

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2020WR028673

### Key Points:

- First complete high-resolution flood hazard analysis of conterminous US flood risk from all major sources (fluvial, pluvial, and coastal)
- In validation tests the model achieved Critical Success Index scores of 0.69–0.82, similar to many local custom-built 2D models
- By 2050, flood hazard increases for the Eastern seaboard and Western states, but decreases or changes little for the center and South-West

### Supporting Information:

- Supporting Information S1

### Correspondence to:

P. D. Bates,  
paul.bates@bristol.ac.uk

### Citation:

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., et al. (2021). Combined modeling of US fluvial, pluvial, and coastal flood hazard under current and future climates. *Water Resources Research*, 57, e2020WR028673. <https://doi.org/10.1029/2020WR028673>

Received 25 AUG 2020

Accepted 10 DEC 2020

## Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates

Paul D. Bates<sup>1,2</sup> , Niall Quinn<sup>2</sup> , Christopher Sampson<sup>2</sup>, Andrew Smith<sup>2</sup>, Oliver Wing<sup>1,2</sup> , Jeison Sosa<sup>2</sup> , James Savage<sup>2</sup>, Gaia Olcese<sup>2</sup> , Jeff Neal<sup>1,2</sup> , Guy Schumann<sup>3</sup> , Laura Giustarini<sup>3</sup>, Gemma Coxon<sup>1</sup> , Jeremy R. Porter<sup>4,5,6</sup>, Mike F. Amodeo<sup>4</sup> , Ziyan Chu<sup>4</sup>, Sharai Lewis-Gruss<sup>4</sup>, Neil B. Freeman<sup>4</sup> , Trevor Houser<sup>7</sup>, Michael Delgado<sup>7</sup> , Ali Hamidi<sup>7</sup> , Ian Bolliger<sup>7</sup> , Kelly E. McCusker<sup>7</sup> , Kerry Emanuel<sup>8</sup>, Celso M. Ferreira<sup>9</sup> , Arslaan Khalid<sup>9</sup>, Ivan D. Haigh<sup>10</sup>, Anaïs Couasnon<sup>11</sup> , Robert E. Kopp<sup>12</sup> , Solomon Hsiang<sup>13</sup>, and Witold F. Krajewski<sup>14</sup> 

<sup>1</sup>School of Geographical Sciences, University of Bristol, Bristol, UK, <sup>2</sup>Fathom™, Square Works, Bristol, UK, <sup>3</sup>RSS-Hydro, Innovation Hub Dudelange, Dudelange, Luxembourg, <sup>4</sup>First Street Foundation, Brooklyn, NY, USA, <sup>5</sup>Quantitative Methods in the Social Sciences, City University of New York, New York, NY, USA, <sup>6</sup>Environmental Health Sciences, Columbia University, New York, NY, USA, <sup>7</sup>Rhodium Group, New York, NY, USA, <sup>8</sup>Lorenz Center, Massachusetts Institute of Technology, Cambridge, MA, USA, <sup>9</sup>Civil, Environmental, and Infrastructure Engineering, George Mason University, Fairfax, VA, USA, <sup>10</sup>School of Ocean and Earth Science, National Oceanography Centre, University of Southampton, European Way, Southampton, UK, <sup>11</sup>Institute for Environmental Studies, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, <sup>12</sup>Department of Earth & Planetary Sciences and Rutgers Institute of Earth, Ocean, and Atmospheric Sciences, Rutgers University, New Brunswick, NJ, USA, <sup>13</sup>Global Policy Lab, University of California Berkeley, Berkeley, CA, USA, <sup>14</sup>IIHR—Hydroscience and Engineering, University of Iowa, Iowa City, IA, USA

**Abstract** This study reports a new and significantly enhanced analysis of US flood hazard at 30 m spatial resolution. Specific improvements include updated hydrography data, new methods to determine channel depth, more rigorous flood frequency analysis, output downscaling to property tract level, and inclusion of the impact of local interventions in the flooding system. For the first time, we consider pluvial, fluvial, and coastal flood hazards within the same framework and provide projections for both current (rather than historic average) conditions and for future time periods centered on 2035 and 2050 under the RCP4.5 emissions pathway. Validation against high-quality local models and the entire catalog of FEMA 1% annual probability flood maps yielded Critical Success Index values in the range 0.69–0.82. Significant improvements over a previous pluvial/fluvial model version are shown for high-frequency events and coastal zones, along with minor improvements in areas where model performance was already good. The result is the first comprehensive and consistent national-scale analysis of flood hazard for the conterminous US for both current and future conditions. Even though we consider a stabilization emissions scenario and a near-future time horizon, we project clear patterns of changing flood hazard ( $3\sigma$  changes in 100 years inundated area of  $-3.8$  to  $+16\%$  at 1° scale), that are significant when considered as a proportion of the land area where human use is possible or in terms of the currently protected land area where the standard of flood defense protection may become compromised by this time.

**Plain Language Summary** We develop a method to estimate past, present, and future flood risk for all properties in the conterminous United States whether affected by river, coastal or rainfall flooding. The analysis accounts for variability within environmental factors including changes in sea level rise, hurricane intensity and landfall locations, precipitation patterns, and river discharge. We show that even for a conservative climate change trajectory we can expect locally significant changes in the land area at risk from floods by 2050, and by this time defenses protecting 2,200 km<sup>2</sup> of land may be compromised. The complete dataset has been made available via a website (<https://floodfactor.com/>) created by the First Street Foundation in order to increase public awareness of the threat posed by flooding to safety and livelihoods.

© 2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

The most damaging flood events have a regional or even continental scale footprint. For example, eight of the top 10 costliest weather and climate disasters between 1980 and 2019 in the US (see <https://www.ncdc.noaa.gov/billions/events>) were floods that affected wide regions, with the most damaging being hurricane Katrina in which flooding affected nine states and resulted in a monetary loss of \$168.8Bn at 2019 prices. Total losses such as this are nationally significant for the countries affected and for insurers, yet the reality is that all flood impacts are felt locally and occur at the level of individual properties. Whether a particular property floods or not depends strongly on its immediate topographic setting, the presence of flood defenses in the locality and the fine scale dynamics of floodplain inundation. Unbiased estimates of large-scale risk therefore need to be built “ground up” from models that have sufficient skill to represent flood hazard at local scales, yet which can be applied over large domains. The traditional approach to this problem taken by regulatory agencies (e.g., the Federal Emergency Management Agency [FEMA] in the US and the Environment Agency in England) is to commission many thousands of local scale hydraulic modeling studies to delimit floodplain areas and then merge these into national maps. However, the cost and time taken to produce a national coverage in this way can be prohibitive, the coverage may not be comprehensive and an amalgamation of local studies by different modeling teams can result in a lack of internal consistency. For example, the total cost (in 2019 dollar values) of FEMA’s national flood mapping program from its inception in 1969–2020 was \$10.6Bn, while covering only 33% of the rivers and streams in the country to variable levels of quality (Association of State Floodplain Managers, 2020).

There is also mounting concern that flood hazards will change significantly in the future due to anthropogenic climate change (Alfieri et al., 2016; Cloke et al., 2013; Hirabayashi et al., 2013; Kundzewicz et al., 2014) that will result in rising sea levels (Field et al., 2012; Oppenheimer et al., 2019) and very likely cause an intensification of the hydrological cycle (Madakumbura et al., 2019; Markonis et al., 2019). Socio-economic shifts will also likely affect vulnerability and exposure (Hallegatte et al., 2013; Jongman et al., 2012; Wing et al., 2018) and to adapt to these changing conditions, analyses are required that cannot only skillfully, comprehensively and consistently predict current hazard at fine spatial resolution, but which can also make robust projections of these quantities into the future. However, assessing changing flood hazards using traditional regulatory approaches based on a patchwork of local models that take decades of time and billions of dollars to develop is effectively impossible.

To address these issues, the last 5 years have seen the extension of inundation modeling approaches to regional, continental or even global scales (Dottori et al., 2016; Hirabayashi et al., 2013; Kulp & Strauss, 2019; Sampson et al., 2015; Wing et al., 2017; Winsemius et al., 2013) for both current and future conditions. These methods are able to produce large scale and internally consistent inundation maps within reasonable timeframes and resource limits, and thus represent a significant advance. They nevertheless all suffer from a number of critical limitations that prevent the development of a fine resolution, comprehensive, and actionable view of flood hazard.

First, all large-scale modeling studies to date examine only a subset of flood drivers. Floods can arise from a number of sources including extreme river flows, rainfall, coastal water levels and waves, and more rarely groundwater or infrastructure failure (e.g., dam collapses). However, at large scales, only Sampson et al. (2015), Wing et al. (2017), and Eilander et al. (2020) have looked at more than one causal mechanism, focusing on either rainfall and river flow or storm surge and river flow. There is thus a pressing need to develop a large-scale multiperil flood risk analysis that includes pluvial, fluvial, and coastal flooding. Low-lying coastal catchments pose a particular challenge in this regard. Here extreme rainfall, river discharge, coastal water levels and waves, often associated with tropical cyclones, can all combine to cause dangerous floods. Moreover, in some cases it is only the joint occurrence of multiple flood drivers simultaneously that results in a hazard (Gallien et al., 2018; Moftakhar et al., 2019; Wahl et al., 2015; Ward et al., 2018). For specific coastal sites, a small number of modeling studies (Couasnon et al., 2018; Lian et al., 2013; Pasquier et al., 2019; van de Hurk et al., 2015; Wing et al., 2019) have started to look at compound events, but as Gallien et al. (2018) point out there is no “integrated framework considering compound and multipathway flooding processes in a unified approach.” At larger (regional to global) scales, studies have begun to look at the joint probability of two (Bevacqua et al., 2019; Couasnon et al., 2020; Hendry et al., 2019; Keef et al., 2009; Wahl et al., 2015; Ward et al., 2018; Zheng et al., 2013) or even three different flood drivers

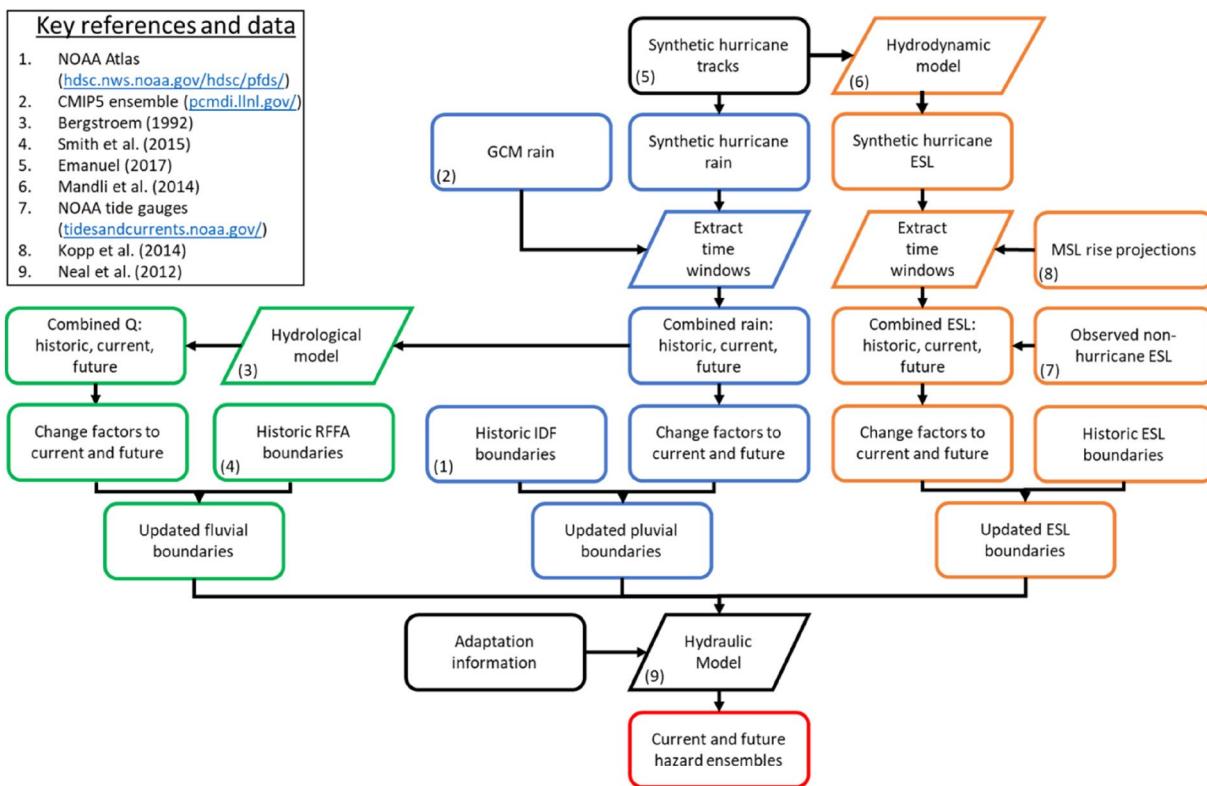
(Svensson & Jones, 2004), however, these studies only examine the occurrence distributions and do not use these as boundary conditions for a model analysis of hazard.

Second, the spatial resolution, exposure datasets, and flow physics in many large-scale inundation models may be too crude to represent property scale flooding. As the intercomparison of global flood models by Trigg et al. (2016) makes clear, many global flood models only simulate dynamic flows at coarse resolution (0.25–0.5°) and then downscale these results using finer scale DEM data to resolutions that are more amenable for hazard analysis. Even then, many global hazard analyses are carried out only at 18–30 arc sec (~540–900 m at the equator) resolution, which cannot resolve local features. Exposure datasets can also be inadequate for the mapping of local risk. For example, Smith et al. (2019) show that current estimates of global flood risk are often made using population maps that distribute people homogenously and unrealistically across large lowland floodplain areas. Although some local misprediction will cancel out when the results generated using these models and datasets are aggregated to national and global scales, considerable bias can still remain as demonstrated by Trigg et al. (2016). To illustrate this, Ward et al. (2013) estimate the proportion of global population exposed to flooding annually to be 169 million, while Alfieri et al. (2016) determine the equivalent quantity to be 54 million. In reality, flooding is a highly localized phenomenon, with marked spatial variability at scales down to 100 m and below (Sanders, 2007). In addition, flood inundation is controlled by complex and nonlinear processes that are not well described by the simple volume spreading and floodplain storage algorithms included in many global flood inundation models.

Third, even the most highly resolved large-scale flood models (e.g., the 30 m resolution US model of Wing et al. [2017]) do not adequately take into account local features that affect flooding. Numerous large-scale modeling studies have highlighted the need for flood adaptation databases (Sampson et al., 2015; Scussolini et al., 2016; Yin et al., 2020), although those currently available have significant limitations. For example, flood defenses in the Wing et al. (2017) model taken from the US Army Corps of Engineers National Levee Database which is thought to be only ~30% complete (American Society of Civil Engineers, 2017). In many places, urbanization and development has caused changes to natural flows and led to the construction of flood control infrastructure that are not yet captured in the inputs to large-scale flood inundation models. This is a result of such adaptions not being recorded in any national-scale database, yet myriad data sources contain fragments of information about interventions in the flood system and these could be more effectively mined to improve our representation of factors that affect local flooding.

Lastly, climate change may be poorly handled in large-scale hazard and risk analyses, and in many studies (e.g., Wing et al., 2018) it is not even considered at all. Many large-scale assessments of future fluvial flood hazard (Alfieri et al., 2016; Hirabayashi et al., 2013; Wobus et al., 2019) are driven with the output of Global Circulation Models that may poorly represent extreme rainfall (Gregersen et al., 2013; Kendon et al., 2014) and which still do not fully resolve tropical cyclones (Emanuel, 2013). This is especially problematic for countries like the US where tropical cyclones play a central role in driving flood losses (Smith et al., 2010) and where hurricane-induced coastal flooding is likely to increase in the future (Marsooli et al., 2019). An additional problem is that current large-scale studies tend to deal only with the change in a single flood driver at a time, for example, sea level rise (Kulp & Strauss, 2019), storm surge (Marsooli et al., 2019) or extreme rainfall (Alfieri et al., 2016), however in coastal catchments all of these factors will be changing the flood hazard.

The lack of a comprehensive analysis of current flood hazard from all sources is a major challenge for policy makers, insurers, and the general public and can lead to poor risk decision making. Moreover, this situation will be exacerbated by anthropogenic climate change, particularly given the lifetime of typical investment decisions, for example, building flood defenses, maintaining critical infrastructure, buying a house or taking out a mortgage. However, the tools, datasets and computing power are now available to address this problem. The purpose of this study is therefore to present the first complete and large-scale analysis of flood hazard arising from all major sources. We select the conterminous US for this study because of the quality of river gauging station, tide level, land elevation, and other data available, and also because the scale and occurrence of tropical cyclones makes this a complex and challenging location to study. In doing so, we seek to address the limitations with previous large-scale flood analyses and provide the first integrated and high-resolution view of US fluvial, coastal and pluvial flood hazard as a single definitive layer for both current and future conditions.



