Large discrepancies between event- and response-based compound flood hazard estimates

Sara Santamaria-Aguilar[1,2], Pravin Maduwantha[1,2], Alejandra R. Enriquez[3], Thomas Wahl[1,2]

[1]Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, FL 32816, USA
[2]National Center for Integrated Coastal Research, University of Central Florida, Orlando, FL 32816, US
[3]School of Geosciences, College of Arts & Sciences, University of South Florida, St Petersburg, FL 33701, USA

Correspondence to: Sara Santamaria-Aguilar (Sara.SantamariaAguilar@ucf.edu)

Abstract. Most flood hazard assessments follow the event-based approach, assuming that the probability of flooding approximates the probability of flood drivers. However, this approach neglects information about the temporal and spatial variability of flood drivers and flood processes such as water propagation inland and its interaction with topography. The response-based approach accounts for these factors by using a large number of flood events that allow the calculation of flood probabilities. Here, we compare differences in flood hazards between the event- and response-based approaches for a case study in Gloucester City (NJ, U.S.). We find that compound events with return periods less than 20 years can produce the 100-year (i.e., 1% annual exceedance probability) flood depths in large areas of the city. This is caused by the temporal and spatial characteristics of these events, such as prolonged high coastal water levels and rainfall fields with higher rainfall rates over urbanized areas. These event characteristics are not included in extreme value models of the flood drivers and are commonly simplified by using a single design event. However, flood hazards largely depend on them, introducing large discrepancies in resulting flood hazards if neglected. The temporal and spatial variabilities of flood drivers need to be incorporated in flood hazard assessments to produce robust estimates.

1 Introduction

Coastal communities worldwide are facing increasing flood hazards from rising sea levels (Taherkhani et al., 2020; Wing et al., 2024) and extreme events such as tropical cyclones (Nederhoff et al., 2024). The rapid development in coastal zones compared to ~~the hinterland~~inland areas is also contributing to increasing the exposure to flooding of people and assets (Cosby et al., 2024), making flooding the costliest hazard for coastal zones. In the U.S. alone, damages from tropical cyclones exceeded $1 trillion since 1980 and account for more than 50% of total disaster costs every year (NCEI, 2024). Therefore, developing adaptation and mitigation strategies to reduce flood impacts and increase the resilience of coastal communities is essential.

The most common framework for estimating coastal flood risks is the one defined by the Intergovernmental Panel on Climate Change (IPCC) in the Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX), in which risks are defined as a function of the hazard, exposure, and vulnerability (IPCC, 2012). In the context of coastal flooding, quantifying the likelihood of coastal flood hazards is thus the first step to estimating flood risks

and impacts. However, there is no standard approach to quantifying flood hazards, resulting in a variety of methods being used, and discrepancies between them are not ~~being~~ well understood. There are two main general approaches to estimating flood hazards, namely event-based and response-based. However, there is also no clear consensus in the literature regarding the terminology used to distinguish these two approaches. The event-based method is often referred to as the "design-storm" or "deterministic" approach. In contrast, the response-based approach has been described using terms such as "probabilistic," "stochastic," "continuous," or "weather-generator-based". Some of these terms (particularly "probabilistic") are also used in other contexts, such as to describe flood maps that incorporate uncertainty in model parameters, which can lead to ambiguity in their interpretation (Alfonso et al., 2016; Di Baldassarre et al., 2010; Bates et al., 2004). Therefore, in this study, we adopt the term "response-based," consistent with its usage in the structural reliability literature (Gouldby et al., 2014; Jane et al., 2022b).

The event-based approach is the most commonly used. It consists of estimating first the probability of the flood driver(s), selecting one event of desired probability (e.g., 1% annual exceedance probability (AEP)), and assuming that the flooding resulting from that event approximates the occurrence probability of the event (i.e. one-to-one relationship). This approach can be applied for the full range of probabilities of events, from low to high, and produce what is also known as probabilistic flood hazard estimates (e.g., Kupfer et al., 2024). However, the most commonly used benchmark event for flood hazards is the 1% AEP, often referred as the 100 year return period or in other words, an event that has a 1 in 100 chance of being equaled or exceeded in any given year. In the U.S., the event-based approach has been widely used and is recommended by ~~-~~the Federal Emergency Management Agency (FEMA) to produce the 1% AEP flood elevations for both coastal and inland flood mapping, which serve as the basis for regulatory floodplain for management and planning (FEMA, 2022). For inland flooding, FEMA applies the event-based approach that starts by defining a design rainfall storm, typically derived from NOAA Atlas 14 which provide rainfall depths for specific probabilities (e.g., 1% AEP, 24-hour storms). The design storms are used in hydrologic models to simulate runoff, with the resulting hydrographs then routed through hydraulic models to estimate flood depths and extents. Similarly for coastal regions, a design event is selected from the distribution of coastal water levels to estimate the 1% AEP regulatory floodplain. In regions affected by tropical cyclones (TCs), FEMA further implements the Joint Probability Method (JPM) to construct a synthetic storm climatology. This involves statistically sampling combinations of key storm parameters (e.g. central pressure deficit, radius to maximum winds, forward speed) based on their joint probability distributions. These synthetic events are then dynamically downscaled to the coast and exceedance probabilities of coastal water levels are calculated based on the probabilities of the storm characteristics. Although the JMP approach might reduce the uncertainties related to estimating the likelihood of low-probability coastal water level events by increasing the sample size of this events, in both cases, the probability of the event is assumed to approximate the probability of flooding (FEMA, 2022). ~~the Atlantic and Gulf coasts by the Federal Emergency Management Agency (FEMA) to produce the 1% AEP (or 100-year) flood elevations, which are used as a regulatory floodplain for management and planning However, s~~Selecting a single event that approximates the 1% AEP floodplain might not be a simple task. On one hand, observational records of flood drivers are typically shorter than 100 years, making it necessary to apply statistical extreme value models to estimate the likelihood of

events and to extrapolate beyond the observed data to characterize low-probability events (such as the 1% AEP) that may not be captured in the historical record. ~~On the one hand, statistical~~ Statistical extreme value models ~~used to derive the likelihood of events~~ focus only on the magnitude of the ~~flood~~ drivers, and the temporal and spatial variability during events are neglected (e.g., Jane et al., 2020; Moftakhari et al., 2019). In the case of coastal water levels, the lack of information about the temporal

