the annual probability of exceedance of a given IM level y, at any location of interest. Mathematically, this probability is calculated by a stochastic time process model. The most used one, due to its simplicity and well-known mathematical properties, is the Homogenous Poisson Process (HPP) for which events occur discretely with independent exponentially distributed inter-arrival times with a mean rate λ_0 that is constant over time. Under these simplifying assumptions, the probability of exceedance over a timespan T can be computed as per Eq. 1.

$$p_T(y) = 1 - \exp\left(-\lambda_0 T p\left(Y \geq y\right)\right) \tag{1}$$

Where Y is the random IM for any given event, T is the timespan of interest, and $p\left(Y \geq y\right)$ is the probability of exceedance of level y for any given event (this probability is constant over time in the HPP).

The hazard is then computed by setting $T = 1$ in Eq. 1 to obtain the annual probability of exceedance. For events with low recurrence λ_0 and low exceedance probability $p\left(Y \geq y\right)$, as is mostly the case in disaster risk analysis, the probability of Eq. 1 can be further simplified as in Eq. 2.

$$p_T(y) \approx \lambda_0 p\left(Y \geq y\right) \tag{2}$$

This is equivalent to the mean rate of exceedance of $Y \geq y$ or its multiplicative inverse, the mean time between occurrences $Tr\left(y\right)$, also known as 'return period' (see Eq. 3). In practical terms, hazard is measured by the annual probability of exceedance or by the mean rate of occurrence of IM level y, which are practically equivalent.

$$\lambda_0 p\left(Y \geq y\right) = \lambda\left(y\right) = \frac{1}{Tr\left(y\right)} \tag{3}$$

Since direct observations of water depths (or IMs in general) during flood events are very rare, probabilistic characterization of Y is usually done via mechanistic fluid dynamics models, here called 'inundation model' or simply 'simulator'. This model depends on a number of observable boundary conditions X that are considered to vary event to event, such as upstream river discharge, rainfall intensity or sea-level rise. It also depends on a set of unobserved calibration parameters β considered constant over events, such as the soil roughness parameters or the channel cross-section geometry (see Eq. 4).

$$Y = S\left(X, \beta\right) \tag{4}$$

In this context, an 'event' is characterized by a magnitude X which describes the impact potential of the phenomenon. For example, in riverine flooding, X can represent the river discharge flow and an event is triggered when it surpasses a given threshold. Analogously, in coastal flooding, X might stand for sea-level extreme rise, or in pluvial flooding where X stands for rainfall intensity. In more complex scenarios, X can be a vector representing multiple quantities, such as flow discharge and volume, rainfall intensity and duration, or a combination of flow discharge and sea-level rise. For the sake of simplicity, this work will focus on scalar X.

Since Y is a function of X, the probability $p\left(Y \geq y\right)$ in Eq. 2 can be computed by conditioning on the probability distribution of the event's magnitude X as given by,

$$p\left(Y \geq y\right) = \int_x \mathbf{1}\left\{S\left(x, \beta\right) \geq y\right\} p\left(x|\theta\right) dx \tag{5}$$

Where $\mathbf{1}\left\{\text{cond}\right\}$ is an indicator function that returns 1 when cond is true and 0 otherwise and θ is a vector of parameters that describe the probability distribution of X.

This expression is useful as long as it is easier to define the probability distribution of events magnitudes $p\left(x|\theta\right)$ than the distribution of IM levels $p\left(Y \geq y\right)$ from data or expert knowledge. As mentioned before, this is the typical case in flood hazard, where we usually have relatively robust historical measurements of river flow discharges or rainfall intensity, but very few of water depths at points of interest in the floodplain.

Introducing Eq. 5 into Eq. 2, we obtain the full expression for the flood hazard,

$$\lambda\left(y\right) = \int \underbrace{\mathbf{1}\left\{S\left(x, \beta\right) \geq y\right\}}_{\substack{\text{Inundation}\\ \text{model}}} \underbrace{\lambda_0 p\left(x|\theta\right)}_{\substack{\text{Events}\\ \text{recurrence}\\ \text{model}}} dx \tag{6}$$

An illustrative scheme of a realization of the described time process is shown in Figure 1. The varying sizes of the blue bubbles reflect the magnitude X_i of the events, while the black bars reflect the IM level (i.e. water depth) for each. According to the HPP model, the time T_i between events follows an exponential distribution with mean rate λ_i, while the time between IM exceedances $T_i\left(y\right)$ (the black bars that cross the dotted red line) follows an exponential distribution with mean rate $\lambda\left(y\right)$ as described in Eq. 6.

It is important to highlight that the model summarized here is replicated when analyzing the hazard for other types of natural phenomena. In the case of seismic hazard, X is the moment magnitude and spatial epicenter location of the earthquake and their probability distribution is typically given by the Gutenberg-Richter law and the source-to-site propagation model is defined by the Ground Motion Prediction Equations (GMPEs) (Baker et al., 2021). For typical hurricane winds hazard, the recurrence model describes the likelihood of the hurricane's central pressure and track, while the propagation model is described through a wind field model (Vickery et al., 2006).

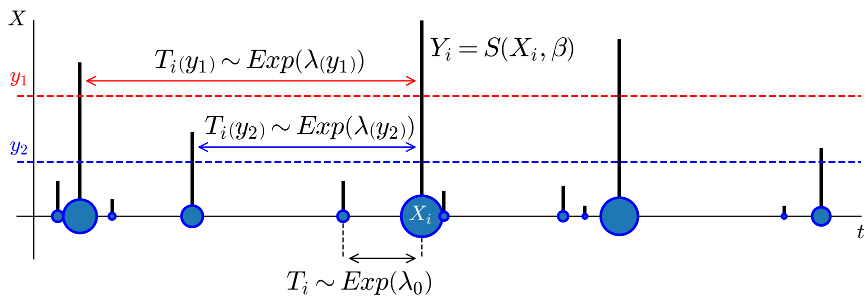

Figure 1. Schematic illustration of a realization of the HPP model, with two arbitrary IM levels y_1 and y_2 to define the hazard

2.2 Including epistemic uncertainties

The hazard problem is, as described by Eq. 6, tightly related to predicting an uncertain event in the future. Thus, it is strictly an uncertainty quantification task. At its core, that expression is a mathematical representation of what is known as 'aleatory uncertainty', here characterized by the exponentially distributed inter-arrival times with mean rate λ_0 and the probability distribution of the event's magnitude $p\,(x|\theta)$.

Aleatory uncertainty is considered an inherent component of the physical process and it does not depend on the amount of knowledge and information the modeller has. However, there are other sources of uncertainty around the estimation of the hazard that are related to our incomplete knowledge about the physical process and data available to characterize it. These are commonly known as 'epistemic uncertainties' (Spiegelhalter & Riesch, 2011).

As described by Spiegelhalter and Riesch (2011), epistemic uncertainty stems mainly from (1) limited information to properly characterize the models and variables involved and (2) limited knowledge to properly describe the true physical processes through the selected models. This more operational description of epistemic uncertainties allows for a more rigorous way of including them in the mathematical model.

Limited information appears in practice, as limited-length data, observation errors, or missing variables. It can, typically, be represented through uncertainty in the parameters that describe the models as data is not sufficient to perfectly identify them. Limited knowledge, on the other hand, is usually represented through simplifying assumptions as the ones in the HPP model, uncertainty in the distribution family chosen for $p\,(x|\theta)$, or the particular physics-based model chosen for S. It can be harder to represent this mathematically, although it has been done through model ensembles (i.e. considering and weighting many possible models) or statistical representations of model deficiencies (Kennedy & O'Hagan, 2001; Balbi & Lallemant, 2023).

Despite the epistemological differences between the two, it is not always clear which sources of uncertainty belong to each category, and it can vary depending on the context. In any case, the most important feature that differentiates aleatory and epistemic uncertainty is the fact that the former cannot be practically reduced since it is an inherent property of the system under analysis. The epistemic uncertainty, on the other hand, can be reduced by further collecting information and improving knowledge. This distinction is crucial when allocating resources for model improvement (Der Kiureghian & Ditlevsen, 2009; Merz & Thieken, 2005).