**Figure 1.** An overview of the modeling framework. Rectangles relate to input or output datasets and parallelograms relate to processes. For clarity, green, blue, and orange symbols are the data and processes associated with construction of fluvial, pluvial, and coastal extreme water level boundary conditions respectively, while red highlights the final product. In this study, we define the historic period as 1980–2000, current as 2010–2030 and provide analysis for two near-future time periods: 2025–2045 and 2040–2060.

In Section 2, we first describe a significant upgrade of the 30 m whole US hydrodynamic model of Wing et al. (2017) that we use to calculate flood hazard layers. We next outline the methods used to determine the current and future magnitude-frequency relationships of the main flood drivers (rainfall, river flows, and coastal water levels, including hurricanes), and in addition examine their joint probability in coastal catchments. These magnitude-frequency relationships are then used as boundary conditions for the hydrodynamic model. Section 2 also describes the steps taken to account for human interventions in the flooding system from traditional hard engineering and nature-based or “green” infrastructure projects. The combination of these three steps then allows us to calculate inundation probabilities from the three major flooding sources down to property tract level. Validation of the system is outlined in Section 3, and discussed in more detail in Section 4. Conclusions are given in Section 5.

## 2. Methods

The overall approach to modeling current and future flood risk used in this study is outlined in Figure 1. Broadly, the method can be divided into four key steps: (i) creation of pluvial, fluvial, and extreme sea level boundary conditions based on historical information; (ii) updating these boundary conditions for current and future scenarios using a change factor approach that increases or decreases historic return period magnitudes; (iii) representation of the impact of flood risk infrastructure and adaptations; and (iv) simulation of the resulting inundation using a state-of-the-art hydraulic model. We first provide a high-level overview of the whole framework, before going on to consider the individual components in more detail.

At the heart of the modeling system lies a two-dimensional, high-resolution hydraulic model based on the LISFLOOD-FP code (Bates et al., 2010; Neal et al., 2012) which can convert boundary conditions of

rainfall, river flow or coastal extreme water level to predictions of flood depth, flow velocity and inundation extent. Following Sampson et al. (2015) and Wing et al. (2017), for inland flooding the boundary conditions for each modeled reach consist of either inflow discharge at the upstream river boundaries and a normal flow condition at the downstream river boundary, or an extreme rainfall defined using Intensity-Duration-Frequency (IDF) curves. We decompose the entire US land surface into many tens of thousands of individual model domains, each of which takes local flood risk infrastructure and adaptions into account, and for each simulate design discharge and rainfall for five different recurrence intervals (5, 20, 100, 250, and 500 years). The model is applied at one arc second ( $\sim 30$  m resolution), with further downscaling (following Schumann et al., 2014) to 10 or 3 m resolution where higher resolution topographic data are available.

Results from these individual models are then combined into single hazard maps per return period representing both pluvial and fluvial flooding across the conterminous US. For coastal catchments, a similar set up is applied except that the downstream boundary is now defined as the extreme water level height at the coast. We also consider the one in 2 years recurrence interval event in coastal catchments, in addition to the five recurrence intervals considered elsewhere. Different to the inland case, we consider the joint occurrence of extreme river flow, rainfall, and coastal water levels (including hurricanes) in coastal catchments in order to allow compound events to be properly represented. The approach taken to determine boundary condition values for the core hydraulic model therefore differs for inland and coastal catchments, and also for historic, current, and future conditions. Within this work we define historic as 1980–2000, current as 2010–2030 and provide analysis for two near-future time periods: 2025–2045 and 2040–2060.

This approach is described in further detail in the following sections. We first describe the inundation modeling framework and then the sets of data and models that are used to create the boundary conditions for this scheme. Finally, we describe the approach taken to adapt the model based on local knowledge of the flood hazard.

## 2.1. Inundation Modeling Framework

The methodology to simulate flood inundation is based on the global fluvial and pluvial flood modeling framework first described by Sampson et al. (2015) and applied over the conterminous US using higher quality and resolution data by Wing et al. (2017). The framework consists of two components: (1) a model builder, which uses a number of different data sources to automatically construct hydraulic models of a specified geographic region; and (2) a hydraulic code based on the LISFLOOD-FP model, which executes the models constructed by the model builder by solving the local inertial form of the shallow water equations in two dimensions across the model domains. The model structure and numerical solver have been described in detail in a series of papers over the last decade (Almeida & Bates, 2013; Almeida et al., 2012; Bates et al., 2010; Neal et al., 2012; Sampson et al., 2013) and these works also report extensive validation tests. In addition, significant work has been undertaken on the numerical optimization of this code for high performance computing architectures via vectorization and parallelization (Neal et al., 2009b, 2010, 2018). These advances show that for gradually varied, subcritical flow the local inertial equations give results that are indistinguishable (given typical input data errors) from the full shallow water equations, but at a fraction of the computational cost.

The model builder described by Wing et al. (2017) performs the following tasks for each geographic area of simulation in an automated manner:

1. Assemble relevant input datasets into harmonized resolution and projection grids (terrain, hydrography, flood defense locations and standards, rainfall and climate characteristics, soil types)
2. Decompose the river network into discrete reaches for simulation using catchment-scale analysis that applies logic rules based on discharge change thresholds and reach length limits (i.e., new reaches are triggered when the upstream area has changed significantly or when a maximum reach length condition is met)
3. Calculate fluvial boundary conditions for each river reach (e.g., river discharge at inflow points and water level at downstream boundaries)
4. Calculate pluvial boundary conditions using IDF curves

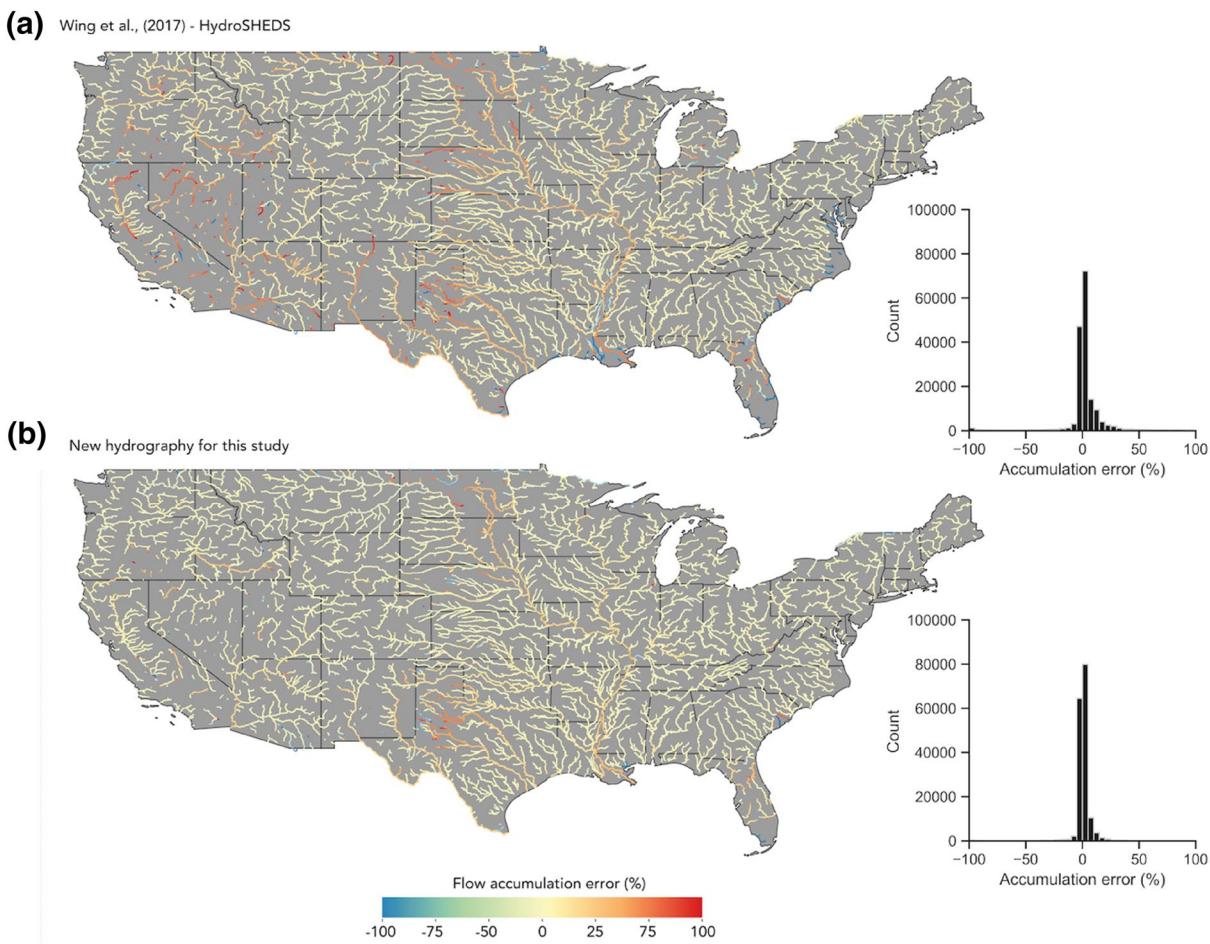
5. Estimate river network characteristics (e.g., river channel widths, depths, and defense standards)
6. Construct high-resolution one arcsecond ( $\sim 30$  m) hydraulic model input files for execution to enable simulation of all reaches across the desired range of return periods
7. Construct server-management files to control batch execution of the hydraulic models (as hundreds of thousands of discrete models will be built across a domain the size of the US)
8. Assemble the individual reach simulation results files into contiguous hazard layers

Wing et al. (2017) subsequently describe the application and validation of this modeling framework across the entire conterminous US. Terrain data were taken from the USGS National Elevation Dataset, with river channels delineated using HydroSHEDS global hydrography data (Lehner et al., 2008). The bathymetry of larger rivers was “burned” directly into the DEM, while smaller rivers were treated as subgrid scale features following Neal et al. (2012). For smaller streams this enables river width to be decoupled from the model grid scale and therefore allows rivers of any size to be represented. Known flood defenses were represented using the US Army Corps of Engineers National Levee Database, although this is known to be substantially incomplete (American Society of Civil Engineers, 2017). The model simulated fluvial flooding in catchments down to  $50 \text{ km}^2$  with boundary conditions derived from a global Regionalized Flood Frequency Analyses (RFFA; Smith et al., 2015). Pluvial flooding was simulated in all catchments using a “rain-on-grid” approach where an appropriate rainfall rate derived from National Oceanic and Atmospheric Administration (NOAA) IDF relationships was input to the model. For full details see Wing et al. (2017).

For the work reported in the current paper, we undertook a major upgrade of this modeling framework to take advantage of newly available methods and datasets. These steps form part of a process of continual enhancement of the approach. In particular, four specific issues have been addressed: (i) upgrading the river hydrography input data; (ii) updating RFFA used to set river flow boundary conditions; (iii) adding a 1D shallow water solver to estimate channel bathymetry based on an assumed bankfull discharge return period; and (iv) implementing a downscaling algorithm to resample 30 m output to 10 m or 3 m resolution (depending on local data availability) in order to better represent property tract level flooding. These four improvements are described in more detail below.

### 2.1.1. Upgrading the River Hydrography Input Data

A new hydrography for the contiguous US was created by combining the USGS National Elevation and National Hydrography Datasets (NED and NHDPlus). The latter is a feature-based database, derived from the NED and other sources, that interconnects and uniquely identifies the stream segments or reaches that make up the US surface water drainage system. The previous version of the US model (Wing et al., 2017) used HydroSHEDS (Lehner et al., 2008) was based on a hydrologically conditioned version of the Shuttle Radar Topography Mission (SRTM) elevation dataset (Farr et al., 2007) and so does not necessarily accurately represent the reality of the entire stream network. For hydraulic models a key requirement is for the hydrography layer to be consistent with the DEM used, as otherwise errors can occur such as streams not being aligned correctly with valleys. Because HydroSHEDS is not consistent with the National Elevation Dataset this led to problems of this type in some areas and was a key limitation of our previous work. We also observed considerable problems with HydroSHEDS in urban areas, as here rivers are highly modified and do not necessarily follow the valley structures in a DEM. Examples of these issues are shown in Figure S1 in the supporting information. The new NHDPlus forms the basis of a solution to this issue as it was created with high accuracy local data on stream locations held by organizations including the United States Geological Survey (USGS), the US Environmental Protection Agency (EPA), the US Forest Service and State authorities. It does not, however, include information on transboundary rivers (such as the Columbia, Rio Grande and Colorado) beyond the US and can miss small streams in catchments  $<100 \text{ km}^2$ . We therefore identified further river locations within the US in addition to NHDPlus using the National Elevation Dataset. This was achieved with a framework which automatically conditions the DEM, calculates flow directions, derives flow accumulations and identifies inland basins. The automatic computational framework allows replication of this methodology with different sources of terrain and river vector datasets at different resolutions. Hydrography for transboundary rivers beyond the US border was infilled with the MERIT-Hydro global hydrography of Yamazaki et al. (2019) which has a number of advantages over HydroSHEDS. The new hydrography was validated using the divergence-routed drainage area “Div-



**Figure 2.** Flow accumulation error as compared with NHDPlus (a) HydroSHEDS (b) this study. The comparison was carried out for rivers larger than  $1,000 \text{ km}^2$  and considered more than 160,000 values. NHDPlus, National Hydrography Datasets.

DASqKm" variable from the NHDPlus dataset. The validation was carried out by upscaling this NHDPlus variable to 1arcmin resolution and benchmarking each cell in river basins larger than  $1,000 \text{ km}^2$ , thus in total  $\sim 160,000$  cells were evaluated across the US. Figure 2 shows the percentage flow accumulation error in HydroSHEDS (a) and the new hydrography (b). The new hydrography shows improvements in areas including northern California, inland basins in Nevada, Nebraska and southern Louisiana. The flow accumulation error in the new hydrography ranged between  $-2.5\%$  and  $2.5\%$  for 75% of rivers (as measured at 126,271 locations).

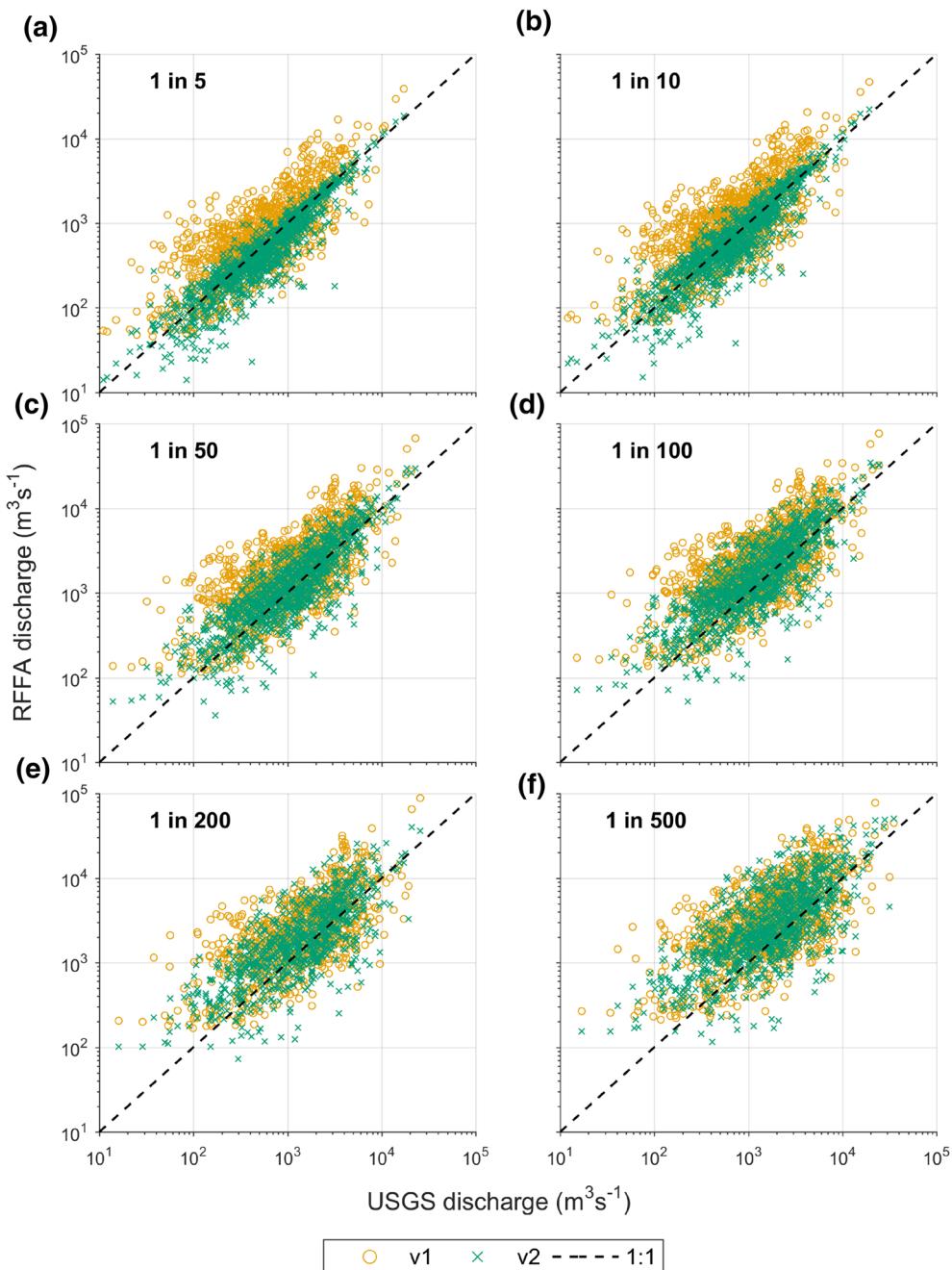
### 2.1.2. An Enhanced Regional Flood Frequency Analysis

