https://doi.org/10.5194/egusphere-2025-3310
Preprint. Discussion started: 17 July 2025

Global sensitivity analysis of large-scale flood loss models

Francesca Pianosi[1,2], Georgios Sarailidis[1,3], Kirsty Styles[3], Philip Oldham[3], Stephen Hutchings[3], Rob Lamb[4,5], Thorsten Wagener[6]

[1]School of Civil, Aerospace and Design Engineering, University of Bristol, Bristol, United Kingdom
[2]Cabot Institute, University of Bristol, Bristol, United Kingdom
[3]JBA Risk Management Limited, Skipton, United Kingdom
[4]JBA Trust, Skipton, United Kingdom
[5]Lancaster Environment Centre, Lancaster University, United Kingdom
[6]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

Correspondence to: Francesca Pianosi (francesca.pianosi@bristol.ac.uk)

Abstract. Flood loss models are increasingly used in the (re)insurance sector to inform a range of financial decisions. These models simulate the interactions between flood hazard, vulnerability and exposure over large spatial domains, requiring a range of input information and modelling assumptions. Due to this high level of complexity, evaluating the impact of uncertain input data and assumptions on modelling results, and therefore the overall model "acceptability", remains a very complex process. In this paper, we advocate for the use of global sensitivity analysis (GSA), a generic technique to analyse the propagation of multiple uncertainties through mathematical models, to improve the sensitivity testing of flood loss models and the identification of their key sources of uncertainty. We discuss key challenges in the application of GSA to large-scale flood models, propose pragmatic strategies to overcome these challenges, and showcase the type of insights that can be obtained by GSA through two proof-of-principle applications to a commercial model, JBA Risk Management's flood loss model, for the transboundary Rhine River basin in Europe, and Queensland in Australia.

1 Introduction

Floods are among the most widespread risks to human lives, infrastructure and property worldwide. Between 2000 and 2019, floods affected 1.65 billion people, accounting for 41% of all people affected by natural disasters (CRED-UNDRR, 2020) or what are increasingly described as "human-made disasters" (see Otto and Raju, 2023). Floods were responsible for 9% of all disaster-related deaths and caused 22% (US$ 651 billion) of all disaster related damages (CRED-UNDRR, 2020). Their impact continues to grow globally (Kreibich et al., 2022) and flood risk is set to increase even further under the combined effect of more extreme weather and urban development into flood-prone areas (Merz et al., 2021). Increasing the penetration of flood insurance and reducing the level of uninsured flood risk – the "protection gap", or proportion of risk that remains uninsured – will be crucial for adaptation to such growing levels of risk (Tesselaar et al 2020; Aerts et al 2024).

To close the protection gap, insurers need to quantify the risks associated with underwritten policies. To do so, they increasingly rely on probabilistic models of natural hazard impacts, which enable them to estimate expected losses in a much more robust way than by extrapolating from limited and- often inaccurate historical loss data. These models, known in the industry as "loss models" or "catastrophe models", are used to inform a range of decisions, from pricing an individual policy to optimizing reinsurance at company level (Mitchell-Wallace et al., 2018). They calculate the total value of expected losses that a insurer may expect to face from a given region over a specified time horizon. This is achieved by combining three core components: i) the hazard module, which estimates the spatial distribution of flood intensities (a measure chosen to represent the strength of the potential for harm, for example flood depths); ii) the exposure module, which locates insured assets and their values; and iii) the vulnerability module, which quantifies the damages to the exposed assets caused by the floods (Grossi and Kunreuther, 2005).

While loss models are continuously revised and upgraded, their evaluation remains an open challenge. Currently there are no prescribed or standardized validation tests[1] for acceptance of a particular flood loss model in the insurance industry (Franco et al., 2020) despite the many limitations in both our understanding and our data around

---

[1] For example, the EU Directive 2009/138/EC "on the taking up and pursuit of the business of Insurance and Reinsurance (Solvency II)" while stating that "*Insurance and reinsurance undertakings shall have a regular cycle of model validation*" (Article 124) does not prescribe any specific approach or test for validation. (https://eur-lex.europa.eu/eli/dir/2009/138/oj)

flood processes and impacts (Lighthill Risk Network, 2019). Such lack of common evaluation approaches is not unique to the insurance sector as "*validation is perhaps the least practised activity in current flood risk research and flood risk assessment*" (Molinari et al., 2019).

In principle, there are different approaches to model evaluation, which broadly fall into two categories (Mertz et al., 2024): "outcome-based" evaluation, where the emphasis is on the modelling outputs, i.e. the loss predictions, and on checking whether they are consistent with either historical observations (hence establishing the model *accuracy*) or other models' predictions (*benchmarking*); and "process-based" evaluation, which focuses on checking the quality of the modelling process itself, rather than its outcome (Figure 1). An example of the latter is sensitivity analysis (or "sensitivity testing"), which involves changing the model's input data, internal parameters and/or assumptions, and checking that the output response to such changes align with the user expectations (in terms of direction of change, relative magnitude etc.). The aim is to establish the model's *fidelity*, i.e. the consistency of the model behaviour with the user's understanding of the simulated processes (Wagener et al. 2022) and use that as a basis for its *legitimacy*. Another aim of sensitivity analysis is to identify the inputs/assumptions that mostly control the output uncertainty and hence prioritise efforts towards improving the model precision (Figure 1). For example, if a particular input is shown to have little influence on the output uncertainty, investing in acquiring better quality data for that input may be unjustified.

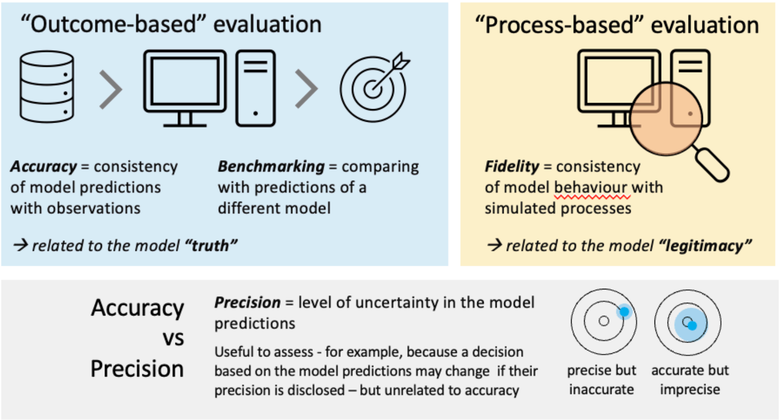

Figure 1 – Conceptual approaches to model evaluation: outcome-based (left) and process-based (right). Sensitivity analysis contributes to the latter, by providing a mechanism to assess model fidelity. While shown here in the context of flood loss models, the framework applies more broadly.

In practice, testing flood loss model accuracy can be very difficult because historical loss data are scarce and, even when available, the associated hazard intensity and exposure may not be fully known. Therefore, the model's inability to reproduce observed losses may be equally attributed to a model deficiency or to gaps and errors in the input data[2]. As for sensitivity analysis, it is routinely used but often limited to investigating the effects of varying input data, typically the exposure portfolio. This is due to the "black-box" nature of most commercial loss models, which do not allow the user to change the parameters and assumptions embedded in the hazard and vulnerability components. Indeed, a recent paper on the state-of-art of insurance loss modelling concluded that "*the assessment of uncertainty all along the modelling chain constitutes the loss modelling framework's notable shortcoming and the one that requires further investigation*" (Déroche, 2023). The need for more transparency in modelling has also been identified as a key priority for improvement of climate risk assessment in relation to the rapidly emerging regulation for climate-related financial disclosure (Arribas et al. 2022). From a methodological perspective, another problem with the sensitivity testing prevailing in the industry is that it follows a "local" approach (Figure 2) whereby output sensitivity is assessed by varying inputs "one-at-the-time" from a default value ("baseline"). There are two major limitations to this approach. First, it does not capture interactions between model inputs that may amplify or dampen the effects of individual input variations – a common feature of complex, non-linear models. Second, its results are conditional on the chosen baseline, leaving the question open of how different the sensitivity assessment would be if using a different baseline (Figure 2).

