ORIGINAL PAPER

The influence of climate model uncertainty on fluvial flood hazard estimation

Lindsay Beevers[1] · Lila Collet[1,2] · Gordon Aitken[1] · Claire Maravat[1] · Annie Visser[1]

Abstract

Floods are the most common and widely distributed natural hazard, threatening life and property worldwide. Governments worldwide are facing significant challenges associated with flood hazard, specifically: increasing urbanization; against the background of uncertainty associated with increasing climate variability under climate change. Thus, flood hazard assessments need to consider climate change uncertainties explicitly. This paper explores the role of climate change uncertainty through uncertainty analysis in flood modelling through a probabilistic framework using a Monte Carlo approach and is demonstrated for case study catchment. Different input, structure and parameter uncertainties were investigated to understand how important the role of a non-stationary climate may be on future extreme flood events. Results suggest that inflow uncertainties are the most influential in order to capture the range of uncertainty in inundation extent, more important than hydraulic model parameter uncertainty, and thus, the influence of non-stationarity of climate on inundation extent is critical to capture. Topographic controls are shown to create tipping points in the inundation–flow relationship, and these may be useful and important to quantify for future planning and policy. Full Monte Carlo analysis within the probabilistic framework is computationally expensive, and there is a need to explore more time-efficient strategies which may result in a similar estimate of the full uncertainty. Simple uncertainty quantification techniques such as Latin hypercube sampling approaches were tested to reduce computational burden.

Keywords Flood inundation · Climate change · Uncertainty quantification · Probabilistic

✉ Lindsay Beevers
  L.Beevers@hw.ac.uk

1 Water Resilient Cities Group, Institute for Infrastructure and the Environment, Heriot-Watt University, Edinburgh, UK

2 Irstea, HYCAR Research Unit, 1 rue Pierre Gilles de Gennes, 92 160 Antony, France

1 Introduction

Floods are the most common and widely distributed natural hazard, threatening life and property worldwide (Jonkman and Vrijling 2008). Flood risk is a function of flood hazard and consequence (IPCC 2014; Balica et al. 2013). Flood hazards result from many different sources (e.g. coastal, fluvial, pluvial or estuarine), whilst the consequences arise from the adverse impacts of flooding on people, property, human health, the environment, cultural heritage and economic activity (Beevers et al. 2016). The UN estimates that 1 Bn people live in areas of potential flood risk and damage caused by floods on a global scale has been significant in recent decades (Jonkman and Vrijling 2008). In the last decade (2007–2017), there have been around 200 significant flood events in Europe, resulting in almost 1000 deaths, affecting 3.9 M people and causing over \$55 Bn worth of damage (2018). In the UK, the estimated cost of flood damage in the UK was £3.2 M, whilst the 2013/2014 events are estimated to have cost the economy £1.3 M. The UK government estimates that each year flooding will cause £1.1 Bn, whilst maintaining current levels of flood defence will require a further £1 Bn per year by 2035 (Sayers et al. 2015). Looking forward, governments worldwide are facing significant challenges associated with flood risk, specifically: increasing urbanization; the current drive to control public expenditure; and the background of uncertainty associated with increasing climate variability under climate change (Guerreiro et al. 2018).

Historically, flood risk management has approached flood hazard prediction in a deterministic manner (Baldassarre and Montanri 2009). The method used routinely in practice uses a deterministic hydraulic (flood) modelling approach that simulates the physical processes controlling flood flows (Baldassarre et al. 2010). An event, characterized by a single set of boundary data, is modelled to produce mapped outputs of the potential depth, velocity and extent of the flood event (Hartanto et al. 2011; Beevers et al. 2012). However, the uncertainty due to the dynamic, stochastic and uncertain nature of the climate, hydrological and river processes (Van Vuren 2005) is generally not considered nor is the underlying model data and choices or the assumptions embedded within the analysis (e.g. return period methodology estimates. Uncertainties in hydraulic (flood) modelling can arise from many sources, and for the purposes of this paper, we follow a structured approach (Warmink et al. 2010; Savage et al. 2016) to these for hydraulic simulation, classifying them as uncertainties associated with:

- the model input: for example, the boundary conditions (e.g. hydrology and flow estimates) or the underlying bathymetry or topographical data;
- the model parameters (e.g. roughness estimates); or
- the choice of the model structure (e.g. the mathematical formulation).

Within a flood modelling framework, uncertainties cascade through the modelling chain, with each contributing to the variability in the final inundation extent. Thus, accounting for uncertainties in model prediction this may be achieved through consideration of multi-model ensembles, perturbed parameter ensembles and/or multi-member (inputs) ensembles. This is often achieved through Monte Carlo simulation (Mundform et al. 2011) providing approximate solutions to a variety of mathematical problems by performing statistical sampling experiments. Whilst such an approach is data intensive and computationally demanding, it is increasingly recognized that there is a growing demand for outputs in the form of probabilistic flood maps (Teng et al. 2017).

With the availability of greater computational power and parallelization of codes, uncertainty analysis in flood model research has become more frequent. Recently, several studies have investigated incorporating uncertainty in models (input and parameters) in hydraulic modelling, flood mapping and inundation analysis (Apel et al. 2004; 2006; Neal et al. 2013; Savage et al. 2014; Mukolwe et al. 2014; Ali et al. 2015; Teng et al. 2017; Winter et al. 2018; Berends, et al. 2018). The sources of uncertainty considered include inflows, observed data and flood frequency analysis, model parameters and underlying topographical data. Assessed uncertainties can be part of a flood risk assessment and highlight the need to quantify input and parameter variability to prevent unnecessary environmental and economic losses or potential flood defence failures (Wobus et al. 2017; Metin et al. 2018; Curran et al. 2019). Some general conclusions have emerged from these studies, suggesting that uncertainty associated with inflow has a significant influence on the flood inundation prediction. This input uncertainty is more significant than the uncertainty associated with model parameterization (Mukolwe et al. 2014) or topographic data (Ali et al. 2015). However uncertainty associated with climate model parameterization has not been explored.

In regard to the input uncertainties (or inflow), one of the biggest challenges to current flood risk management is the uncertainty associated with climate change (Collet et al. 2017) and the impact it may have on future floods (hazard and exposure) in terms of increasing frequency, duration and magnitude (Collet et al. 2017; 2018a, b; Thober et al. 2018; Visser-Quinn et al. 2019). For example, assuming a non-stationary climate under climate change:

- What is the likely magnitude of a future flood event (return period) in a particular location, and how may that differ from flood events at present (Collet et al. 2018a)?
- How will this affect the future flood footprint of a given event in the future?
- Will this uncertainty be greater than the inherent uncertainty associated with flood frequency estimation at present (e.g. Winter et al. 2018; Collet et al. 2017), or will parameter uncertainty in the model become more important?

To date, no study has attempted to answer these questions. Previous work by the authors (Collet et al. 2017, 2018a, b) used the Future Flows Hydrology (FFH) database (Prudhomme et al. 2013), a set of spatially coherent daily flow projections, derived from an eleven-member Perturbed Parameter Ensemble (PPE) for 282 catchments across Great Britain. This work focused on analysing the change in extreme run-off for a range of return period events across Great Britain. National scale maps were created displaying the uncertainty related to climate model and probabilistic distributions for flood events from 1 in 10 to 1 in 200 year return periods, with 1 in 100 and 1 in 200 year return period events being the IL industry standard for flood management and general planning purposes (e.g. housing developments). Return periods were determined via two extreme value methods, generalized extreme value (GEV) and generalized Pareto (GP), for two time periods, a baseline (1961–1990) and the future (2070–2099). The analysis evaluated the uncertainty associated with the extreme value (EV) distributions and climate model parameters. For the same time period, the GP and GEV methods saw estimates of run-off with different mean values but similar uncertainty bounds, whilst between time periods, there was an increasing trend in both. Collet et al. (2017, 2018a, b) concluded that methods investigating the future flood hazard and extent should account for uncertainty associated with both the climate model parameters and the EV distribution.

Building on this previous work, the aim of this paper is to determine *the relative role climate model uncertainty has in the prediction of future flood extents*. To do this, the

run-off estimates must be used as hydrograph inputs to numerical hydraulic models to calculate flood inundations (cascading the uncertainty through the modelling process). This is achieved through the application of a probabilistic framework which *explicitly captures the uncertainty associated with climate model parameterization* alongside flood modelling uncertainty (extreme value estimation and hydraulic model roughness specifically).

The potential to reduce computational effort using numerical approaches is also considered. The method is demonstrated through application to a Scottish case study for two future return period events (1:10 year and 1:100 year) with focus on the higher magnitude 1:100 year event. Two return period events were considered to address the uncertainty of fitting extreme value statistics on limited time series (30 year time slices).