Second, we implemented a significant upgrade to the regionalized flood frequency analysis used to estimate river discharge boundary conditions. The previous approach of Smith et al. (2015) used annual maxima data from 703 selected gauging stations from the World Meteorological Organization Global Runoff Data Center (WMO GRDC) and 242 stations from the USGS, along with three globally available catchment descriptors: area, annual average rainfall, and slope. A hybrid-clustering approach was then used to divide the observed gauging stations into homogeneous regions based on the Köppen-Geiger climate classification (Ramachandra Rao & Srinivas, 2006). Finally, an index flood approach (Griffis & Stedinger, 2007) was used to estimate the Mean Annual Flood (MAF) and extreme value distributions for each region. This allowed computation of the magnitude of any return period flow anywhere on the global river network solely from knowledge of the three global catchment descriptors. While this approach worked surprisingly well over the US (cf. the validation study of Wing et al., 2017), it ignores obvious sources of high-quality

local data that could lead to valuable improvement. To further improve this method, we upgraded the algorithm to use a much larger set of annual maxima data from 6,902 USGS river gauges. To best utilize the far denser USGS gauge network, gauges were carefully partitioned in space to estimate the index flood. This partitioning was undertaken using a hierarchical approach whereby the model moves from smaller scale catchments, to larger catchment zones, until a set of pooled gauge requirements are satisfied. These requirements were based on upstream catchment area, defining whether gauges with both small and large upstream areas were present in the pooling group. Once the catchment zone has enough gauges, representative of the required range of catchment areas, the gauges were pooled together into a single group. Therefore, only gauges from the same river or hydrological zone were used in the estimation of the index flood for a particular site. However, for the estimation of the growth curves, gauges were pooled together across all catchment zones to increase the sample size. A validation of the updated RFFA was undertaken, comparing the discharge values underpinning Wing et al. (2017) (v1 of the RFFA) and the updated RFFA in this paper (v2), against independently derived flood frequency curves contained in the USGS StreamStats data. StreamStats data were obtained for 1095 USGS gauging stations, and contain discharge frequencies computed in accordance with Bulletin 17C guidelines (England et al., 2019). Figure 3 displays the results of this validation, presenting the errors from the v1 and v2 RFFA approaches against USGS defined discharge estimates. Overall, we see substantial improvement in model skill, particularly for higher frequency events. For example, the 5 years flow magnitude is computed with 25% error on average compared with 54% in v1. 80% of sites match the 5 years USGS flow within 50%, compared to only 46% of sites in v1. Improvements in model skill, particularly at low return periods, are unsurprising given the updated index flood approach that utilizes local gauge data where possible. Model skill at high return periods is also improved, with absolute errors in 100 years flow quantification reducing from 75% to 47% and biases brought down from 72% to 29%. Given that errors in return period flows estimated from river gauges can themselves be considerable, particularly where the gauge time series is significantly shorter than the target return period (Benson, 1958), the RFFA performance against the StreamStats benchmark can be considered reasonable.

### 2.1.3. Using a 1D Shallow Water Solver to Estimate Channel Bathymetry

Third, we implemented a new method to estimate more accurately the size of river channels based on the work of Neal et al. (in review). No consistent national-scale channel bathymetry datasets exist anywhere in the world, yet channel dimensions are a critical control on how water moves through a landscape. Neal et al. (2012) showed that hydraulic models that do not have river channels explicitly defined, and which simply route water over a DEM, cannot correctly simulate wave propagation over long reaches. This is because of the different wave speeds in the channel and on the floodplain and the exchange of momentum between these areas through the shear zone that is set up at their interface (Turner-Gillespie et al., 2003). The river channel is usually the main conveyor of discharge and will interact with the floodplain in a complex manner as water moves both from and to the channel according to local variations in topography and friction. Despite the importance of such processes, many current large-scale hydraulic models do not include river channels explicitly because of the lack of available data. Given the need to include channels and the lack of wide-area bathymetry databases, an approximation must be used that best represents the relationship between channel dimensions and discharge. Our previous work determined channel conveyance based on classical geomorphic principles by assuming a bankfull discharge return period of one in 2 years for natural channels (Dunne & Leopold, 1978), and greater values where known flood defenses increase the standard of protection. Manning's equation was then used with the magnitude of bankfull discharge defined by the RFFA, the river width and assumed friction to calculate an appropriate channel depth. This has the advantage of implicitly bias correcting some of the error in the RFFA and ensuring that, even if bankfull discharge is misestimated, floodplain inundation commences at the correct flow frequency. When channel area and discharge are estimated independently, mismatches can very quickly arise: either the channel can be too big to contain the assumed bankfull discharge and inundation does not happen when and where it should, or the opposite occurs and too much inundation is the result. The key step therefore is to tie together the estimate of bankfull discharge taken from the RFFA and the channel cross sectional area by fixing the value of the (unknown at this scale) channel depth. In this way, any biases in the estimation of bankfull discharge in the RFFA can be largely eliminated, although errors in the RFFA estimates of return period flows greater than bankfull discharge will obviously still remain.



**Figure 3.** Validation of RFFA methods applied by Wing et al. (2017), v1, and the updated version, v2, used in this study. The figure displays the errors in estimated discharge, computed using benchmark data from USGS StreamStats data, for the 1 in 5 (a), 10 (b), 50 (c), 100 (d), 200 (e), and 500 years (f) events. RFFA, Regionalized Flood Frequency Analyses; USGS, United States Geological Survey.

However, using Manning's equation in this procedure has the significant flaw that it fails to account for backwater effects and thus does not represent realistic water surface profiles. Channel dimensions, and thus conveyance, can be underestimated and, as a result, too much water can be transferred to floodplain sections with this approach. This effect is particularly apparent for low return period flows where the channel conveyance is a much larger proportion of the total discharge. For this reason, the validation tests against the one in 100 years return period FEMA maps reported in Wing et al. (2017) did not properly diagnose this issue. We here solve this problem following the approach described in detail in Neal et al. (in review)

by replacing the Manning's equation with a 1D gradually varied flow model to estimate bathymetric depth along the channel network. This approach solves the following equations:

$$\frac{\partial h}{\partial x} = \frac{(S_o - S_f)}{1 - Fr^2} \quad (1)$$

where  $h$  is the channel depth,  $Fr$  is the Froude number,  $S_o$  is the channel bed slope, and  $S_f$  is the friction slope defined as:

$$S_f = n^2 \left( \frac{Q}{wh} \right)^2 \left( \frac{wh}{2h + w} \right)^{\frac{4}{3}} \quad (2)$$

where  $n$  is the Manning's roughness coefficient ( $sm^{-1/3}$ ),  $w$  is the channel width (m), and  $Q$  is the discharge ( $m^3 s^{-1}$ ). This allows a channel bed profile to be estimated that is consistent with both the DEM and the discharge estimates for that channel, creating a set of channel network geometries that behave according to the assumptions of bankfull return period made by the model user.

#### 2.1.4. Implementing a Downscaling Algorithm to Better Represent Property Tract Level Flooding

Lastly, to take advantage of higher-resolution terrain data where this is available, and thereby improve predictions down to property tract level, a downscaling algorithm was implemented. For explicit numerical methods, the computational cost of a simulation increases by approximately an order of magnitude for every halving of the model grid size. As a result, undertaking dynamical simulations at  $\sim 3$  m resolution, which is typical of the current best available terrain data in the US, compared to the  $\sim 30$  m resolution of the current model would require  $>1,000\times$  more compute time. Given that the simulation of a set of return period hazard maps at  $\sim 30$  m resolution for the whole US currently takes  $\sim 2$  months on a server with 2,000 cores, decreasing the grid resolution to  $\sim 3$  m would result in a run time of  $>160$  years using the same computational resource. However, water surfaces are effectively flat at 30 m scale during typical flood flows. This means that if higher-resolution terrain data are available, it is a simple matter to re-sample the water surface elevation computed at 30 m resolution to produce finer scale water depths and inundation extents. The effectiveness of this approach was shown very early in the development of two-dimensional inundation modeling by Horritt and Bates (2001a, 2001b), and here we implement an approach based on the more recent work of Schumann et al. (2014). First, the water surface elevations at  $\sim 30$  m (native) resolution are taken from the continental US model. To allow inundation to propagate out from the edge of the native resolution data, the model water surface elevations at the edge of the simulated inundation footprint are extended by one native resolution ( $\sim 30$  m) cell. The native resolution extended water surface elevations are then resampled to the same resolution as the higher-resolution DEM using bilinear resampling. This yields a water surface elevation mask and a high-resolution DEM, and it is then a simple matter to create a high-resolution water depth grid from these layers. To avoid discontinuities at the edge of the simulated floodplain related to the extension of the native resolution grid, an isolated depth mask is applied. This identifies all cells that are disconnected from the floodplain, using a connected cell threshold of 36, corresponding to four cells in the coarse resolution data; the flow depths in all water bodies in the high-resolution depth data that are not connected to other water bodies by more than this value are set to zero. The computational advantage of downscaling arises from the fact that it only needs to be undertaken once at the end of the simulation based on the maximum flood extent, and not at every model time step. Thus, while a single downscaling "pass" over a 30 m model grid covering the whole US is not trivial, it is able to recover much of the information that would be contained in a much higher-resolution dynamical simulation at a fraction of the total computational cost.

These four developments constitute a very significant upgrade to the inland US inundation model of Wing et al. (2017) and enhance the ability of the model to define flood hazard down to property tract level and for a wider range of flow return periods. In Section 3, we undertake revalidation of this updated approach, to quantify the improvement in performance over Wing et al. (2017) and also test the ability of the model to better simulate high-frequency flooding where the previous model version has been shown to underperform.

## 2.2. Boundary Conditions

This section describes in detail the approach taken to produce estimates of pluvial, fluvial, and coastal extreme water level boundary conditions, for historic, current, and future conditions. Figures S2–S4 provide summary diagrams showing the spatial structure in the coastal water level, precipitation and river discharge change factors produced.

### 2.2.1. Pluvial Boundary Conditions

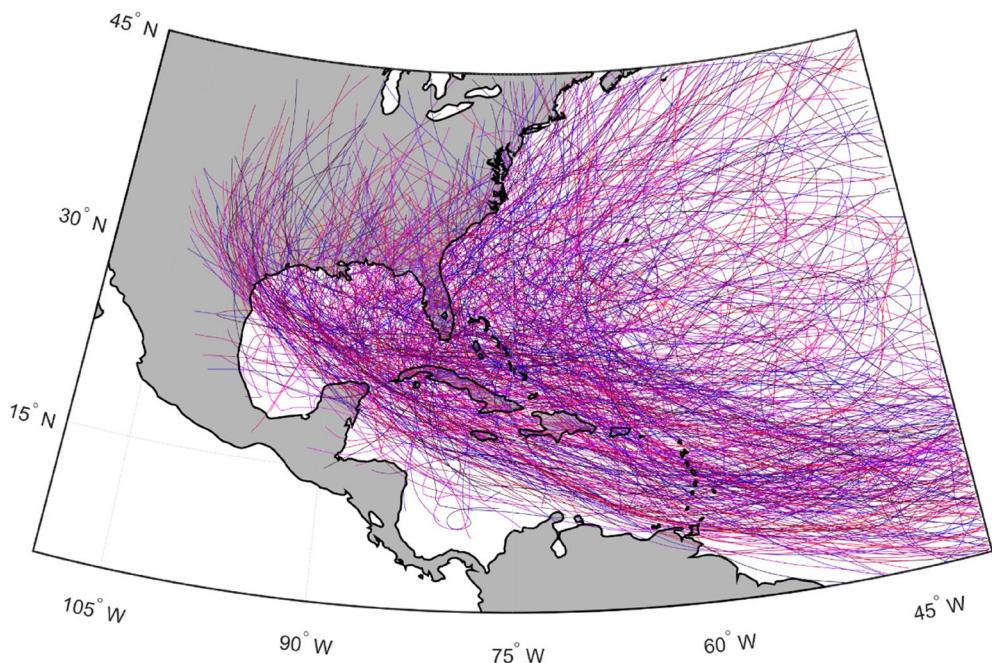
Pluvial boundary conditions are represented by IDF functions defining the intensity and duration of event rainfall and are used to force the hydraulic model using a rain-on-grid approach. For historical simulations IDF functions are taken from the NOAA Atlas 14 precipitation frequency estimates (available at <https://hdsc.nws.noaa.gov/hdsc/pfds/>). To represent climatological changes that have occurred between the historical time period captured within NOAA observations and current and future time windows, updates to the boundary conditions are required. Updating the existing IDF functions is achieved through the creation of “change factors” (scaling factors that increase or decrease historic precipitation return period magnitudes) based on Global Circulation Model (GCM) and synthetic hurricane precipitation predictions.

The method first extracts time series of daily precipitation from each of the 21 GCM ensemble members within the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP, <https://cds.nccs.nasa.gov/nex-gddp/>). NEX-GDDP comprises projections for RCP 4.5 and RCP 8.5 for the period from 1950 through to 2100 and is reported to have good skill in representing precipitation extremes (Chen et al., 2017; Kumar et al., 2020; Zhao et al., 2020). The spatial resolution of the dataset is 0.25° (~25 km × 25 km) and the downscaling approach used is based on bias-correction spatial disaggregation (Thrasher et al., 2012; Wood et al., 2004). Next, a large synthetic catalog of hurricane activity and associated storm precipitation was simulated using the approaches defined in Emanuel et al. (2008) and Feldmann et al. (2019). The hurricane simulations (see example in Figure 4) utilized seven downscaled GCM scenarios from leading modeling centers. Within each time window a 55,000 years synthetic time series of sampled precipitation events at 0.25° resolution was generated from each of the seven synthetic hurricane track sets, and the annual maximum (AMAX) rainfall recorded for each year. This length of time series was chosen to ensure adequate sampling of very extreme events. Finally, we integrated the GCM and hurricane precipitation outputs by fitting Generalized Extreme Value (GEV) distributions to the GCM data, and then sampling from these distributions to create an equivalent 55,000 years GCM AMAX time series at each of the ~13,000 GCM grid points for each time window. For each synthetic year, the largest AMAX value in either the GCM or hurricane time series was retained, resulting in a 55,000 years AMAX time series at 0.25° resolution for each time window.

These data were then used to produce estimates of precipitation change factors between historic and current/future conditions at each of the 13,000 sample points. Specifically, where hurricane activity was influential upon the precipitation tails, we calculate the percentage change in the 1 in 10 years return period 24 h rainfall event, while in nonhurricane regions the percentage change in the 98th precipitation quantile was used. These metrics were used due to the differences in the data availability. Within nonhurricane regions, within a given time window, only 20 years of daily data was available from the GCM runs. Therefore, the lower 98th quantile was deemed a more reliable metric. Within areas where hurricanes were significant contributors to distribution tails, this limitation was removed due to the extensive synthetic catalog available and a lower frequency metric, more closely associated with large loss causing events, could be used. Median, upper (75th percentile) and lower (25th percentile) change factors were then given as the difference between the historic and current/future conditions.