The inclusion of epistemic uncertainties greatly increases the complexity of the problem from an analytical and computational standpoint. In particular, this work focuses on the inclusion of epistemic uncertainty through model's parameters as a broad representation of limited data to define the models. In this context, hazard as calculated in Eq. 6 can be understood as being conditional to a given set of models and parameters as given by,

$$\lambda\left(y|\beta,\lambda_0,\theta\right)=\int \mathbf{1}\left\{S\left(x,\beta\right)\geq y\right\}\lambda_0 p\left(x|\theta\right)dx \tag{7}$$

One way of incorporating uncertainty regarding the values of the set of parameters $\{\beta,\lambda_0,\theta\}$ is provided by Bayesian decision theory. In this context, an appropriate estimate of the hazard should take into account the consequences of over or underpredicting its true value. Fawcett and Green (2018) discusses this when estimating return period levels for environmental extreme events, and they suggest the use of the predictive posterior return level as a point estimate that reliably incorporates epistemic uncertainty.

The posterior predictive estimate is obtained by averaging the conditional hazard of Eq. 7 over all possible values of the parameters $\{\beta,\lambda_0,\theta\}$ weighted by their posterior distribution $p\left(\Theta|\text{data}\right)$ as per Eq. 8.

$$\lambda\left(y|data\right)=\int_{\Theta}\lambda\left(y|\lambda_0,\beta,\theta\right)p\left(\lambda_0,\beta,\theta|\text{data}\right)d\lambda_0 d\beta d\theta \tag{8}$$

In the Bayesian framework, the posterior distribution of the parameters $p\left(\lambda_0,\beta,\theta|\text{data}\right)$, is the probability conditional on the available data and modeller's prior knowledge obtained by means of Bayes' Theorem (Gelman et al., 2013). The posterior distribution is proportional to the probability of observing the data given a set of parameters, also known as 'likelihood function', multiplied by the probability of a given set of parameters before incorporating the data, also known as 'prior distribution' (see Eq. 9). This can be colloquially described as the modeller's knowledge (i.e. prior distribution) times the information contained in the observations (i.e. the likelihood function).

$$p\left(\lambda_0,\beta,\theta|\text{data}\right)\propto p\left(\text{data}|\lambda_0,\beta,\theta\right)p\left(\lambda_0,\beta,\theta\right) \tag{9}$$

Equation 8 allows us to compute the hazard curve of water depths (or the IM chosen for the analysis) for any given point in space. Or inversely, compute the water depth for a given rate or return period. Hence, we can develop the T-years flood hazard map by marginally computing the y level for that return period from the expression.

In the following sub-sections we describe the methodology to estimate the posterior probability distributions of the parameters from Eq. 9, and the simulation methodology to estimate the hazard as per Eqs. 7 and 8.

2.2.1 A Bayesian recurrence model

In the framework described above, the recurrence model aims to quantify the mean rate of occurrence λ_0 of events, and characterize the distribution of the magnitude X and its parameters θ for any given event. These can be obtained from the statistical analysis of time series of past events. Poisson Point Process theory, as a generalization of the well-known Peaks-Over-Threshold (POT), methodology provides a robust framework for this (Bezak et al., 2014). Extreme events are individualized from historical records of daily discharges, by selecting an appropriate minimum threshold u and time separation to ensure independence. This results in a dataset of observed independent times between events $\hat{T}$ and a dataset of observed event's magnitudes $\hat{x}$. A Generalized Pareto Distribution (GPD) is then used as the probability model for the exceedances above the threshold threshold $x-u$, and the inter-arrival time between events is assumed to be exponentially distributed. This is,

$$T \sim Exp\left(t|\lambda_0\right) \quad (10)$$
$$X-u \sim \mathcal{GPD}\left(x|\xi,\sigma\right) \quad (11)$$

Where λ_0 is the mean rate of arrival and ξ and σ are the shape and scale parameters that define the probability distribution, and u an appropriately defined threshold.

Hence, two probability models are required to describe the occurrence of events. An exponential distribution model of parameter λ_0 for the time between arrivals, and a GPD of parameters $\theta = \{\xi, \sigma\}$ for the magnitude of each event. Bayesian statistics provide an ideal framework to compute uncertainty in model's parameters that are consistent with the modeller's prior knowledge and proposed model (Bousquet & Bernardara, 2021).

For the dataset of observed interarrival times $\hat{T}$, the likelihood function is simply a product of n exponential densities, where n is the number of observations. The posterior distribution of the mean rate λ_0 can be obtained by assuming a weakly informative $Gamma\left(1/2, 1/2\right)$ prior distribution. A weakly informative prior distribution is relatively flat in the entire range of plausible values for the parameter. For this choice of prior, the posterior distribution has a closed form solution as given in Eq. 12 since it is a conjugate pair for the Exponential likelihood (Gelman et al., 2013).

$$p\left(\lambda_0|\hat{T}\right) = Gamma\left(n+1/2, t+1/2\right) \quad (12)$$

Where n is the number of events and $t = \sum_{\forall i} \hat{T}_i$ is the total number of years in the series.

The most likely value of the rate of occurrence given the observations is given by the mode of the posterior probability of Eq. 12, also known as the Maximum A-Posteriori (MAP) estimate. The MAP estimate and the mean value of the posterior distribution for λ_0 are given by Eqs. 13 and 14 respectively.

$$\lambda_0^* = \left(n-1/2\right)/\left(t+1/2\right) \quad (13)$$
$$\overline{\lambda_0} = n/t \quad (14)$$

For the dataset of observed event's magnitudes $\hat{x}$, the likelihood function is given by a product of n GPD densities. There is no conjugate model for this likelihood, but a non-informative prior can built for the shape and scale parameters following Castellanos and Cabras (2007) (Eq. 15).

$$p\left(\xi,\sigma\right) \propto \sigma^{-1}\left(1+\xi\right)^{-1}\left(1+2\xi\right)^{-1/2} \quad (15)$$

Valid for $\xi > -0.5$ and $\sigma > 0$.

Then, the un-normalized expression for the posterior can be obtained by Bayes theorem as per Eq. 16, and the predictive posterior distribution for X can be subsequently computed as per Eq. 17. In both cases, there is no analytic solution, and samples from the distribution can be done via standard Markov Chain-Monte Carlo (MCMC) methods (Gelman et al., 2013). These can also be used to compute posterior mean and MAP estimates for ξ^*, σ^*.

$$p\left(\xi,\sigma|\hat{\mathbf{x}}\right) \propto p\left(\xi,\sigma\right)\prod_{i=1}^{n}\left(1+\hat{\mathbf{x}}_i\xi/\sigma\right)^{-(1+\xi)/\xi} \quad (16)$$

$$p\left(x-u|\hat{\mathbf{x}}\right) = \int \mathcal{GPD}\left(x-u|\xi,\sigma\right)p\left(\xi,\sigma|\hat{\mathbf{x}}\right)d\xi d\sigma \quad (17)$$

2.2.2 *Probabilistic inundation model*

The inundation model is, in this context, a computational solver of some simplified version of the fluid dynamics equations that depend on variable inputs X and calibration parameters β. Epistemic uncertainties might come from lack of sufficient information to calibrate the parameters, observation errors, mechanistic simplifying assumptions, and numerical simplification of the equations solver (Kennedy & O'Hagan, 2001). We assume here, due to simplicity, that for a given simulator $S(X, \beta)$ these can be represented by uncertainty in the model's calibration parameters β. More complex procedures to include model uncertainty can be used to include uncertainty in the model structure. For example, a formally probabilistic calibration procedure that includes model structural uncertainty as an additive Gaussian Process is discussed in Balbi and Lallemant (2023).

Parameter's distributions can be obtained using nominal probability models from expert's knowledge (Kalyanapu et al., 2012; Stephens & Bledsoe, 2020) or statistically calibrated ones generally via the GLUE methodology (Di Baldassarre et al., 2010; G. T. Aronica et al., 2012; Kiczko et al., 2013; Romanowicz & Kiczko, 2016; Zahmatkesh et al., 2021). In this work, epistemic uncertainty will be represented by probability distributions in the roughness parameters only, for the floodplain and for the channel, considering all other inputs as constant regarding the calibration procedure. These distributions will be obtained by means of the GLUE framework, where all possible sets of parameters are assigned a normalized score (i.e. pseudo-likelihood) from an appropriately selected scoring rule. In the case of flood extent binary observations (as in the case study developed in this work), it is typical to use the F-score (as per Eq. 18), a variant of the classical Jaccard Index (G. T. Aronica et al., 2012; Papaioannou et al., 2017).

$$F(\boldsymbol{\beta}) = \frac{A - B}{A + B + C} \tag{18}$$

Where A is the number of correctly predicted pixels, B the number of over-predicted pixels (predicted flooded, observed non-flooded), and C is the number of under-predicted pixels (predicted non-flooded, observed flooded).