2 Data and models

2.1 Case study

The River Don (Fig. 1a), Aberdeenshire, north-east Scotland, serves as the case study catchment. It has a total catchment area of approximately 1300km^2, and the river flows from the Cairngorms National Park to the North Sea at Aberdeen. Previous work by the authors (Collet et al. 2017, 2018a) identified the River Don as a catchment likely to see an increased flood hazard as a result of climate change.

This work focuses on a 5 km reach of the river downstream of the Parkhill flow gauge (station number: 11001, NRFA), which flows through Dyce, a suburb of Aberdeen. The gauging station has flow records since 1950 and shows frequent flood events (most recently in 2016); the highest water levels on record occurred on 8 January 2016. This is a relatively short and hydraulically simple reach of the river, with a topographically variable floodplain, allowing the exploration of complex uncertainties in a systematic manner.
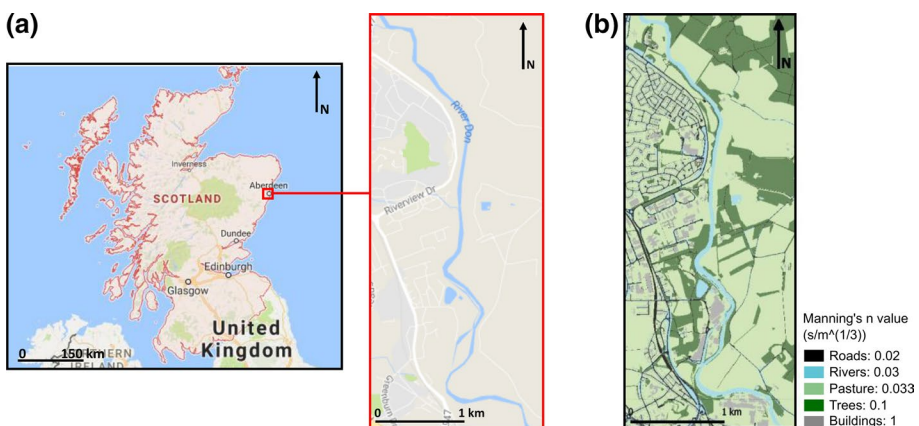

Fig. 1 Case study catchment; **a** location and extent of the flood model; **b** land use and roughness parameters

2.2 Hydraulic model: LISFLOOD-FP

LISFLOOD-FP (Bates and De Roo 2003) is selected as the hydraulic model, due to its computational efficiency, and reduced physics formulation. The LISFLOOD-FP model requires input data in the form of boundary conditions, a digital elevation model, channel geometry and friction coefficients. The upstream boundary was located at the Parkhill gauging station, with a free boundary located 5 km downstream of the gauge as the outflow of the domain (Fig. 1). The Scottish Environment Protection Agency (SEPA 2015) provided the digital elevation model (1 m LiDAR) as well as 13 river cross sections from which channel geometry was derived. A land cover map was determined using OS data at a 5 m resolution, with the associated Manning's n determined from the literature (Chow 1959); see Table 1.

The LISFLOOD-FP model was built using the sub-grid channel solver for main channel flow and the acceleration solver for floodplain flow. The model was calibrated for November 2002, October 2002 and September 1995 flood events [daily gauged flow and digitized maximum flood extent maps (SEPA)] and validated to National Flood Hazard Maps for 1:10, 1:200 and 1:1000 return periods. Roughness (Manning's n) was adjusted in order to maximize agreement between observed and simulated maximum flood extents using the fit statistic Fstat (the overlap between observed and predicted as a proportion of wet cells); Fstat ranges from 0 to 1, with 1 indicating perfect agreement (Goudet 1995).

2.3 Flow projections

Following Collet et al. (2017), this study utilized daily flow projections from the FFH database, an 11-member perturbed parameter ensemble. It is through this perturbing of parameters that climate model parameter uncertainty may be explored. Flow projections were extracted for the Don River at Parkhill gauging station. As is standard in climate impact assessment (Allen et al. 2018), this study focuses on two 30-year climate normals, a 1961–1990 baseline and the 2080s (2069–2098). To ensure consistency in the analysis, only flow projections are utilized. The focus is on the change in extreme events between the time periods, i.e. the changes in peak flow and uncertainty due to climate change.

3 Methods

This paper aims to understand the role of climate model parameter uncertainty in future flood inundation projections through a case study catchment (Sect. 2.1). In order to achieve this aim, the role of uncertainty associated with climate model parameters was explored in

Table 1 Roughness coefficients, an input to the hydraulic model LISFLOOD-FP, for the study area

Category	Land cover (%)	Manning's n
Paths/roads	6	0.02
River/canal	7	0.025–0.05
Pasture, no brush	59	0.025–0.05
Heavy growth trees	21	0.08–0.12
Buildings	7	1.00

context with other sources of uncertainty by sampling across the distributions of inputs for the hydraulic model following four structured tests (Sect. 3.1 and Fig. 2):

a. Climate model uncertainty (CMU), as quantified in Collet et al. (2017) (Fig. 2a). The empirical CDF of the return period estimates for the 11-member ensemble;
b. Extreme value (structure and parameter) uncertainty (EVDU) for two different extreme value models, the generalized extreme value (GEV) and generalized Pareto (GP) distributions (Fig. 2b). Parameter uncertainty is considered through the 95% confidence intervals. After Collet et al. (2017), in order to isolate this uncertainty from the climate model uncertainty, only the ensemble median return period estimate was considered;
c. Hydraulic model parameter uncertainty (HMPU; Fig. 2c), the roughness (Manning's n) values for the floodplain and channel definition. This test fixes the Q_{peak} used to the median estimate across the ensemble;
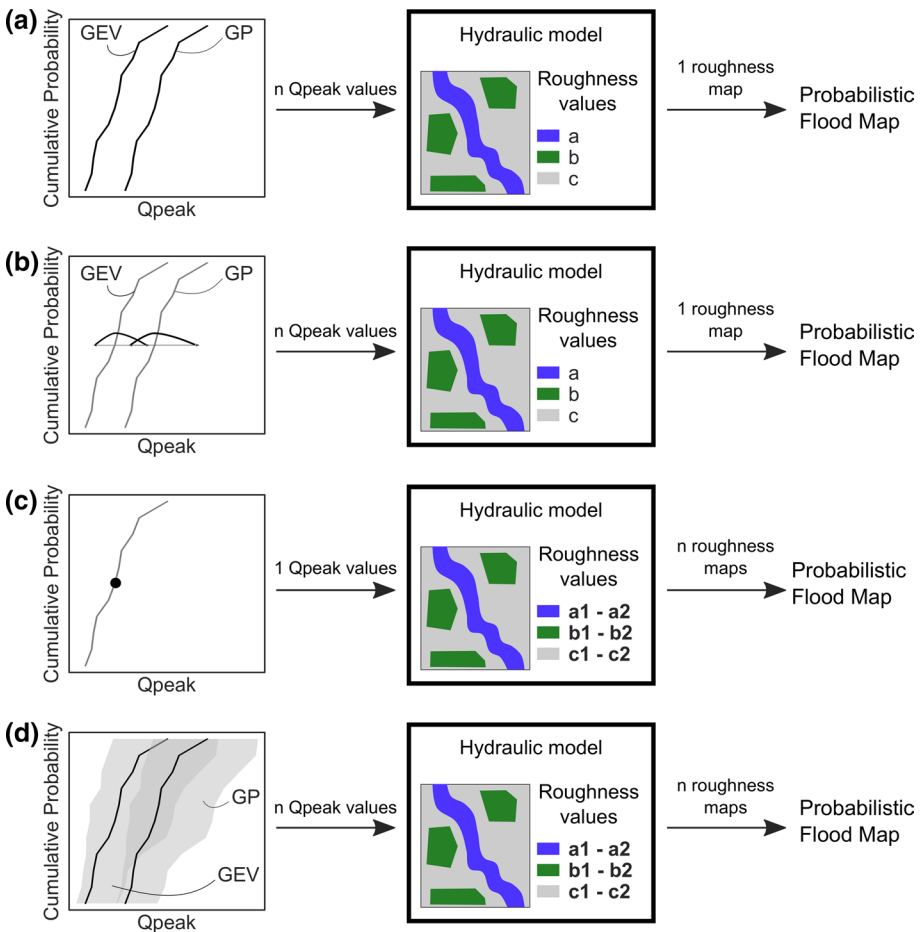

Fig. 2 Uncertainty quantification framework: **a** regional climate model parameter uncertainty (CMU); **b** extreme value distribution parameter and structure uncertainty (EVDU); **c** hydraulic model parameter uncertainty (HMPU); and **d** total uncertainty; n: number of sampled values, varies depending on the chosen probabilistic framework

d. Total uncertainty (Fig. 2d), to determine the relative uncertainty associated with each source. Total uncertainty is measured as a combination of tests a–c, where all three factors are varied.