---

[2] Some have argued that this problem is intrinsic to numerical modelling of earth systems, due to their open nature, and thus establishing the model "truth" is inherently impossible. Model evaluation should instead focus on establishing that the model "does not contain known or detectable flaws and is internally consistent" (Oreskes et al. 1994).

In this paper, we argue that flood loss models should be made more transparent to their users and that their sensitivity testing should be made more robust. In particular, we suggest that local, one-at-the-time sensitivity analysis should be replaced by a global sensitivity analysis (GSA) approach. GSA is a methodology to systematically investigate how simultaneous variations in the inputs of a model (including parameters, forcing inputs, or initial conditions) affect the model outputs. A typical result of GSA is a set of sensitivity indices each measuring the relative contribution of every varied input to the uncertainty in each model output (Saltelli et al., 2000). Note that, different to local approaches, in GSA all inputs are varied simultaneously within their variability space without any baseline being required (Figure 2).

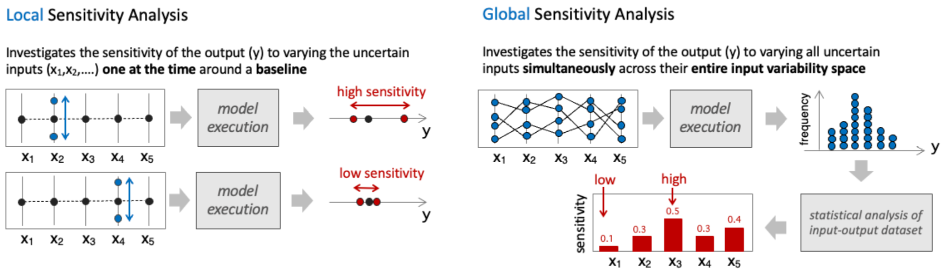

Figure 2 – Approaches to Sensitivity Analysis: local vs global.

While GSA is increasingly used for investigating uncertainty propagation in hydrological modelling (e.g. Song et al. 2015) and, to a lesser extent, flood inundation modelling (e.g. Savage et al 2016), applications to flood loss models have been few and limited to relatively small spatial domains (e.g. de Moel et al (2014) for a 40 by 60 km region in the Netherlands) or simplified strategies to understand flooding potential at global scale (Devitt et al., 2023). This may be due to a lack of awareness of the advantages of global over local sensitivity analysis – a common issue across many modelling sectors (Saltelli et al. 2019) – but also to the specific difficulties that would be encountered in applying GSA to large-scale flood loss models. These include three key challenges: the potentially very large number of input uncertainties that one could analyse; the “uncertainty about the uncertainties”, i.e. the fact that, for many of those input uncertainties, the analyst may not be able to state what variations would be reasonable and what would not; and the computational burden of running the model hundreds or even thousand times, when each model execution may take from several minutes to hours.

The goal of this paper is to discuss how we can begin to tackle these challenges and showcase the type of results that could be obtained through two proof-of-principle applications to a commercial model – JBA Risk Management’s (hereafter JBA’s) flood loss model – applied to two large domains: the transboundary Rhine River basin in Europe and Queensland in Australia. In the next two sections, we will describe the key working principles of large-scale flood loss models and of the GSA methodology for context. We will then discuss in more detail the key challenges in applying GSA to large-scale flood loss models and possible solutions. Last, we will show the results of two illustrative applications of GSA to JBA’s flood loss model, a large-scale commercial flood loss model used by insurers and reinsurers for local- and portfolio-level risk assessments (further information about the model can be found in previous studies where the model and underpinning data have been exploited, e.g. Kay et al., 2018; D’Ayala et al., 2020; Becher et al., 2023; Darlington et al., 2024, Galloway et al., 2025). The first application to the Rhine River basin in Europe will showcase how GSA can be used to understand the relative importance in hazard, vulnerability and exposure uncertainties in a consistent way across a large spatial domain, as a way to identify priorities for model improvement. The second application to Queensland state in Australia will showcase how GSA provides a framework to consistently compare present-day uncertainties in vulnerability and exposure with uncertainty about future climate.

2. High-level description of a large-scale Flood Loss Model

Flood loss models use multiple components to estimate flood risk in terms of monetary losses induced by floods (Figure 3). In essence, the loss $L_{s,t}$ for a single exposed asset s and a single flood event k is given by

$$L_{s,k} = EV_s \times DR(FD_{s,k}), \quad (1)$$

where EV_s is the asset's value, $FD_{s,k}$ is the flood depth of the event k at the asset s, and $DR(.)$ is the damage ratio, given by a vulnerability curve (or "depth-damage function") that returns a fraction between 0 and 1 as a function of the flood depth. For a given spatial domain, Equation (1) is applied for all the exposed assets and all flood events in a synthetic event catalogue spanning a long time horizon (10,000 years in our study). Such a catalogue enables estimation of the full distribution of annual losses, including both frequent and rare events. In this context, probabilities are often expressed in terms of a return period (RP), where, for instance, a 200-year event corresponds to a probability p=0.005 via the relationship RP=1/p.

Insurance sector users are typically interested in two summary statistics of the annual loss distribution: the average annual loss (AAL) and the loss exceedance (LE(p)), which indicates the annual loss that is exceeded with probability p (e.g., the 200-year loss). Loss statistics are typically aggregated spatially, such as into CRESTA zones (https://about.cresta.org/), which are standardised regions used in the insurance industry (Grossi and Kunreuther, 2005).

The *AAL* for a CRESTA zone *C* is calculated as

$$AAL_C = \sum_{s\in C} \sum_k p_{s,k} L_{s,k}, \qquad (2)$$

where $p_{s,k}$ is the annual probability of event k affecting asset s.

To quantify tail risk, the loss exceedance curve (EP curve) is constructed by simulating total annual losses $L_{C,Y}$ for each zone *C* and year *Y* in the catalogue, then empirically estimating their exceedance probabilities. Formally, the exceedance value $LE_C(p)$ is defined by

$$LE_C(p) = X \quad such\ that \quad \mathbb{P}\big(L_{C,Y} > X\big) = p, \qquad (3)$$

This is the p-th upper quantile of the empirical distribution of annual total losses across the simulated years. Each annual total loss that builds the distribution is calculated as

$$L_{C,Y} = \sum_{s\in C} \sum_{k\in K_Y} L_{s,k}, \qquad (4)$$

where K_Y is the set of events in year *Y*.

Interpolation of flood depths. The flood depth at each location and for each flood event ($FD_{s,k}$) could, in principle, be estimated by running a flood inundation model against the river flow and rainfall forcing inputs characterising that event. However, running the inundation model for millions of different flood events at large scales is computationally prohibitive. To overcome this problem, the flood inundation model is first used to derive a set of flood hazard maps for a limited set of return periods and then flood depths for any other return period are obtained by interpolating through the flood depths of the appropriate hazard maps (see Fig. 3). The JBA flood loss model used in this paper uses 6 hazard maps with return periods of 20, 50, 100, 200, 500 and 1,500 years.

Flood events set. The flood events considered for the calculations of AAL and LE curves are pre-computed and stored in a flood event set, which contains millions of plausible flood events representative of a 10,000-year sample. Each event comprises an event severity per location, defined as a return period of flow or precipitation at that location. Event sets are typically generated using both historical records of observed events and synthetic events generated by a mix of physically-based and statistical models (Grossi and Kunreuther, 2005, Lamb et al., 2010, Keef at al., 2013).