70 evolution of the event has been commonly simplified using different approaches such as selecting one historical event as a "design event" and rescaling its time series to the desired magnitude, e.g., matching the 1% AEP event (Dawson et al., 2005; Peña et al., 2023; Wadey et al., 2015); assuming a triangular or sine shape (e.g., Vousdoukas et al., 2016, Moftakhari et al., 2019); or defining a mean hydrograph shape from hindcast data (Dullaart et al., 2023). However, neglecting the temporal variability of coastal water levels can introduce large uncertainties in estimated flooding (Kupfer et al., 2024; Quinn et al.,

75 2014; Santamaria-Aguilar et al., 2017). Similarly, for rainfall and river discharge, traditional approaches defined a single "design storm" or "design event" to represent the temporal and spatial patterns of these drivers (i.e. a representative event structure). However, some recent studies have shown that relying on a single 'design storm', overlooking the variability in event structure across multiple storms, can underestimate flood hazards and associated impacts (Baer, 2025; Perez et al., 2024). Furthermore, when flooding results from multiple drivers (e.g., tropical cyclones producing both storm surge and heavy

80 rainfall), various combinations of driver magnitudes may share the same probability yet lead to differing flood depths and extents (see e.g. Peña et al., 2023). On the other hand, flooding (i.e., the response) also depends on other factors beyond the flood driver characteristics, such as topography and associated water dynamics.

In contrast, the response-based approach can account for all these factors to produce more robust flood hazard estimates (Baer, 2025; Perez et al., 2024). This approach involves simulating flooding from many events, enabling the calculation of empirical

85 flood depth distributions at different points in the floodplain. However, the response-based approach also has limitations. First, a large set of events is needed, which is unavailable in observed records that rarely span more than a few decades ~~up to century~~ (Ponte et al., 2019). Therefore, synthetic event datasets generated through dynamical modelling and/or complex statistical frameworks are necessary (Gori et al., 2020; Kim et al., 2023; Maduwantha et al., 2025). Second, this approach is computationally more demanding and hence it has been rarely used in the past, but it is becoming more feasible due to advances

90 in computing power (Gori et al., 2020) and new computationally efficient flood models (Bates et al., 2005; Leijnse et al., 2021). Although the response-based approach provides more robust estimates of flood depths at the household level (or at single points), the corresponding flood extent does not represent the floodplain of a single event, which might be needed for some applications such as emergency management, government budgeting for natural disasters, and insurance market. ~~For floodplain management and planning, FEMA only recommends the use of the response-based approach to produce the 1%~~

95 ~~AEP floodplain of all Great Lakes coastal Flood Risk Projects and the Pacific Coast, where the wave runup is considered the main hazard and thus different combinations of sea states (i.e., wave climate) and water levels (i.e., combination of astronomical tide and storm surge) can lead to similar total water However, c~~Coastal flooding often occurs from a combination of different drivers such as storm surges, wave runup, tides, heavy precipitation, and river discharge; so-called compound events. In fact, the risk of compound flooding from storm surges and rainfall is larger in the Atlantic and Gulf coasts of the

100 U.S. (Wahl et al., 2015)., for which However, FEMA recommends using the event-based approach has not planned to incorporate compound flood modelling in coastal regions or transition to a full response-based approach to estimate coastal compound flood hazards. Although FEMA provides some guidelines to map the 1% AEP floodplain, Mapping Partners can deviate from the guidelines if they consider it appropriate (FEMA, 2022). Thus, choosing between an event-based or response-based approach to estimate flood hazard is a decision that can be made. However, this choice is challenging to make in advance

105 since it is unclear how closely the 1% AEP event (in terms of the flood drivers) approximates the 1% AEP flood (in terms of the response). To our knowledge, the differences in flood hazard estimates between these two approaches have not yet been evaluated, specifically when flooding arises from compound events. have only been evaluated for rainfall flooding (Baer, 2025; Perez et al., 2024; Winter et al., 2020), but remain unexplored for compound coastal flooding. For the latter, selecting a single 1% AEP design event is particularly challenging, as multiple combinations of flood drivers can yield the same joint

110 exceedance probability. This challenge has sometimes been addressed through the introduction of ambiguous constructs, such as the "most likely" event, which attempts to identify a representative scenario among equally probable combinations based on the density of observed events (Jane et al., 2022; Moftakhari et al., 2019; Salvadori et al., 2011).

Here, we explore how the degree of linearity in the relationship between events of 1% chance of occurring any year and flooding of equal probability, closely the 1% AEP event approximates the 1% AEP flood from compound events of

115 precipitation and estuarine water levels in a case study for Gloucester City, New Jersey. We first assess the variability in flooding from different synthetic 1% AEP events of equal probability but different magnitudes, and temporal and spatial evolutions, to quantify the uncertainties related to using a single design event for estimating flood hazards. Then, we compare flood extents and depths from the 1% AEP events with the response 1% AEP flood. Finally, we investigate which individual compound events can cause the response 1% AEP flood depth in different parts of the study area.

120 **2 Study site**

Our study site is Gloucester City, New Jersey, a small municipality located in the Delaware estuary (Fig. 1) frequently affected by pluvial and coastal flooding (CDM Smith, 2023). We selected this study site based on the exploratory scoping analysis of Helgeson et al. (2024) for place-based convergence research. Gloucester City is bordered by water on multiple sides, with the Delaware River to the west and Newton Creek and Little Timber Creek to the north and south, respectively. The catchments

125 of these two creeks are relatively small (147.45 Km^2), extending slightly beyond the city's administrative boundaries and draining into the Delaware River to the north and south of Gloucester City. Alongside the confluence of Newton Creek, Little Timber Creek, and the Delaware River, the city's low-lying terrain, with elevation <10 meters above NAVD88, makes it especially susceptible to compound flooding from rainfall and elevated estuarine water levels, including storm surges, tides, and river discharge. In addition, these catchments are highly urbanized (See Fig. S2) and the sewer and stormwater systems

130 are combined. The municipalities in these catchments have faced long-standing issues with repetitive flooding of streets and

properties caused by inadequate stormwater drainage systems and tidal influences in the outflow discharging systems (CDM Smith, 2023).

The FEMA Risk Map and Report, dated in 1979 and updated in 2016, defines a coastal and riverine 1% AEP floodplain (i.e., Special Flood Hazard Area, SFHA) that covers large areas of the city, including five essential facilities (FEMA, 2016).