The tests were applied for the 1:10- and 1:100-year return period flood events for the baseline (1961–1990) and future (2070–2099) time periods (to understand the influence of a non-stationary climate), across the 11 equally plausible hydroclimatological realizations for the Parkhill gauge station.

The tests were applied through two probabilistic frameworks (Sect. 3.2). The first, a full Monte Carlo analysis, is a computationally expensive approach where a large number of randomly sampled realizations are assumed to represent the distributions in their entirety. Whilst computational power is increasing, there is still a need to explore more time-efficient strategies which may result in a similar estimate of the full uncertainty. As a first step, this paper also explored a simple stratified sampling approach in order to reduce this computational burden.

3.1 Uncertainty tests

3.1.1 Climate model parameter uncertainty (CMU)

In order to determine the 1:10- and 1:100-year return period flood events (Q_{peak}), GEV and GP distributions (Coles 2001) were fitted to annual maxima (AMAX) and peak-over-threshold (POT) flows, respectively. See Collet et al. (2017) for the detailed methodology. For both return periods (1:10 and 1:100) and each extreme value model (GEV and GP), Q_{peak} values were computed for the equally probable ensemble members, per time period ($N = 11 \times 2$). Using the empirical CDF constructed by the 11 values of N, a total of n Q_{peak} values were randomly sampled across the distribution (where n is determined by the probabilistic framework outlined in Sect. 3.2, i.e. full Monte Carlo or LHS). These n Q_{peak} values served as the input to the hydraulic model (see Fig. 2a). The input hydrograph for the LISFLOOD-FP model was scaled for each simulation using the sampled Q_{peak}, generating corresponding hydrographs that were used to run the hydraulic model with the spatially distributed roughness (Manning's n) values held constant.

3.1.2 Extreme value model (structure and parameter) uncertainty (EVDU)

For the N ensemble members and time periods, the median Q_{peak} value was determined across both EV distributions, after Collet et al. (2017). This median represents the structural uncertainty, whilst 95% confidence intervals capture parameter uncertainty. This range was randomly sampled following a normal distribution, generating n hydrographs which represent the boundary conditions in the hydraulic model. The spatially distributed roughness (Manning's n) values were held constant in the hydraulic model (see Fig. 2b).

3.1.3 Hydraulic model parameter uncertainties (HMPU)

Assuming a uniform distribution, the HMPU was determined by sampling across the Manning's n parameter space (range—see Table 1), producing n maps. The input hydrograph to the hydraulic model was held constant, corresponding to the median value of the GEV distribution across the 11 ensemble members (see Fig. 2c).

3.1.4 Total uncertainty

Total uncertainty simulations were then completed where all three factors from tests a–c were varied: (1) CMU, the uncertainty in Q_{peak} from across the empirical CDF; (2) EVDU, where the median Q_{peak} and 95% confidence intervals represent the structure and parameter uncertainty, respectively; and (3) the HMPU, captured through maps of roughness value.

3.2 Probabilistic frameworks

3.2.1 Full probabilistic approach: Monte Carlo analysis

The full probabilistic approach adopts a Monte Carlo approach (Mundform et al. 2011). For each uncertainty test (Sect. 3.1), factor(s) are varied by sampling across the respective range of values with an associated probability distribution. A pseudo-random number generator (PRNG) was used to sample ($n = 10,000$) from the prescribed distributions for each test, for a total of 80,000 simulations (2 time periods $\times$ 4 tests $\times$ n).

3.2.2 Reducing the computational burden: stratified sampling: Latin hypercube sampling

Latin hypercube sampling (LHS) divides the distributions into regions of equal probability and randomly samples a value from each region. This increases the sampling efficiency by ensuring uniform coverage of the parameter space whilst using fewer samples (Helton and Davis 2003). Theoretical studies have shown that the sampling error for LHS, $O\left(\frac{1}{\sqrt{n}}\right)$, is quadratically faster than randomly sampled Monte Carlo simulations, $\left(\frac{1}{n}\right)$; simply put, LHS can achieve the same accuracy as n^2 Monte Carlo simulations (Aistleitner et al. 2012). However, in practice, as the dimensionality of the input increases, the accuracy reduces (Huntington and Lyrintzist 1998). An extension of LHS, orthogonal sampling, has been employed in this study as this ensures that each subspace is evenly sampled.

The distributions for each test were binned, $n = \{50, 125, 250, 500, 1000\}$. Each bin was sampled once, with 10 replicates performed to test sample size variability. These replicates were tested to determine the minimum number of samples required for convergence (mean and standard deviation of the flooded area in the full Monte Carlo analysis).

4 Results

4.1 Model calibration

The final calibrated model has a reasonable fit and was found to perform best for higher (out of bank) events, making it suitable for this study (Table 2), where significant flow events (1: 10-year and 1:100-year return period) are chosen for analysis. Water levels at specified cross sections were within ± 0.2 m. The calibrated model was validated against

Table 2 Hydraulic model calibration and validation results: fit between observed and modelled extent

Calibration		Validation	
Event	Fitting stat (*F*stat)	Return period (National Flood Hazard Maps)	Fitting stat (*F*stat)
November 2002	0.64	1:10-year	0.48
October 2002	0.66	1:200-year	0.74
September 1995	0.61	1:1000-year	0.76

the more extreme out-of-bank events (1:200-year and 1:1000-year events), for which data were available, performing well and confirming reasonable calibration of the model.

4.2 Probabilistic framework: full Monte Carlo analysis

The outcomes from the full Monte Carlo analysis are first considered. The 1:100-year return period event results are presented in the main paper, whilst the results for the
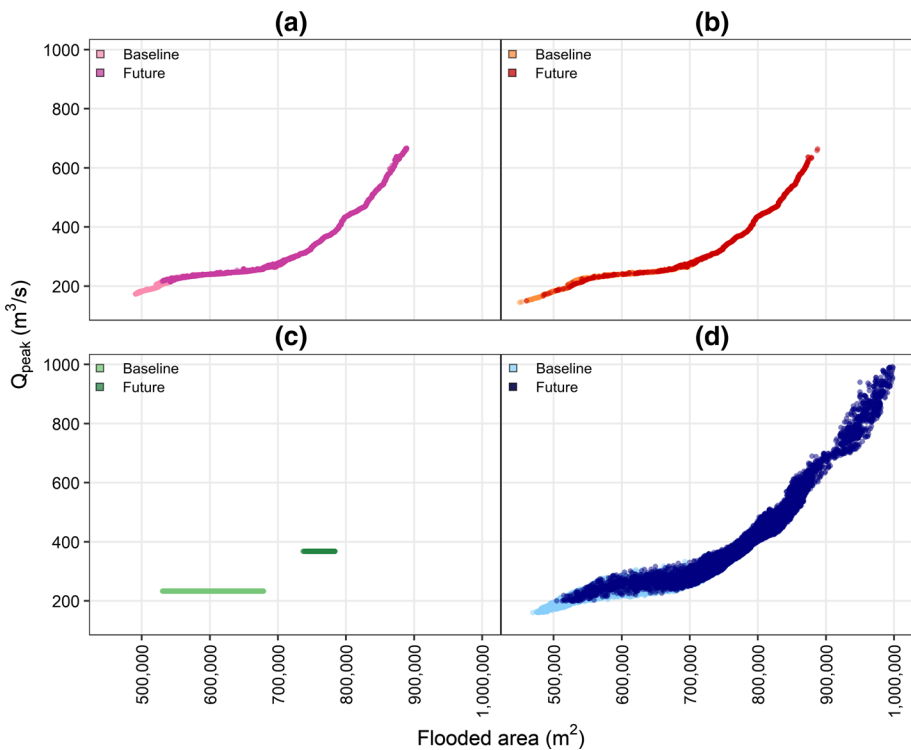
Fig. 3 Flooded area versus Q_{peak} per test for both baseline (lighter colour) and future (darker colour) time periods: **a** the climate model parameter uncertainty (CMU); **b** the extreme value (model structure and parameter) uncertainty (EVDU); **c** the hydraulic model parameter uncertainty (HMP); and **d** the total uncertainty
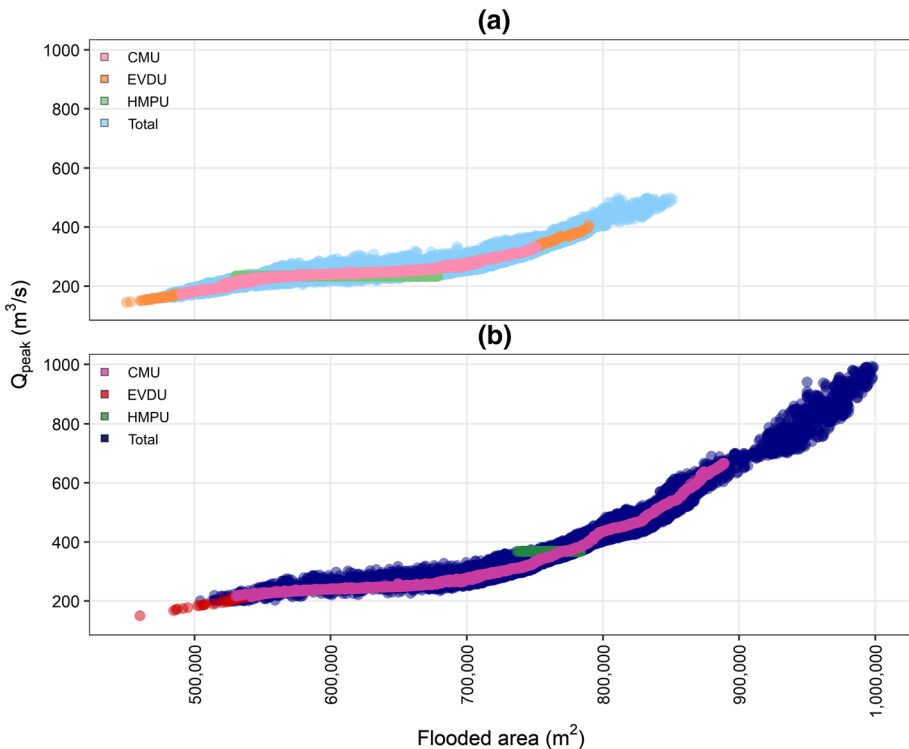
Fig. 4 Aggregated flooded area versus Q_{peak} for the **a** baseline and **b** future