Vulnerability curve. The flood depth at each location and for each flood event ($FD_{s,t}$) is then used to calculate the damage ratio through a vulnerability curve ($DR(.)$ in Eq. (1)). Out of several different possible shapes (for a review and comparison of vulnerability curves see e.g. Jongman et al. (2012) or Cammerer et al. (2013)), the JBA flood loss model used in this paper employs step functions (see Fig. 3), which corresponds to assuming that sufficiently similar flood depths all lead to the same damage ratio.

Exposure portfolio. Finally, the estimated damage ratio value is multiplied by the value of the exposed asset ($EVs\ $) which is retrieved from the exposure portfolio. The exposure portfolio contains all the relevant information regarding the exposed assets (e.g. location, total value etc.). Exposure portfolios can be provided at different levels

https://doi.org/10.5194/egusphere-2025-3310
Preprint. Discussion started: 17 July 2025

of spatial resolution. The most detailed portfolios include information about the exact location (expressed with latitude and longitude coordinates) of every asset, along with details about the building characteristics and insured value. Lower resolution portfolios include information about the total number of buildings and their 'typical' characteristics within a certain spatial unit. Such spatial unit can be a CRESTA (Catastrophe Risk Evaluation and Standardizing Target Accumulations) zone or a province or, in the case of most aggregate portfolios, an entire country. The model then applies a disaggregation algorithm to locate the assets in each spatial unit based on other higher-resolution data (typically population density).

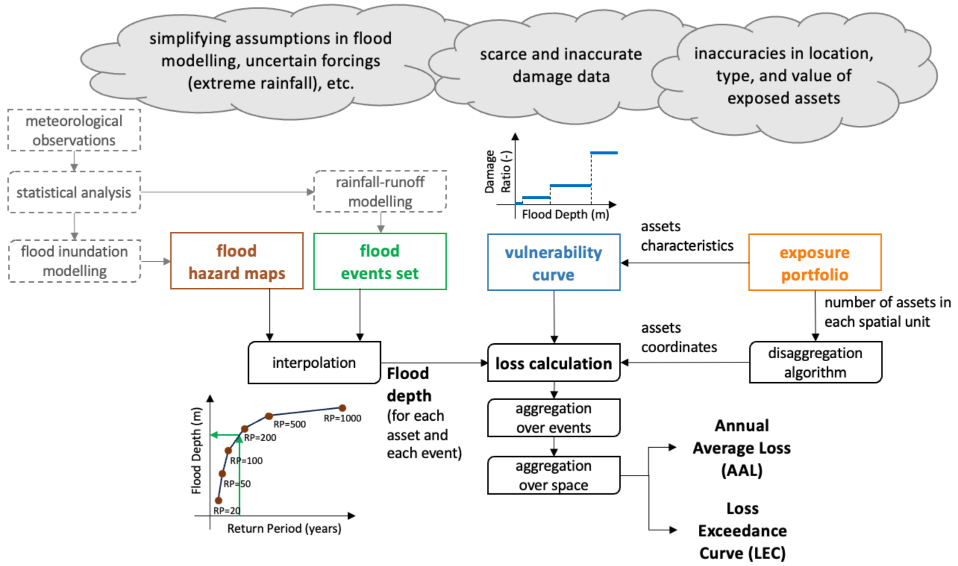

Figure 3 - The typical structure of flood loss models used in the (re)insurance sector and their main sources of uncertainty. Coloured boxes highlight the level in the modelling chain where we will apply input variations in our 'backward-from-the-end' Global Sensitivity Analysis (see Section 4.1).

3 The global sensitivity analysis (GSA) approach

GSA is a methodology to systematically investigate how variations in the inputs of a model (including parameters, forcing inputs, or initial conditions) propagate into variations in the model outputs.

In brief, GSA comprises four key steps (Figure 4). First, the analyst selects the uncertain model inputs that will be subject to the analysis and characterizes their uncertainty. Note that uncertain inputs can be both numerical, such as forcing data and parameters, and non-numerical, such as discrete modelling choices and assumptions. Second, a prescribed number of input combinations are generated through statistical sampling (e.g. random sampling or Latin hypercube sampling) from the input distributions defined in the first step. Third, the model is run for each input combination, and one or more output metrics are calculated for simulation. Here, the output metrics are the average annual loss and the loss exceeded with given probabilities of Eq. (2) and (3). Fourth, the input–output dataset is analysed to quantify the uncertainty in the output metrics, e.g. in the form of a frequency distribution or a set of quantiles, and to quantify the relative contribution of the uncertain inputs to the output uncertainty, e.g. in the form of sensitivity indices. Sensitivity indices typically vary between 0 and 1, with higher values denoting greater contribution to output uncertainty.

The definition of the sensitivity indices, and the calculation procedure to approximate their values from the input–output dataset, vary depending on the chosen GSA method. For example, some methods use correlations between inputs and output samples to measure sensitivity, while other methods measure sensitivity to an input through the reduction in output variance (or some other statistic) when fixing that input. More details about different GSA methods and the implementation of the above four steps can be found in, e.g. Saltelli et al. (2000) and Pianosi et al (2016).

https://doi.org/10.5194/egusphere-2025-3310
Preprint. Discussion started: 17 July 2025

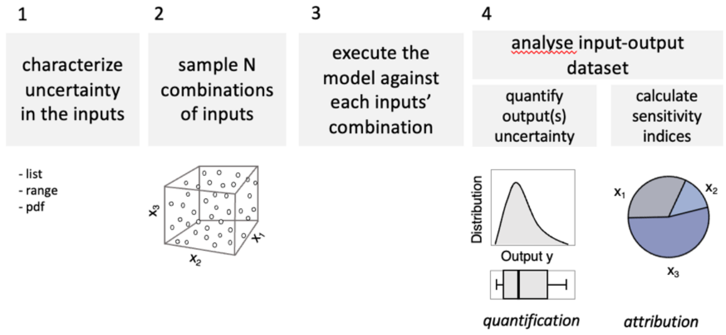

Figure 4 - The steps of global sensitivity analysis for the quantification and attribution of uncertainty in model output(s).

4. Challenges in applying GSA to large-scale flood loss models and possible solutions

When applying GSA to a complex flood risk model like the one described in Section 2, several challenges arise. In the following sections, we discuss these challenges and possible ways forward, which are then demonstrated in our example application.

4.1 Reducing the number of input uncertainties

The first challenge in the application of GSA to a complex modelling chain like that underpinning a flood risk model is that the number of input uncertainties that could be propagated and analysed is potentially very large. This is problematic as it increases the analytical burden of characterising input uncertainties (step 1 in Figure 2) and the computational burden of executing the model (step 3), as the number of input combinations required to approximate sensitivity indices grows quickly with the number of uncertain inputs (Sarrazin et al. 2016).

To overcome this challenge, here we suggest that instead of propagating all possible uncertainties "forward-from-the-start" of the modelling chain, one can follow a structural dimension reduction through a "backward-from-the-end" approach and perturb only the inputs to the last component of the modelling chain, that is, the loss calculation (Figure 3). By doing so one only has four input uncertainties: (i) the flood hazard maps used for interpolation of flood depths ($FD_{s,k}$ in Eq. 1); (ii) the event set; (iii) the vulnerability curve $DR(.)$; and (iv) the exposed assets value EV_s.

Framing the GSA in this way makes it computationally tractable while still fulfilling the key aim of determining the dominant control of the uncertainty in the flood losses. Once this has been established, the root causes can be further analysed, if needed, by going "one step up" the modelling chain and conducting a new GSA of the model component responsible for the dominant uncertainty. In other words, the idea is that instead of trying to conduct one comprehensive analysis of the entire modelling chain, which may be computationally intractable, one could start with a smaller analysis of the end of the modelling chain and use the results to focus subsequent rounds of equally tractable analyses on the components that have been proven to matter most.