### 2.2.2. Fluvial Boundary Conditions

Fluvial boundary conditions (the exceedance probability of flows entering each unique river reach within the hydraulic model) are defined using the RFFA approach described in Section 2.1. These were updated from historic to current (2020 centered) and future (2035 and 2050 centered) magnitudes using change factors created by transforming the GCM precipitation time series discussed above into discharge estimates in ~700 river catchments across the US using rainfall-runoff modeling. These sites have high-quality data and represent rivers of various sizes and discharge magnitudes across nine different Köppen-Geiger climate zones. The selected catchments were chosen from USGS records which were first screened using a Kendall trend test to uncover temporal trends in the observed discharge records. Gauge records with significant



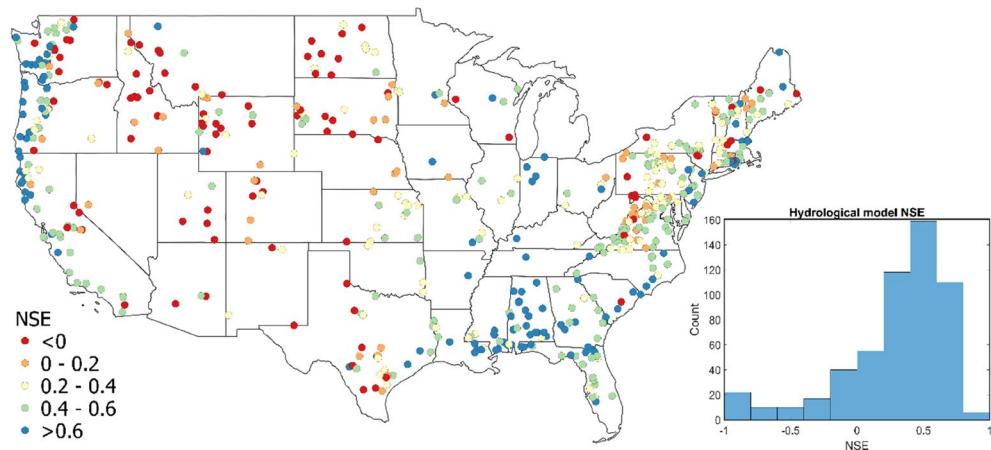
**Figure 4.** An example subset of 2,200 synthetic hurricane tracks from the historic reanalysis. Different colors for the hurricane tracks are used to make these easier to visualize.

trends were not included within this analysis (after Quinn et al., 2019). Further tests searching for sudden shifts (often termed change points) in the observed time series (e.g., due to gauge movement) were conducted, and sites with a significant change point were removed to ensure homogeneity in the observational records used to calibrate the hydrological models.

For most of the selected basins, a record of observed discharges of at least 30 years was available and used for calibration of a lumped hydrological model (HBV). HBV was chosen as it has minimum input data requirements (precipitation and temperature), simulates most components of the hydrological cycle (e.g., evaporation, groundwater, soil moisture, snow water equivalent, etc.) and has been used extensively in climate impact studies across a wide variety of catchments and scales (Akhtar et al., 2008; Bergström et al., 2001; Cloke et al., 2013; Hinzman & Kane, 1991; Seibert, 1997; Smith et al., 2014; Teutschbein et al., 2011). It is computationally efficient, which is vital when carrying out simulations at large scales and with multiple scenarios and models. HBV was first run with observed meteorological data and calibrated with observed discharge, and then the calibrated model was forced with the GCM precipitation time series for historic, current and future conditions assuming current land use to create the change factors.

For each selected basin, HBV was initialized using the global regionalized parameterization provided in Beck et al. (2016) and used to simulate historical daily discharge time series at outlet points of all the selected basins. For the observed meteorology calibration runs, HBV forcing data for each basin, including daily precipitation, minimum and maximum temperature, were taken from the Daymet data product on a 1 km<sup>2</sup> gridded surface (Thornton et al., 2016). Daymet data are derived from a collection of algorithms designed to interpolate and extrapolate from daily meteorological observations to produce gridded estimates of daily weather parameters, including minimum and maximum temperature and precipitation.

Calibration of the rainfall-runoff model was performed separately for each basin. For each basin, 2,000 sets of model parameter values were generated randomly around the closest validated parameter set provided in Beck et al. (2016) and simulated using the model. This produced a range of discharge time series for each basin. The best performing HBV model was selected for each basin using the Nash-Sutcliffe Efficiency (NSE) coefficient. The procedure focused on high flows, therefore, the NSE analysis considered only the observed flows  $\geq$  to the 99th percentile. Unsurprisingly, lower performance was generally observed in small basins, in semi-arid climates and in snowmelt dominated high-latitude basins. There was no notable dif-



**Figure 5.**  $>Q99$  flow NSE achieved at hydrological model catchments. Red sites showing  $\text{NSE} < 0$  represent those excluded from the final utilized set. NSE, Nash-Sutcliffe Efficiency.

ference between the coastal and inland basins simulated. An NSE of  $>0$  (representing model prediction of flows  $\geq Q99$  as the same or better than using the mean of flows  $\geq Q99$ ) was used as an exclusion threshold, resulting in 447 models being retained (Figure 5). A low NSE threshold was deemed appropriate here due to the high levels of uncertainty from observed rainfall (McMillan et al., 2012) and discharge records (Coxon et al., 2015). Once calibration was complete within each catchment, each of the precipitation datasets (GCM with synthetic hurricanes added) were used as inputs to produce 55,000 years discharge time series from which exceedance probability curves could be extracted. Using the same approach as defined in Section 2.2.1, this information was then used to provide a series of change factors, which were then interpolated across all US catchments using the regionalization approach based on catchment characteristics defined in Section 2.1. For simplicity, we use a single change factor for all return periods and this limitation of the approach should be addressed in future work.

### 2.2.3. Coastal Extreme Water Level Boundary Conditions

Coastal boundary condition levels for the “historic” time window were constructed using data from the NOAA tide gauge network. Observed water levels and predicted tide time series from this database were extracted and filtered based on the availability of records containing at least 25 years of data between 1980 and 2015. These were then detrended to account for MSL rise by fitting a polynomial to annual mean sea levels, smoothed using a 10 years window. There is a trade-off in the length of record used between ensuring sufficient sampling of extreme events and avoidance of nonstationarity in the storm surge characteristics, with 25 years chosen as a reasonable compromise. Sixty-eight gauge stations (20 along the west coast and 48 in the Gulf of Mexico and east coasts) were retained for analysis. For the Gulf and East coasts, where synthetic hurricane model outputs were available, NOAA tide gauge records were removed during historic hurricane events included in the IBTrACS historical tropical cyclone track data set available from NCAR to avoid any over-sampling. Next, the skew surge (the deviation between the recorded high water level and the predicted astronomical tide) was extracted, and sampling from the historic skew surge minus hurricane events and empirical tidal distributions was used to generate 55,000 years of synthetic extreme water levels at each selected NOAA gauge site. This was done to provide a data set of matching length that could be combined with the synthetic hurricane events and enable the estimation of the full marginal distributions (as opposed to purely the low-frequency tails defined by the hurricane events). Extreme water level boundary conditions were updated to correct for current (2020) mean sea levels using the median RCP4.5 mean sea level estimates provided by Kopp et al. (2014) as a simple addition of the difference between time periods. Finally, spatial variability in extreme water level between NOAA gauge sites was estimated using analysis of a hindcast coastal water level time series from the Global Tide Surge Model (GTSM) of Muis et al. (2016). This procedure is described in detail in supporting information S1.

The above steps yield 55,000 years of synthetic nonhurricane extreme water levels at each selected NOAA gauge site which can be combined with the extreme water levels generated by the 55,000 years of synthetic

hurricane events. These were determined using the pressure and wind fields predicted by the hurricane model to force a GeoCLAW model of the US Eastern seaboard. GeoCLAW uses finite volume methods to solve the two-dimensional depth-averaged shallow water equations over an adaptive mesh to enable high computational efficiency, well suited to modeling at continental scales (Mandli & Dawson, 2014). Grid resolutions in this model range from  $0.25^{\circ}$  offshore down to  $\sim 430$  m close to the coast. Surge in GeoCLAW is, in part, determined by a specification of the wind field at each time point and relies on an implementation of the Chavas et al. (2015) wind model. The wind drag is modeled using the linear form specified in Garratt (1977), with a maximum drag coefficient set to 0.0035. Consistent with Chavas et al. (2015), the radius of maximum wind speed is sampled for each simulated storm from a log-normal distribution fitted to historical observational data. Modeled levels in GeoCLAW have been shown to produce an  $r^2$  value of 0.85 against historical event high water marks, while comparison of 100 years model elevations against FEMA Base Flood Elevations (BFEs) within the south Atlantic and Gulf showed a median discrepancy of  $-0.06$  m, with residuals following a normal distribution.

The above procedure provides a greatly extended distribution of plausible storm water levels from which an estimate of return period magnitudes at each coastal boundary condition point were generated. Each estimate was shifted higher and lower based on the RCP4.5 2020 mean sea level bounds at the nearest NOAA gauge provided by Kopp et al. (2014) resulting in 21 return period estimates (low, medium, and high estimates for each of the seven hurricane storm track sets), the median of which was used as the best estimate of 2020 coastal boundary condition return periods. The 25th and 75th bounds of this distribution were used to provide uncertainty ranges within the estimate of the coastal boundary conditions.

#### 2.2.4. Sampling From the Hazard Catalog in Coastal Catchments

The fluvial, pluvial, and coastal boundary conditions defined above were used to force the hydraulic model to produce resulting flood hazard layers in coastal catchments. Typically, flood hazard is often considered independently for each driving process at a given return period, and the maximum depth used. However, this approach ignores the potential for compound events leading to more extreme flooding than the maximum of the individual components. This research attempts to characterize a portion of the large parameter space associated with potential compound flooding, focusing on the fluvial and coastal extreme water level drivers.

To do this, we generate 90 hazard layers within each catchment, representing the binary (drivers either coincide or not) interaction of fluvial and coastal boundary conditions across nine return periods (1, 5, 20, 50, 100, 250, 1,000, 5,000, and a nonevent “0” bin where a given variable will not be interacting with the others), and separate, noninteracting pluvial layers. Binary approaches were used as it was not computationally feasible to represent the vast parameter space associated with all potential timing combinations between driving variables. Similarly, pluvial layers were excluded from the compound flooding hazard layer generation due to computational constraints associated with adding another variable and the disproportionately long runtime associated with pluvial simulations.

Once the underlying hazard layers (including those considering compound flooding explicitly), the rate at which we sample each hazard layer, and therefore, the prevalence of compound flooding, is determined by gathering clusters of observed fluvial, pluvial and coastal water level datasets around the coastline, following the approaches of previous studies such as Wahl et al. (2015) and Ward et al. (2018). The dependence between the distribution tails of the drivers during nontropical cyclone events is characterized and used to generate 55,000 years of synthetic events (after Quinn et al., 2019). The synthetic tropical cyclone events are then added to this event set, resulting in a 55,000 years time series of events, each with return period magnitudes for each driver at the cluster sites. Working through these time series informs how many of each underlying hazard layer to sample throughout the synthetic period. For instance, within a synthetic multivariate event where each driver was estimated to be at the 10 years return period, the underlying hazard layer in which fluvial and coastal extreme water levels were jointly modeled would be selected, while the 10 years pluvial layer would then be overlain and the max depths taken at each pixel for that event. In each case, the associated event depth is stored for every pixel, from which a median, upper (75th percentile) and lower (25th percentile) depth-frequency curve for each coastal model pixel can be generated. Full details of the method are given in supplementary information section S2.

### 2.3. Local Adaption of the National Model

Changes to natural flows and infrastructure projects can significantly affect local patterns of flooding. While  $\sim 30$  m simulations of flow pathways and downscaling to even higher resolutions can capture the local topographic setting and fine scale dynamics of inundation, many human interventions in the flooding system are undocumented. With only partial national databases of such measures available, large-scale flood models have, to date, largely failed to take these into account. Nevertheless, many informal sources of data on flood defenses exist, and here we exploit these sources to adapt the national model to represent known local measures with the potential to reduce floods. Flood adaptation projects can generally be categorized as: (i) traditional hard engineering or “gray” infrastructure, such as levees, dams, hardened ditches, etc.; or (ii) nature-based soft or “green” infrastructure projects designed to mimic nature, capturing and slowing the advance of floodwaters (e.g., wetland creation, living shorelines, mangrove planting, etc.). Policy-driven, non-structural flood adaptation also exists; however, these approaches change exposure or vulnerability rather than the hazard so are not considered here. Only adaptation measures that are currently built and operational were included in the model, and these were represented in various ways depending on the precise nature of the intervention. The adjustments included: (i) increasing channel bankfull capacity; (ii) reducing flooding by an assumed amount within the estimated service area behind an adaptation structure; (iii) increasing infiltration rates; and (iv) changing local elevations. The following categories of adaption measures were addressed.

**Levees:** Levee data incorporated into the flood model came from US Army Corps of Engineers National Levee Database (NLD), supplemented with other sources given this is known to be only  $\sim 30\%$  complete (American Society of Civil Engineers, 2017). Further levee projects were manually identified from news articles, USACE technical manuals, correspondence with local engineers and water managers, and visual inspection of modeled flood data. Various automated data checks performed on an initial run of the model, such as the identification of urban areas lying within implausibly high-frequency modeled flood plains, were used to prioritize locations for the manual data search. Standards of protection and service areas were assigned to the structures where this value was known, or reasonable assumptions made for levees where documentation was limited. Water was then prevented from entering the service areas of levees for events whose return period was less than or equal to the standard of protection. This was performed within the hydraulic model, rather than as a postprocessing step, to ensure that the dynamic effects of reduced floodplain storage were captured. For events whose return period exceeds the defense standard, levees were assumed to have failed entirely as there was insufficient data to characterize other failure modes (such as breach or overtopping) at the national scale.

**Dams:** The National Inventory of Dams (NID, <https://nid.sec.usace.army.mil/>) was used to determine the location and function of dams, and for each the design standard and extent of their downstream service area was determined from either the dam’s Emergency Action Plan or through reasonable assumptions where this information was unavailable. Channel bankfull flow return period within the service area was then increased in the model to this design level. Flooding caused by high tides and storm surges along tidal rivers can also be prevented by dams. Major dams in coastal regions block the influx of tidal floodwaters upstream of the dam’s location on the river. The NID was used to identify dams in coastal areas, and service areas upstream of the dam were manually digitized to represent the area where tidal waters cannot feasibly reach, based on the elevation of the dam. Within this area flooding was assumed to be reduced by 100% and a maximum return period of protection was assigned, if known.

**Other Gray adaption:** FEMA Flood Insurance Studies (FIS) provide a detailed narrative of “principal flood problems” and “flood-protection measures” at the county level throughout the US. These reports, and additional county level data sources, were used to investigate the presence of any additional flood-protection schemes within a county. The reports also revealed the existence of levees and dams not accounted for within the relevant national databases. FIS reports often provide the standard of protection for the structure, but where this did not exist reasonable assumptions were made in order to assign an appropriate return period.

**Green infrastructure:** While not a main target for this first attempt at creating a local flood adaption database, where storm water basins, catch basins, retention ponds, rain gardens, etc. were identified their effect on flooding was included in the model, for example, by changing local infiltration capacity. Subsequent versions of the model will attempt to address the impact of these effects in greater detail.

*Coastal adaptions:* Accounting for the diverse set of approaches to coastal flood risk prevention, from structural flood-protection measures to natural solutions, is likely to increase the accuracy of the final flood hazard layers. News reports, public meeting minutes, foundation funded studies, university papers, environmental, and planning nonprofit reports all provide insights into local histories of coastal flooding and the construction of flood-protection infrastructure. Within these reports it was possible to find information on the community or area served by the intervention and the type of flooding scenario it was designed to mitigate against. Based on this information, protected areas were digitized, and standards of protection were assigned. The type of protection measure was also considered in deciding how to adjust the model. For example, some adaptation infrastructure completely protects an area (e.g., levees), while other methods only reduce water depths (e.g., pumps).