1:10-year return period event are presented in the supplementary material. For the 1:100-year return period event, Fig. 3 depicts the flooded area versus Q_{peak} per test; the test results are aggregated in Fig. 4, allowing comparison of the baseline and future uncertainties. Change in flooded area, per test, is further considered in Fig. 5 through a series of frequency distributions. The change in output mean values (Delta mean) is indicated for each uncertainty, further measures of change are provided in Table 3. Finally, Fig. 6 presents the probabilistic flood maps for the total uncertainty (test *d*); see supplementary information for the probabilistic flood maps for tests *a–c*. The results consider the influence the different sources of uncertainty provide, before exploring the impacts of considering all sources together. Results are presented in supplementary material for the 1:10-year return period event, as shown in Figs. 10, 11 and 12.

4.2.1 Sources of uncertainty: tests a–c

Figures 3 and 5 and Table 3 consider each source of uncertainty, as well as the total uncertainty, separately (1:100-year return period event). It can be seen that, in the case of climate model parameter uncertainty (a), the range of uncertainty is smallest on the baseline, with Δ-mean increasing by 20% in the future (Fig. 5; bias between the samples). With a Δrange of almost 40% (Table 3), the overall climate model parameter uncertainty can be seen to increase. Similar changes in mean (19%, Table 3) are observed for the EV model structure and parameter uncertainty (b). A change in distribution from the baseline (bimodal) to the
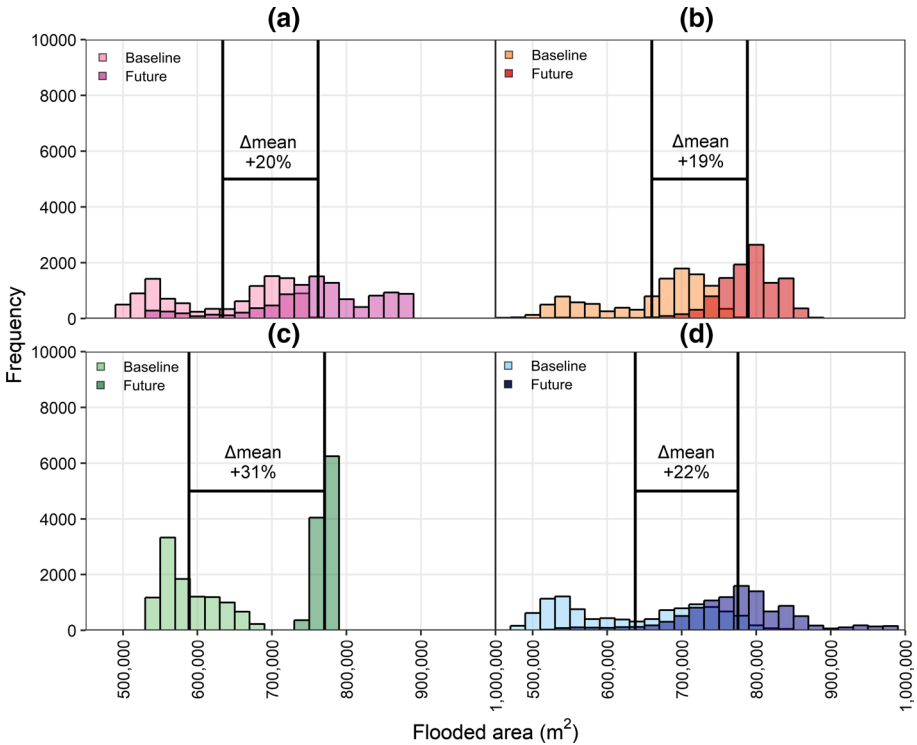
Fig. 5 Frequency distribution of flooded area and Δ-mean from the baseline (lighter colour) to future (darker colour) time periods (see also Table 3). The sources and total uncertainty are considered across the panels: **a** climate model parameter (CMU); **b** extreme value (EVDU); **c** hydraulic model parameter (HMPU); and **d** total

Table 3 Percentage change to series statistics from the baseline to the future for a 1:100 year return period

Source of uncertainty	Δmean	Δmin	Δmax	Δrange
CMU	20.13	7.99	18.43	38.15
EVDU	19.29	2.71	12.46	23.98
HMPU	30.95	38.91	19.28	-63.90
Total	21.69	7.50	17.76	30.37

future (normal) can be observed in Fig. 5b. This is likely due to topographic thresholds present in the baseline simulations but not observed in future outputs due to water levels higher than these features. Interestingly, the results of the 1:10-year return period event (supplementary material) reinforce the importance of the topographic control in this case study.

Despite this, the increase in overall uncertainty is lower than for the climate model parameter (Δrange = 24%, Table 3). Test c saw Q_{peak} held constant as the roughness maps were varied; accordingly, the plot in Fig. 3c is a horizontal line. The change in roughness under test c presents a different picture. The Δ-mean (31%) is higher than in tests *a* and *b*, whilst the
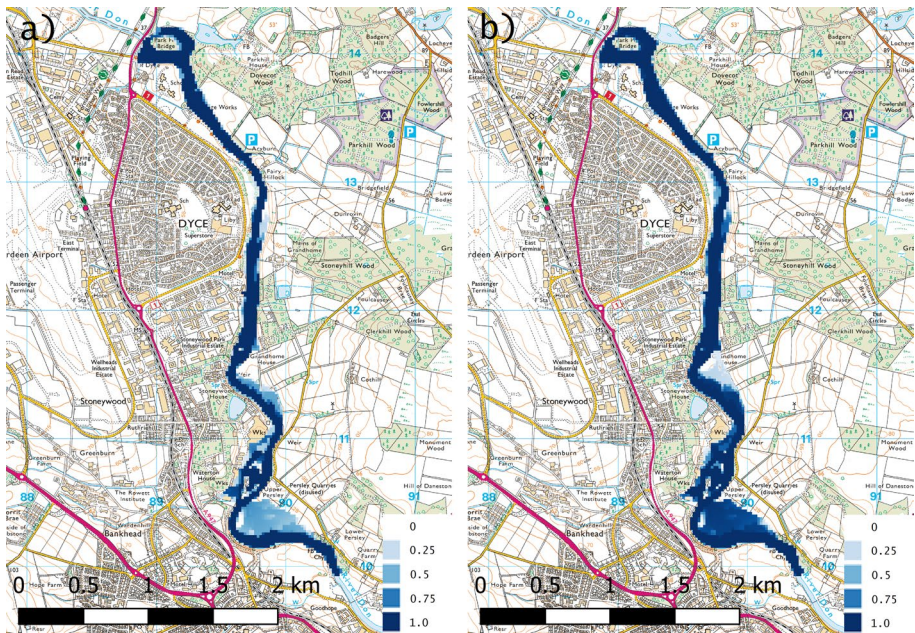
Fig. 6 Probabilistic flood map showing total uncertainty (test d) for the baseline (**a**) and the future (**b**)

Δ-range sees a large reduction (64%). Overall, hydraulic model parameter uncertainty contributes the least to the range of modelled flooded area for both the baseline and future time periods, whilst the source of uncertainty contributing the most to the range of output is marginally EVD (over CMU). This finding is irrespective of return period explored (1:10 vs. 1:100), an interesting outcome particularly given the assumption that EVD uncertainty increases with higher return period estimates. This would appear to agree with previous studies, e.g. Apel et al. (2018).