4.2 Generating physically-plausible input samples

Having embraced a "backward-from-the-end" approach, a subsequent challenge is how to produce random yet physically plausible samples for complex spatially distributed input uncertainties such as flood hazard maps. For example, perturbing flood depths at individual locations independently from one another may produce physically implausible hazard maps, while defining spatially consistent perturbation is possible in principle but difficult in practice.

One approach to maintain spatial consistency could be to put together a set of hazard maps produced by different inundation models (or models with different set-up choices) and randomly sample from this set with uniform discrete probability (i.e. where each possible choice gets the same probability of being sampled). A similar idea has been used in the context of heatwaves rather than flood risk assessment (Dawkins et al., 2023).

A second option, which we will use in the Rhine River Basin application of Sec. 5.1, is to exploit the fact that each hazard map and flood event in the event set is associated with a return period (see Sec. 2) and thus randomly sampling the return period leads to retrieving a different map/event and thus producing different flood depths for each model execution, while maintaining spatial consistency over the simulated domain.

A third option to perturb flood depths consistently across the spatial domain, which we will use in the Queensland application (Sec. 5.2), is to specify a distribution $F_s(.)$ for the flood depth at each location, then generate one perturbation factor (p) via uniform sampling over [0,1], and finally obtain a perturbed flood depth at each location by inversion of the cumulative location-specific flood depth distribution: $FD_s = F_s^{-1}(p)$.

4.3 Characterising poorly bounded input uncertainties

Another challenge in applying GSA to large-scale loss models is the characterisation of the selected input uncertainties. In fact, the scarce quantity and quality of data on, for example, extreme floods and flood damages, makes it difficult to rigorously specify probability distributions for most input uncertainties (Apel et al. 2024). This is critical as obviously the results of the GSA – i.e. the relative importance attributed to every input uncertainty – strongly depend on the chosen distributions (e.g. Paleari and Confalonieri (2016)).

To handle this problem, we propose to conceptualise the probability distributions of the inputs not as statements of what is likely, but rather as statements of what cannot be excluded. In practice, this means using uniform distributions with very wide ranges: the choice of uniform distributions translates our inability to discriminate between probabilities of different input values; the use of wide ranges allows for testing input values that may be largely different from “default” set-up, but still physically plausible.

One can then perform a “sensitivity analysis of the sensitivity analysis” and investigate how sensitivity results are affected by the ranges’ definition. For example, one could reduce the range of the input that was found to be dominant and check by how much that range needs reducing before the input stops being dominant. In other words, instead of asking: “*what is the relative importance of input X given its (assumed) uncertainty distribution*?” one can ask: “*how much do we need to reduce uncertainty in input X before it becomes unimportant (or another input becomes more important)*?”

4.4 Reducing the number of model executions

The last challenge is that of keeping the overall computational burden of the analysis within the limits of available resources. While continuous increase in computing power may ease this problem in the long-term, the requirement of executing the model hundreds or potentially thousands of times for the forward propagation of input uncertainties (step 3 in Fig. 4) remains a bottleneck to the application of GSA to large-scale flood loss models.

One way to mitigate the issue is to use a GSA method as frugal as possible, such as the method of Morris (1991) or to use a surrogate modelling approach such as polynomial chaos expansion (Sudret, 2008). Here, we use the PAWN method (Pianosi and Wagener, 2015) which has been shown to effectively estimate sensitivity indices with a relatively low number of model evaluations (Pianosi and Wagener, 2018). In this method, the sensitivity of the output y to the input factor x_i is quantified by measuring the distance between the unconditional cumulative distribution function (CDF) of y that is obtained by varying all inputs simultaneously, and the conditional CDF obtained when all inputs vary but x_i.

Operationally, given the input-output samples dataset, the PAWN sensitivity indices $\widehat{S_i}$ are approximated by splitting the range of variation of each input uncertainty x_i into n equally-spaced intervals I_k, using

$$\hat{S}_i = mean_{k=1,\dots,n}\, KS(I_k), \qquad (5)$$

$$KS(I_k) = max_y \left|\hat{F}_y(y) - \hat{F}_{y|x_i}(y|x_i \in I_k)\right|, \qquad (6)$$

where $\hat{F}_y(y)$ and $\hat{F}_{y|x_i}(y|x_i \in I_k)$ are the unconditional and conditional CDFs of the output y and KS is the Kolmogorov-Smirnov statistic that measures the distance between the CDFs. The sensitivity indices vary between 0 and 1: the higher the value of the sensitivity index the more sensitive the output to the input uncertainty. The code to implement PAWN in Matlab, R and Python is available in the SAFE Toolbox (Pianosi et al. 2015; SAFE 2025).

To further reduce the number of model evaluations required by the GSA, we also propose a simple bootstrap-based approach to robustly determine the dominant uncertainty even if using a particularly small sample size. With this approach, the calculation of the PAWN sensitivity indices is repeated for a prescribed number of times using bootstrap resamples of the original input-output dataset. For each bootstrap resample, the dominant input uncertainty is identified as the one with highest sensitivity index. Then, for each input uncertainty, we can calculate the frequency with which it was identified as dominant across bootstrap resamples. Last, input uncertainties are ranked in order of decreasing frequency of being dominant. If the input uncertainty with highest frequency exceeded the second highest by at least a prescribed difference (0.3 in our case), we deem that input uncertainty as dominant. If instead the difference between the first and second highest frequencies is less than 0.3, we deem both the first and second ranked as dominant uncertainties.

The key point here is recognising that sensitivity indices are just a means to achieving the actual relevant result, which is the ranking of input uncertainties from most to least important, and that a sufficiently robust ranking can be obtained even if indices are not precisely approximated due to the small sample size used (Sarrazin et al. 2016).

5. Application examples

In this section, we showcase some of the solutions for applying GSA to a large-scale flood loss model as described in Sec. 4, using two application examples.

The first application is the flood loss model of the Rhine River basin in Europe, covering a spatial extent of approximately 185,000 km^2 and including major European cities of eight different countries (Switzerland, Liechtenstein, Austria, France, Germany, Belgium, Luxembourg and the Netherlands). An exposure portfolio was constructed at CRESTA level using market-informed values. We analysed the propagation of uncertainty in the value of residential buildings, vulnerability curves and the return period of the flood events and hazard maps, with the goal of identifying the dominant controls of uncertainty in the loss predictions across the spatial domain, and thus understand which model component – hazard, vulnerability or exposure – should be prioritized in future efforts for reducing uncertainty, and where.

The second application is the flood loss model of Queensland in Australia, covering a spatial extent of over 1,700,000 km^2 and including over 300,000 assets. Here, we focussed on understanding to what extent the quality of the exposure portfolio matters with respect to other sources of uncertainty – including uncertainty in future climate. The motivation is that while substantial attention has been devoted to improving the hazard component of flood loss models (including incorporating climate change), exposure data may vary widely in granularity and reliability. For instance, reinsurers frequently receive aggregated portfolios from insurers, which lead to loss of critical information about building location and characteristics. To perform our analysis, we created a coordinate level portfolio informed by the PERILS Industry Exposure & Loss Database (https://www.perils.org/) and then created three more portfolios at increasing level of aggregation. We then used GSA to quantify the relative importance of using one portfolio over the other with respect to other uncertainty sources and modelling choices – including the choice between different climate-conditioned event sets.

5.1 Rhine River basin case study

To identify the minimum and maximum plausible values for each input uncertainty, we reviewed the scientific literature on flood risk in the Rhine River basin and identified past studies that quantified the variability of flood return periods, depth-damage ratios and residential building values for either specific locations or sub-regions within the basin. We synthetised this information by taking the average of minimum and maximum values found in past studies, as summarised by Sarailidis (2023, Chapter 3) and summarised in Table 1.