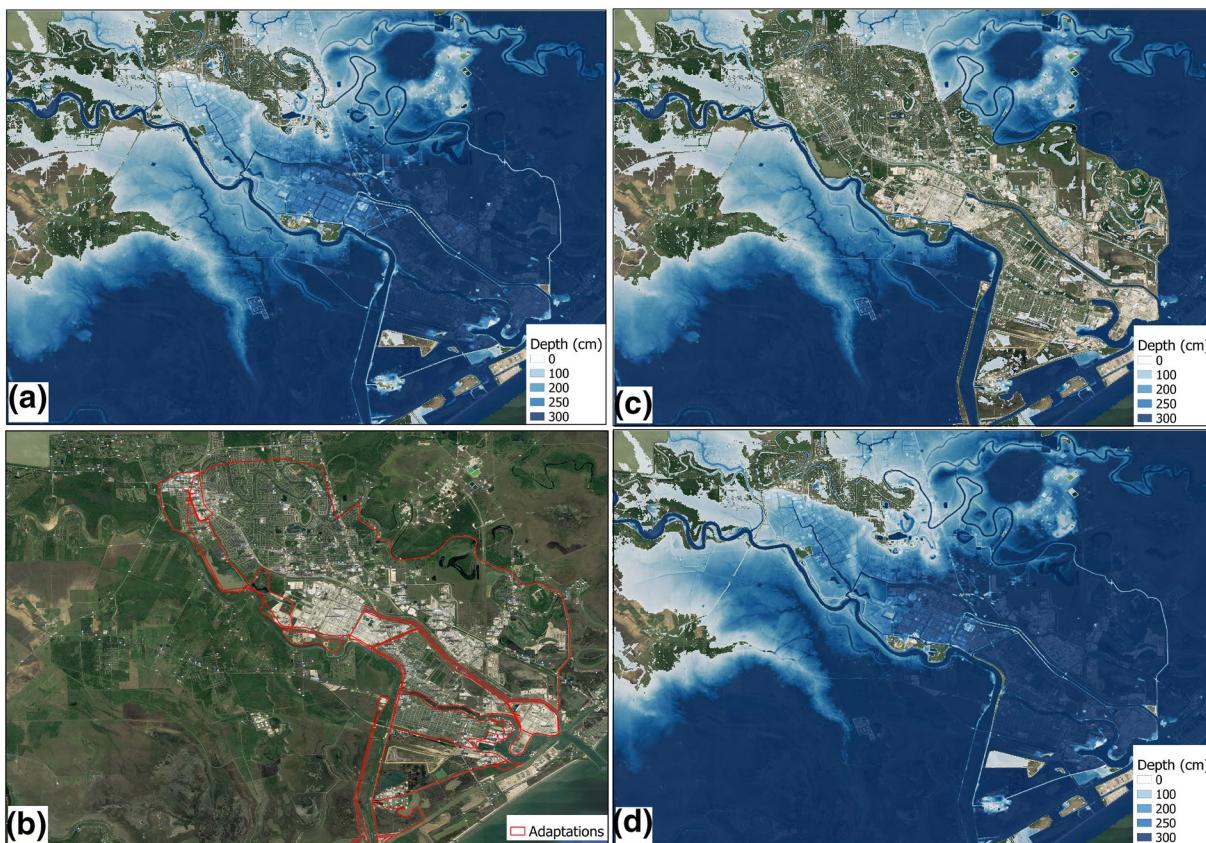
The product of this labor-intensive process was a formally structured postGIS database of local adaption measures that could be incorporated into the modeling (see Figure 6 for an example). This database will undoubtedly contain many omissions and errors, not least because the only public data available to characterize many adaptation features is qualitative and thus requires subjective interpretation when integrating into the model. However, at the very least, it should avoid the worst cases of inundation being predicted where local information clearly shows that this should not exist. Critically, it also establishes a formal standard workflow through which further information on local flooding adaptions can be incorporated into the database as it becomes available, for example, through public feedback on the resulting hazard layers.

### 3. Results

#### 3.1. Model Validation

Validation of the modeled hazard layers utilized relevant data from a variety of sources, including Iowa Flood Center (IFC, Krajewski et al., 2017) local model predictions, National Flood Insurance Program (NFIP) claims information and FEMA 100 years zones. Four summary quantitative metrics were used; the Hit Rate (HR), the False Alarm Ratio (FAR), the Critical Success Index (CSI), and the Error Bias (EB). In summary, the hit rate tests the proportion of wet benchmark data that was replicated by the model, ignoring whether the benchmark flood boundaries were exceeded; the false alarm ratio indicates the proportion of wet modeled pixels that are not wet in the benchmark data; the critical success index accounts for both over-prediction and underprediction, but ignores easy to predict true negatives, and can range from 0 (no match between modeled and benchmark data) to 1 (perfect match between modeled and benchmark data); while the error bias indicates whether the model has a tendency to overpredict (values  $> 0.5$ ) or underpredict ( $< 0.5$ ). Further details can be found in Wing et al. (2017). In particular, by restricting the pattern comparison to areas that may be wet in either the model or observations the CSI only considers the land area that could be broadly defined as the “functional floodplain.” For this reason, it is a rigorous metric for comparing flood inundation models (although still imperfect, see Stephens et al., 2014). All pattern comparisons were undertaken at the model native resolution of 30 m.

Outputs from high-quality, local flood models provided by the Iowa Flood Center (IFC) were used to evaluate the continental model high-frequency and low-frequency flood layers (Figure 7). The IFC maps were produced using state-wide HEC-RAS one-dimensional modeling for multiple return period events, with boundary conditions derived by extreme value analysis of data from USGS gauging stations or using standard USGS regression equations for ungauged streams. Water surface profiles were generated via step backwater calculations using surveyed or LiDAR-derived channel cross sections which were then intersected with 1 m resolution LiDAR-derived DEMs to produce a flood extent map for each return period flow. Further details on the Iowa Flood Center mapping procedures can be found in Gilles et al. (2012). The Iowa Flood Center 5 and 100 years flood zones were compared with corresponding layers from the old and new US models using the four metrics outlined above. At the 5 years magnitude the new US model shows a marked improvement in CSI (0.69) and FAR (0.15), relative to the original Wing et al. (2017) assessment (0.63 and 0.26, respectively). The HR is shown to decrease slightly from 0.80 in the original model to 0.77 in the new dataset while the EB swaps from a slight tendency to overprediction (EB 0.58) to underprediction (EB 0.38). At the 100 years level little change is seen between the original and new model statistics. The HR increases slightly from 0.91 to 0.93, FAR decreases from 0.09 to 0.07, CSI improves from 0.84 to 0.87, while EB remains at 0.49. Overall, these results indicate that the quality of the model assessed by Wing

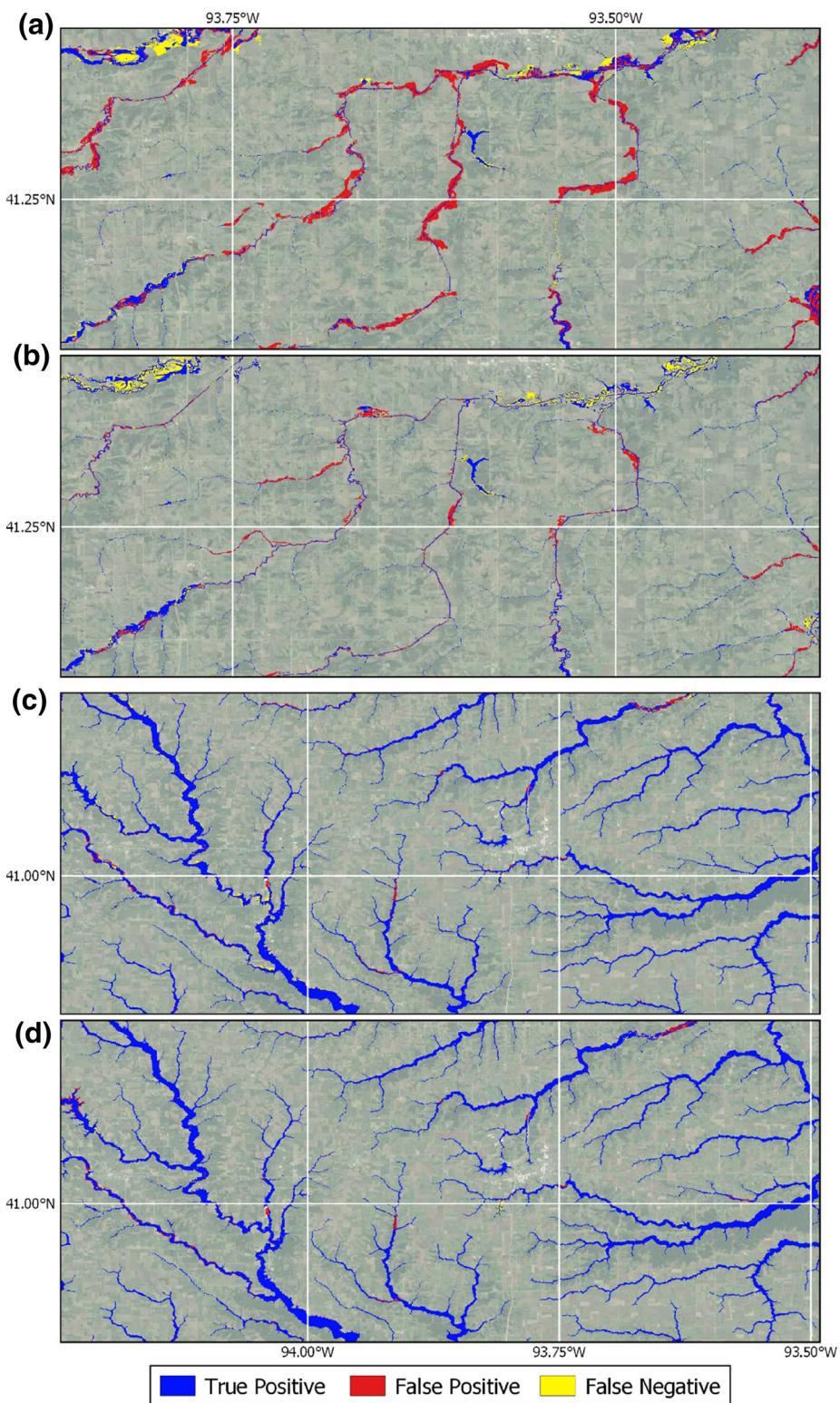


**Figure 6.** The application of adaptation information to Freeport, Texas. Panels (a) shows the 2020 median inundation prior to adaptations; (b) indicates the locations of adaptation structures in the region, which are estimated to withstand a 100 years event magnitude; (c) shows the 100 years hazard layer, after taking infrastructure adaptions into account (note that pluvial inundation is readded after removals under the assumption that levee features will not protect against this driver); (d) shows the 2050 defended layer in which the increase in event magnitude is predicted to surpass the existing defense standard.

et al. (2017) has not changed significantly during low-frequency events, but shows marked improvements in the higher frequencies due to the updates implemented as part of this research (see summary of data in Table 1).

To place these results in context, a CSI value of 0.5 indicates that the model correctly predicts only 50% of the functional floodplain area as being either wet or dry. Reasonable local models built and validated with good quality data, for example, suitably resolved and accurate terrain information, gauged flows, and high-resolution flood imagery, routinely achieve CSI values of 0.7–0.8 (Aronica et al., 2002; Bates & De Roo, 2000; Bates et al., 2004; Horritt & Bates, 2001a, 2001b). In other words, they correctly predict inundation over 70%–80% of the floodplain. This is generally accepted to indicate reasonable model skill, largely because errors in routinely used model validation data, such as satellite and airborne imagery and post-event water level surveys (Fewtrell et al., 2011; Horritt et al., 2001), place practical limits on what can be conclusively discriminated in such tests. Local models that are built with considerable time input by skilled operators and using the very highest quality boundary condition, terrain and validation data can sometimes achieve CSI scores >0.9 in both rural (Bates et al., 2006) and urban areas (Neal et al., 2009a), although this is rare. A CSI of 0.9 also represents an effective upper limit for this metric given irreducible uncertainties in, for example, river flow measurements.

A further examination of the performance of the higher frequency portion of the modeled hazard was conducted by contrasting the number of assets within the modeled 5 years national flood layer against the rate at which claims have been made within NFIP records. In the US, all properties in the 100 years floodplain should, theoretically at least, be insured through the NFIP, which as a result provides cover for about five million properties. Properties in the 5 years floodplain thus have a very high chance of appearing in the



**Figure 7.** Example comparisons of the Iowa Flood Center (IFC) flood zones and the original and new model predictions for an area of central southern Iowa south of Des Moines. Panels (a and b) show the old and new model layers compared with 5 years flood zones from the IFC, while panel c (old) and d (new) consider the 100 years layers. Significant improvements are shown for higher frequency flows.

**Table 1**

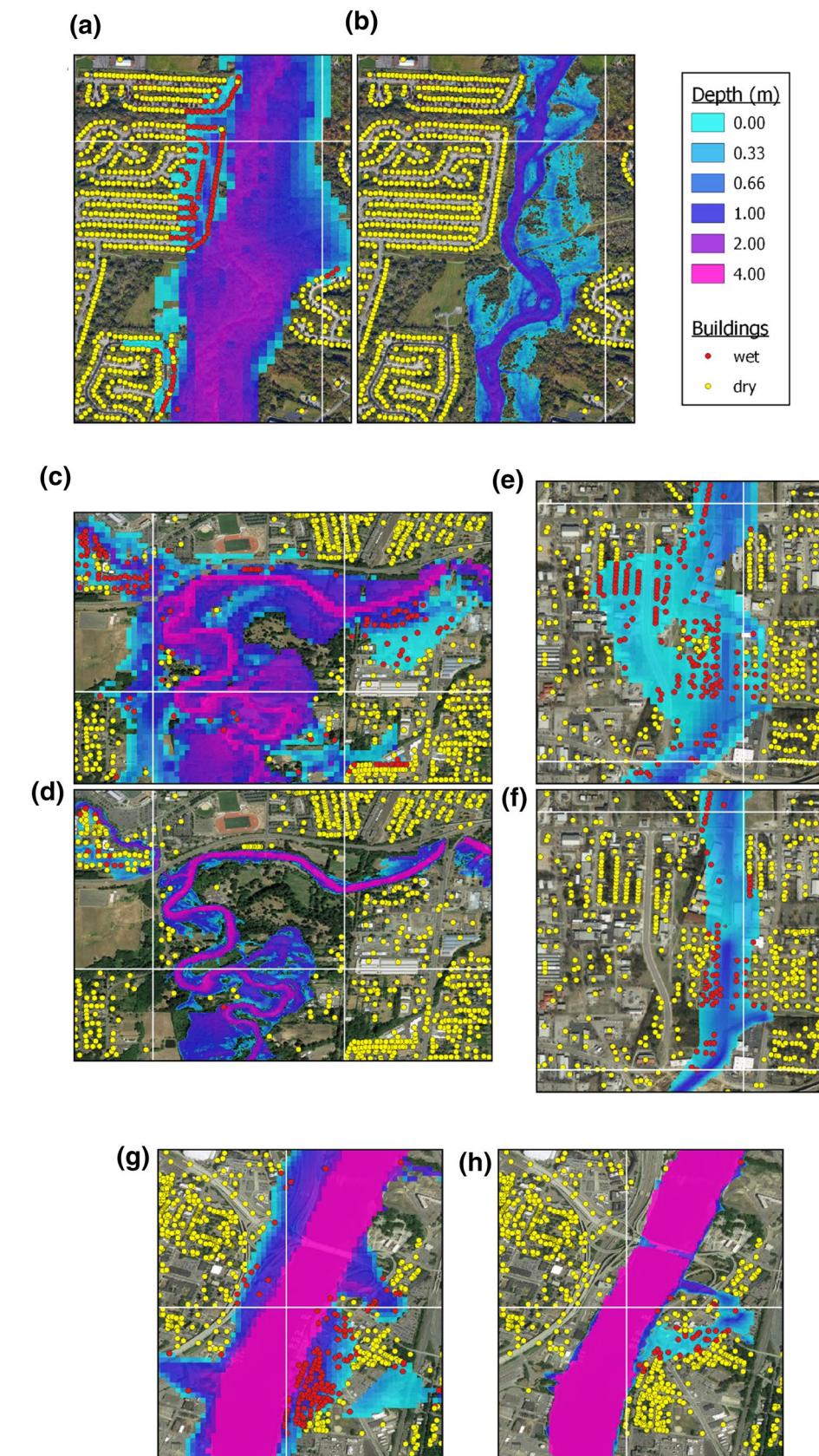
Summary of validation Test Performance Metrics for the Old (Wing et al., 2017) and New (This Paper) US Flood Inundation Models

Metric	Hit rate (HR)		False alarm ratio (far)		Critical success index (CSI)		Error bias (EB)	
	Old	New	Old	New	Old	New	Old	New
Iowa Flood Center 5 years	0.80	0.77	0.26	0.15	0.63	0.69	0.58	0.38
Iowa Flood Center 100 years	0.91	0.93	0.09	0.07	0.84	0.87	0.49	0.49
FEMA 100 years	0.87	0.85	0.14	0.09	0.76	0.78	0.53	0.38
FEMA 100 years, coastal only	0.73	0.92	0.11	0.12	0.67	0.82	0.26	0.62

NFIP data, and the sample size is appropriately large. The claims rate in the NFIP loss data is therefore a representative sample of the rate at which properties are flooded in the modeled 5 years floodplain. Analysis of the NFIP records found that there was a greater than 20% chance that 1.76% of the policies in place would make a claim in any given year, while the new 5 years hazard model (i.e., an analogous 20% annual exceedance probability) contains 1.70% of US building stock within it, a significant decrease from 3.33% in the original model predictions. This suggests that frequency with which the new US model inundates US buildings is consistent with the claims rate evident in the relatively short period of the NFIP record. While the NFIP claim rate arises from a mixture of high-frequency and low-frequency events instead of the single event frequency in the hazard model, the comparison does usefully add to our understanding of model skill. Figure 8 provides a number of local example areas in which the update to the model channel specifications and RFFA has led to improvements in the high-frequency estimations. In general, we would expect few properties to be flooded as frequently as every 5 years and Figure 8 demonstrates that the new model leads to significant decreases in building exposure during such events.