135 Between 1974 and 2016, Gloucester City was subject to five federally declared flood-related disasters. Despite this, only 94 properties were enrolled in the National Flood Insurance Program (NFIP) as of 2016, according to data from OpenFEMA. Of these, 76 properties were located within the SFHA. Based on the National Structure Inventory, Gloucester City contains a total of 3,341 single-family homes, 148 of which are situated within the SFHA. Until 2016, there have been five federal disaster declarations for flooding in Gloucester City, but only 118 properties have policy coverage under the National Flood Insurance

140 Program (NFIP). Gloucester City has been facing problems with repetitive localized pluvial and coastal flooding for years (Fig. S43 of Supplementary Material), further exacerbated by an inadequate stormwater drainage system (Smith, 2023). This is also highlighted in the FEMA Risk Map, in which an intersection of the city outside the SFHA is marked together with a photo of flooding from an event in 2009 (FEMA, 2016. Fig. S21).

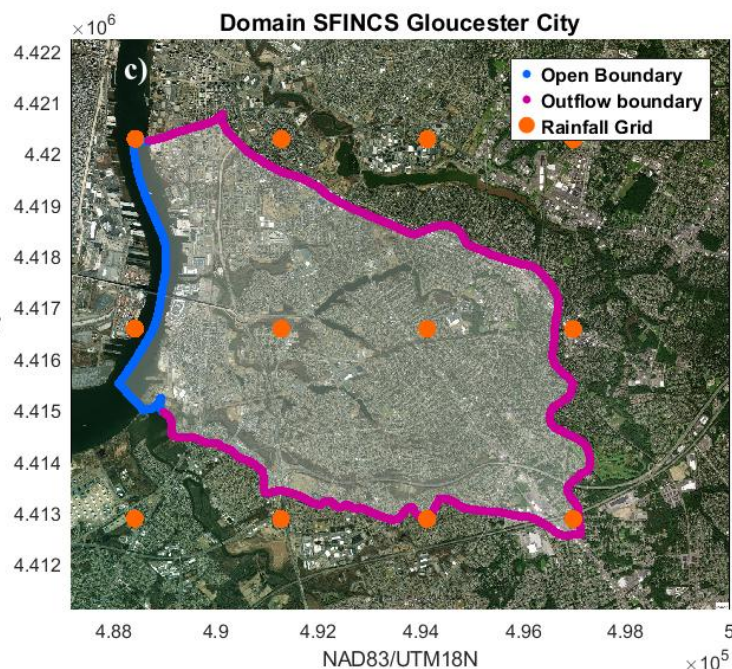
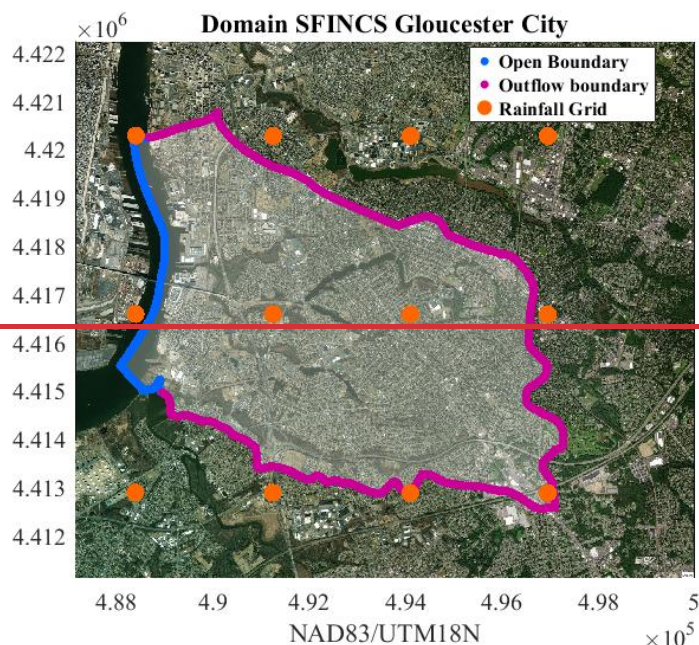
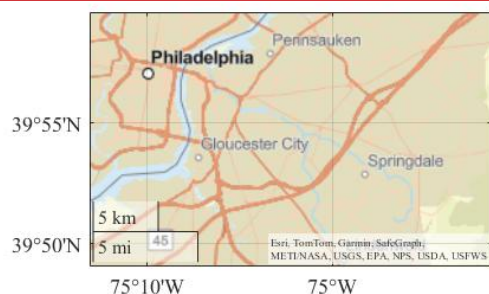



Figure 1. Location of Gloucester City (NJ, U.S.) within the inner part of the Delaware estuary (a-b). The map ~~on the right~~c) shows the flood model domain covering the catchments of Newton Creek and Little Timber Creek that surround the study site of Gloucester City. The blue line shows the location of the open boundary of the flood model along the Delaware River, the purple line is the inland outflow boundary, and the orange dots are the grid nodes of the rainfall forcing. [NAD83/UTM18N. ©Esri]

3 Data and Methods

We investigate differences in flooding between the event- and response-based approaches by simulating flooding from a large number (5,000) of compound events that allow estimating the empirical distribution of flooding and comprise several ~~1% AEP~~ events that have a 1% chance of happening in any year. We created a catalog of 5,000 synthetic compound events (more details on those events are provided in Section 3.1) following the framework of Maduwantha et al. (2025), which provides storm tide hydrographs and rainfall fields.

The joint probabilities of these events were calculated using the multivariate statistical framework of Maduwantha et al. (2024). We use the reduced-complexity flood model SFINCS (Super-Fast INundation of CoastS) to simulate pluvial and coastal flooding from the synthetic events in the study area (Fig. 1). Details of the flood model configuration and input data are described in Section 3.2, and the model validation is presented in Section 3.3.