Varying the roughness (test c) results in a range of flood extents across similar Q_{peak} values (roughness influencing the water level and thus extent). This adds a width or fuzziness to the flooded area ~ Q_{peak} relationship (Fig. 4a, b) and results from the consideration of two different factors (dimensions) within the uncertainty analysis.

It is interesting to notice from Fig. 3 that, independent of the test, inflexion points exist in the Q_{peak} ~ flooded area relationship, where the relationship between these variables changes. These points are observed throughout all four tests (for both baseline and future) and for both return period simulation ensembles, at around $Q_{peak} = 230$ m³/s and around $Q_{peak} = 270$ m³/s. The relationship with flooded area flattens out between these two Q_{peak} values (230–270 m³/s), indicating that the flooded area is increasing more rapidly with an increasing Q_{peak}. It is very likely that these inflexions are related to topographic controls in the case study area. These observed relationships are context specific, and the details are non-transferable between case studies.

4.2.2 Total uncertainty—test d

Test d (Figs. 3d, 5d) explored the total uncertainty through consideration of the full parameter space. From Table 3, it can be seen that Δmean and Δrange are within the range of the individual sources of uncertainty (tests a, b and c). Two dimensions of uncertainty, the hydraulic model parameter (HMPU) and the input uncertainty (Q_{peak}), arising from both climate model and extreme value distribution uncertainties (CMU and EVDU) are shown to result in fuzziness in the relationship. Figure 6 shows clearly the impact on the flood area and the change in probability of occurrence from the baseline to the future time periods. These probabilistic flood maps show areas which become much more likely to be flooded in the future given the same return period event, as well as some areas which are newly inundated in the future.

4.3 Probabilistic framework: stratified sampling LHS results

Figure 7 shows the results of the convergence of the mean and standard deviation for the LHS study for the four tests over two time periods (baseline and future). In the first three tests, where the sources of uncertainty are considered separately, the sample mean
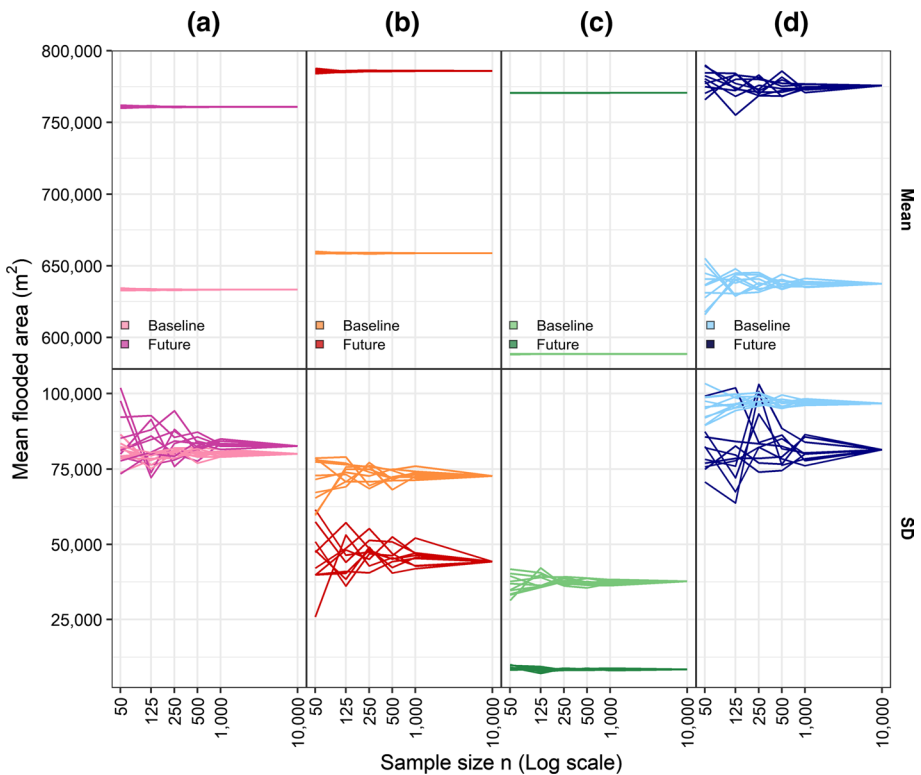

Fig. 7 Convergence of sample mean and standard deviation for Latin hypercube stratified sampling technique (towards full Monte Carlo analysis results): **a** climate model (CMU), **b** extreme value (EVDU), **c** hydraulic model parameter (HMPU), **d** total uncertainty

converges for $n = 1000$, thereby suggesting that the $n = 1,00$ runs are able to capture the full uncertainty from the Monte Carlo analysis ($n = 10,000$). This is seen clearly on the baseline where the uncertainty range in the inputs is lower than for the future. For the hydraulic model parameters, the uncertainty range is relatively small; despite this, the uncertainty is well captured. By comparison, when the total uncertainty is explored using the full parameter space, the LHS stratified sampling performs poorly. This is likely due to the increase in dimensionality (three components, collapsed into two parameters) influencing the efficacy of stratified sampling; this is further considered in the discussion.

A total of ten ensembles were considered for each $n = \{50, 125, 250, 500, 1,000\}$ sample size. For the total uncertainty, the frequency distribution for each sample size is presented for the baseline (Fig. 8) and future (Fig. 9). For the baseline (Fig. 8), it can be seen that the shape of the distribution (bimodal) is well replicated from $n = 125$; however, with high variability across the ensemble, the associated uncertainty is high. At $n = 1000$, the variability is significantly reduced and in the variability associated with the ensemble, ability to replicate the shape has reduced. For the future time period, this reduction in variability is more distinct given the greater range of flooded area. In the future (Fig. 9), the sample size required for suitable replication of the frequency distribution increases to a minimum of $n = 250$. Improvements in the replication of the tails of the distribution (900,000–1,000,000 m^2) can be seen at $n = 500$, whilst the ensemble can be seen to move towards convergence at $n = 1000$. From Figs. 8 and 9, it is clear that the LHS stratified sampling is able to provide a reasonable representation of the full Monte Carlo analysis. However, it is also apparent that $n = 100$ is insufficient to capture the full range of uncertainty (as n^2 MC simulations; Aistleitner et al. 2012).

5 Discussion

5.1 Main results: uncertainty sources

For the full Monte Carlo simulations (for both return period events), the results for the baseline and future suggest that the dominant source of uncertainty is the estimated input hydrograph. This is a reflection of both the uncertainties associated with the parameterization of the regional climate model (CMU) and the extreme value method (EVDU). These findings are consistent for both return periods tested and with recent studies by Mukolwe et al. (2014) and Ali et al. (2015). It can thus be inferred that the relative scale of input hydrograph uncertainty to hydraulic model parameter uncertainty (CMU and EVDU > HMPU) is transferable between sites. However, the specifics of the relationship between Q_{peak} and flooded area are location and context specific, with local topographic controls, such as floodplains controlling the processes; thus, each case study location will deliver different findings. In such cases, where wetting and drying of cells becomes critical and therefore more sensitive to the modelling process (parameterization, technique, mathematical formulation), the relationship may become more complex.