The third evaluation test for the model utilized data obtained from FEMA across the entirety of the US. Specifically, the vast (>2 million) set of 100 years FEMA flood zone polygons were compared with the old and new modeled 100 years hazard layers. This comparison is not completely straightforward as it is rarely possible to discriminate between areas which have been modeled and in which no flooding is predicted and areas which simply have not been modeled at all. Moreover, the national model simulates pluvial flooding, whereas the FEMA layers do not. In the national model the “rain-on-grid” approach is used to simulate both pluvial flooding and also the fluvial flooding in small headwater catchments, so separating true surface water inundation from fluvial flooding in headwater valley bottom areas is challenging. To address these issues, and thereby ensure a fair comparison, a number of preprocessing steps for the FEMA polygons and national model layers are required. First, permanent water pixels were discarded from the analysis to ensure we do not unfairly inflate the count of true positives. Second, owing to the low coverage of smaller streams and lack of surface water modeling in the FEMA catalog (and the inability of a buffer to resolve this problem adequately, see Wing et al., 2017), false positives arising from the “rain-on-grid” model component were discarded. This assumes that false positives are (at least potentially) indicative of pluvial flooding that the FEMA layer does not include and thus any comparison involving them would be unreliable. True positives in the “rain-on-grid” layer can, by contrast, be reasonably compared to fluvially flooded areas in the FEMA polygons. While no solution to this issue is perfect, the above approach appears to be the closest “like for like” comparison that can be achieved. False positives are therefore generated only by coastal and fluvial models, only within a 1 km buffer of the FEMA Special Flood Hazard Areas, and only in areas FEMA claim to have studied. The results indicate that the new model outperforms the original in terms of FAR, CSI (0.14–0.09 and 0.76–0.78, respectively), while HR shows a slight decrease from 0.87 to 0.85 and the EB value of 0.38 indicates a movement toward underprediction relative to the FEMA. Examples of the new model 100 years hazard and the underlying FEMA estimates are given in Figure 9 and performance metrics are summarized in Table 1.

Restricting this analysis to the coastal region, designated here as the area within 1 km of the 100 years extreme sea level inundation limit, reveals a large improvement in CSI from 0.67 to 0.82 over the original model (where no storm surge element was included), while hit rates increase from 0.73 to 0.92. While it might seem surprising that the previous version of the national model did have some skill in storm surge



affected areas, this is because inland floods in these areas were modeled. Coastal zone inundation depths in the previous model will clearly be wrong, however. As one might expect, the inclusion of the storm surge element to the modeling framework reduces the significant underprediction in the original model (EB 0.26), with the new model showing a slight tendency for overprediction (EB 0.62), see Table 1.

Overall, the model validation results show the enhanced US model described in this paper to be a marked improvement of the previous version described in Wing et al. (2017). The previous national model version already had a level of skill consistent with studies where a local model has been compared to satellite SAR data (Aronica et al., 2002; Bates & De Roo, 2000; Bates et al., 2004; Horritt & Bates, 2001a, 2001b), so even modest improvements in the validation metrics, particularly when these are computed by comparison with extensive national or state-wide datasets, are significant. Improvements are particularly apparent for higher frequency flows and in coastal zones, and for all flood drivers the continental model approaches the prediction skill that is possible with local models. This is especially the case given that these local model benchmarks have their own uncertainties and are affected by irreducible uncertainties, such as errors in the estimation of low-frequency return period flows from (relatively) short duration records. It is likely that in many places the continental scale model is indistinguishable from local scale results once such uncertainties are taken into account. With the model skill thus established, the following section considers what the results tell us about the changing nature of flood hazards in the conterminous US.

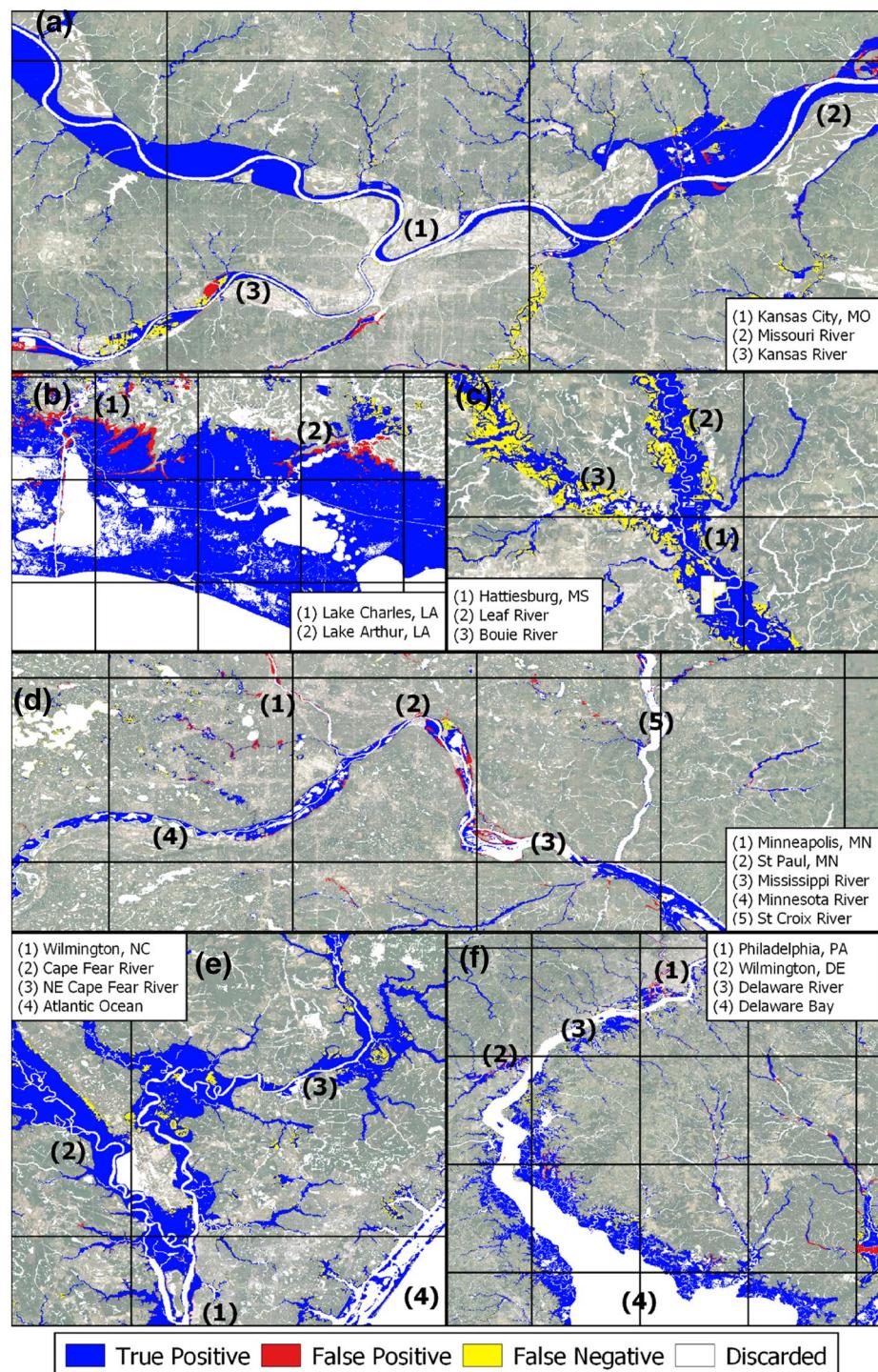
### 3.2. Characterizing Current and Future Flood Hazard in the Conterminous US

This research provides a national assessment of the current and future flood hazard facing the conterminous US. In the current (2020) time window, the model reveals that 1,007,000 km<sup>2</sup> of land lies within the median 100 years magnitude flood zone (~13% of the total land area). For the 5, 20, and 250 years hazard layers, the area of inundation is 376,500, 703,500, and 1,135,100 km<sup>2</sup>, respectively. This estimate of the 100 years flood area is significantly greater than the 572,000 km<sup>2</sup> identified by FEMA, but in line with other estimates such as that by Woznicki et al. (2019) who report an estimated 100 years combined fluvial and coastal floodplain area of 980,450 km<sup>2</sup>. The 25th and 75th percentile bounds on the modeled 100 years hazard area are 945,000 and 1,072,300 km<sup>2</sup>, respectively ( $\pm 6.1\%$  and 6.5% of the median, respectively).

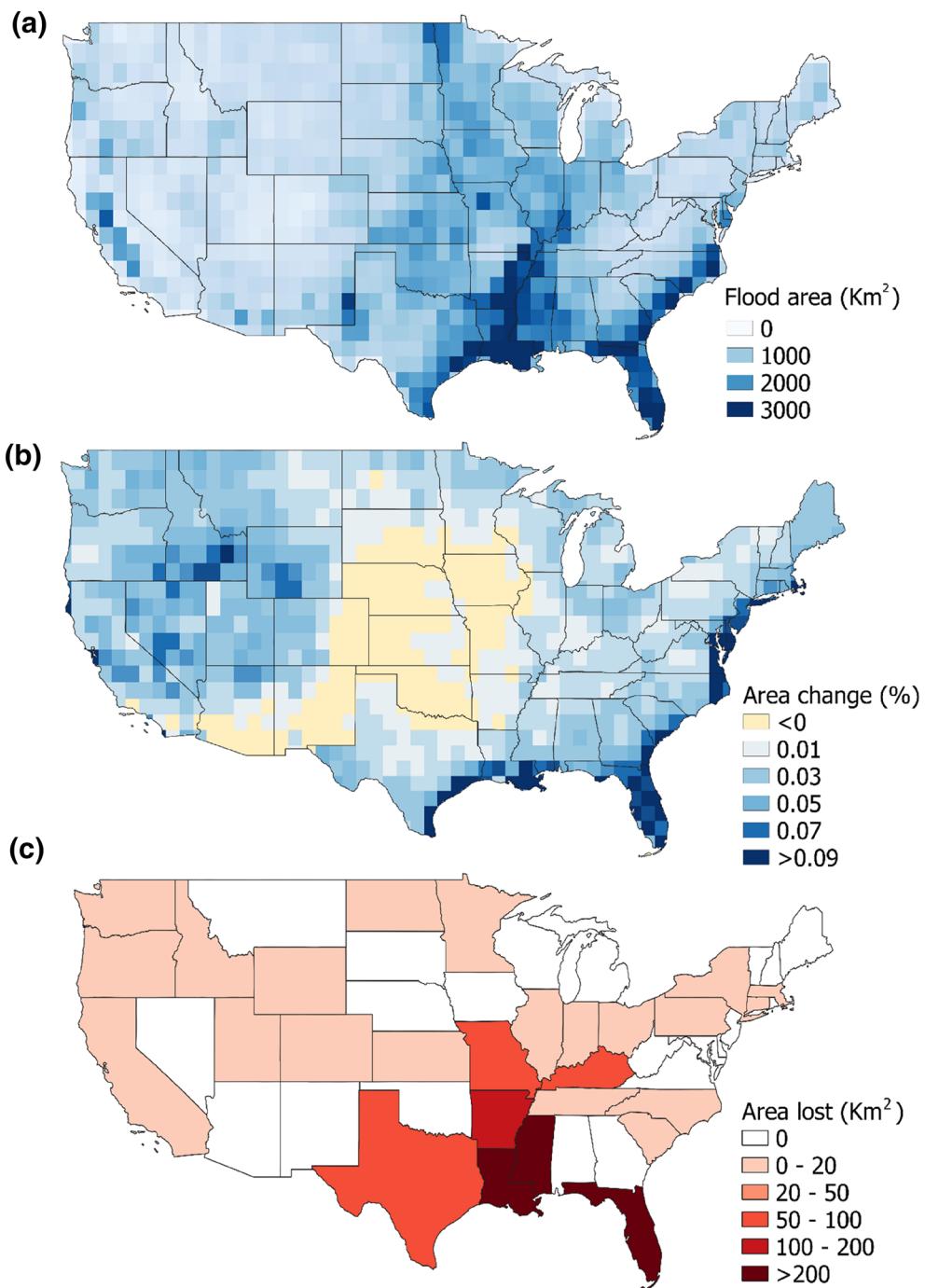
Figure 10a plots the area of inundation within the 100 years flood zone in 2020 discretized over 1° tiles. The image captures the broad spatial pattern of flood hazard across the US, with the Gulf and eastern seaboard containing a relatively large proportion of the highly inundated tiles, as does the region loosely following the lower Mississippi, through states such as Louisiana, Arkansas, Mississippi, and Missouri. The least flood prone regions are, unsurprisingly, found in the arid southwestern US through states such as Nevada, Utah and Arizona. Figures S5a and S6a plot the equivalent information for the 20 and 250 years events, respectively, while Figures S7a, S8a, and S9a plot the same information at 0.1° (~10 km) resolution.

Looking to the 2050 time window, given the RCP4.5 climate scenario, the median estimate of the 100 years flood zone increases to 1,037,200 km<sup>2</sup>. This is an increase of 30,200 km<sup>2</sup> (~3%) relative to the 2020 median, with 25th and 75th bounds of 935,800 and 1,135,800 km<sup>2</sup>. These results indicate that not only is the median estimate of the hazard increasing into the future, but also that the spread around that median is larger than in the 2020 ensemble. Interestingly, the change in median 100 years flood area between the 2020 and 2050 scenarios is less than the uncertainty given by the 2020 100 years interquartile bounds. This indicates that although it is important to consider future hazards, characterizing the current hazard given the climatological uncertainty is itself extremely difficult. These aggregated measures will conceal important local sensitivity to climate and changes in water depth, even if overall changes in flood extent are modest. This will be particularly true for higher magnitude events where floods may already be valley filling. For the 5, 20, and 250 years layers the increase in flood area is 7%, 4%, and 2.5%, respectively, indicating the rate of expansion in flooded area increases with decreasing event magnitude and supporting the above hypothesis.

**Figure 8.** Example site comparisons between the original and new 5 years hazard layers. Panels (a, c, e, and g) show the old model layers at sites in Columbus (OH), Corvallis (OR), Huntsville (AL), and Albany (NY), respectively. Panels (b, d, f, and h) show the corresponding layers in the new model layers. In all cases the number of properties designated as at risk during one in 5 years return period events is substantially reduced.



**Figure 9.** Example comparisons of the new model 100 years hazard layers and the underlying FEMA polygons for: (a), Kansas City, MO; (b), western coastal LA; (c), south eastern MS; (d), Twin Cities, MN; (e), southern coastal NC; (f), Delaware River estuary containing areas of NJ, DE, PA. Discarded pixels are those relating to permanent water and pluvial false positives. Grid lines on each figure represent  $0.25^\circ$  ( $\sim 27$  km) squares. FEMA, Federal Emergency Management Agency.



**Figure 10.** (a) The area of inundation in the 2020 100-years median layer, discretized to  $1^\circ$  tiles; (b) the percentage change in inundated area between the 2020 and 2050 median 100 years layers; (c) the area per state currently benefitting from  $\geq 1$  in 100 years standard of protection flood defenses in the 2020 median layer that is inundated by a 100 years event in 2050.

Figure 10b plots the changing spatial distribution of the 100 years flood zone by 2050. Equivalent plots for the 20 and 250 years events are given in Figures S5b and S6b, while Figures S7b, S8b, and S9b show the same information in terms of absolute changes in inundation at  $0.1^\circ$  ( $\sim 10$  km) resolution. Here, most  $1^\circ$  tiles show an increase in the area of the 100 years floodplain, but a section of the central and southwestern US shows area changes that are either negative or close to zero consistent with other climate projections

(Melillo et al., 2014). The average change in flooded area at 1° scale is +2.8%, with two and three standard deviation ranges of 0.0% to 7.8% and -3.8% to 16%, respectively. At 0.1° resolution, the equivalent two and three standard deviation ranges are -1.7% to 11.5% and -9.8% to 58% indicating that over smaller regions the percentage change in flooded area can be much more significant. The greatest proportional increases in flood area lie within the coastal regions of the Gulf and eastern coast and the Western states (e.g., California, Nevada, Idaho), while central states (Arizona through to Nebraska) actually show regions of decreased hazard. When considered in relation to existing adaptation structures, the results reveal that ~2,200 km<sup>2</sup> (or 7%) of the 100 years flood zone extension lies within regions currently protected to a 100 years or greater standard (Figure 10c). Given as a proportion of the 2020 dry area within the adaptation zones, we see that across US states the loss of currently protected regions can be greater than 10%.

#### 4. Discussion

This paper describes a major update to the US flood inundation model of Wing et al. (2017) and its extension to incorporate all major flood drivers and to represent the change of flood hazards into the future. Significant technical changes have been made to the model to: (i) improve the river channel hydrography data; (ii) better represent surface water profiles during high-frequency events; (iii) enhance the regionalized flood frequency methods used to calculate extreme event magnitudes; (iv) use downscaling where terrain data quality allows to recover property tract level hazard information; and (v) include the impact of local interventions in the flooding system.