3.1 Synthetic compound events

We need a large sample of compound events to estimate the response-based flood hazard, in which the probability of certain flood thresholds being exceeded is calculated for each model cell based on the empirical distribution. We use both the multivariate statistical framework of Maduwantha et al. (2024) and the event generation approach of Maduwantha et al. (2025) to derive a catalog of synthetic compound events, including information on rainfall fields and coastal water levels along the Delaware River at Gloucester City. Maduwantha et al. (2024) developed a new multivariate statistical framework to estimate joint probabilities of rainfall and non-tidal residuals (NTR) using copulas, accounting for the dependencies between these two flood drivers but also stratifying the extreme events by the different storm types that generate them, namely tropical cyclones and non-tropical cyclones (Fig. 2), since these show different statistical characteristics. Non-tropical events dominate the low return levels, while tropical cyclones have a stronger effect on large return levels, such as the ones associated with ~~the 1% AEP event~~events of 1% chance occurring in any given year. Accounting for the different statistical characteristics of events caused by these different storm types, the joint probability analysis avoids mischaracterization of both low and high-return level events.

For the catchments of our study site, Maduwantha et al., (2024) used around 120 years of in-situ rainfall and coastal water level measurements to estimate the joint probabilities of the flood drivers, namely rainfall and NTR. Since our study site is located in the mid-estuarine region of the Delaware Estuary, the NTR reflects contributions from both fluvial discharge and coastal storm surge, as well as their nonlinear interactions. We opted not to disaggregate the NTR into riverine and coastal components due to the substantial complexity of their coupled dynamics and the additional challenges this would introduce into the structure and parameterization of the multivariate statistical model.~~ They found that~~Maduwantha et al., (2024) found the largest dependency between NTR peak and rainfall exists for 18-hour rainfall accumulation. Since single-point rainfall might not be representative of the entire catchment, they also used 40 years of 4km gridded rainfall data from the Analysis of Period of Record for Calibration (AORC, Kitzmiller, 2018) of the corresponding catchments to obtain spatial rainfall

information and average catchment values. Observed compound events were identified using the Peaks Over Threshold (POT) approach combined with a two-sided conditional sampling method. Thresholds were set to capture an average of five events per year, providing a balance between sufficient sample size and an appropriate representation of the tail distribution. These thresholds also maximized the statistical dependence between variables. Additionally, the conditional sampling method includes events where one variable is not extreme, allowing for coverage of the full range of driver magnitudes, including those that may not lead to flooding. Further details about the multivariate statistical framework can be found in Maduwantha et al. (2024).

Maduwantha et al. (2025) developed an approach to generate synthetic compound events based on the joint probability distribution from the previous analysis and by considering the temporal and spatial information of historical events. From the joint probability distribution, they derived a sample of 5,000 events, ensuring that the proportion of observed tropical and non-tropical events is retained in the synthetic data (Fig. 2). Dynamic flood models such as SFINCS also require information about the temporal evolution of events, namely time series of both coastal water levels and rainfall fields. For that, Maduwantha et al. (2025) used the time series of historical events to generate new time series for the synthetic event set. For each synthetic event, the time series of a historical event is selected randomly accounting for their proximity in the joint probability space, and thus accounting for differences in the temporal and spatial characteristics of these events depending on their magnitude. The historical event is then rescaled to the desired magnitude of the synthetic event. The rescaled NTR time series is then combined with a mean sea-level value and a tidal curve while accounting for seasonality. The NTR hydrograph (i.e., time series) and the selected tidal curve are combined by selecting the lag from the observed events in order to account for the tide-surge interaction. Likewise, the synthetic rainfall field is combined with the synthetic water level hydrograph selecting a time lag between peaks based on the observed historical events. The synthetic compound events were validated by comparing observed and simulated distributions of key event characteristics (e.g. magnitude of the peaks, duration, times lags, intensities) and dependencies among them, finding a good agreement between observed and simulated events. Further details about the methodology used to generate the synthetic compound events can be found in Maduwantha et al. (2025).

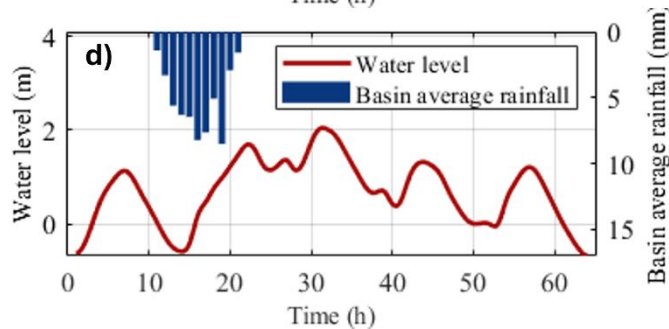
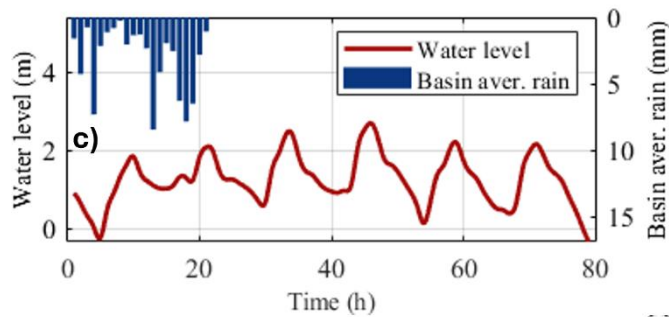
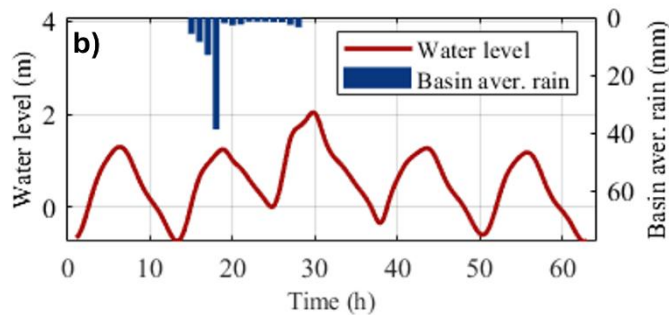
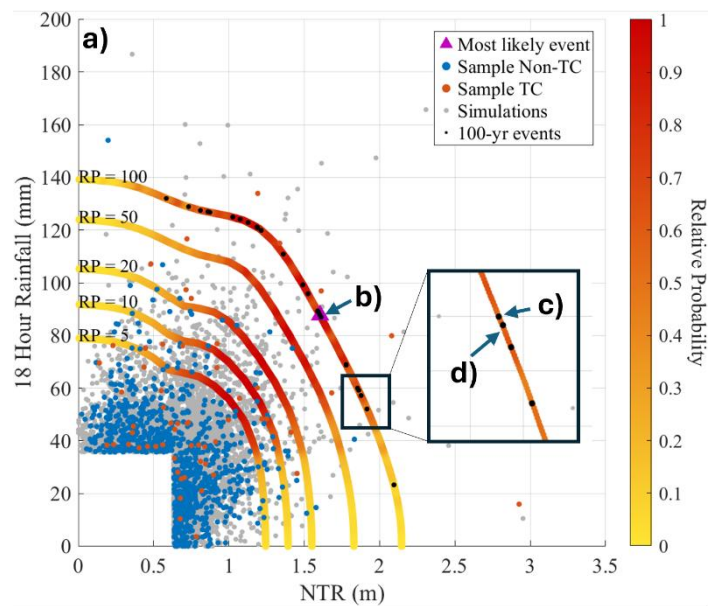