The results presented in Figs. 3 and 4 suggest that there are certain inflexions within the Q_{peak} and flooded area relationship, where the relationship changes (230 m^3/s and 270 m^3/s). Here, there is a range of values of Q_{peak} for which flooded area extent does not vary much despite a relatively larger increase in Q_{peak}. The identification of these thresholds may be possible with fewer simulations depending on the complexity of the modelled stretch; however, probabilistic methods are required to fully represent the
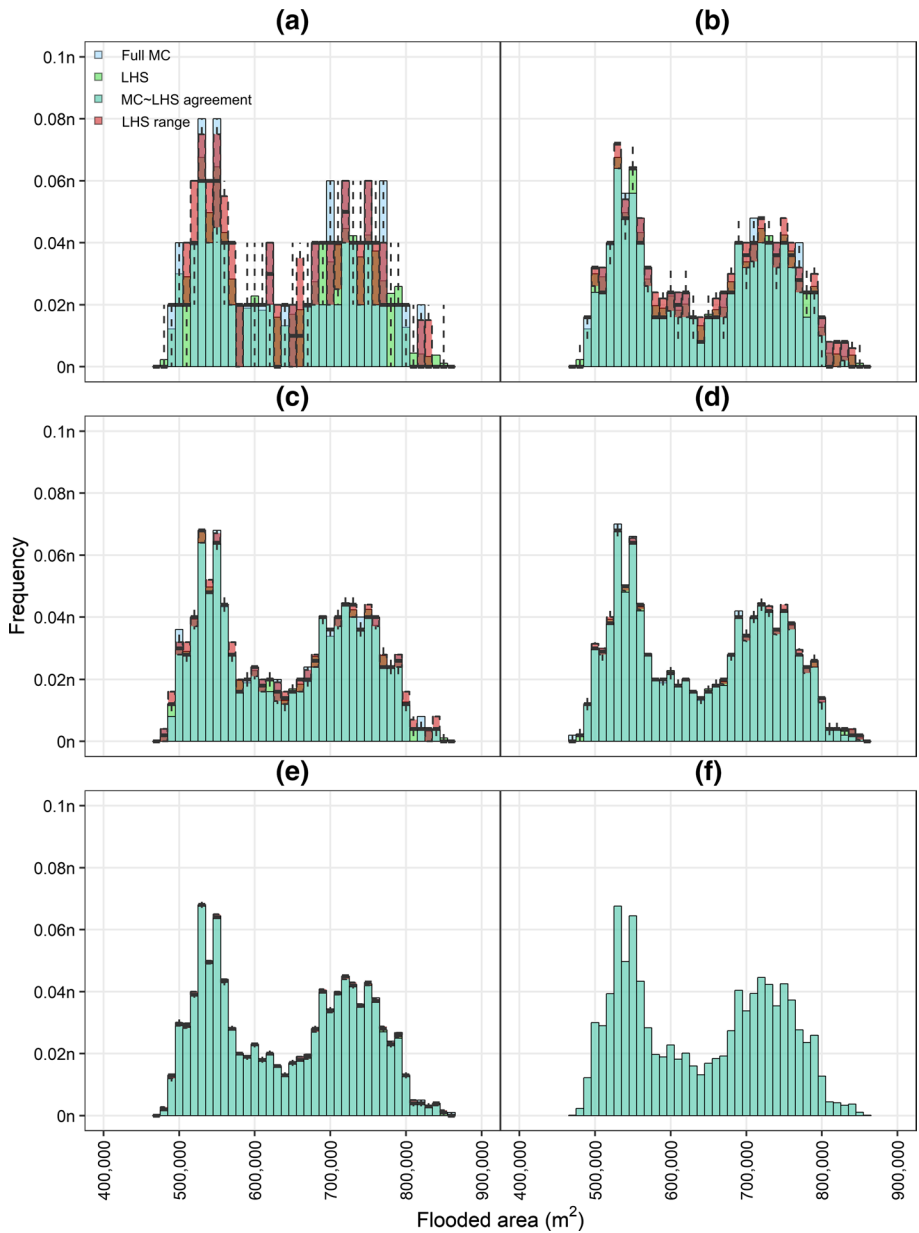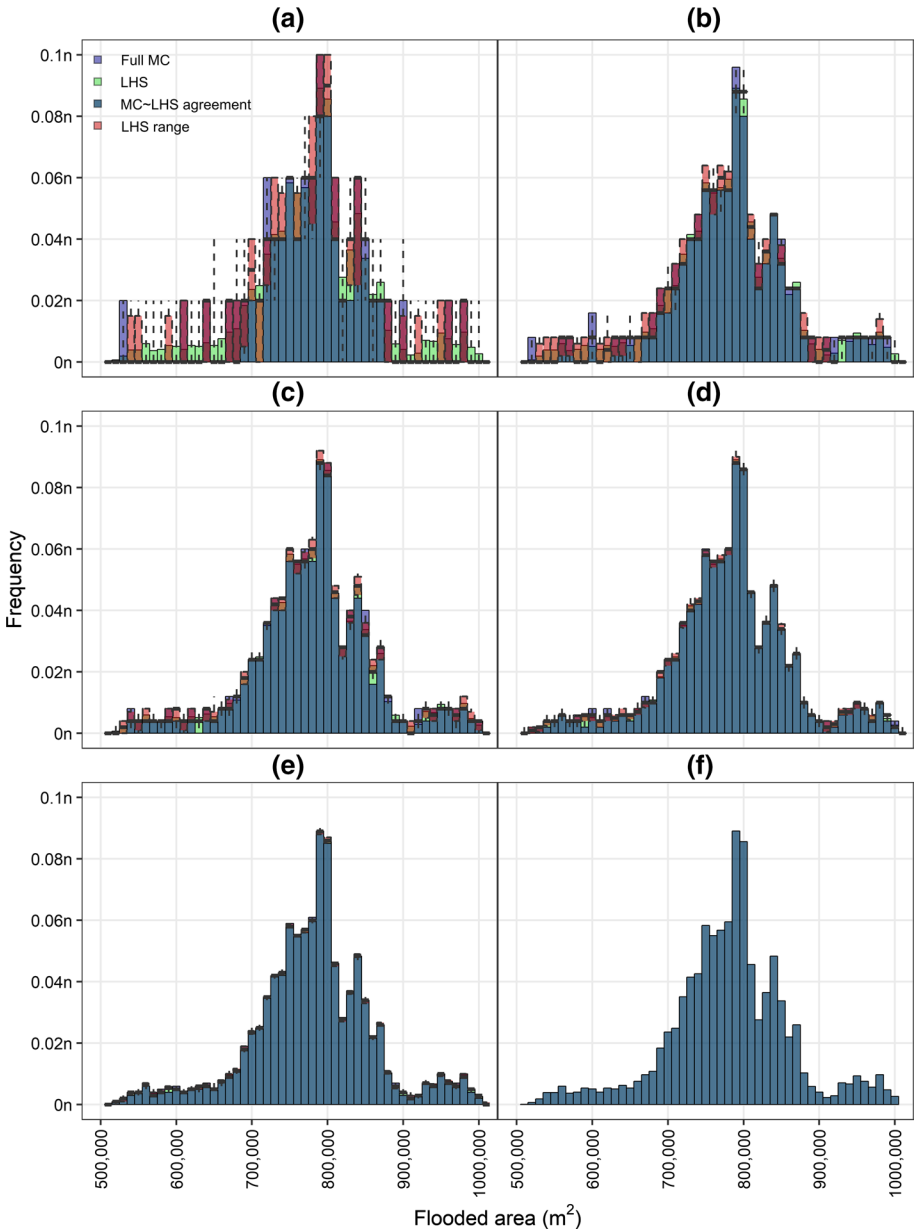
Fig. 8 Frequency distribution of the total uncertainty on the baseline by sample size $n = \{50, 125, 250, 500, 1000, 10000\}$ across **a–f**. The uncertainty captured by the full MC analysis, the ensemble mean and ensemble range are represented in light blue, green and orange/brown, respectively

relationship and provide useful insights for planners and consultants. For example, in the UK this could inform the specification of flood zones in policy guidance from the Environment Agency, Natural Resources Wales or the Scottish Environmental Protection Agency.

Fig. 9 Frequency distribution of the total uncertainty for the future projections by sample size $n = \{50, 125, 250, 500, 1000, 10000\}$ across **a–f**. The uncertainty captured by the full MC analysis, the ensemble mean and ensemble range are represented in dark blue, green and orange/brown, respectively

The quantification of this relationship may allow for the use of surrogate models. Surrogate modelling is a technique that approximates a more computationally expensive model whilst honouring the underlying physics. A small sample size of well-chosen data points is used to model the response of the original model, creating an emulator that requires fewer simulations

to quantify the uncertainty. Surrogate models have already demonstrated their applicability to hydrological uncertainty quantification (Roy et al. 2018; Moreno-Rodenas et al. 2018) with the most common surrogate models being polynomial chaos expansion (PCE) and Gaussian processes (GP). The reduced sample size allows for more numerically intensive models to be used with little increase in computational burden, producing more accurate and efficient flood hazard assessments.

Non-stationarity in climate change impacts to flood inundation extents.

Climate ensembles project changes in climatic variables such as temperature and precipitation (IPCC 2014), the two main drivers of the hydrological cycle. Research into the effects of climate change on the hydrological system has been ongoing in excess of 20 years (Augustin et al. 2008; Christierson et al. 2012). Although highly uncertain, recent work in Great Britain and the UK indicates an important change in future extreme flows (Collet et al. 2017; 2018a, b); spatial variability has also been highlighted as an important factor (Thober et al. 2018; Visser-Quinn et al. 2019).

To understand the impact on population, exposure and subsequent risk, there is a clear need to characterize and quantify this uncertainty, cascade this into flood inundation assessment using flood models and compare consistently between the present climate (baseline) and in the future (e.g. the 2080s) climate model estimates. The results of this study show that input hydrographs dominate the uncertainty in flood modelling and interestingly showed that uncertainty associated with climate model parameterization is at least as important to capture as the uncertainty associated with the extreme value distribution fit, irrespective of return period estimate. The range of possible inundation extents which capture the uncertainty in input parameters is significant for both return periods on the baseline and future, which suggests that probabilistic flood maps have an important role in future flood inundation assessments.