The details of the calibration procedure using the GLUE framework can be found in Balbi and Lallemant (2023), but they can be summarized in four steps:

1. Sample a large set of β from their prior distribution
2. Compute the F-score for the sampled β
3. Reject all 'non-behavioral' models using some thresholding criteria: $F < f^*$
4. Standardize the resulting F-scores so that they are all positive and integrate to 1

The model fit for each value of β is, then, a measure of its uncertainty, or pseudo-posterior probability (as it is not strictly obtained from a probabilistic likelihood). In addition, the MAP value for parameter β is the one that yields the best fit (i.e. largest F-score).

2.3 Numerical implementation

Computing the integral from Eq. 8 requires numerical procedures since no analytic solution exist for the posterior distributions of the parameters just described. Since many parameters are involved in its computation, a summary of variables and symbols can be found in Table 1.

Equation 8 can be slightly simplified, however, by noting that $\lambda(y|\lambda_0, \beta, \theta)$ is linear on λ_0 and its posterior distribution (Eq. 12) is independent from the rest of the pa-

Table 1. Summary of variables and symbols

Variable	Description
X	Flood event's magnitude (i.e. river peak discharge)
Y	Flood IM (i.e. flood depth) at a given point in space
λ_0	Mean rate of occurrence of events $X \geq 0$
$\lambda(y)$	Mean rate of occurrence of events $Y \geq y$
β	Inundation simulator calibration parameters
$S(X, \beta)$	Inundation simulator
ξ	$\mathcal{GPD}$ shape parameter
σ	$\mathcal{GPD}$ scale parameter
u	Threshold value for X POT model
$\overline{(.)}$	Mean value of parameter
$(.)^*$	MAP value of parameter (mode of its posterior distribution)

rameters as given by Eq. 19.

$$\lambda(y|data) = \overline{\lambda_0} \int_{\beta,\xi,\sigma} p(Y \geq y|\beta,\xi,\sigma)\, p(\beta,\xi,\sigma|\text{data})\, d\beta d\xi d\sigma \tag{19}$$

Where $\overline{\lambda_0}$ is the mean of the posterior distribution of λ_0 given by Eq. 14.

Standard Monte-Carlo (MC) integration techniques can be employed to compute such integral. Conceptually the task is straightforward: we need to sample from the posterior distribution of Y, by sampling first from the posterior of X. This can be done through the following steps:

1. Sample N values from the posterior distribution of β (as per Sect. 2.2.2) and $\{\xi, \sigma\}$ (from Eq. 16)
2. For each sample of $\{\xi_i, \sigma_i\}$, sample x_i from Eq. 17
3. For each sample x_i and β_i, compute water depth at all points of interest from the simulator $y_i = S(x_i, \beta_i)$
4. Estimate the mean rate of exceedance of y as:

$$\lambda(y|\text{data}) \approx \frac{\overline{\lambda_0}}{N} \sum_i \mathbf{1}\{S(x_i, \beta_i) \geq y\}$$

The number of samples N required for the simulation depends on the percentile of the curve (i.e. return period) we are trying to estimate and the precision desired. For example, to estimate the 100 years return level y_{100} we need to estimate an exceedance probability $p(Y \geq y_{100}) = (100\lambda_0)^{-1}$. According to the standard theory of empirical estimates of probabilities, based on the Central Limit Theorem (CLT), we can obtain an approximate minimum number of simulations for a 95% confidence interval as per,

$$N > \frac{1.96}{\varepsilon^2}\sqrt{\frac{1-p}{p}} \tag{20}$$

Where p is the actual probability being estimated (not exactly known) and ε is the width of the relative interval (Melchers & Beck, 2018).

On the other hand, the computation of the hazard in the classical approach (as given by Eq. 5), where no epistemic uncertainties are considered, is much simpler:

1. Obtain the return levels of the GPD distribution for each return period Tr of interest , and for fixed parameters ξ^* and σ^* as given by (see Appendix A for de-

tails),

$$x_{Tr} = u + \begin{cases} \frac{\sigma^*}{\xi^*} \left\{ (\lambda_0^* \cdot Tr)^{\xi^*} - 1 \right\}, & \text{if } \xi^* \neq 0 \\ \sigma^* \ln (\lambda_0^* \cdot Tr), & \text{if } \xi^* = 0 \end{cases}$$

2. For each Tr, compute water depth at all points of interest for fixed parameters β^* from the simulator $y_{Tr} = S(x_{Tr}, \beta^*)$

That is, we compute x for different return periods of interest (also known as return levels), and then evaluate the inundation model at each. It is important to note, that this two-step approach can be followed but for an entire ensemble of posterior realizations of parameters $\{\beta_i, \xi_i, \sigma_i\}$, to obtain an estimate of the predictive posterior estimate including epistemic uncertainties. This method requires many times more the number of calls to the inundation simulator S relative to the four-step procedure described above, but has the advantage of obtaining credible intervals for the estimate.

3 Case study

The proposed methodology described in Sect. 2 is applied here in a real-world case study, with the purpose of analyzing the influence of the inclusion of epistemic uncertainties in the recurrence model of river discharges (as per Sect. 2.2.1) and in the inundation model (as per Sect. 2.2.2).

3.1 Models and data

The case study is based on a short reach on the upper river Thames in Oxfordshire, England, just downstream from a gauged weir at Buscot (Fig. 2). The river at this reach has an estimated bankfull discharge of 40 m/s^3 and drains a catchment of approximately 1000 km^2. The topography was obtained from stereophotogrammetry at a 50 m scale with a vertical accuracy of ±25 cm, obtained from large-scale UK Environment Agency maps and surveys. This reach has also been study previously in G. Aronica et al. (2002), J. W. Hall et al. (2011) and Balbi and Lallemant (2023)

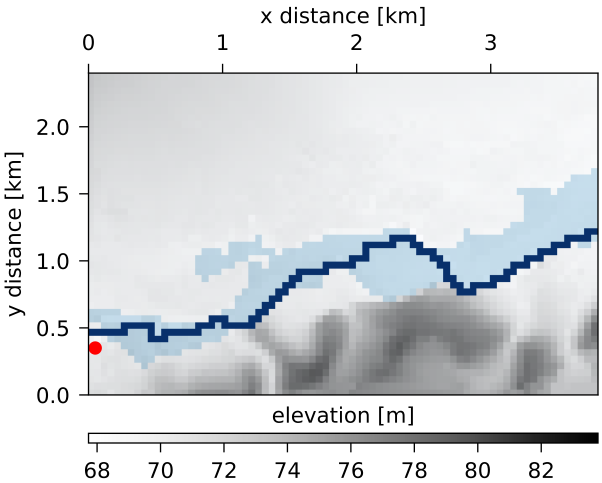

Figure 2. Floodplain topography at Buscot, SAR imagery of 1992 flood event (light blue), channel layout (dark blue) and gauge station location (red dot).

The events are characterized by the river discharge flow only. To develop the recurrence model for events, a publicly available daily discharge data series at Buscot weir was obtained from the UK National River Flow Archive (see Fig. 3). The series spans from 19 years from 1980 to 1998 with some minor gaps that are not expected to affect the extreme statistics analysis to perform.

On the other hand, for the calibration of the inundation model, a satellite observation of the flood extent of 1-in-5 year event occurred during December 1992 was used (see Fig. 2). The satellite SAR(synthetic aperture radar) image of the flood was captured 20 hours after the flood peak when discharge was at a level of $73m^3/s$ (G. Aronica et al., 2002). The resolution of the image is $50m$.

The computational inundation model used is the raster-based Lisflood-fp model (Neal et al., 2012). Lisflood-fp couples a 2D water flow model for the floodplain and a 1D solver for the channel flow dynamics. Its numerical structure makes it computationally efficient and suitable for the many simulations needed for probabilistic flood risk analysis and model calibration.