Table 1 - Input uncertainties analysed in the GSA demonstration for the Rhine River basin, their characterisation and underpinning assumptions and implications.

Input uncertainty	**Sampled variable**	**Characterization of uncertainty**	**Aspects captured by the characterization**	**Aspects not captured by the characterization**
Exposure portfolio	EV: exposed value (€/building)	Country-specific ranges (*)	Uncertainty in the asset value	Uncertainty in the location and type of asset
Hazard maps	RP: return period of hazard maps (years)	+/- 50% ranges around default values (***)	Uncertainty in the forcings of the flood inundation model used to produce hazard maps	Parametric and structural uncertainty in the flood inundation model
Flood event set	RP: return period of flood events (years)	+/- 50% ranges around original values	Uncertainty in the assigned frequency of each individual event	Parametric and structural uncertainty in the event generation model
Vulnerability curve	DR: Damage Ratio (-)	Depth-dependent ranges (**)	Parameter uncertainty in chosen vulnerability curve	Structural uncertainty around appropriate choice of the vulnerability curve (which shape and which independent variables – e.g. beyond flood depth)

(*) Country-dependent ranges of exposed value (thousands €/building):		(**) Depth-dependent ranges of damage ratio (-):		(***) Ranges of return periods around original values (years):	
Country	EV_{min}–EV_{max}	*Depth range (m)*	$DR_{min} - DR_{max}$	*Original (years)*	RP_{min}–RP_{max}
Switzerland	64–119	0–0.2	0–0.11	2	1–3
Liechtenstein	69–130	0.2–0.5	0–0.14	5	2.5–7.5
Austria	66–128	0.5–0.7	0–0.18	10	5–15
France	66–129	0.7–1	0–0.22	15	7.5–22.5
Germany	61–116	1–1.7	0.01–0.3	20	10–30
Luxembourg	57–110	1.7–2.34	0.03–0.34	50	25–75
Belgium	58–114	2.34–3	0.08–0.42	100	50–150
Netherlands	61–119	>3	0.22–0.68	200	100–300
				500	250–750
				1500	750–2250

We used Latin hypercube sampling to generate 400 combinations of the 4 uncertain inputs and executed the model against each of these input combinations. Sensitivity indices were calculated using the PAWN method and dominant input uncertainties identified with the bootstrapping-based approach described in Sec. 4.3.

Figure 5 shows an example of the type of results that can be produced by the GSA. The map on the left shows the Rhine River basin with spatial units (CRESTA zones) coloured according to the dominant input uncertainty on average annual loss (AAL) in that units (units coloured in white did not include any exposed assets). The bar plot summarises the total number of units in which each input uncertainty is dominant for AAL estimation. It shows that damage ratio is the dominant uncertainty almost everywhere across the spatial domain. This result aligns with previous sensitivity analyses of flood loss models for smaller regions of the Rhine River basin (e.g. De Moel et al 2014) as well as an industry report containing opinions of practitioners (Lighthill Risk Network 2019) that identified uncertainty in the vulnerability curves as a key contributor to uncertainty in AAL estimates and a priority to improve precision of flood loss estimates. The right panel in Figure 5 further shows that the uncertainty in the vulnerability curves (i.e. the variability ranges of the damage ratios) needs to be reduced by 50% to see the second most important input uncertainty, i.e. the flood event set, become dominant in as many spatial units.

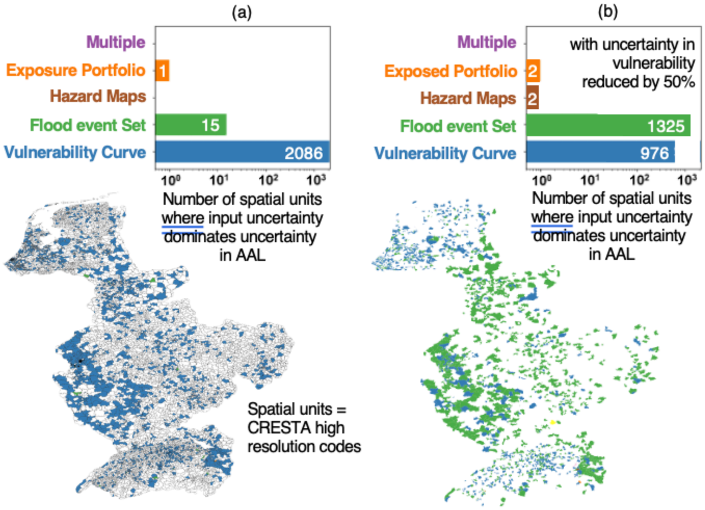

Figure 5 - Example results of a global sensitivity analysis of the Rhine River Basin showing dominant input uncertainties for average annual loss across the spatial domain for (a) original uncertainty ranges as in Table 1 and (b) reduced uncertainty ranges for the vulnerability curves.

Figure 6 shows a similar analysis for another output metric typically used by (re)insurers: the loss exceeded (LE) with six different return periods. It shows that, for small return periods (RP = 10 years or 20 years), the LE uncertainty is dominated by either the vulnerability curve or the event set uncertainties. At larger return periods, the vulnerability curve becomes the dominant uncertainty in more spatial units. This can be explained by the fact that losses with low return periods are produced by small and frequent flood events where the impact of localized features is greater, and thus the uncertainty in flood depths and/or individual asset value is more important. Loss levels with large return periods instead are caused by large inundations where many assets experience similar flood depths and thus flood loss uncertainty is mainly driven by the uncertainty in the damage ratios.

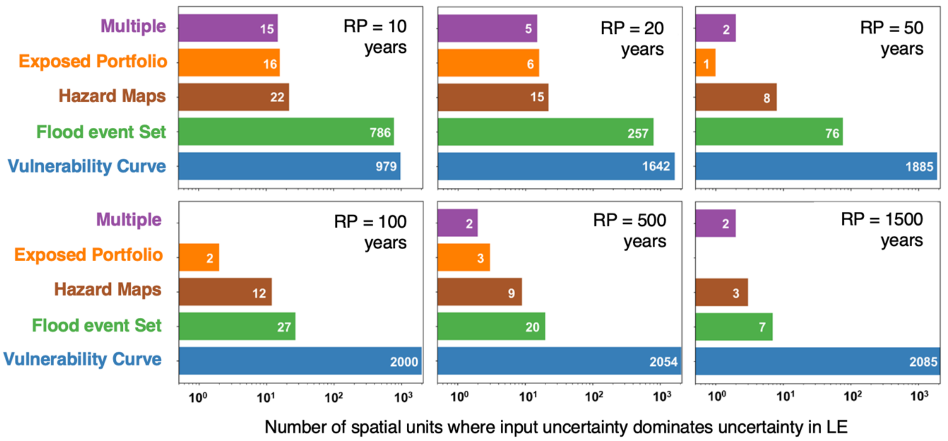

Figure 6 - Example results of a global sensitivity analysis of the Rhine River basin showing dominant input uncertainties on loss exceeded (LE) with different return periods (RPs).

These findings are somewhat consistent with previous literature. For example, in their analysis of the variance of the European flood model, Kaczmarska et al. (2018) also found that the vulnerability component is increasingly important at greater return periods of exceeded losses. Local-scale studies at different reaches (e.g. Apel et al. 2008) or subbasins (Winter et al 2018) of the Rhine River basin instead found slightly contrasting results: the former found that LE at low return period is dominated by uncertainty in frequency of flood events and the latter found that it is still dominated by uncertainty in vulnerability.