Presently, the UK utilizes standard hard edge flood maps. Derivation of probabilistic flood maps (such as Fig. 6) for a 1:100-year return period flood extent could serve to illustrate the projected change in flood probability and thus exposure through a model domain. Whilst this is computationally expensive and may be more difficult to communicate, probabilistic maps can provide a new level of information for use in future adaptation planning. Work over the last decade (Baldassarre and Montanri 2009) has increasingly called for this change in the way flood mapping is considered. Recent studies (Teng et al. 2017) predict that, going forward, probabilistic mapping will be in greater demand as the industry, insurance companies and policy makers alike grapple with how best to capture and convey uncertainty in projections within the context of flood legislation such as the EU Floods Directive (binding legislation for EU members states).

With an increase in mean flooded area for total uncertainty of between 18 and 22% (based on the two return periods explored in this paper), this work shows that it is critical to capture the change to inflows as a result of a non-stationary climate in future flood assessments and acknowledge the uncertainty associated with its projection. Thus, it seems sensible to suggest the non-stationarity of climate and its influence on extreme events in the future; alongside the uncertainty associated with these projections needs to be captured more routinely and explicitly in flood risk assessments, through probabilistic flood hazard maps.

5.2 Reducing the computational cost of uncertainty studies

This study has highlighted the ability of LHS to reduce computational cost with minimal to no loss of accuracy. Figure 7 shows that, for each source of uncertainty considered, a sample size of $n = 500$ was able to converge to the mean and standard deviation of the full

Monte Carlo simulation. For the total uncertainty, where both Manning's n and the input hydrograph are varied, a sample size of $n = 1000$ was required (Figs. 8 and 9) to highlight variability in the LHS ensemble, suggesting $n = 1000$ provides the better replication of the full uncertainty. This was shown to be true for non-normal distributions (bimodal) in particular, with thresholds in the topography influencing tipping points (i.e. either wet or dry).

The theory indicated that LHS is quadratically faster, achieving the same accuracy as n^2 Monte Carlo simulations (Aistleitner et al. 2012). The empirical findings in this study suggest that in order to achieve adequate replication of the mean, standard deviation, range and frequency of the flooded area distribution, a sample size of $n = 1000$ is required (i.e. $\left(\frac{n}{10}\right)^2$). This is still computationally intensive, and the added complication is the dimensionality of the uncertainty parameters considered in this study. Therefore, there is a critical need to investigate ways in which the computational cost can be reduced. In this study, a reduced physics numerical engine (LISFLOOD-FP) was used to drive the analysis. The benefit of this approach is the speed of numerical calculation. Whilst reducing the number of runs by a factor of ten is useful, there is a clear need for further work. However, in industry the majority of flood models use more computationally intensive fully 2D flood models (e.g. TuFLOW, TELEMAC, MIKE21) which are not well suited to this probabilistic approach. Developments in other fields may provide alternative solutions. One example is Markov chain Monte Carlo (MCMC), an uncertainty quantification method that uses Markov chain to sample from the input distribution. The Markov chain is essentially a random walk through the parameter space with each step depending only upon the previous value (memoryless). The most common MCMC method is the Metropolis–Hastings algorithm (Chib and Greenberg 1995) which requires a reduced number of input values compared to full Monte Carlo. The MCMC has been shown to converge faster than Monte Carlo (Elsakout et al. 2015). With this increased rate of convergence, the application of more complex hydraulic models requiring larger simulation times is more feasible. The computational cost may be further reduced through application of Multi-Level Markov Chain Monte Carlo (MLMCMC), where multiple grids are used to estimate quantities of interest, whilst applying a Markov chain approach to sampling to ensure computational efficiency (Elsakout et al. 2015).

5.3 Limitations of the study

There are some limitations which are worth noting for this study and which warrant further consideration in future research. This study utilized data from the FFH database for a single emissions scenario, SRES A1B. This data is derived from the UKCP09 projections, a perturbed physical ensemble based on a single climate model (HADCM3). Projections from a hydroclimatological modelling chain that utilizes a multi-model ensemble may yield different results (e.g. the EDgE projections, Thober et al. 2018). Other important considerations for future work include:

- Modelling with the most recent CMIP5 projections;
- Comparison across multiple emissions scenarios (e.g. RCP2.6 and RCP8.5);
- Modelling across a range of topographical diverse catchments or longer river reaches, which explores the dynamic interaction of floodplain inundation, flood defence failures and flood management within river reaches—adding hydraulic complexity into the analysis.

This is particularly relevant for studies looking at surrogate methods (Sect. 5.1) and alternative uncertainty quantification methods (Sect. 5.3). Whilst this study demonstrated a reasonable reduction in computational cost through the application of a LHS approach, fully 2D codes, as used in industry, require more than a tenfold reduction in simulations. Alternative approaches need to be explored in order to open up the possibility of probabilistic flood assessments as part of routine practice, particularly in industry.

6 Conclusions

This study has explicitly captured the uncertainty associated with climate model parameterization and the flood modelling process (extreme value estimation and hydraulic model roughness specifically) using a probabilistic framework and investigated the influence of climate change projections on future flood hazard predictions. The findings clearly demonstrate the need to capture climate change projections in future flood inundation estimation. The results suggest that the uncertainties associated with hydrological uncertainty, or the flow input, result in the most substantial uncertainty in the flood inundation extent. Certainly, hydraulic parameter uncertainty is considerably less influential on the spread of inundation extents. These findings are consistent with recent studies (Mukolwe et al. 2014; Ali et al. 2015).

Uncertainty in the inflow parameters is important to the eventual outcome (inundation extents), and thus, there is a need to routinely include this uncertainty into flood inundation studies. From the findings of this study, the uncertainty associated with extreme value distribution models can result in a reasonably similar range of inundation extents as the uncertainty arising from climate model parameterization; however, more importantly non-stationarity in climate (i.e. the impact of climate change) results in a considerable shift in the mean of inundation extents, as well as the range. Thus, alongside capturing the uncertainty in hydrological inputs and cascading this through hydraulic models, there is an urgent need to include non-stationarity in the climate.

Tipping points in the relationships between the peak flows and the resulting inundation extents can be useful for policy and in practice (e.g. planning control); however, these will be very context specific. These tipping points are strongly controlled by topographic features in the flood study areas, and this study constrained this issue through the selection of a short, reasonably simple hydraulic reach. For example, large floodplains will influence the relationship between flood peak and inundation area. Further investigation of such relationships is needed through more case study assessments.

Finally, in practice the ability to run repeated hydraulic simulations to capture the full uncertainty in the input and parameter space is unlikely. Due to computational burden, capturing the full uncertainty is likely to be a luxury afforded in the research environment only. Thus, there is a real need to investigate the opportunities to reduce computational burden by reducing the number of hydraulic model runs necessary whilst still capturing the full range on possible inundation extents. This paper has explored a simple first approach to reducing simulations through the use of a stratified sampling approach (LHS). This method, whilst found to be effective, still has not reduced the number of runs sufficiently for it to be a routinely viable option in practice. Therefore, more sophisticated methods are needed to quantify uncertainty when considering uncertain parameter space in more than one dimension. This paper has suggested several different methods that could be explored,

by looking to other disciplines for inspiration, and is recommended as an avenue for future research.