A simplified rectangular cross-section is used for the channel with a constant width of 20m for the entire reach and a varying height of around 2m. The observed event is defined by the boundary condition of a fixed input discharge of $x = 73m/s^3$ at the geographic location of the gauging station shown in Fig. 2, and by an assumed downstream boundary condition of a fixed water level of approximately 90cm above the channel bed height. The short length of the reach and the broadness of the hydrograph imply that a steady-state hydraulic model is sufficiently accurate for the calibration (G. Aronica et al., 2002).

The model's parameters used for calibration are the Manning's roughness parameters for the channel r_{ch} and for the floodplain r_{fp}, both considered spatially uniform in the domain of analysis. That is, $\beta = \{r_{ch}, r_{fp}\}$. For the calibration method described in Sect. 2.2.2, the inundation model was ran for a fixed observed discharge of $73m^3/s$ and for a uniform prior for both parameters in the range $0.01 - 0.15$.

3.2 Computational implementation

The statistical models and simulation method described in Sect. 2.3 were implemented in Python 3.X language (Van Rossum & Drake, 2009), using a 10-core Intel i9-10700k processor computer. Each evaluation of the inundation model $S(X, \beta)$ takes approximately 4s. For the calibration of the inundation model 19,600 evaluations of the simulator were needed to cover the entire grid of β values, and around 100,000 evaluations were needed for the hazard simulation procedure from Sect. 2.3.

4 Results

4.1 Discharge recurrence model

To define the events and their magnitudes, a threshold of $u = 12m^3/s$ and minimum distance between clusters of 7 days (i.e. there has to be 7 days of values below the threshold for two events to be considered as separate events) were selected aiming to satisfy the conditions required by POT standard theory (Bousquet & Bernardara, 2021):

- The minimum threshold for which the modified scale and shape parameters of the fitted GPD of the exceedances remain constant for higher thresholds.
- The resultant threshold exceedances (cluster's peaks) should form an independent sample.

A total of 73 clusters were identified in 18.8yrs of data as shown in Fig. 3. That is, on average, 3.9 events per year, and it shows the relative advantage of this type of analysis versus the standard annual maximum approach for which there would only be 18 data points. The posterior distribution of the mean rate λ_0, as given by Eq. 12, has a mode (i.e. MAP) $\lambda_0^* = 3.8yrs^{-1}$ and a mean value $\overline{\lambda_0} = 3.9yrs^{-1}$. Statistical graphical tests, shown in Fig. 4, showed that the resulting series of extreme discharges can be considered independent and that the time between events has a good fit to the exponential model as assumed in the HPP model.

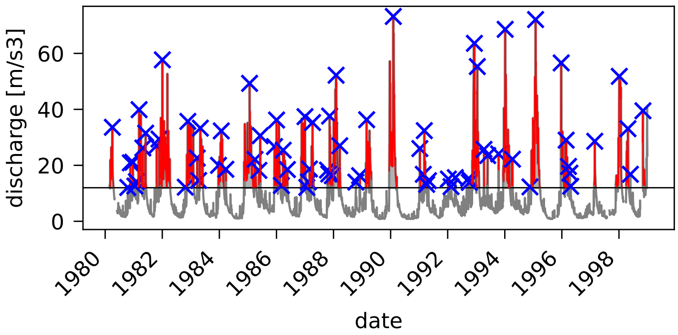

Figure 3. Daily discharge for Buscot and identification of flood events by clustering with a 12 m/s^3 threshold and 7 days of minimum return. Blue cross indicates event's peak discharge.

Samples of the posterior distribution of the GPD parameters (Eq. 16) were drawn by a standard MCMC algorithm of 4 chains of 15,000 samples each. Discarding the first half of sample from each as burn-in stage, convergence of the chains was assessed by verifying that Gelman-Rubin R-scores remains below 1.01 (Gelman et al., 2013). Goodness-of-fit tests showed a good agreement of the exceedances with the GPD. The shape parameter ξ is centered around -0.05 while the scale parameter σ is centered around 16.5, both with a relatively small skewness (see Fig. 5). The MAP values for the parameters practically coincide with these values.

For each posterior sample of ξ and σ, the probability distribution of the discharges follows a GPD. With the ensemble of distributions for each sample, we computed the mean curve (i.e. the predictive posterior distribution of X) and the 90% confidence posterior intervals. These are shown as return period curves in Fig. 6 as computed by,

$$Tr\,(x) = 1/\lambda_0^* p\,(X \geq x) \tag{21}$$

A 'deterministic' hazard curve was also computed using the MAP values for the GPD parameters, following the classical approach. It can be seen that epistemic uncertainties have the effect of increasing the discharges for a given return period, and that this effect increases with increasing return period. This is intuitive, as larger return periods are more uncertain with limited-length data. Similar results have been obtained before (Merz & Thieken, 2005; Romanowicz & Kiczko, 2016; Fawcett & Green, 2018).

4.2 Inundation model

The statistical calibration of the inundation model was done by using a uniform grid for both parameters in the range $(0.01, 0.15)$ with a step of 0.01, and a threshold

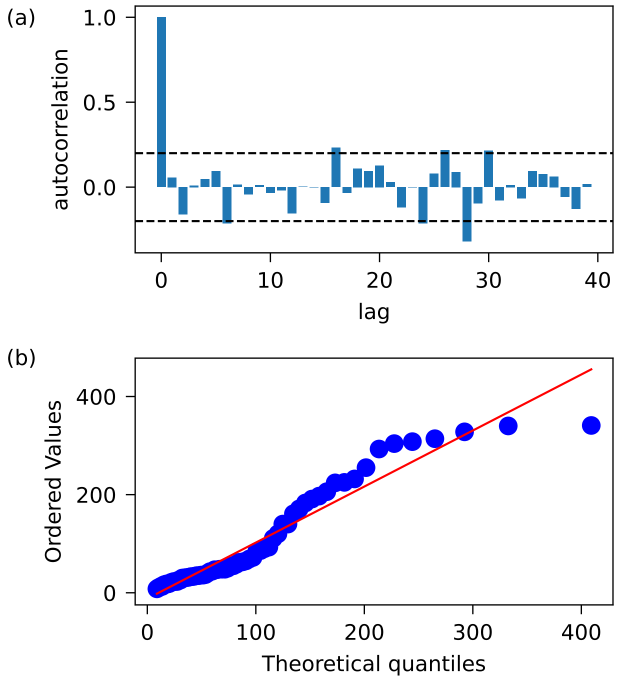

Figure 4. (a) Autocorrelation plot of discharge series; (b) Probability plot of interarrival times compared to the exponential distribution.

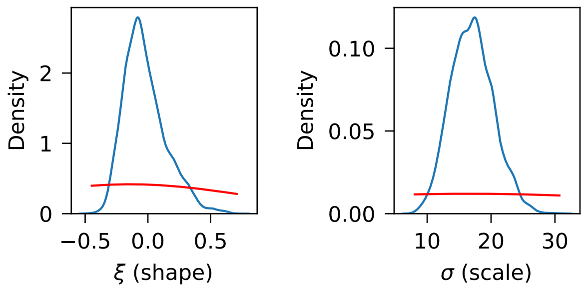

Figure 5. Posterior (blue) and prior (red) distributions of the parameters of the frequency model (blue)

of 0.5 to filter out non-behavioral models. This resulted in a total of 543 accepted simulations out of 19,600. The bivariate pseudo-posterior distribution for the roughness parameters is shown in Fig. 7. The set of parameters that yields the maximum F-score (i.e. MAP parameters) were $r_{ch} = 0.029, r_{fp} = 0.045$ giving $F = 0.54$.