Figure 2. a) Joint probabilities of non-tidal residual (NTR) and 18hr rainfall accumulation. Blue dots show the historical non-tropical cyclone events, red dots show the historical tropical cyclone events, and grey dots the 5,000 synthetic events generated for this study. All synthetic events (grey points) have assigned water level hydrographs and rainfall fields to be used as boundary conditions for SFINCS. b) c) and d) show as an example the time series of three 1% AEP (100-year) events (black dots along the 100-year isoline); b) shows the time series of the "most likely" event, marked as a purple triangle in a); c) shows the 1% AEP (100-year) event that produces the largest flood from all 1% AEP (100-year) events (black dots in a); and d) shows the 1% AEP (100-year) event that produces the smallest flood from all 1% AEP (100-year) events. Water levels are referenced to NAVD88.

Of the 5,000 synthetic compound events used, 25 lie along the 100-year (1% AEP) isoline (i.e. with a 1% probability of happening any given year; Fig. 2). To further investigate differences in flood hazard estimations between approaches, we also define a "design event" from all the 100-year events following the "most likely" approach for multivariate events (Jane et al., 2022; Moftakhari et al., 2019). This approach selects one event in the isoline based on the ~~Based on the~~ density of observed events along ~~the isolines~~it, ~~we can define the relative probability of these events~~ (Salvadori et al., 2011b), ~~to~~ identifying this event as the most representative scenario ("most likely") ~~event along the isoline~~among the equally probable combinations along the isoline.

The water levels at the Delaware River boundary of the model are also affected by the tidal variability, which is periodic and thus its probability is not included in the multivariate extreme method of Maduwantha et al. (2024) for stochastic variables. We estimate the likelihood of tidal levels based on the predicted tides of the 19-year period from 2003 to 2021 to include long-term tidal variations such as the perigean and nodal cycles (4.4 and 18.6 years). Predicted tidal levels are generated based on the annual harmonic analysis performed by Maduwantha et al. (2024) including nodal corrections estimated from astronomical parameters (see Codiga (2011) for further information about the tidal harmonic analysis using UTide). We focus only on the likelihood of high tide peaks since flooding is more likely at these levels, but it is important to notice that the synthetic events are generated by combining the NTR peak and the high tide peak accounting for the historical distribution of time lags, and thus accounting for tide-surge interactions (Maduwantha et al., 2025).

The tidal regime in our study region is mixed semidiurnal, with two high tides per day, but one is higher than the other. We calculate the Mean Higher High Water (MHHW) level following the definition by the National Oceanic and Atmospheric Administration (NOAA) to provide an average level of the largest tidal level that happens once a day. MHHW is estimated as the average of the higher high water peaks of each day over a specified period, which in our case is the 19-year period from 2003 to 2021 (instead of the National Tidal Datum Epoch (1983-2001) used by NOAA), in order to provide an updated estimate of MHHW and better representing present-day tidal conditions (Fig. 3). Although the largest variability of tidal levels is at daily scale, tidal high waters also vary at fortnightly, seasonal, and interannual time scales. Therefore, we also estimate the mean spring tidal levels as the average of the largest high-water levels every 14 days and the average "king tide" as the mean of the annual largest tide over the 19-year period (Fig. 3).

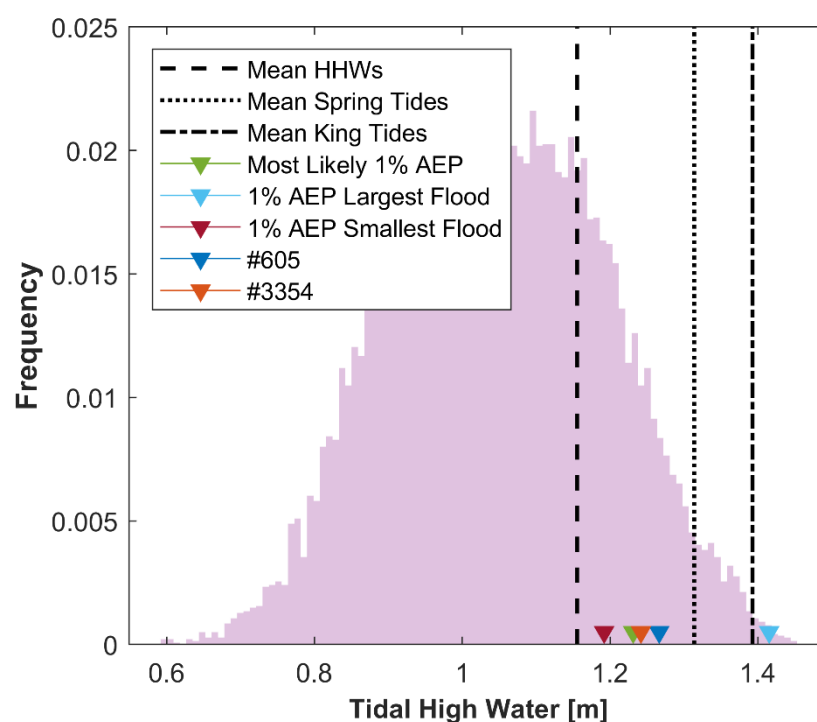
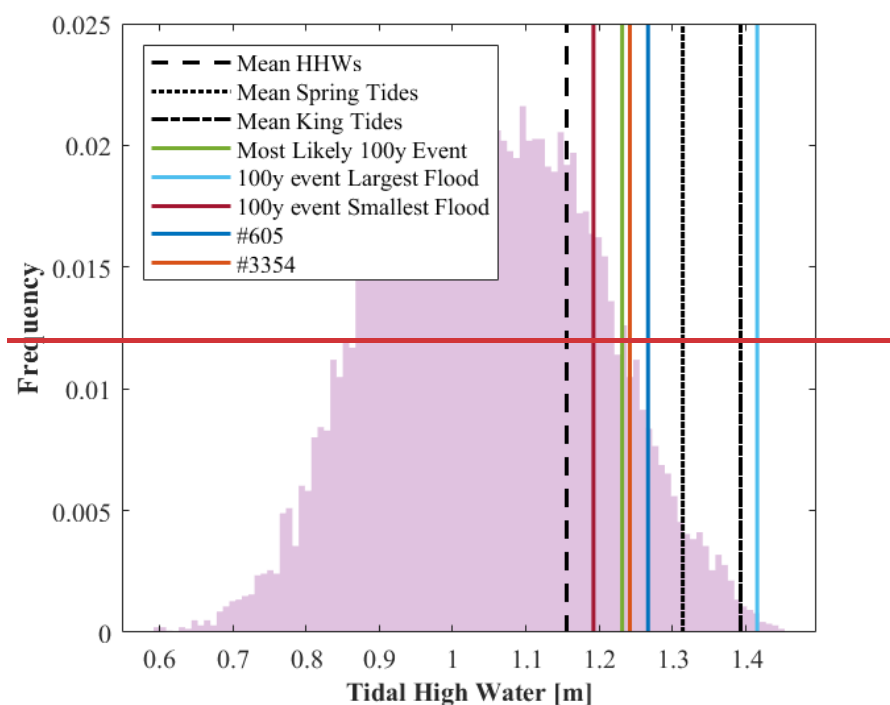