References

Aistleitner C, Hofer M, Tichy R (2012) A central limit theorem for Latin hypercube sampling with dependence and application to exotic basket option pricing. Int J Theor Appl Finance 15(7):1250046. https://doi.org/10.1142/S021902491250046X

Ali MA, Solomatine D, Baldassarre G (2015) Assessing the impact of different sources of topographic data on 1-D hydraulic modelling of floods. Hydrol Earth Syst Sci 19:631–643. https://doi.org/10.5194/hess-19-631-2015

Allen MR, Dube OP, Solecki W, Aragón-Durand F, Cramer W, Humphreys S, Kainuma M, Kala J, Mahowald N, Mulugetta Y, Perez R, Wairiu M, Zickfeld K (2018) Framing and context. In: Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR, Pirani A, Moufouma-Okia W, Péan C, Pidcock R, Connors S, Matthews JBR, Chen Y, Zhou X, Gomis MI, Lonnoy E, Maycock T, Tignor M, Waterfield T (eds) Global warming of 1.5 °C. An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty

Apel H, Thieken AH, Merz B, Blöschl G (2004) Flood risk assessment and associated uncertainty. Nat Hazards Earth Syst Sci 4:295–308

Apel H, Thieken AH, Merz B, Blöschl G (2006) A probabilistic modelling system for assessing flood risks. Nat Hazards 38:79–100. https://doi.org/10.1007/s11069-005-8603-7

Augustin N, Beevers L, Sloan W (2008) Predicting river flows for future climates using an autoregressive multinomial logit model'. Water Resour Res. https://doi.org/10.1029/2006WR005127

Bates PD, De Roo APJ (2000) A simple raster-based model for flood inundation simulation. J Hydrol 236:54–77

Baldassarre G, Montanri A (2009) Uncertainty in river discharge observations: a quantitative analysis Hydrol. Earth Syst Sci 13:913–921

Baldassarre G, Schumann G, Bates P, Freer J, Beven K (2010) Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. Hydrol Sci J 55(3):364–376

Balica SF, Beevers L, Popescu I, Wright NG (2013) Parametric and physically based modelling techniques for flood risk and vulnerability assessment: a comparison. J Environ Model Softw 41(3):81–92. https://doi.org/10.1016/j.envsoft.2012.11.002

Beevers L, Douven W, Lazuardi H, Verheij H (2012) Cumulative impacts of road developments in floodplains. Transp Res Part D Transp Environ 17(5):398–404

Beevers L, Walker G, Strathie A (2016) A systems approach to flood vulnerability. Civil Eng Environ Syst 33(3):199–213. https://doi.org/10.1080/10286608.2016.1202931

Berends K, Warmink JJ, Hulscher SJMH (2018) Efficient uncertainty quantification for impact analysis of human interventions in rivers. Environ Model Softw 107:50–58. https://doi.org/10.1016/j.envsoft.2018.05.021

Chib S, Greenberg E (1995) Understanding the metropolis–hastings algorithm. Am Stat 49(4):327–335. https://doi.org/10.2307/2684568

Chow VT (1959) Open channel hydraulics. McGraw-Hill, New York

Christierson BV, Vidal JP, Wade SD (2012) Using UKCP09 probabilistic climate information for UK water resource planning. J Hydrol 425:48–67. https://doi.org/10.1016/j.jhydrol.2011.12.020

Coles S (2001) An Introduction To Statistical Modelling Of Extreme Values. Springer, London

Collet L, Beevers L, Prudhomme C (2017) Assessing the impact of climate change on extreme flows across Great Britain. Water 9(2):103. https://doi.org/10.3390/w9020103

Collet L, Beevers L, Stewart M (2018a) Decision-making and flood risk uncertainty: statistical dataset analysis for flood risk assessment. Water Resour Res. https://doi.org/10.1029/2017WR022024

Collet L, Formetta G, Harrigan S, Prudhomme C, Beevers L (2018b) Future hot-spots for hydro-hazards in Great Britain: a probabilistic assessment. Hydrol Earth Syst Sci 22(10):5387–5401

Curran A, de Bruijn KM, Klerk WJ, Kok M (2019) Large scale flood hazard analysis by including defence failures on the Dutch river system. Water 11:1732. https://doi.org/10.3390/w11081732

Elsakout DMA, Christie MA, Lord GJ (2015) Multilevel Markov Chain Monte Carlo (MLMCMC) for uncertainty quantification. In: SPE North Africa technical conference and exhibition, 14–16 September, Cairo, Egypt Society of Petroleum Engineers

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Guha-Sapir D (2018) EM-DAT: The Emergency Events Database - Universite catholique de Louvain (UCL) - CRED, - www.emdat.be, Brussels, Belgium (Accessed: 22/11/2018).

Guerreiro S, Dawson RJ, Kilsby C, Lewis E, Ford A (2018) Future heat-waves, droughts and floods in 571 European cities. Environ Res Lett. https://doi.org/10.1088/1748-9326/aaaad3

Hartanto IM, Beevers L, Popescu I, Wright NG (2011) Application of a coastal modelling code in fluvial environments. Environ Model Softw 26(12):1685–1695

Helton JF, Davis J (2003) Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Saf 81:23–69

Huntington DC, Lyrintzist S (1998) Improvements to and limitations of Latin hypercube sampling. Prob Engng Mech 13(4):245–253

IPCC (2014) Climate change 2014: synthesis report. In: Pachauri K, Meyer LA (eds) Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change [core writing team]. IPCC, Geneva, Switzerland

Jonkman S, Vrijling J (2008) Loss of life due to floods. J Flood Risk Manag 1:43–56

Metin AD, Dung NV, Schröter K, Guse B, Apel H, Kreibich H, Vorogushyn S, Merz B (2018) How do changes along the risk chain affect flood risk? Nat Hazards Earth Syst Sci 18(3089–3108):2018. https://doi.org/10.5194/nhess-18-3089-2018

Moreno-Rodenas AM, Bellos V, Langeveld JG, Clemens FHLM (2018) A dynamic emulator for physically based flow simulators under varying rainfall and parametric conditions. Water Res 142:512–527

Mukolwe MM, Di Baldassarre G, Werner MGF, Solomatine DP (2014) Flood modelling: parameterisation and inflow uncertainty. In: Proceedings of the ICE—water management, vol 167, pp 51–60

Mundform DJ, Schaffer J, Kim M-J, Shaw D, Thongteeraparp A, Supawan P (2011) Number of replications required in Monte Carlo simulation studies: a synthesis of four studies. J Mod Appl Stat Methods 10(1):4. https://doi.org/10.22237/jmasm/1304222580

Neal J, Keef C, Bates P, Beven K, Leedal D (2013) Probabilistic flood risk mapping including spatial dependence. Hydrol Process 27:1349–1363

Prudhomme C, Haxton T, Crooks S, Jackson C, Barkwith A, Williamson J, Kelvin J, Mackay J, Wang L, Young A, Watts G (2013) Future Flows Hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. Earth Syst Sci Data 5:101–107. https://doi.org/10.5194/essd-5-101-2013

Roy NPT, El Moçayd A, Ricci S, Jouhaud N, Goutal M, De Lozzo M, Rochoux C (2018) Comparison of polynomial chaos and Gaussian process surrogates for uncertainty quantification and correlation estimation of spatially distributed open-channel steady flows. Stoch Env Res Risk Assess 32(6):1723

Savage J, Bates P, Freer J, Neal J, Aronica G (2014) The impact of scale on probabilistic flood inundation maps using a 2D hydraulic model with uncertain boundary conditions. In: Vulnerability, uncertainty, and risk, pp 279–289

Savage J, Bates P, Freer J, Neal J, Aronica GT (2016) When does spatial resolution become spurious in probabilistic flood inundation predictions? Hydrol Process 30(13):2014–2032. https://doi.org/10.1002/hyp.10749

Sayers PB, Horritt M, Penning-Rowsell E, McKenzie A (2015) Climate change risk assessment 2017: projections of future flood risk in the UK. Research undertaken by Sayers and Partners on behalf of the Committee on Climate Change. Published by Committee on Climate Change, London

SEPA (2015) Provision of numerical model. Personal communication

Thober S, Kumar R, Wanders N, Marx A, Pan M, Rakovec O, Samaniego L, Sheffield J, Wood EF, Zink M (2018) Multi-model ensemble projections of European river floods and high flows at 1.5, 2, and 3 degrees global warming. Environ Res Lett 13(1):014003

Teng JAJ, Jakeman J, Vaze BFW, Croke D, Dutta S, Kim s(2017) Flood inundation modelling: A review of methods, recent advances and uncertainty analysis, Environmental Modelling & Software 90:201–216, https://doi.org/10.1016/j.envsoft.2017.01.006.

Van Vuren S (2005) Stochastic modelling of river morphodynamic. PhD thesis, Delft University of Technology, IOS Press, Amsterdam, Netherlands. ISBN: 90-407-2605-1

Visser-Quinn AG, Beevers L, Collet L, Formetta G, Smith K, Wanders N, Thober S, Pan M, Kumar R (2019) Spatio-temporal analysis of compound hydro-hazard extremes across the UK. Adv Water Resour. https://doi.org/10.1016/j.advwatres.2019.05.019

Warmink J, Janssen J, Booij M, Krol M (2010) Identification and classification of uncertainties in the application of environmental models. Environ Model Softw 25(12):1518–1527

Winter B, Schneeberger K, Huttenlau M, Stotter J (2018) Sources of uncertainty in a probabilistic flood risk model. Nat Hazards 91:431. https://doi.org/10.1007/s11069-017-3135-5

Wobus C, Gutmann E, Jones R, Rissing M, Mizukami N, Lorie M, Mahoney H, Wood AW, Mills D, Martinich J (2017) Climate change impacts on flood risk and asset damages within mapped 100-year floodplains of the contiguous United States. Nat Hazards Earth Syst Sci 17(2199–2211):2017. https://doi.org/10.5194/nhess-17-2199-2017