4.3 Flood hazard

Close to 15,000 thousand posterior samples of water depth Y at all points in the region were obtained, following the simulation procedure described in Sect. 2.3. This allowed to empirically estimate the posterior exceedance recurrence $\lambda\left(Y \geq y|\text{data}\right)$ for ev-

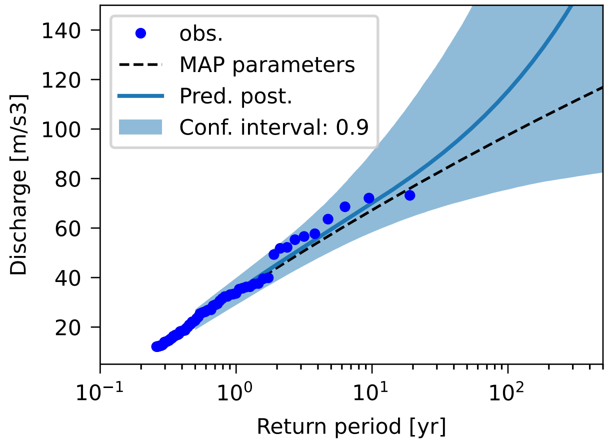

Figure 6. Hazard curves for river discharge

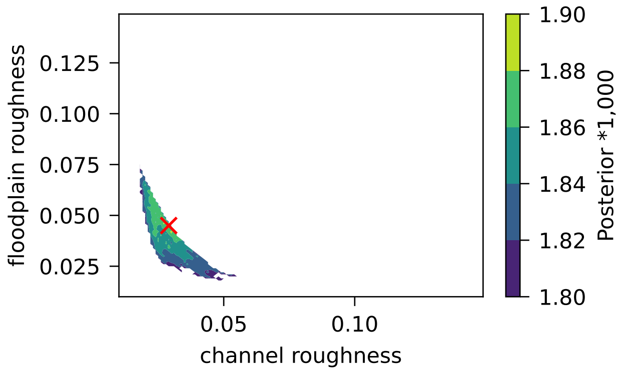

Figure 7. Bivariate posterior distribution for inundation model parameters

ery pixel and, consequently, the hazard curve. The number of simulations ran implies that exceedance probabilities as small as 10^{-4} can be estimated with a 10% interval according to Eq. 20. For a mean value $\overline{\lambda_0} = 3.8yrs^{-1}$, this is equivalent to RPs of up to $1,500yrs$.

Figure 8 shows the flood depth hazard curves for different points in the floodplain. The posterior predictive curves are compared with the classical approach that uses the deterministic discharge hazard curve (dotted black curve in Fig. 6) and the MAP parameters for the inundation simulator. In every case, it can be seen that the flood depth values of the posterior predictive curves increase faster than the classical approach as the return period grows, in a similar fashion observed in the discharge hazard model of Fig. 6. Furthermore, the two curves are very similar for lower return periods.

The 100yrs flood hazard map was developed by computing the 100yrs flood depth from the posterior predictive curves at each point (see Fig. 9). This map is compared in plot (a) of Fig. 10 with the traditional hazard map computed using the best inundation model with the deterministic estimate of the 100yrs discharge. The increasing flood depth for the posterior predictive map can be seen to be replicated for every pixel in the region of analysis with the exception of some isolated pixels right next to the channel.

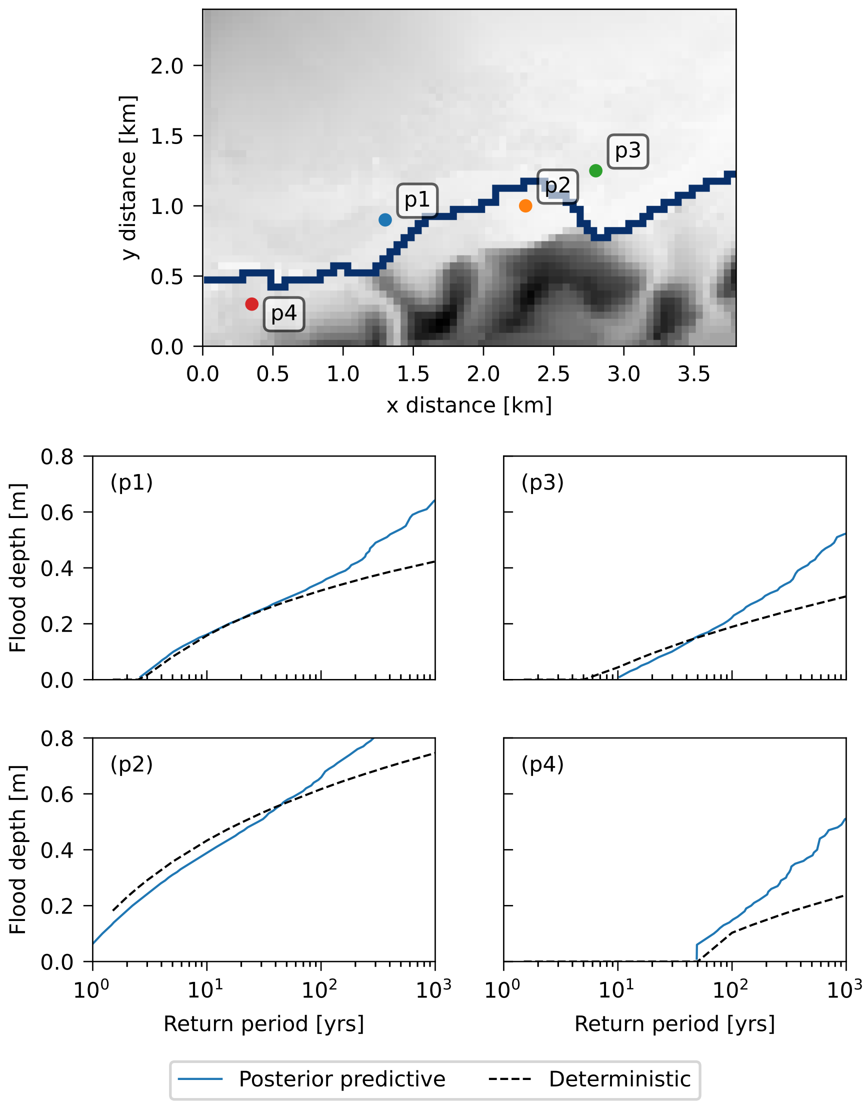

Figure 8. Hazard curves for flood depths at different locations in the floodplain

This effect, as in the discharge hazard curve, is exacerbated with increasing return periods as can be seen in plot (b) and (c) of the figure for the 250yrs and 500yrs maps comparison.

5 Discussion

5.1 On the influence of epistemic uncertainties in hazard estimates

From an engineering design perspective, the water height used to design a structure for a specified safety level (i.e. return period) will be larger when including epistemic uncertainties. Results show that this is true for every pixel in the region of analysis as can be seen from the maps in Fig. 10 and the curves in Fig. 11 for a specific pixel. That is, when including our lack of knowledge and information on the process, we need to be more conservative in design to ensure an appropriate level of reliability. Furthermore, this trend increases with the return period, and is practically negligible for more

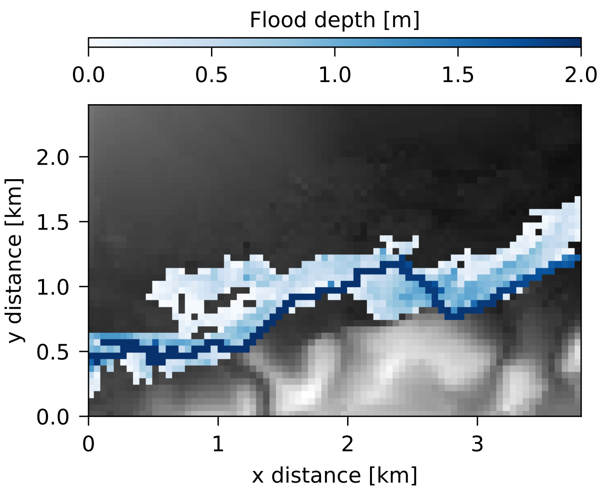

Figure 9. 100yrs flood hazard map using posterior predictive flood depth at each point

recurrent events. This result is somewhat intuitive as we usually have less knowledge on rare events, and similar conclusions have been obtained by other researchers in hydrological hazards (Merz & Thieken, 2005; Fawcett & Green, 2018).

To further understand and generalize this result, however, we need to understand the relative influence of epistemic uncertainties on the discharge recurrence model and on the inundation model. To do this, we obtained the hazard curves while including epistemic uncertainties one model at a time, as shown in Fig. 11. It can be seen that the effect of more conservative water depths for any given return period is entirely due to epistemic uncertainties in the recurrence model. This trend reflects the heavier tails of the posterior distribution of the discharges (as also seen in Fig. 6) that mainly represents uncertainty due to the limited-length observed time series used to build the model. That is, using an 18-years data record, there are practically no observations of much higher return periods which is reflected in the larger uncertainty.