Figure 3 Histogram of tidal high water levels over the last 19 year period from 2003 to 2021. Black lines show the tidal levels of MHHW, Mean Spring Tide, and Mean King Tide (defined as the largest annual tide). Coloured lines show as an example the largest high tide levels of the tidal curves selected for the synthetic events shown in Fig. 2b-d, and two additional synthetic events (#605 and #3354) discussed in the results section.

245 Previous studies pointed to periods of increases in both high-tide flooding (Thompson et al., 2021) and extreme coastal flooding (Enriquez et al., 2022) caused by the nodal and perigean modulations of high-tide levels. Although these modulations are at longer time scales (4.4 and 18.6 years), the next peaks of both cycles will occur between 2025 and 2034 for diurnal and semidiurnal regimes. Since the tidal regime in our study site is mixed semidiurnal, the peaks of these two long-term tidal cycles are expected to occur within that period. To evaluate potential impacts of the long-term tidal modulations on the compound

250 flood analyses, we estimate the 4.4- and 18.6-year tidal cycles following the approach of Enriquez et al. (2022) for the tide-gauge records of Philadelphia. We fit a least-squares regression to the annual king tidal levels (Eq. 1) of the last 60 years of record as suggested by Haigh et al. (2011).

$$H(t) = \beta + \beta_1(t) + \beta_2 \cos\left(\frac{2\pi}{4.4}t\right) + \beta_3 \sin\left(\frac{2\pi}{4.4}t\right) + \beta_4 \cos\left(\frac{2\pi}{18.6}t\right) + \beta_5 \sin\left(\frac{2\pi}{18.6}t\right) \tag{1}$$

255

Where H(t) are the king tides of each year t, β_0 is a constant term, β_1 is the linear term, β_2 and β_3 are the amplitudes of the perigean cycle and β_4 and β_5 are the amplitudes of the nodal cycle. Based on the fitted regression, we estimate the amplitudes of both the perigean and nodal cycles, and the timing of the next peak of both cycles for our study region; this provides a better estimation of the present-day probability of large astronomical tides.

260

3.2 Flood Model

We use the dynamic flood model SFINCS (Super-Fast INundation of CoastS), which was designed specifically for simulating flooding from multiple flood drivers (Leijnse et al., 2021), since we are interested in capturing interactions between rainfall and coastal water levels as well as the effects of spatio-temporal variability of compound events on the flood response. SFINCS

265 is a reduced-complexity flood model that balances computational efficiency with accuracy, making it a perfect candidate to simulate thousands of events at a reduced computational cost.

The municipality of Gloucester City is encircled by the catchments of Newtown Creek and Little Timber Creek, for both of which discharge data is unavailable. Therefore, we define the SFINCS model domain (Fig. 1) to cover the catchments of these two creeks by their 14-digit hydrologic units from the NJDEP Bureau of GIS (Table S1). This domain encompasses all runoff

270 that could potentially lead to pluvial flooding in the study area or fluvial flooding from the creeks. We define an open boundary along the Delaware River where the water level boundary conditions are given. The inland boundaries of the model domain are defined as "outflow" to allow any water flow to exit the domain. We use the subgrid approach of SFINCS with a dual resolution of 10 m (758,904 cells) and 1 m (147,450,698 cells) and the Digital Elevation Model (DEM) of Coastal National

Elevation Database (CoNED) from the U.S. Geological Survey (Fig. S1), which has a horizontal resolution of 1 m and a

275 vertical accuracy of 10 cm (Danielson et al., 2016). This DEM at 1 m is aggregated using the median to 10 m in ArcGIS pro-3.2.0. We use spatially varying surface roughness based on land cover data from the NJDEP Bureau of GIS (Table S1), converting land classifications into Manning's coefficients based on guidance from the U.S. Army Corps of Engineers (USACE, 2021). Water level boundary conditions are provided as time series at the location of the Philadelphia tide-gauge and the model interpolates them along the open boundary of the Delaware River. Rainfall forcing is applied as spatially varying

280 fields, with the same resolution as the AORC data (Fig. 1); SFINCS interpolates these onto the model grid resolution. The model is run with the advection term neglected, solving the local inertia equations. We use the GPU version of SFINCS and run the 5,000 simulations on an Intel(R) Core (TM) i7-13700KF CPU and NVIDIA GeForce RTX 4080 GPU. The outputs of the simulations at 10 m resolution are downscaled to 1 m resolution using MATLAB 2023a.

285 ### 3.3 Flood Model Validation

Validation and calibration of flood models is a difficult task due to the common lack of observed flood data worldwide (Merz et al., 2024; Molinari et al., 2019). This is especially true for under-resourced regions; but the lack of observed flood data is also an issue in developed countries and more noticeable in the case of pluvial flood events, which are the most frequent in our study area of Gloucester City (Hino and Nance, 2021).