5.2 Queensland case study

Table 2 describes the input uncertainties considered in the Queensland case study. Similarly to the Rhine River basin case study, we considered uncertainties in the vulnerability curve parameters and in the hazard maps (though this time using a different approach to perturbing the flood depths – see Table 2 and Sec. 4.2). The key differences with the previous case studies are two. First, for the event set, we do not perturb the characteristics (e.g. return periods) of the events in one set, but we sample from five different sets, each one built under a different climate change scenario. Second, for the exposure portfolio, again we do not perturb characteristics (e.g. building value) of the assets in one portfolio but we sample from different portfolios built through intersecting two different characteristics: the level of spatial aggregation (from low to high we consider: coordinate level, CRESTA, province, and state level) and the level of detail of the building characteristics (we consider two cases: complete information including the line of business, number of storeys and height of first floor for each building, and a partial information including the line of business only). These portfolios are representative of different levels of information that may be accessible to different users of flood loss models in the risk analysis sector. For example, insurers might have coordinate-level portfolios of their clients, whereas reinsurers may only have access to CRESTA-level portfolios, and an international risk reduction agency may only have access to a province-level or even state-level portfolios.

Table 2 - Input uncertainties analysed in the GSA demonstration for Queensland, their characterisation and underpinning assumptions and implications.

Input uncertainty	Sampled variable	Characterization of uncertainty	Aspects captured by the characterization	Aspects not captured by the characterization
Exposure portfolio	Identifier of the aggregation level to be used	List of 4 options: portfolio at (i) coordinate level, (ii) CRESTA level, (iii) province level, (iv)	Uncertainty in the assets location	Uncertainty in the assets value

		state level		
	Identifier of the level of details about building information	List of 2 options: (i) complete (line of business, number of storeys and first floor height) or (ii) line of business only	Uncertainty in the assets characteristics	
Hazard maps	Perturbation factor (-)	Varies between 0 and 1.	Any source of uncertainty affecting flood depth	Any source of uncertainty affecting flood inundation extent
Event set	Identifier of the climate-conditioned event set to be used	List of five options: (i) baseline, (ii) under RCP4.5 scenario in 2050 and (iii) in 2080, (iv) under RCP8.5 in 2050 and (v) in 2080	Uncertainty in nature of climate change (e.g. weather patterns etc) and future level of warming	All epistemic uncertainties in modelling chain that generated climate and hydrology conditions for given level of warming
Vulnerability curves	Perturbation factor (-)	Varies between 0 and 1.	Parameter uncertainty in chosen vulnerability curve	Structural uncertainty around appropriate choice of the vulnerability curve (which shape and which independent variables – e.g. beyond flood depth)

Figure 7 shows an example of the type of insights delivered by the GSA. Each scatter plot shows the values of the output metric (average annual loss) against the 210 samples of each input uncertainty. The more clearly a scatter plot shows a pattern when moving along the horizontal axis, which represents the variability range of an input, the greater the influence of that input. For discrete inputs, we use boxplots to help visualise the distribution of output values for each discrete choice of input value. The damage ratio (first scatter plot from the left) is the most influential input, with a clear pattern of higher losses with increasing values of damage ratios. The strength of such input-output relationship is confirmed by the value of the PAWN sensitivity index, reported at the top of the scatter plot. The second most influential input is the choice of the portfolio resolution, with more aggregate portfolio generally leading to higher losses. Predicted losses are slightly higher when the portfolio includes less building information (third scatter plot) and when hazard intensity is increased (fourth). The choice of the event set has the least influence on the loss predictions.

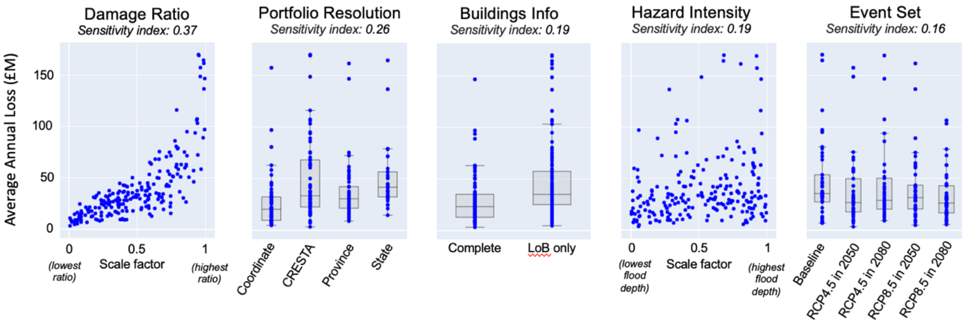

Figure 7 - Example results of a global sensitivity analysis of the flood loss model for Queensland, Australia showing scatter plots of model output (average annual loss, AAL) against the five input uncertainties. For discrete inputs, box plots are overlayed to represent the 25th, median, and 75th percentile of the output sample distributions at each discrete input values. The title of each scatter plot reports the PAWN sensitivity index for each input.

The findings confirm the key role of uncertainty in the vulnerability component for predicted losses, consistent with results from the Rhine River basin application. They also show that the spatial resolution of the exposure portfolio has an important impact on loss estimates.

In this case study, more aggregate (i.e., lower resolution) portfolios lead to overestimated losses, assuming the coordinate-level portfolio as the reference or "truth". To understand why, we analysed the characteristics of these different portfolios and identified two key sources of systematic bias (Figure 8 – left). First, the disaggregation algorithm spreads assets more widely than in the coordinate-level portfolio. This increases the number of assets exposed to any given flood event. Second, aggregation leads to loss of detail in building characteristics. In each spatial unit of an aggregated portfolio, all assets are assigned the same characteristics. This tends to increase the attributed vulnerability relative to the coordinate-level portfolio. For example, in the coordinate-level portfolio, 77% of the buildings are residential and therefore classified as more vulnerable than the 23% commercial ones. In contrast, the residential portion is 96% in the CRESTA-level portfolio, 94% at the province level, and 100% at the state-level. Similar patterns are seen in floor height data. In the coordinate-level portfolio, 32% of the buildings have a first-floor height below 0.4 m –a characteristic associated with higher vulnerability. This increases to 49% in the CRESTA-level portfolio and 100% in the state-level one. Conversely, buildings with first floor height above 0.5 m – a characteristic associated with lower vulnerability – decline from 35% in the coordinate-level portfolio to 19% at the CRESTA-level, and 0% at the province and state levels. In summary, the disaggregation algorithm introduces systematic bias in both exposure and vulnerability when moving from detailed to aggregate exposure portfolios.

Figure 7 also suggests a very limited role of future climate uncertainty with respect to other sources of uncertainty. This may be explained by the fact that the climate projections used here suggest both increasing and decreasing trends in precipitation and river flows in different sub-regions of Queensland (Figure 8), and therefore these opposite trends may offset each other when looking at overall losses across the region. Here, the flood event sets were generated based on the outputs of a single climate model, MRI-ESM2-0 (Yukimoto et al. 2019), and incorporating event sets generated from other models could capture a wider range of uncertainty.

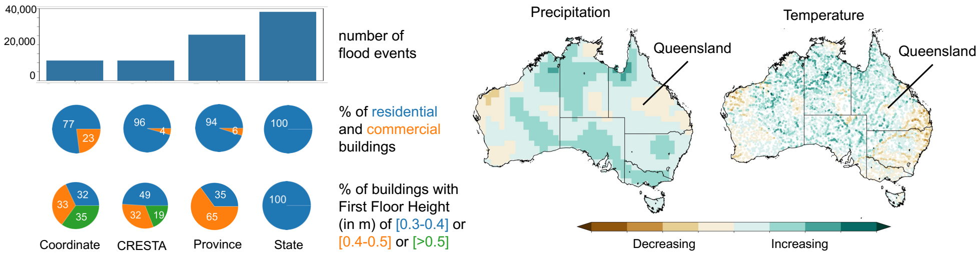

Figure 8 – Exposure and building characteristics in the four portfolios considered in the GSA application in Queensland (left) and projected changes in climate and hydrology under the RCP4.5 in 2080 scenario.