Uncertainty on the inundation model parameters, on the other hand, seems to have a relatively constant decreasing effect over return periods. That is, it gives lower (i.e. less conservative) water depths for a given recurrence relative to the classical approach. This is related to the shape of the posterior distribution of the parameters, but also on the non-linear nature of the S transformation. Thus, the influence of epistemic uncertainty in the inundation model parameters is related to the type of observations used for calibration (i.e. binary flood extent observations in this case), the statistical procedure used (i.e. F-score pseudo-likelihood), and also the non-linearity of the inundation simulator itself. Given that water levels are relatively constrained by the topography, it is not expected that the inundation simulator presents a radically high non-linearity. Thus, the highly skewed shape of the roughness parameters posterior distribution (Fig. 7), obtained from the GLUE method, with respect to the MAP values might be the main driver responsible for the underestimation in flood hazard. However, further analysis is required in order to deeply understand how this influence varies in different settings (i.e. different observations, different calibration methods and inundation models), and understand if this effect can be magnified in some contexts.

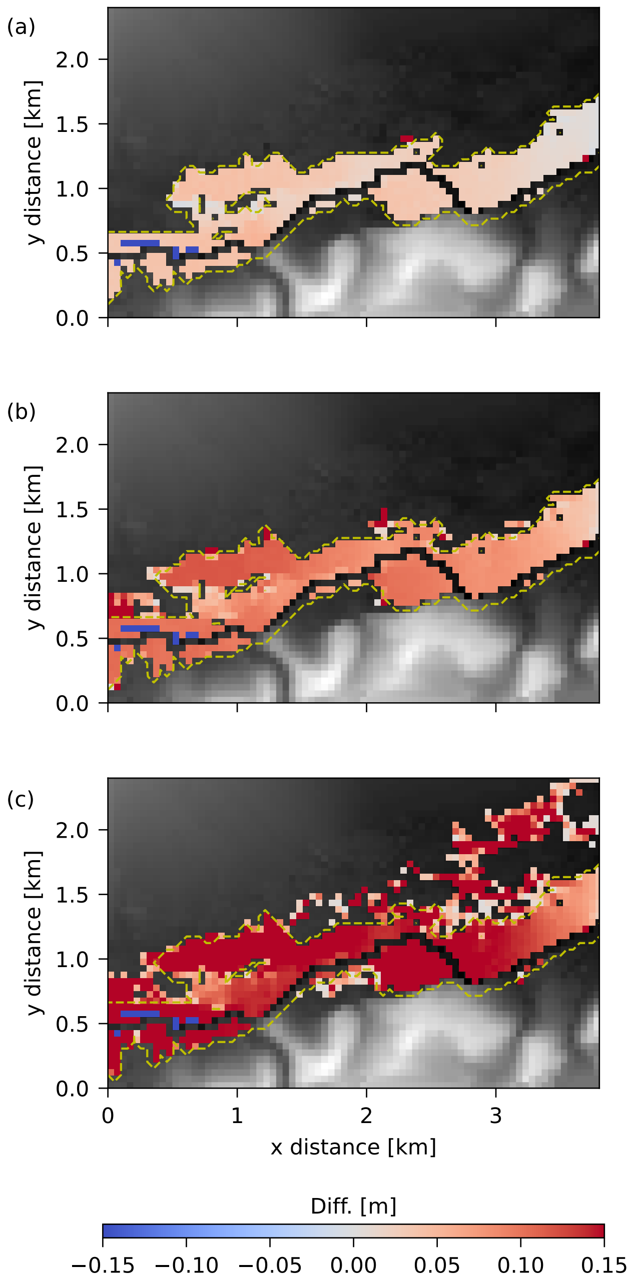

Figure 10. Difference maps between posterior predictive estimates and deterministic estimates for (a) 100yrs, (b) 250yrs, (c) 500yrs

5.2 On the usefulness of the output

From a risk analysis standpoint, we might be interested in computing some damage measure, or any higher-level decision metric, that reflects the impact of the flood in human communities. The framework is analogous to the one described here, but instead we are interested in the distribution of the decision metric Z over all the potential hazard events. We can straightforwardly compute this from the probability distribution of

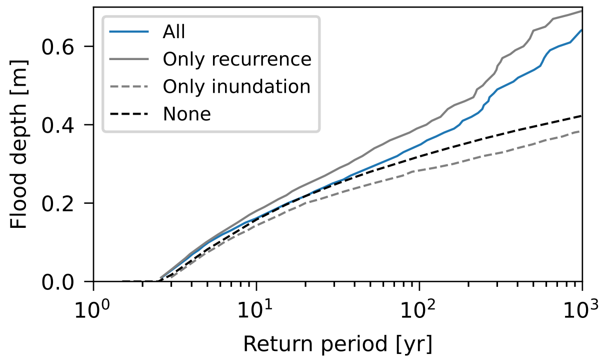

Figure 11. Hazard curves for water depth at location $x = 1.30km, y = 1.05km$

the IM, as in Eq. 22, since the vulnerability model is generally dependent on the y level.

$$\lambda(z) \approx \int \underbrace{p\left(Z \geq z|y\right)}_{\substack{\text{Vulnerability} \\ \text{model}}} \; \underbrace{\lambda_0 p\left(y\right)}_{\substack{\text{Hazard} \\ \text{model}}} \; dy \tag{22}$$

Equation 22 shows that we actually need the recurrence of water depth y (and eventually other IMs like duration, or velocity) at any site of interest in order to compute the risk. For most modern vulnerability models then, a probability of flood map for a given return period is not useful since it does not provide the required information. The hazard maps, as developed in this work, provide a reliable estimate of the recurrence $\lambda_0 p(y)$ while also accounting for epistemic uncertainties. Specifically the maps show the water depths for a given return period that can be transformed into an exceedance probability as per Eq. 2.

It is important to note that the maps reflect marginal probabilities and do not take into account spatial correlation in the flood process, as they are built by individually computing the hazard curves at each point. In other words, the resulting hazard maps (as in Fig. 9) do not show a real flood event. For this reason, the hazard maps are useful for site-specific hazard, and eventually risk analysis, but not for analyzing spatially-distributed assets. However, these maps were built from an ensemble of simulated flood maps as per the MC simulation approach described in Sect. 2.3. This ensemble is a direct output of the inundation simulator, and thus, can be used to estimate damage for spatially distributed exposure while propagating epistemic uncertainties in the hazard model.

5.3 On the separation of aleatory and epistemic uncertainties

The posterior predictive estimate for the hazard proposed in this work combines in one metric both the aleatory and epistemic uncertainty. While this is important for risk-based decision making, it does not directly help exploiting the main distinction between the two: that epistemic uncertainties can potentially be reduced by further collecting data or improving our models.

Qualitatively speaking, the departure of the posterior predictive curves from the deterministic hazard curve (Fig. 11) can be interpreted as a measure of the relevance of epistemic uncertainties on top of aleatory uncertainties. In this sense, it can be noted that for the case study developed, epistemic uncertainty on the future occurrence of discharge events seems to be more impactful than uncertainty on the inundation model parameters; at least for return periods of 100 years and above.

Developing uncertainty bounds for the hazard curve estimates is the typical way of showing the sensitivity of the hazard estimates on the model's parameters, and their overall relevance. This was done for the hazard curve of discharges in Fig. 6. Doing the same for the flood depths would require to compute a hazard curve for each posterior sample of parameters with the consequent added computational demand. An ensemble of 100 hazard curves was computed and is shown in Fig. 12 together with the posterior predictive (i.e. the mean curve of the ensemble) and the deterministic hazard curve.

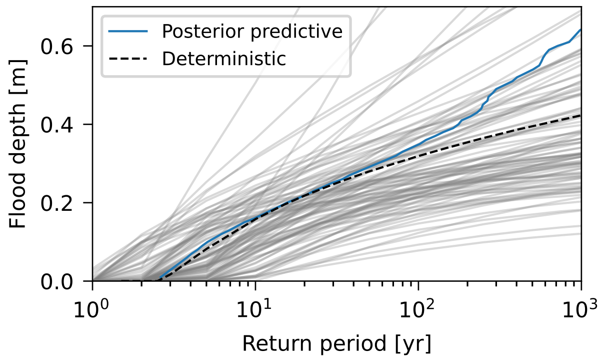

Figure 12. Ensemble of posterior samples of the hazard curves for water depth at location $x = 1.30km, y = 1.05km$

Figure 12 shows that uncertainty in the model's parameters can yield hazard curves that are vastly different from one another, resulting in large uncertainty bounds. This type of analysis can encourage modellers to obtain more data and/or refine the knowledge of the models used in order to reduce it.