290 To address the challenges of flood model validation in data-scarce environments, we conducted an extensive search for observational data in Gloucester City. We evaluated a wide range of sources, including satellite imagery, high-water marks, FEMA reports, NOAA's Storm Events database, local news articles, and crowd-sourced platforms such as MyCoast and Twitter. Although conventional data sources provided limited information, we identified and simulated three documented flood events (2009, 2019, and 2020) for model validation. As part of this process, we evaluated the model outputs when including

295 infiltration, finding an overestimation of infiltration and underprediction of flooding using the Curve Number method. Based on this and the high imperviousness of the urban area, infiltration was excluded from the final model configuration. Additionally, the known flood-prone areas identified by local authorities (Smith, 2023) were well captured by our 10-year flood hazard map. This multi-source approach provides the most robust validation feasible for this site given current data availability. A detailed description of the model validation is included in the Supporting Material.

300

Different types of observations can be used to validate and calibrate flood extent and water depths, such as high-water marks and satellite images. In Table S2, we list the different sources we considered in this study to validate the flood model; it shows that little information is available about historical flooding in Gloucester City. Satellite images (or aerial images) are the only type of observed data that can measure the flood extent (Bates, 2004), and although these can be very useful for validating and

305 calibrating large-scale flood models against large flood events (Tellman et al., 2021), for small-scale applications their resolution is too coarse (Table S2, row 1). For validating and calibrating water depth outputs of flood models, high-water

marks are commonly used. In the U.S., the U.S. Geological Survey, in partnership with FEMA, is the agency responsible for collecting high water marks after major flood events (Table S2, row 2). However, no high water marks from USGS are available for Gloucester City. Another source of flood-related information is the NOAA Storm Event dataset, which documents significant weather events from 1950 onwards. However, no flood events have been recorded for Gloucester City (Table S2, row 3). Comparison against FEMA Special Flood Hazard Areas (SFHA) has also been used as a measure of flood model performance (Bates et al., 2021). However, we are including pluvial flooding, which is neglected when delineating the FEMA SFHA, and thus the comparison of our 1% AEP flood map to the FEMA SFHA will not provide a suitable measure of model performance. In addition, some studies have pointed to inaccuracies in the 1% AEP FEMA SFHA (Flores et al., 2023; National Academies of Sciences, 2019). Nevertheless, we found useful information for one particular flood in Gloucester City in the Flood Risk Map of FEMA (FEMA, 2016; Table S2, row 4). This map is paired with a photo of a flooded intersection that is located outside of the SFHA and marked on the map with a star (Fig. S1). This flood happened in 2009, and based on the water level records and precipitation data, the flooding occurred due to a short but heavy rainfall event, with a maximum precipitation of 33.5 mm/h. We simulated this event using two approaches: one neglecting infiltration and another accounting for infiltration based on the Curve Number method included in SFINCS. We find that including infiltration underestimates flooding since that simulation results in no flooding in the intersection where flooding was documented (see Pollack et al., 2024 for further details). Therefore, we decided to neglect infiltration in the configuration of SFINCS to perform all simulations in this study.(Nederhoff et al., 2024)Neglecting infiltration, the average water depth resulting from the simulation of that event in the intersection is 36 cm, a value similar to the water depth (approx. height of the curbside) shown in the FEMA Flood Risk Map (FEMA, 2016). We also searched for reported flooding in Gloucester City in local news, finding a couple of articles regarding a flood event that occurred in June of 2019 (Table S2, row 5). That event flooded a few areas of the city and led to power outages. In one of the articles, flooding of ~91 cm (3 feet) was reported in the basement of the Historical Society Museum. We simulated this event, finding that our flood model results in flooding of the area where the museum is located with average water depths of 68 cm (2.23 feet).

Since almost no information about historical flood hazards was found in the most commonly used sources, we also considered crowd-sourced platforms such as social media and the citizen science platform MyCoast: New Jersey. Crowd-sourced observed flood data such as photos and geo-localized tweets have been shown to be useful for flood model validation as they often include information about the location of the flood and water depth levels can be extracted from photos (de Bruijn et al., 2019; Wang et al., 2018)(de Bruijn et al., 2019; Wang et al., 2018)(de Bruijn et al., 2019; Wang et al., 2018)(de Bruijn et al., 2019; Wang et al., 2018). We checked the Global Flood Monitor dataset (de Bruijn et al., 2019)(de Bruijn et al., 2019)(de Bruijn et al., 2019)(de Bruijn et al., 2019), which has an algorithm that automatically searches for flood keywords within tweets posted in any location of the world. In this database, we found nine tweets reporting flooding within the region of interest. However, only one of them (which occurred in 2020) contained photos and information about the exact location of the flooding reported (Table S2, row 6). In one of these photos, a traffic barrel is shown in reference to the flooding, which allows us to estimate a flood depth of approx. 60 cm based on standard traffic barrels of 5 sheeting of 15.24 cm each. We simulated that flood event

4 Results

4.1 Differences in flood response from events with the same (joint) probability

We simulate flooding from 25 events with a 1% chance of occurring in any given year (1% AEP) that have different combinations of the magnitude of rainfall and NTR peak, but also different temporal and spatial evolutions and are combined with different tidal curves (Fig. 2). We find that the floodplain of each of these 1% AEP events is different (Fig. S7), resulting in very large differences in both flood extent and depth between some of the events (Fig. S7). In Figure 4-a, we show the frequency of flooding at each cell ($1m^2$) from all 25 1% AEP events; a frequency of 1 indicates that the particular area is flooded from all 25 events and a frequency near zero indicates that this area is only flooded from one or few of the 25 events. Certain areas scattered throughout the municipality experience flooding during all 1% AEP events with a 1% chance of occurring in any given year (1% AEP; green areas in Fig. 4-a). However, a larger area is flooded only by a few of the 1% AEPse events, which shows that selecting only one 1% AEP event with 1% chance of occurring in any given year (1% AEP) for estimating flood hazard can introduce large uncertainties in exposure (and subsequently risk). Areas flooded only from a few 1% AEP events are mainly along the Delaware River and creeks, where both flood drivers interact. In contrast, pluvial hot spots, i.e. regions that are not hydrologically connected to the Delaware River or creeks and thus rainfall is the only flood driver, exhibit notably lessshow almost no variability in flood extents and water depths between the different 1% AEP events of 1% chance of happening in any given year (. The variability of flooding in pluvial hotspots is more clearly observed by comparing the flood hazard maps across the 25 events presented in Figure S7 of the Supporting Material.