Validation tests show the model performance for high-frequency events and coastal areas to be significantly improved, while at the same time the already good performance for low-frequency fluvial and pluvial events have not been compromised (and in fact is also slightly better). CSI values for these tests range from 0.69 to 0.82, and this is close to the values of 0.7–0.8 achieved by local 2D models built with reasonable quality terrain and gauge data and compared to satellite SAR and air photo images of inundation (Aronica et al., 2002; Bates & De Roo, 2000; Bates et al., 2004; Horritt & Bates, 2001a, 2001b). While comparing CSI values achieved using different sets of benchmark data and at different sites is not straightforward, this does give some confidence that the national model is performing well given typical, and sometimes irreducible, uncertainties, for example, in the methods to estimate the magnitude of extreme flows. Indeed, we are probably reaching a limit in terms of the ability of available data to conclusively discriminate between competing model formulations, and further step changes in flood inundation modeling science may require the development of large sets of much more rigorous model validation information.

There is no such thing as a perfect inundation prediction and future work should attempt to quantify model uncertainties in more detail. Ideally, robust prediction would utilize ensembles of model runs to characterize the impact of parameter, boundary condition and model structural uncertainty in order to produce probabilistic (and not binary) flood hazard maps. However, for computational reasons this is still a distant aspiration. The set of simulations described in this paper took ~2 months on a cluster of ~2,000 cores for each time period, so the compute cost, and indeed the whole model development process, is more akin to that of a climate model than the hydraulic models typically used in engineering studies. National-scale supercomputing could likely allow ensembles of this model in the same way that large sets of climate model runs are created, and this is a credible future direction for the research.

While taking 2 months to compute a single climate scenario for the whole US is nontrivial, this time period is far less than the 51 years that has elapsed since the US National Flood Insurance Act of 1968 that FEMA has taken to complete floodplain mapping for only 33% of the rivers and streams in the country (Association of State Floodplain Managers, 2020). The results in this paper were also achieved at a substantially lower cost than the ~\$10.6Bn that FEMA has so far spent on their mapping process. The Association of State Floodplain Managers (2020) estimate that completing FEMA mapping for the whole US would likely cost between \$3.2 BN and \$11.8Bn, with a further annual investment of \$107M to \$480M to keep these maps up to date. Even then, due to the observed data used to determine model boundary conditions, this would only represent historic average conditions. Simulation of current or future scenarios, as is done here, would require a complete re-run of all the local models. Given these timescales and costs it is clear that using col-

lections of local models to determine climate change impacts on flood hazard, or even just keeping historic average predictions up to date, is effectively impossible. Instead, it is possible to envisage an approach that uses a model like the one developed in this paper to provide a baseline national-scale coverage that can be used to triage areas of high risk. High-quality local models can then be constructed for these zones with the resources that this frees up.

Future changes in inundated area by 2050 are modest, with maximum increases of around 7% in any given  $1^{\circ}$  tile in the conterminous US. However, this almost certainly masks significant local changes and will be a much higher proportion of the land area on which human settlement is possible. By 2050 under RCP4.5 changes in extreme precipitation patterns and average sea level are still relatively modest, yet the overall pattern and direction of travel are clear even though the underlying climate model variability is large. Modest changes at  $1^{\circ}$  scale are also unsurprising given we project less than 30 years into the future and use the somewhat conservative RCP4.5 climate scenario. RCP4.5 assumes greenhouse gas emissions peak in 2040 and then decline, which may be optimistic given current rates of emission change. An alternative way of looking at this is to say that even by the near future and under a stabilization emissions scenario we see changes in flood inundation that will be locally significant. A projection for end of the century assuming higher rates of radiative forcing would show much more alarming changes, albeit with higher levels of uncertainty. Moreover, it is likely that increases in exposure will not scale linearly with changing hazard, and further research should seek to intersect the hazard layers with current and future socio-economic data and projections.

The analysis also shows that  $\sim 2,200 \text{ km}^2$  of land currently benefitting from flood defenses built to 100 years standard will be at risk from 100 years events by 2050. While this specific number needs to be treated with some caution given the sparsity and limited quality of available flood defense information, the overall pattern of a nontrivial region of currently protected land becoming exposed in the relatively near future as a result of climate change is likely to hold. This is important given the time and cost involved in planning and installing new flood defenses. Given only 30 years are available till 2050, the thinking around where flood adaptation measures should be implemented or enhanced needs to start now. The planning system in the US also needs to be cautious about sanctioning development in areas which may become flood zones in the future.

## 5. Conclusions

In recent years, large-scale flood inundation models have evolved rapidly from global schemes applied at coarse resolution with limited physics to full numerical solutions of forms of the shallow water equations solved over grids of less than 100 m spatial resolution. Validation tests of early global inundation models showed CSI values for prediction of 100 years flood extents in the range 0.43–0.67 (Alfieri et al., 2013; Sampson et al., 2015; Ward et al., 2017), indicating that most models could only correctly predict inundation over around half of the functional floodplain. Moreover, this performance was achieved despite such large events likely being valley filling and therefore relatively easy to simulate. This paper represents the next step in this evolution, increasing the performance of such models to CSI values in the range 0.69–0.82, where the lower limit of the range represents much harder to predict 5 years events. The application range of large-scale inundation models has therefore been extended in this paper to both high-frequency and low-frequency events, and also to compound flooding in coastal catchments. The result is the first comprehensive and consistent national-scale analysis of flood hazard for the conterminous US for both current and future conditions. We deliberately consider a relatively conservative scenario (RCP4.5) for a near-future time horizon (2050) and even so project clear patterns of changing flood hazard, that may be significant when considered in terms of either the land area available for development or in terms of the currently protected land area where the current standard of flood defense protection may be compromised by this time. Inundation extent changes from 2020 to 2050 summed over  $1^{\circ}$  tiles range from  $-3.8$  to  $+16\%$  (at  $3\sigma$ ), but these aggregate values derive from the net effect of local changes that can be much more significant (see Figures 6c and 6d for example) and may put at risk over  $2,200 \text{ km}^2$  of currently defended land. Future work needs to consider the implications of these new results for exposure and loss in order to develop sound and timely risk mitigation measures.

## Data Availability Statement

All model analyses in this study were conducted on behalf of the First Street Foundation (<https://firststreet.org/>) and are available to view on their Flood Factor website (<https://floodfactor.com/>). The hydraulic model LISFLOOD-FP is available for noncommercial research purposes by contacting Paul Bates at the University of Bristol ([paul.bates@bristol.ac.uk](mailto:paul.bates@bristol.ac.uk)). The US Army Corps of Engineers National Levee Database is available from <https://levees.sec.usace.army.mil/>. The USGS NED can be accessed from <https://ned.usgs.gov/>. The IFC benchmark inundation maps are available from the Iowa Flood Information System (<https://ifis.iowafloodcenter.org/>). The USGS gauge data used in the study are available online at <https://waterdata.usgs.gov/nwis>. The US National Inventory of Dams is available at <https://nid.sec.usace.army.mil/>. The US National Hydrography Dataset (NHDPlus) is available at <https://www.usgs.gov/core-science-systems/ngp/national-hydrography/access-national-hydrography-products>. NOAA IDF data are available from the NOAA Precipitation Frequency Data Server (<http://www.hpsc.nws.noaa.gov/hpsc/pfds>). MERIT-Hydro is available from [http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT\\_Hydro/](http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/). Global Runoff Data Center discharge data are available from [https://www.bafg.de/GRDC/EN/Home/homepage\\_node.html](https://www.bafg.de/GRDC/EN/Home/homepage_node.html). USGS StreamStats data are available from <https://streamstats.usgs.gov/ss>. NOAA atlas 14 data are available from <https://hpsc.nws.noaa.gov/hpsc/pfds/>. NASA Earth Exchange Global Daily Downscaled Projections can be obtained at <https://cds.nccs.nasa.gov/nex-gddp/>. The HBV Light model can be downloaded from <https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model.html>. Daymet data are available from <https://daymet.ornl.gov/>. IBTrACS data are available from <https://www.ncdc.noaa.gov/ibtracs/>. NOAA tide gauge data are available from <https://www.ngdc.noaa.gov/hazard/tide.shtml>. NOAA WAVEWATCH III hindcasts are available from <https://polar.ncep.noaa.gov/waves/hindcasts/>. FEMA Flood Insurance Studies and flood maps are available from <https://msc.fema.gov/portal/home>.

## Acknowledgments

Paul Bates was supported by a Royal Society Wolfson Research Merit Award. The work in this study was in part supported by UK Natural Environment Research Council grant NE/S006079/1. Ivan Haigh was supported by UK Natural Environment Research Council Grant NE/S010262/1. Anaís Couasnon acknowledges funding from The Netherlands Organisation for Scientific Research (VIDI Grant No. 016.161.324) and the EU COST Action DAMOCLES (CA17109) travel grant for ECS. Robert E. Kopp was supported by grants from the National Science Foundation (ICER-1663807) and the National Aeronautics and Space Administration (80NSSC17K0698) and from the Rhodium Group (for whom he has previously worked as a consultant) as part of the Climate Impact Lab consortium.

## References

Akhtar, M., Ahmad, N., & Booij, M. J. (2008). The impact of climate change on the water resources of Hindukush–Karakorum–Himalaya region under different glacier coverage scenarios. *Journal of Hydrology*, 355(1), 148–163. <https://doi.org/10.1016/j.jhydrol.2008.03.015>

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., De Roo, A., Peter, S. N., et al. (2016). Global projections of river flood risk in a warmer world. *Earth's Future*, 5(2), 171–182. <https://doi.org/10.1002/2016EF000485>

Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., & Feyen, L. (2013). Advances in Pan-European flood hazard mapping. *Hydrological Processes*, 28(13), 4067–4077. <https://doi.org/10.1002/hyp.9947>

Almeida, G. A. M., & Bates, P. (2013). Applicability of the local inertial approximation of the shallow water equations to flood modeling. *Water Resources Research*, 49(8), 4833–4844. <https://doi.org/10.1002/wrcr.20366>

Almeida, G. A. M., Bates, P., Freer, J. E., & Souvignet, M. (2012). Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. *Water Resources Research*, 48(5), W05528. <https://doi.org/10.1029/2011WR011570>

American Society of Civil Engineers (2017). *Infrastructure report card: Levees*. Retrieved from <https://www.infrastructurereportcard.org/cat-item/levees/>

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

Association of State Floodplain Managers (2020). *Flood mapping for the nation: A cost analysis for completing and maintaining the Nation's NFIP flood map Inventory*. Madison, WI: Association of State Floodplain Managers. Retrieved from [https://asfpm-library.s3-us-west-2.amazonaws.com/FSC/MapNation/ASPFM\\_MaptheNation\\_Report\\_2020.pdf](https://asfpm-library.s3-us-west-2.amazonaws.com/FSC/MapNation/ASPFM_MaptheNation_Report_2020.pdf)

Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 236(1), 54–77. [https://doi.org/10.1016/S0022-1694\(00\)00278-X](https://doi.org/10.1016/S0022-1694(00)00278-X)

Bates, P. D., Horritt, M. S., Aronica, G., & Beven, K. (2004). Bayesian updating of flood inundation likelihoods conditioned on flood extent data. *Hydrological Processes*, 18(17), 3347–3370. <https://doi.org/10.1002/hyp.1499>

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modeling. *Journal of Hydrology*, 387(1), 33–45. <https://doi.org/10.1016/j.jhydrol.2010.03.027>

Bates, P. D., Wilson, M. D., Horritt, M. S., Mason, D. C., Holden, N., & Currie, A. (2006). Reach scale floodplain inundation dynamics observed using airborne synthetic aperture radar imagery: Data analysis and modeling. *Journal of Hydrology*, 328(1), 306–318. <https://doi.org/10.1016/j.jhydrol.2005.12.028>

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Miralles, D. G., McVicar, T. R., Schellekens, J., & Bruijnzeel, L. A. (2016). Global-scale regionalization of hydrologic model parameters. *Water Resources Research*, 52(5), 3599–3622. <https://doi.org/10.1002/2015WR018247>

Benson, M. A. (1958). *Characteristics of frequency curves based on a theoretical 1,000-year record (Rep. 58-13)*. Reston, VA: U.S. Department of the Interior, Geological Survey, Water Resources Division. <https://doi.org/10.3133/ofr5813>

Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Pettersson, A., & Rummukainen, M. (2001). Climate change impacts on runoff in Sweden: Assessments by global climate models, dynamical downscaling and hydrological modeling. *Climate Research*, 16(2), 101–112. <https://doi.org/10.3354/cr016101>

Bevacqua, E., Maraun, D., Voudoukou, M. I., Voukouvalas, E., Vrac, M., Mentaschi, L., & Widmann, M. (2019). Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change. *Science Advances*, 5(9), eaaw5531. <https://doi.org/10.1126/sciadv.aaw5531>

Chavas, D. R., Lin, N., & Emanuel, K. (2015). A model for the complete radial structure of the tropical cyclone wind field. Part I: Comparison with observed structure. *Journal of the Atmospheric Sciences*, 72(9), 3647–3662. <https://doi.org/10.1175/JAS-D-15-0014.1>

Chen, H.-P., Sun, J.-Q., & Li, H.-X. (2017). Future changes in precipitation extremes over China using the NEX-GDDP high-resolution daily downscaled data-set. *Atmospheric and Oceanic Science Letters*, 10(6), 403–410. <https://doi.org/10.1080/16742834.2017.1367625>

Cloke, H. L., Wetterhall, F., He, Y., Freer, J. E., & Pappenberger, F. (2013). Modeling climate impact on floods with ensemble climate projections. *Quarterly Journal of the Royal Meteorological Society*, 139(671), 282–297. <https://doi.org/10.1002/qj.1988>

Couasnon, A., Eilander, D., Muis, S., Veldkamp, T. I. E., Haigh, I. D., Wahl, T., et al. (2020). Measuring compound flood potential from river discharge and storm surge extremes at the global scale. *Natural Hazards and Earth System Sciences*, 20(2), 489–504. <https://doi.org/10.5194/nhess-20-489-2020>

Couasnon, A., Sebastian, A., & Morales-Nápoles, O. (2018). A copula-based Bayesian network for modeling compound flood hazard from riverine and coastal interactions at the catchment scale: An application to the Houston Ship Channel, Texas. *Water*, 10(9), 1190. <https://doi.org/10.3390/w10091190>

Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., & Smith, P. J. (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water Resources Research*, 51, 5531–5546. <https://doi.org/10.1002/2014WR016532>

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., & Feyen, L. (2016). Development and evaluation of a framework for global flood hazard mapping. *Advances in Water Resources*, 94, 87–102. <https://doi.org/10.1016/j.advwatres.2016.05.002>

Dunne, T., & Leopold, L. B. (1978). *Water in environmental planning*. San Francisco, CA: W.H. Freeman and Co.