5.4 On the modelling methodology

The inclusion of epistemic uncertainties in the computation of flood hazard result in a non-uniqueness of the parameters (and models eventually) used to compute the hazard curves. Thus, there is not a single discharge for a given return period, and there is not a single flood map for any given input discharge. As a result, there is no direct translation between the discharge for a given return period and its corresponding flood map. The water depth for a given return period, as per Eq. 19, has now contributions from all possible discharges which is a more accurate representation of knowledge (and uncertainty).

There are many methodologies to include this uncertainty in the modelling process, and we have chosen to rely on rigorous probabilistic modelling based on a Bayesian framework. The Bayesian methodology allows to consistently include modeller's knowledge and data from different sources into posterior estimates of probabilities. In particular, this work has limited the epistemic uncertainties to uncertainty over the parameters of some models appropriately chosen (i.e. the GPD model for the discharges and the Lisflood simulator for the inundation model), but further prior distributions can be set over different models without affecting the workflow of the method proposed.

Finally, other deeper sources of uncertainty cannot be discarded in risk analysis and include what Spiegelhalter and Riesch (2011) named 'indeterminacy' and 'ignorance'. The former are associated with known limitations in understanding and modelling ability, while the latter is associated with unknown limitations of understanding. Different approaches have been proposed to deal with epistemic uncertainties that are not fully quantifiable including these deeper sources of uncertainties; these were not treated in this work and the reader is referred to (Spiegelhalter & Riesch, 2011; Goldstein, 2011; Beven, 2014).

5.5 On the computational challenges

From a computational perspective, the inclusion of epistemic uncertainties via the Bayesian posterior predictive distribution of IMs means that we need to perform lots of simulations of the inundation model in order to compute the hazard curves. This contrasts with the traditional approach where we only need to run the inundation model once to obtain the 100yrs flood hazard (see Sect. 2.3). Furthermore, the number of simulations needed grow with the return period according to Eq. 20.

This can result in an unfeasible computational burden, particularly for very large return periods. Ongoing advances in computer technology, on the other hand, are producing faster computers that will make simulation approaches like the one proposed, easier to deal as time progresses. In this context, it is expected that this type of numerical analysis will become more common in future research.

In any case, it is also important the implementation of efficient simulation techniques in order to reduce the computational cost. There are mainly two families of techniques that can be implemented in order to reduce the computational time required to compute the desired probabilities: using a more efficient simulation algorithm that targets the desired return period faster, such as importance sampling (Zio, 2013); and making each run of the inundation model faster by using a statistical emulator (Jiang et al., 2020).

6 Conclusions

We propose and develop in this work, a methodology to compute flood hazard curves and maps including epistemic uncertainty on the model's parameters. The framework aims, not only to consistently include this uncertainty into robust hazard estimates, but also to produce useful output for risk analysis and engineering decision-making.

Rather than computing probabilities for an uncertain T-yrs peak discharge event, we propose to compute the flood depth distributions from all possible events within a given time-frame. The posterior distribution of the models' parameters were computed to include epistemic uncertainty, and the average recurrence-rate over these distributions (i.e. predictive posterior distribution) was used as a point estimate for the hazard. As a result, the flood hazard maps developed provide information related to the water-depths for different recurrence rates (or return periods) that can be readily used for further damage analysis including the forward propagation of epistemic uncertainties. The mathematical notation was kept intentionally generic to encourage its use in other natural hazards applications.

The results of the real-world case study developed, showed significant differences between the hazard estimates considering epistemic uncertainties and the classical ones without. In particular, the results show that not considering epistemic uncertainties might underestimate the water-depths hazard at any point in the region of analysis, resulting in less reliable decisions and designs. Furthermore, this tendency aggravates with larger return periods, that tend to be the main focus in many risk analysis. A deeper analysis shows that uncertainty in the prediction of future discharges as a product of a short-length observation series, is the main driver responsible for the underestimation in flood hazard. A similar pattern can be observed in the recurrence curves for the peak discharges in line with some results from the literature (Merz & Thieken, 2005; Romanowicz & Kiczko, 2016; Fawcett & Green, 2018).

On the other hand, the influence of the uncertainty in the inundation model parameters (i.e. floodplain and channel roughness) proved to be less significant. Results show that this uncertainty seems to have a constant effect over return periods, towards the conservative side; that is, it yields lower hazard values than using the classical approach. This seemingly non-intuitive influence requires further studies to understand how it generalizes to different applications and contexts (i.e. different calibration approaches, different computational simulators, different available data), although a potential explanation is the highly skewed shape of the posterior distribution (see Fig. 7) relative to the MAP parameters used in the classical approach.

On the computational aspects, the proposed numerical methodology requires the simulation of thousands of inundation maps. This ensemble can be straightforwardly used in spatially distributed vulnerability models covering a continuous range of return periods. It comes at the expense, however, of a much larger computational burden than the classical approach where only a few runs for selected return periods are needed. Technological advances are expected to rapidly reduce the development-time of this simulation approach, particularly in applications that are practically feasible today. Additionally, this should also encourage researchers to find ways of optimizing the computation of the estimates by using more efficient sampling algorithms or cheaper emulators of the inundation model.

To summarize, we have shown that the inclusion of epistemic uncertainties can significantly modify the estimates of hazard estimates that are later used for risk assessments and damage mitigation decision-making. This is particularly exacerbated for rare and big events (i.e. longer return periods). A framework that can consistently include these in robust probabilistic outputs is, we believe, therefore a major advance for risk analysis.

Appendix A The GPD distribution

This appendix summarizes the probability functions for the Generalized Pareto Distribution (GPD), to avoid confusion in the definition and meaning of the parameters used.

$$p\left(x|\xi,\sigma\right) = \begin{cases} \left(1+\frac{\xi}{\sigma}y\right)^{-\frac{\xi+1}{\xi}}, & \text{if } \xi\neq 0 \\ \exp\left(-\frac{y}{\sigma}\right), & \text{if } \xi = 0 \end{cases} \tag{A1}$$

$$F\left(x|\xi,\sigma\right) = \begin{cases} 1-\left(1+\frac{\xi}{\sigma}y\right)^{-1/\xi}, & \text{if } \xi\neq 0 \\ 1-\exp\left(-\frac{y}{\sigma}\right), & \text{if } \xi = 0 \end{cases} \tag{A2}$$

Where the support of X is $x \geq 0$ when $\xi \geq 0$, and $0 \leq x \leq -\sigma/\xi$ when $\xi < 0$.

Return levels x_{Tr} for this distribution can be computed from,

$$Tr = \frac{1}{1-F\left(x_{Tr}|\xi,\sigma\right)} \tag{A3}$$

Open Research Section

All the data used as input in this work (DEM, daily discharge series, flood extent observation) is publicly available and the sources were mentioned in the manuscript. The open-source software Lisflood-fp was used as inundation simulator (http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/). All the simulations and figures were performed Python 3.X (https://www.python.org/) scripts developed by the authors. All the code used to develop the results and figures mentioned in this manuscript, as well as the input data and the Lisflood-fp binaries used, are publicly available on GitHub https://github.com/mbalbi/epistemic_flood_hazard.git.

Acknowledgments

Jim W. Hall, from Oxford University, is thanked for freely providing the satellite observation for the 1992 flood event at the reach under study. Paul Bates and the LISFLOOD-FP development team are thanked for their support during the simulator learning curve and for the free access to their software, including manuals and tutorial examples, which this case study was drawn from. The research was partially funded by the School of Engineering of the University of Buenos Aires, Argentina, through a Peruilh doctoral scholarship. This project is supported by the National Research Foundation, Prime Minister's Office, Singapore under the NRF-NRFF2018-06 award.