Outlook and Conclusions

This paper has discussed challenges and possible ways forward in the application of global sensitivity analysis to flood loss models across large spatial domains. We have discussed and demonstrated several strategies to make GSA tractable even for computationally expensive models such as large-scale flood loss models. Thanks to these strategies, we were able to extract meaningful information from Monte Carlo simulations based on a few hundred model runs. Still, given the large spatial domain covered, these runs required up to 3 days to run on relatively powerful machines: the computational facilities of the Advanced Computing Research Centre at the University of Bristol for the Rhine River basin application, and JBA's machine (an Intel i7 3.4GHz, with 16 cores, and 128 GiB RAM) for the Queensland application.

The two example applications reported here, to the Rhine River basin in Europe and Queensland state in Australia, consistently signal that uncertainty in the quantification of vulnerability is the dominant source of uncertainty in our flood loss estimates, within the scope of the input space studied. In the Rhine River basin application, the second most important source of uncertainty (of those analysed) was the frequency of floods in the event sets. In the Queensland case study, we also found systematic biases induced by the loss of information when moving to more aggregate and less detailed exposure portfolios. These insights could be useful for model developers to set priorities for future acquisition of higher-quality data or refinement of model components; or they can be leveraged by model users – for instance in (re)insurance – to advocate for access to higher resolution exposure data.

Our sensitivity results are conditional on the choices made in the experimental design, and in particular which input variables were varied and within which ranges. Indeed, our experience suggests that the selection of the input uncertainties and their characterisation is a critical step, if not the most critical, in the application of GSA to flood loss models. Here, we used a mix of literature review and expert opinion to define the distributions of input uncertainties that seem plausible for our case studies. Structured approaches to elicit probabilities from experts () could be very useful in this context. More research on this possibly-overlooked aspect of uncertainty and sensitivity analysis will certainly be needed in the future, as also advocated by Lo Piano et al. (2022) and Page et al (2023). At a minimum, GSA results should always be presented along with a clear summary of the key assumptions made in characterising input distributions, clarifying what sources of uncertainty are being captured and what are left out – as we did for example in Tables 1 and 2 of this paper.

Equally important in the communication of GSA results is to be clear about the fact that, being a stress test of the model, not of the system being modelled, GSA can only tell us about the behaviour of the model within the scope of its underpinning assumptions, and not whether those assumptions are correct. Put differently, and returning to the distinction made in Figure 1 – GSA can help to tackle the precision of the model but not its accuracy. For example, we have found that reducing uncertainty in damage curves would be of paramount importance in making the loss predictions more precise in the Rhine River case study, but this does not imply it would make them more accurate if the use of mean damage ratios oversimplifies the reality of damage distributions and other variables beyond flood depth would be needed to better predict damage (Lighthill Risk Network, 2019, Wing et al 2020).

Overall, GSA can (and we think should) play an essential role in the diagnostic evaluation of models (Wagener et al. 2022) and in guiding investments towards improving data quality and model components where they will have the greatest effect. We hope this paper will provide motivation as well as practical ideas to foster the application of GSA to flood risk models and contribute to increasing their transparency and legitimacy.

Data and code availability

Access to the JBA catastrophe model and data used in the GSA experiments presented in this paper can be made available upon reasonable request to JBA Risk Management: hello@jbarisk.com. The MRI-ESM2-0 climate model output data used to condition the flood event sets used in the Queensland experiments is available through https://esgf-ui.ceda.ac.uk/cog/search/cmip6-ceda/ .

Author contribution

FP, GS, TW, RL, SH designed the experiments of the Rhine River Basin application and GS conducted the experiments. FP, GS, KS and PO designed the experiments of the Queensland application and GS conducted the experiments. FP, TW and KS supervised the work. FP prepared the manuscript with contributions from all co-authors.

Competing interests

The authors declare that they have no conflict of interest

Acknowledgement

We gratefully acknowledge the support and permission to use the industry exposure and loss data from PERILS AG to inform the portfolios used in the Queensland case study presented in this paper. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted the sixth phase of the Coupled Model Intercomparison Project (CMIP6) (https://esgf-ui.ceda.ac.uk/cog/search/cmip6-ceda/). We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. The climate change event sets

underpinning the GSA experiments for Queensland case study were generated using output from the climate model Meteorological Research Institute Earth System Model, Version 2.0 (MRI-ESM2-0) from the Meteorological Research Institute (MRI), Japan

Financial support

This work was supported by the Engineering and Physical Sciences Research Council in the UK via the Water Informatics: Science and Engineering (WISE) Centre for Doctoral Training [grant number EP/L016214/1], JBA Trust (project W19-1717), and by an Innovate UK Knowledge Transfer Partnership between the University of Bristol and JBA Risk Management Limited (KTP 13266). T.W. also acknowledges support from the Alexander von Humboldt Foundation in the framework of the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Reseach (BMBF).

References

Aerts, J.C.J.H., Bates, P.D., Botzen, W.J.W. et al. Exploring the limits and gaps of flood adaptation. Nat Water 2, 719–728 (2024). https://doi.org/10.1038/s44221-024-00274-x.

Apel, H., Thieken, A. H., Merz, B., & Blöschl, G. (2004). Flood risk assessment and associated uncertainty. Natural Hazards and Earth System Sciences, 4(2), 295-308.

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

Arribas, A., Fairgrieve, R., Dhu, T. et al. Climate risk assessment needs urgent improvement. Nat Commun 13, 4326 (2022). https://doi.org/10.1038/s41467-022-31979-w.

Becher, O., Pant, R., Verschuur, J., Mandal, A., Paltan, H., Lawless, M., Raven, E., Hall, J. (2023). A multi-hazard risk framework to stress-test water supply systems to climate-related disruptions. *Earth's Future*, 11(1). https://doi.org/10.1029/2022EF002946.

Cammerer, H., Thieken, A. H., Lammel, J. (2013). Adaptability and transferability of flood loss functions in residential areas. Natural Hazards and Earth System Sciences, 13(11). https://doi.org/10.5194/nhess-13-3063-2013.

CRED-UNDRR (2020) The human cost of disasters: an overview of the last 20 years (2000-2019). https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019.

Darlington, C., Raikes, J., Henstra, D., Thistlethwaite, J., Raven, E. (2024) Mapping current and future flood exposure using a 5 m flood model and climate change projections. Natural Hazards and Earth System Sciences, 24(2), 699-714. https://doi.org/10.5194/nhess-24-699-2024.

D'Ayala, D., Wang, K., Yan, Y., Smith, H., Massam, A., Filipova, V., Pereira, J. J. (2020). Flood vulnerability and risk assessment of urban traditional buildings in a heritage district of Kuala Lumpur, Malaysia. *Natural. Hazards and Earth System Sciences,* 20(8), 2221-2241. https://doi.org/10.5194/nhess-20-2221-2020.

Dawkins, L.C., Bernie, D.J., Pianosi, F., Lowe, J.A., Economou, T. (2023) Quantifying uncertainty and sensitivity in climate risk assessments: Varying hazard, exposure and vulnerability modelling choices, *Climate Risk Management*, 40, 100511, https://doi.org/10.1016/j.crm.2023.100511.