Eilander, D., Couasnon, A., Ikeuchi, H., Muis, S., Yamazaki, D., Winsemius, H., & Ward, P. J. (2020). The effect of surge on riverine flood hazard and impact in deltas globally. *Environmental Research Letters*, 15, 104007. <https://doi.org/10.1088/1748-9326/ab8ca6>

Emanuel, K. A. (2013). Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *Proceedings of the National Academy of Sciences of the United States of America*, 110(30), 12219–12224. <https://doi.org/10.1073/pnas.1301293110>

Emanuel, K., Sundararajan, R., & Williams, J. (2008). Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bulletin of the American Meteorological Society*, 89(3), 347–368. <https://doi.org/10.1175/BAMS-89-3-347>

England, J. F., Jr., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas, W. O., Jr., Veilleux, A. G., et al. (2019). *Guidelines for determining flood flow frequency—Bulletin 17C* (Rep. 4-B5, Version 1.1: May 31, 2019). Reston, VA: U.S. Geological Survey. <https://doi.org/10.3133/tm4B5>

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45, RG2004. <https://doi.org/10.1029/2005RG000183>

Feldmann, M., Emanuel, K., Zhu, L., & Lohmann, U. (2019). Estimation of Atlantic tropical cyclone rainfall frequency in the United States. *Journal of Applied Meteorology and Climatology*, 58(8), 1853–1866. <https://doi.org/10.1175/JAMC-D-19-0011.1>

Fewtrell, T. J., Neal, J. C., Bates, P. D., & Harrison, P. J. (2011). Geometric and structural river channel complexity and the prediction of urban inundation. *Hydrological Processes*, 25(20), 3173–3186. <https://doi.org/10.1002/hyp.8035>

Field, C. B., Barros, V., Stocker, T. F., & Dahe, Q. (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the Intergovernmental Panel on climate change*. Cambridge, UK: Cambridge University Press. Retrieved from [https://www.ipcc.ch/pdf/special-reports/srex\\_SREX\\_Full\\_Report.pdf](https://www.ipcc.ch/pdf/special-reports/srex_SREX_Full_Report.pdf)

Gallien, T. W., Kalligeris, N., Delisle, M.-P. C., Tang, B.-X., Lucey, J. T. D., & Winters, M. A. (2018). Coastal flood modeling challenges in defended urban backshores. *Geosciences*, 8(12), 450. <https://doi.org/10.3390/geosciences8120450>

Garratt, J. R. (1977). Review of drag coefficients over oceans and continents. *Monthly Weather Review*, 105(7), 915–929. [https://doi.org/10.1175/1520-0493\(1977\)105<915:RODCOO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1977)105<915:RODCOO>2.0.CO;2)

Gilles, D., Young, N., Schroeder, H., Piotrowski, J., & Chang, Y.-J. (2012). Inundation mapping initiatives of the Iowa Flood Center: State-wide coverage and detailed urban flooding analysis. *Water*, 4(1), 85–106. <https://doi.org/10.3390/w4010085>

Gregersen, I. B., Sørup, H. J. D., Madsen, H., Rosbjerg, D., Mikkelsen, P. S., & Arnbjerg-Nielsen, K. (2013). Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. *Climatic Change*, 118(3), 783–797. <https://doi.org/10.1007/s10584-012-0669-0>

Griffis, V. W., & Stedinger, J. R. (2007). Log-Pearson type 3 distribution and its application in flood frequency analysis. I: Distribution characteristics. *Journal of Hydrologic Engineering*, 12(5), 482–491. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:5\(482\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(482))

Hallegatte, S., Green, C., Nicholls, R. J., & Corfee-Morlot, J. (2013). Future flood losses in major coastal cities. *Nature Climate Change*, 3(9), 802–806. <https://doi.org/10.1038/nclimate1979>

Hendry, A., Haigh, I. D., Nicholls, R. J., Winter, H., Neal, R., Wahl, T., et al. (2019). Assessing the characteristics and drivers of compound flooding events around the UK coast. *Hydrology and Earth System Sciences*, 23(7), 3117–3139. <https://doi.org/10.5194/hess-23-3117-2019>

Hinzman, L. D., & Kane, D. L. (1991). Snow hydrology of a headwater Arctic basin: 2. Conceptual analysis and computer modeling. *Water Resources Research*, 27(6), 1111–1121. <https://doi.org/10.1029/91WR00261>

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., et al. (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9), 816–821. <https://doi.org/10.1038/nclimate1911>

Horritt, M. S., & Bates, P. D. (2001a). Effects of spatial resolution on a raster based model of flood flow. *Journal of Hydrology*, 253(1), 239–249. [https://doi.org/10.1016/S0022-1694\(01\)00490-5](https://doi.org/10.1016/S0022-1694(01)00490-5)

Horritt, M. S., & Bates, P. D. (2001b). Predicting floodplain inundation: Raster-based modeling versus the finite-element approach. *Hydrological Processes*, 15(5), 825–842. <https://doi.org/10.1002/hyp.188>

Horritt, M. S., Mason, D. C., & Luckman, A. J. (2001). Flood boundary delineation from Synthetic Aperture Radar imagery using a statistical active contour model. *International Journal of Remote Sensing*, 22(13), 2489–2507. <https://doi.org/10.1080/01431160116902>

Jongman, B., Ward, P. J., & Aerts, J. C. J. H. (2012). Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change*, 22(4), 823–835. <https://doi.org/10.1016/j.gloenvcha.2012.07.004>

Keef, C., Svensson, C., & Tawn, J. A. (2009). Spatial dependence in extreme river flows and precipitation for Great Britain. *Journal of Hydrology*, 378(3), 240–252. <https://doi.org/10.1016/j.jhydrol.2009.09.026>

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., & Senior, C. A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, 4(7), 570–576. <https://doi.org/10.1038/nclimate2258>

Kopp, R. E., Horton, R. M., Little, C. M., Mitrovica, J. X., Oppenheimer, M., Rasmussen, D. J., et al. (2014). Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. *Earth's Future*, 2(8), 383–406. <https://doi.org/10.1002/2014EF000239>

Krajewski, W. F., Ceynar, D., Demir, I., Goska, R., Kruger, A., Langel, C., et al. (2017). Real-time flood forecasting and information system for the state of Iowa. *Bulletin of the American Meteorological Society*, 98(3), 539–554. <https://doi.org/10.1175/BAMS-D-15-00243.1>

Kulp, S. A., & Strauss, B. H. (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. *Nature Communications*, 10(1), 1–12. <https://doi.org/10.1038/s41467-019-12808-z>

Kumar, P., Kumar, S., Barat, A., Sarthi, P. P., & Sinha, A. K. (2020). Evaluation of NASA's NEX-GDDP-simulated summer monsoon rainfall over homogeneous monsoon regions of India. *Theoretical and Applied Climatology*, 141(1), 525–536. <https://doi.org/10.1007/s00704-020-03188-2>

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., et al. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, 59(1), 1–28. <https://doi.org/10.1080/02626667.2013.857411>

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10), 93–94. <https://doi.org/10.1029/2008EO100001>

Lian, J. J., Xu, K., & Ma, C. (2013). Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: A case study of Fuzhou City, China. *Hydrology and Earth System Sciences*, 17(2), 679–689. <https://doi.org/10.5194/hess-17-679-2013>

Madakumbura, G. D., Kim, H., Utsumi, N., Shiogama, H., Fischer, E. M., Selander, Ø., et al. (2019). Event-to-event intensification of the hydrologic cycle from 1.5 °C to a 2 °C warmer world. *Scientific Reports*, 9(1), 1–7. <https://doi.org/10.1038/s41598-019-39936-2>

Mandli, K. T., & Dawson, C. N. (2014). Adaptive mesh refinement for storm surge. *Ocean Modelling*, 75, 36–50. <https://doi.org/10.1016/j.ocemod.2014.01.002>

Markonis, Y., Papalexiou, S. M., Martinkova, M., & Hanel, M. (2019). Assessment of water cycle intensification over land using a multisource global gridded precipitation dataset. *Journal of Geophysical Research: Atmospheres*, 124(21), 11175–11187. <https://doi.org/10.1029/2019JD030855>

Marsooli, R., Lin, N., Emanuel, K., & Feng, K. (2019). Climate change exacerbates hurricane flood hazards along US Atlantic and Gulf Coasts in spatially varying patterns. *Nature Communications*, 10(1), 1–9. <https://doi.org/10.1038/s41467-019-11755-z>

McMillan, H., Krueger, T., & Freer, J. (2012). Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrological Processes*, 26(26), 4078–4111. <https://doi.org/10.1002/hyp.9384>

Melillo, J. M., Richmond, T. (T. C. .), & Yohe, G. W. (2014). *Climate change impacts in the United States: The third national climate assessment.* U.S. Global Change Research Program. <https://doi.org/10.7930/J0Z31WJ2>

Moftakhar, H., Schubert, J. E., AghaKouchak, A., Matthew, R. A., & Sanders, B. F. (2019). Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries. *Advances in Water Resources*, 128, 28–38. <https://doi.org/10.1016/j.advwatres.2019.04.009>

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H., & Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications*, 7, 11969. <https://doi.org/10.1038/ncomms11969>

Neal, J. C., Bates, P. D., Fewtrell, T. J., Hunter, N. M., Wilson, M. D., & Horritt, M. S. (2009a). Distributed whole city water level measurements from the Carlisle 2005 urban flood event and comparison with hydraulic model simulations. *Journal of Hydrology*, 368(1), 42–55. <https://doi.org/10.1016/j.jhydrol.2009.01.026>

Neal, J., Dunne, T., Sampson, C., Smith, A., & Bates, P. (2018). *Optimisation of the two-dimensional hydraulic model LISFOOD-FP for CPU architecture.* Environmental Modelling and Software.

Neal, J. C., Fewtrell, T. J., Bates, P. D., & Wright, N. G. (2010). A comparison of three parallelization methods for 2D flood inundation models. *Environmental Modelling & Software*, 25(4), 398–411. <https://doi.org/10.1016/j.envsoft.2009.11.007>

Neal, J., Fewtrell, T., & Trigg, M. (2009b). Parallelization of storage cell flood models using OpenMP. *Environmental Modelling & Software*, 24(7), 872–877. <https://doi.org/10.1016/j.envsoft.2008.12.004>

Neal, J., Hawker, L., Savage, J., Durand, M., Bates, P. D., & Sampson, C. (in review). *Representation of river channels in large scale flood inundation models.* Water Resources Research.

Neal, J., Schumann, G., & Bates, P. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, 48(11), W11506. <https://doi.org/10.1029/2012WR012514>

Oppenheimer, M., Glavovic, B., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, M., et al. (2019). Sea level rise and implications for low-lying islands, coasts and communities. In H.-O. Pörtner, et al. (Eds.), *Special report on the ocean and cryosphere in a changing climate.* IPCC. Retrieved from <https://www.ipcc.ch/srocc/chapter/chapter-4-sea-level-rise-and-implications-for-low-lying-islands-coasts-and-communities/>

Pasquier, U., He, Y., Hooton, S., Goulden, M., & Hiscock, K. M. (2019). An integrated 1D–2D hydraulic modeling approach to assess the sensitivity of a coastal region to compound flooding hazard under climate change. *Natural Hazards*, 98(3), 915–937. <https://doi.org/10.1007/s11069-018-3462-1>

Quinn, N., Bates, P. D., Neal, J., Smith, A., Wing, O., Sampson, C., et al. (2019). The spatial dependence of flood hazard and risk in the United States. *Water Resources Research*, 55(3), 1890–1911. <https://doi.org/10.1029/2018WR024205>

Ramachandra Rao, A., & Srinivas, V. V. (2006). Regionalization of watersheds by hybrid-cluster analysis. *Journal of Hydrology*, 318(1), 37–56. <https://doi.org/10.1016/j.jhydrol.2005.06.003>

Sampson, C. C., Bates, P. D., Neal, J. C., & Horritt, M. S. (2013). An automated routing methodology to enable direct rainfall in high resolution shallow water models. *Hydrological Processes*, 27(3), 467–476. <https://doi.org/10.1002/hyp.9515>

Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., & Freer, J. E. (2015). A high-resolution global flood hazard model. *Water Resources Research*, 51(9), 7358–7381. <https://doi.org/10.1002/2015WR016954>

Sanders, B. F. (2007). Evaluation of on-line DEMs for flood inundation modeling. *Advances in Water Resources*, 30(8), 1831–1843. <https://doi.org/10.1016/j.advwatres.2007.02.005>

Schumann, G. J.-P., Andreadis, K. M., & Bates, P. D. (2014). Downscaling coarse grid hydrodynamic model simulations over large domains. *Journal of Hydrology*, 508, 289–298. <https://doi.org/10.1016/j.jhydrol.2013.08.051>

Scussolini, P., Aerts, J. C. J. H., Jongman, B., Bouwer, L. M., Winsemius, H. C., de Moel, H., & Ward, P. J. (2016). FLOPROS: An evolving global database of flood protection standards. *Natural Hazards and Earth System Sciences*, 16(5), 1049–1061. <https://doi.org/10.5194/nhess-16-1049-2016>

Seibert, J. (1997). Estimation of Parameter Uncertainty in the HBV Model. Paper presented at the Nordic Hydrological Conference (Akureyri, Iceland—August 1996). *Hydrology Research*, 28(4–5), 247–262. <https://doi.org/10.2166/nh.1998.15>

Smith, A., Bates, P. D., Wing, O., Sampson, C., Quinn, N., & Neal, J. (2019). New estimates of flood exposure in developing countries using high-resolution population data. *Nature Communications*, 10(1), 1814. <https://doi.org/10.1038/s41467-019-0928-y>

Smith, A., Freer, J., Bates, P., & Sampson, C. (2014). Comparing ensemble projections of flooding against flood estimation by continuous simulation. *Journal of Hydrology*, 511, 205–219. <https://doi.org/10.1016/j.jhydrol.2014.01.045>

Smith, A., Sampson, C., & Bates, P. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, 51(1), 539–553. <https://doi.org/10.1002/2014WR015814>

Smith, J. A., Villarini, G., & Baeck, M. L. (2010). Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the eastern United States. *Journal of Hydrometeorology*, 12(2), 294–309. <https://doi.org/10.1175/2010jhm1242.1>

Stephens, E., Schumann, G., & Bates, P. (2014). Problems with binary pattern measures for flood model evaluation. *Hydrological Processes*, 28(18), 4928–4937. <https://doi.org/10.1002/hyp.9979>

Svensson, C., & Jones, D. A. (2004). Dependence between sea surge, river flow and precipitation in south and west Britain. *Hydrology and Earth System Sciences*, 8(5), 973–992. <https://doi.org/10.5194/hess-8-973-2004>

Teutschbein, C., Wetterhall, F., & Seibert, J. (2011). Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dynamics*, 37(9), 2087–2105. <https://doi.org/10.1007/s00382-010-0979-8>

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S., & Cook, R. B. (2016). *Daymet: Daily surface weather data on a 1-km grid for North America, version 3*. Oak Ridge, TN:ORNL DAAC. <https://doi.org/10.5334/ORNLDAC/1328>

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, 16(9), 3309–3314. <https://doi.org/10.5194/hess-16-3309-2012>

Trigg, M. A., Birch, C. E., Neal, J. C., Bates, P. D., Smith, A., Sampson, C. C., et al. (2016). The credibility challenge for global fluvial flood risk analysis. *Environmental Research Letters*, 11(9), 094014. <https://doi.org/10.1088/1748-9326/11/9/094014>

Turner-Gillespie, D. F., Smith, J. A., & Bates, P. D. (2003). Attenuating reaches and the regional flood response of an urbanizing drainage basin. *Advances in Water Resources*, 26(6), 673–684. [https://doi.org/10.1016/S0309-1708\(03\)00017-4](https://doi.org/10.1016/S0309-1708(03)00017-4)

van den Hurk, B., van Meijgaard, E., de Valk, P., van Heeringen, K.-J., & Gooijer, J. (2015). Analysis of a compounding surge and precipitation event in the Netherlands. *Environmental Research Letters*, 10(3), 035001. <https://doi.org/10.1088/1748-9326/10/3/035001>

Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, 5(12), 1093–1097. <https://doi.org/10.1038/nclimate2736>

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., et al. (2018). Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. *Environmental Research Letters*, 13(8), 084012. <https://doi.org/10.1088/1748-9326/aad400>

Ward, P. J., Jongman, B., Aerts, J. C. J. H., Bates, P. D., Botzen, W. J. W., Loaiza, A. D., et al. (2017). A global framework for future costs and benefits of river-flood protection in urban areas. *Nature Climate Change*, 7(9), 642–646. <https://doi.org/10.1038/nclimate3350>

Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., Beek, R. van, Bierkens, M. F. P., et al. (2013). Assessing flood risk at the global scale: model setup, results, and sensitivity. *Environmental Research Letters*, 8(4), 044019. <https://doi.org/10.1088/1748-9326/8/4/044019>

Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erickson, T. A. (2017). Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resources Research*, 53(9), 7968–7986. <https://doi.org/10.1002/2017WR020917>

Wing, O. E. J., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Joseph, F., & Morefield, P. (2018). Estimates of present and future flood risk in the conterminous United States. *Environmental Research Letters*, 13(3), 034023. <https://doi.org/10.1088/1748-9326/aaac65>

Wing, O. E. J., Sampson, C. C., Bates, P. D., Quinn, N., Smith, A. M., & Neal, J. C. (2019). A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model. *Journal of Hydrology X*, 4, 100039. <https://doi.org/10.1016/j.hydroa.2019.100039>

Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., & Bouwman, A. (2013). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences*, 17(5), 1871–1892. <https://doi.org/10.5194/hess-17-1871-2013>

Wobus, C., Zheng, P., Stein, J., Lay, C., Mahoney, H., Lorie, M., et al. (2019). Projecting changes in expected annual damages from riverine flooding in the United States. *Earth's Future*, 7(5), 516–527. <https://doi.org/10.1029/2018EF001119>

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, 62(1), 189–216. <https://doi.org/10.1023/B:CLIM.0000013685.99609.9e>

Woznicki, S. A., Baynes, J., Panlasigui, S., Mehaffey, M., & Neale, A. (2019). Development of a spatially complete floodplain map of the conterminous United States using random forest. *The Science of the Total Environment*, 647, 942–953. <https://doi.org/10.1016/j.scitotenv.2018.07.353>

Yin, J., Jonkman, S., Lin, N., Yu, D., Aerts, J., Wilby, R., et al. (2020). Flood risks in sinking delta cities: Time for a reevaluation?. *Earth's Future*, 8(8), e2020EF001614. <https://doi.org/10.1029/2020EF001614>

Zhao, C., Gong, J., Wang, H., Wei, S., Song, Q., & Zhou, Y. (2020). Changes of temperature and precipitation extremes in a typical arid and semiarid zone: Observations and multi-model ensemble projections. *International Journal of Climatology*, 40(12), 5128–5153. <https://doi.org/10.1002/joc.6510>

Zheng, F., Westra, S., & Sisson, S. A. (2013). Quantifying the dependence between extreme rainfall and storm surge in the coastal zone. *Journal of Hydrology*, 505, 172–187. <https://doi.org/10.1016/j.jhydrol.2013.09.054>