References

Ahmadisharaf, E., Kalyanapu, A. J., & Bates, P. D. (2018, September). A probabilistic framework for floodplain mapping using hydrological modeling and unsteady hydraulic modeling. *Hydrological Sciences Journal*, *63*(12), 1759–1775. doi: 10.1080/02626667.2018.1525615

Apel, H., Merz, B., & Thieken, A. H. (2008, June). Quantification of uncertainties in flood risk assessments. *International Journal of River Basin Management*, *6*(2), 149–162. doi: 10.1080/15715124.2008.9635344

Aronica, G., Bates, P. D., & Horritt, M. S. (2002, July). Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes*, *16*(10), 2001–2016. doi: 10.1002/hyp.398

Aronica, G. T., Candela, A., Fabio, P., & Santoro, M. (2012, January). Estimation of flood inundation probabilities using global hazard indexes based on hydrodynamic variables. *Physics and Chemistry of the Earth, Parts A/B/C*, *42–44*, 119–129. doi: 10.1016/j.pce.2011.04.001

Baker, J., Bradley, B., & Stafford, P. (2021). *Seismic Hazard and Risk Analysis* (1st ed.). Cambridge University Press. Retrieved 2022-06-23, from https://www.cambridge.org/core/product/identifier/9781108348157/type/book doi: 10.1017/9781108425056

Balbi, M., & Lallemant, D. C. B. (2023, March). Bayesian calibration of a flood simulator using binary flood extent observations. *Hydrology and Earth System Sciences*, *27*(5), 1089–1108. doi: 10.5194/hess-27-1089-2023

Beven, K. (2014, March). A Framework for Uncertainty Analysis. In *Applied Uncertainty Analysis for Flood Risk Management* (pp. 39–59). IMPERIAL COLLEGE PRESS. doi: 10.1142/9781848162716_0003

Bezak, N., Brilly, M., & Šraj, M. (2014, May). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, *59*(5), 959–977. doi: 10.1080/02626667.2013.831174

Bharath, R., & Elshorbagy, A. (2018, November). Flood mapping under uncertainty: A case study in the Canadian prairies. *Natural Hazards*, *94*(2), 537–560. doi: 10.1007/s11069-018-3401-1

Bousquet, N., & Bernardara, P. (Eds.). (2021). *Extreme Value Theory with Applications to Natural Hazards: From Statistical Theory to Industrial Practice.* Cham: Springer International Publishing. doi: 10.1007/978-3-030-74942-2

Candela, A., & Aronica, G. T. (2017, March). Probabilistic Flood Hazard Mapping Using Bivariate Analysis Based on Copulas. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *3*(1). doi: 10.1061/AJRUA6.0000883

Castellanos, M. E., & Cabras, S. (2007, February). A default Bayesian procedure for the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, *137*(2), 473–483. doi: 10.1016/j.jspi.2006.01.006

Der Kiureghian, A., & Ditlevsen, O. (2009, March). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105–112. doi: 10.1016/j.strusafe.2008.06.020

Di Baldassarre, G., Schumann, G., Bates, P. D., Freer, J. E., & Beven, K. J. (2010). Flood-plain mapping: A critical discussion of deterministic and probabilistic approaches. *Hydrological Sciences Journal*, *55*(3), 364–376. doi: 10.1080/02626661003683389

Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., & Brath, A. (2013, August). Probabilistic flood hazard mapping: Effects of uncertain boundary conditions. *Hydrology and Earth System Sciences*, *17*(8), 3127–3140. doi: 10.5194/hess-17-3127-2013

Falter, D., Schröter, K., Dung, N. V., Vorogushyn, S., Kreibich, H., Hundecha, Y., ... Merz, B. (2015, May). Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. *Journal of Hydrology*, *524*, 182–193. doi: 10.1016/j.jhydrol.2015.02.021

Fawcett, L., & Green, A. C. (2018, August). Bayesian posterior predictive return levels for environmental extremes. *Stochastic Environmental Research and Risk Assessment*, *32*(8), 2233–2252. doi: 10.1007/s00477-018-1561-x

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition.* CRC Press.

Goldstein, M. (2011, October). External Bayesian Analysis for Computer Simulators. In J. M. Bernardo et al. (Eds.), *Bayesian Statistics 9* (pp. 201–228). Oxford University Press. doi: 10.1093/acprof:oso/9780199694587.003.0007

Grimaldi, S., Petroselli, A., Arcangeletti, E., & Nardi, F. (2013, April). Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling. *Journal of Hydrology*, *487*, 39–47. doi: 10.1016/j.jhydrol.2013.02.023

Hall, J., & Solomatine, D. (2008, June). A framework for uncertainty analysis in flood risk management decisions. *International Journal of River Basin Management*, *6*(2), 85–98. doi: 10.1080/15715124.2008.9635339

Hall, J. W., Manning, L. J., & Hankin, R. K. (2011). Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research*, *47*(5), 5529. doi: 10.1029/2009WR008541

Jiang, P., Zhou, Q., & Shao, X. (2020). *Surrogate Model-Based Engineering Design and Optimization.* Singapore: Springer Singapore. doi: 10.1007/978-981-15-0731-1

Kalyanapu, A., Judi, D., McPherson, T., & Burian, S. (2012, March). Monte Carlo-based flood modelling framework for estimating probability weighted flood risk: Monte Carlo-based flood modelling framework. *Journal of Flood Risk Management*, *5*(1), 37–48. doi: 10.1111/j.1753-318X.2011.01123.x

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(3), 425–464. doi: 10.1111/1467-9868.00294

Kiczko, A., Romanowicz, R. J., Osuch, M., & Karamuz, E. (2013, December). Maximising the usefulness of flood risk assessment for the River Vistula in Warsaw. *Natural Hazards and Earth System Sciences*, *13*(12), 3443–3455. doi: 10.5194/nhess-13-3443-2013

Melchers, R. E., & Beck, A. T. (2018). *Structural reliability analysis and prediction* (Third edition ed.). Hoboken, NJ: Wiley.

Merz, B., & Thieken, A. H. (2005, July). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, *309*(1-4), 114–132. doi: 10.1016/j.jhydrol.2004.11.015

Neal, J., Keef, C., Bates, P., Beven, K., & Leedal, D. (2013, April). Probabilistic flood risk mapping including spatial dependence. *Hydrological Processes*, *27*(9), 1349–1363. doi: 10.1002/hyp.9572

Neal, J., Schumann, G., & Bates, P. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, *48*(11), 1–16. doi: 10.1029/2012WR012514

Nuswantoro, R., Diermanse, F., & Molkenthin, F. (2016). Probabilistic flood hazard maps for Jakarta derived from a stochastic rain-storm generator. *Journal of Flood Risk Management*, *9*(2), 105–124. doi: 10.1111/jfr3.12114

Papaioannou, G., Vasiliades, L., Loukas, A., & Aronica, G. T. (2017, April). Probabilistic flood inundation mapping at ungauged streams due to roughness coefficient uncertainty in hydraulic modelling. *Advances in Geosciences*, *44*, 23–34. doi: 10.5194/adgeo-44-23-2017

Pregnolato, M., Galasso, C., & Parisi, F. (2015, July). A compendium of existing vulnerability and fragility relationships for flood : Preliminary results. In *12th International Conference on Applications of Statistics and Probability in Civil Engineering.* Vancouver, Canada. doi: 10.14288/1.0076226

Romanowicz, R. J., & Kiczko, A. (2016, July). An event simulation approach to the assessment of flood level frequencies: Risk maps for the Warsaw reach of the River Vistula: Event Simulation Approach to Flood Risk Assessment. *Hydrological Processes*, *30*(14), 2451–2462. doi: 10.1002/hyp.10857

Spiegelhalter, D. J., & Riesch, H. (2011, December). Don't know, can't know: Embracing deeper uncertainties when analysing risks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1956), 4730–4750. doi: 10.1098/rsta.2011.0163

Stephens, T. A., & Bledsoe, B. P. (2020, March). Probabilistic mapping of flood hazards: Depicting uncertainty in streamflow, land use, and geomorphic adjustment. *Anthropocene*, *29*, 100231. doi: 10.1016/j.ancene.2019.100231

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual.* Scotts Valley, CA: CreateSpace.

Vickery, P. J., Lin, J., Skerlj, P. F., Twisdale, L. A., & Huang, K. (2006, May). HAZUS-MH Hurricane Model Methodology. I: Hurricane Hazard, Terrain, and Wind Load Modeling. *Natural Hazards Review*, *7*(2), 82–93. doi: 10.1061/(ASCE)1527-6988(2006)7:2(82)

Zahmatkesh, Z., Han, S., & Coulibaly, P. (2021, January). Understanding Uncertainty in Probabilistic Floodplain Mapping in the Time of Climate Change. *Water*, *13*(9), 1248. doi: 10.3390/w13091248

Zio, E. (2013). *The Monte Carlo Simulation Method for System Reliability and Risk Analysis.* London: Springer London. doi: 10.1007/978-1-4471-4588-2