Devitt, L., Neal, J., Coxon, G., Savage, J., & Wagener, T. (2023). Flood hazard potential reveals global floodplain settlement patterns, Nat. Commun., 14, 2801.

de Moel, H., Bouwer, L. M., Aerts, J.C.J.H. (2014) Uncertainty and sensitivity of flood risk calculations for a dike ring in the south of the Netherlands. *Science of the Total Environment*, *473–474*. https://doi.org/10.1016/j.scitotenv.2013.12.015

Déroche, M.S. (2023) Invited perspectives: An insurer's perspective on the knowns and unknowns in natural hazard risk modelling, Nat. Hazards Earth Syst. Sci., 23, 251–259, https://doi.org/10.5194/nhess-23-251-2023

Franco, G., Becker, J. F., & Arguimbau, N. (2020). Evaluation methods of flood risk models in the (re)insurance industry. Water Security, 11. https://doi.org/10.1016/j.wasec.2020.100069

Galloway, E., Massam, A., Allard, J., Oldham, P., Sarailidis, G., Catto., J., Germon-Duret, C., Young, P. (2025). *Journal of Catastrophe Risk and Resilience.* https://doi.org/10.31223/X5C42Z

Grossi P., Kunreuther H. (2005). Catastrophe Modeling: A New Approach to Managing Risk. In *Catastrophe Modeling: A New Approach to Managing Risk*. Kluwer Academic Publishers. https://doi.org/10.1007/b100669

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C. J. H., Ward, P. J. (2012): Comparative flood damage model assessment: towards a European approach, *Nat. Hazards Earth Syst. Sci.*, 12, 3733–3752, https://doi.org/10.5194/nhess-12-3733-2012.

Kaczmarska, J., Jewson, S., & Bellone, E. (2018). Quantifying the sources of simulation uncertainty in natural catastrophe models. *Stochastic Environmental Research and Risk Assessment*, *32*(3). https://doi.org/10.1007/s00477-017-1393-0

Kay, L. A., Booth, N., Lamb, R., Raven, E., Schaller, N., Sparrow, S. (2018). Flood event attribution and damage estimation using national-scale grid-based modelling: Winter 2013/2014 in Great Britain. *International Journal of Climatology*, 38(14), 5205-5219, https://doi.org/10.1002/joc.5721.

Keef, C., Tawn, J. A., & Lamb, R. (2013). Estimating the probability of widespread flood events. Environmetrics, 24(1), 13-21. https://doi.org/10.1002/env.2190

Kreibich, Heidi, et al. (2022) The challenge of unprecedented floods and droughts in risk management. Nature 608.7921: 80-86.

Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., Dunning, P., Batstone, C. (2010). A new method to assess the risk of local and widespread flooding on rivers and coasts. Journal of Flood Risk Management , 3 (4). https://doi.org/10.1111/j.1753 318X.2010.01081.x

Lighthill Risk Network. (2019). *Flood Research Needs of the (Re)insurance sector.* Available at: https://lighthillrisknetwork.org/reports/. Last visited: 30 August 2024.

Lo Piano, S., Sheikholeslami, R., Puy, A., Saltelli, A. (2022). Unpacking the modelling process via sensitivity auditing. Futures, 144, 103041.

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., ... & Macdonald, E. (2021). Causes, impacts and patterns of disastrous river floods. Nature Reviews Earth & Environment, 2(9), 592-609.

Merz, B., Blöschl, G., Jüpner, R., Kreibich, H., Schröter, K., and Vorogushyn, S. (2024): Invited perspectives: Safeguarding the usability and credibility of flood hazard and risk assessments, *EGUsphere* [preprint], https://doi.org/10.5194/egusphere-2024-856.

Mitchell-Wallace, K., Jones, M., Hillier, J., Foote, M. (2017): Natural catastrophe risk management and modelling: A practitioner's guide. John Wiley & Sons.

Molinari, D., De Bruijn, K.M., Castillo-Rodríguez, J.T., Aronica, G.T., and Bouwer, L.M. (2019): Validation of flood risk models: Current practice and possible improvements. International Journal of Disaster Risk Reduction, 33, 441-448. https://www.sciencedirect.com/science/article/pii/S221242091830596X

Morris, M.D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics, 33 (2), 161–174.

A. O'Hagan, C. Buck, A. Daneshkhah, J.R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakley, T. Rakow (2006). Uncertain Judgements: Eliciting Experts' Probabilities, Wiley.

Otto, F.E.L., Raju, E. (2023) Harbingers of decades of unnatural disasters. *Commun Earth Environ* **4**, 280 (2023). https://doi.org/10.1038/s43247-023-00943-x.

Page, T., Smith, P., Beven, K., Pianosi, F., Sarrazin, F., Almeida, S., Holcombe, L., Freer, J., Chappell, N., Wagener., T. (2023) The CREDIBLE Uncertainty Estimation (CURE) toolbox: facilitating the communication of epistemic uncertainty. Hydrology and Earth System Sciences 27(13): 2523-2534.

Paleari, L., Confalonieri, R. (2016), Sensitivity analysis of a sensitivity analysis: We are likely overlooking the impact of distributional assumptions, Ecological Modelling, 340, 57-63. https://doi.org/10.1016/j.ecolmodel.2016.09.008.

Pianosi, F., Wagener, T. (2015a). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling and Software, 67.* https://doi.org/10.1016/j.envsoft.2015.01.004

Pianosi, F., Sarrazin, F., Wagener, T. (2015b), A Matlab toolbox for Global Sensitivity Analysis, Environmental Modelling & Software, 70, 80-85. https://doi.org/10.1016/j.envsoft.2015.04.009

Pianosi, F., Wagener, T. (2018). Distribution-based sensitivity analysis from a generic input-output sample. *Environmental Modelling and Software*, *108*. https://doi.org/10.1016/j.envsoft.2018.07.019.

SAFE (2025). https://safetoolbox.github.io/ Last accessed 10/06/2025

Saltelli A., Aleksankina, K., Becker, W., Fennell, P. Ferretti, F. Holst, N. Li, S. Wu, Q. (2019) Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices, Environmental Modelling & Software, 114, 29-39, https://doi.org/10.1016/j.envsoft.2019.01.012

Sarailidis, G. (2023), Uncertainty quantification and attribution in flood risk modelling, PhD Thesis, University of Bristol, https://hdl.handle.net/1983/df451ec1-2e9d-42ee-8977-dc10bac48ca3.

Sarrazin, F. J., Pianosi, F., Wagener, T. (2016). Global Sensitivity Analysis of environmental models: Convergence and validation. *Environmental Modelling and Software*, 79, 135-152. https://doi.org/10.1016/j.envsoft.2016.02.005

Savage, J., Pianosi, F., Bates, P., Freer, J., Wagener, T. (2016). Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model. Water Resources Research, 52(11), 9146-9163.

Song, X., Zhang, J.,Zhan, C., Xuan, Y., Ye, M., Xu, C. (2015) Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. *J. Hydrol.*, 523, 739-757, https://doi.org/10.1016/j.jhydrol.2015.02.013

Sudret, B. (2008) Global sensitivity analysis using polynomial chaos expansions, Reliab. Eng. Sys. Safety, 93, 964-979

Tesselaar, M., Botzen, W.J.W., Robinson, P.J. Jeroen, Aerts, C.J.H., Zhou, F. (2022) Charity hazard and the flood insurance protection gap: An EU scale assessment under climate change, *Ecological Economics*, 193, https://doi.org/10.1016/j.ecolecon.2021.107289

Wagener T., Reinecke R., Pianosi F. (2022), On the evaluation of climate change impact models, Wiley Interdiscip. Rev. Clim. Chang., 13 (2022), Article e772. https://doi.org/10.1002/wcc.772

Wing, O.E.J., Pinter, N., Bates, P.D. et al. (2022) New insights into US flood vulnerability revealed from flood insurance big data. Nat Commun 11, 1444. https://doi.org/10.1038/s41467-020-15264-2

Winter, B., Schneeberger, K., Huttenlau, M., & Stötter, J. (2018). Sources of uncertainty in a probabilistic flood risk model. *Natural Hazards*, *91*(2). https://doi.org/10.1007/s11069-017-3135-5.

Yukimoto, S. et al. (2019) The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component, Journal of the Meteorological Society of Japan. Ser. II, 2019-051, https://doi.org/10.2151/jmsj.2019-051