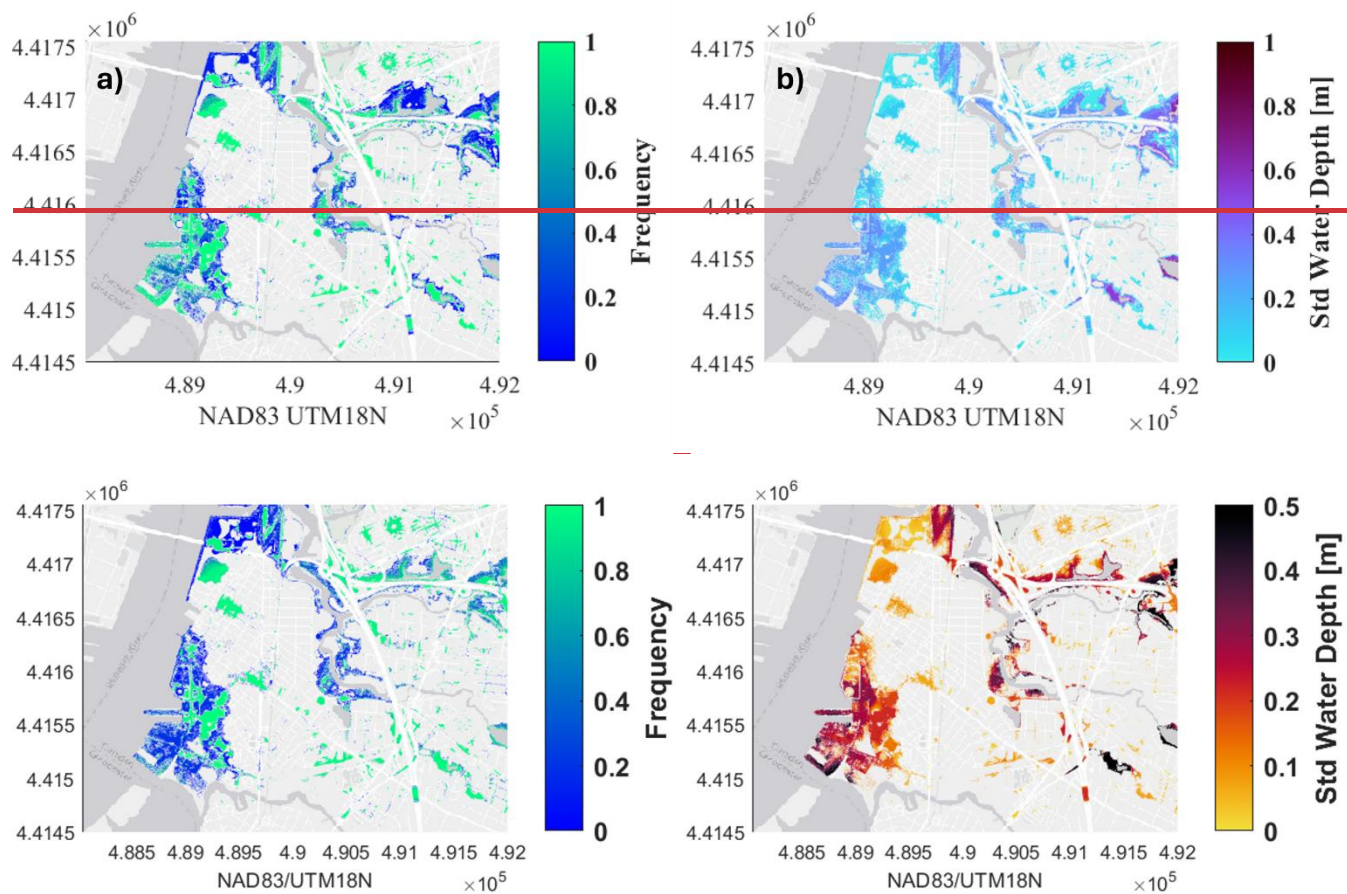



Figure 4 a) Frequency of flooding at each model cell (1m²) from the 25 events selected along the 100-year (1% AEP) isoline. b) The standard deviation of water depth (m) at each model cell (1m²) between the 25 events selected along the 100-year (1% AEP) isoline. [NAD83/UTM18N. ©Esri]

We also analyze the variability in water depths using the standard deviation between the flooding from all 25 ~~1% AEP~~ events 1% chance of occurring in any given year -in all model cells (Fig. 4-b). Larger standard deviations in water depth exist in regions where all ~~1% AEP~~ events of 1% chance of occurring in any given year -produce flooding, with maximum values of ~0.8 m. Larger variability of water depths also exists in coastal regions, where both flood drivers interact, while small variations occur in pluvial hotspots (see also Fig. S7).

The largest and smallest flooding, in terms of flood extent and volumes, are produced by events that have almost the same 18h accumulation rainfall and NTR peak, 59.18 mm and 1.86 m and 57.18 mm and 1.88 m respectively, and thus lie very close to each other on the 100-year (1% AEP) isoline (Fig, 2 c and d). However, the NTR hydrographs of these events are different, with one of them lasting for several hours with sustained large water levels (Fig. 2c) while the other is shorter and with lower

water levels (Fig. 2d). In addition, the two events are combined with different tidal curves with high-tide levels that differ by more than 20 cm, an average MHHW and a larger than average king tide (see Fig. 3). The 1% AEP event that produces the smallest flooding is combined with a tidal curve with high-tide levels similar to MHHW, but the 1% AEP event that generates

390   the largest flood is combined with a tidal curve that reaches values larger than the average king tide (Fig. 3). These factors cause the water level hydrographs of these events to differ in their temporal evolution and to have water level peaks that differ by ~0.5 m, which combined leads to the large differences in flood response. Although the tidal curve combined with the 1% AEP event that produces the largest flood might appear "extreme", the analysis of the long-term modulations of the tide reveals that this king tide level was reached several years earlier in the current nodal cycle (Fig. 5). By extending the fitted long-term

395   modulation, we show that the tides are currently in the ascending phase of both nodal and perigean cycles, with a peak expected in 2026. As a result, the likelihood of the tidal level of this particular synthetic event is higher over the coming years.

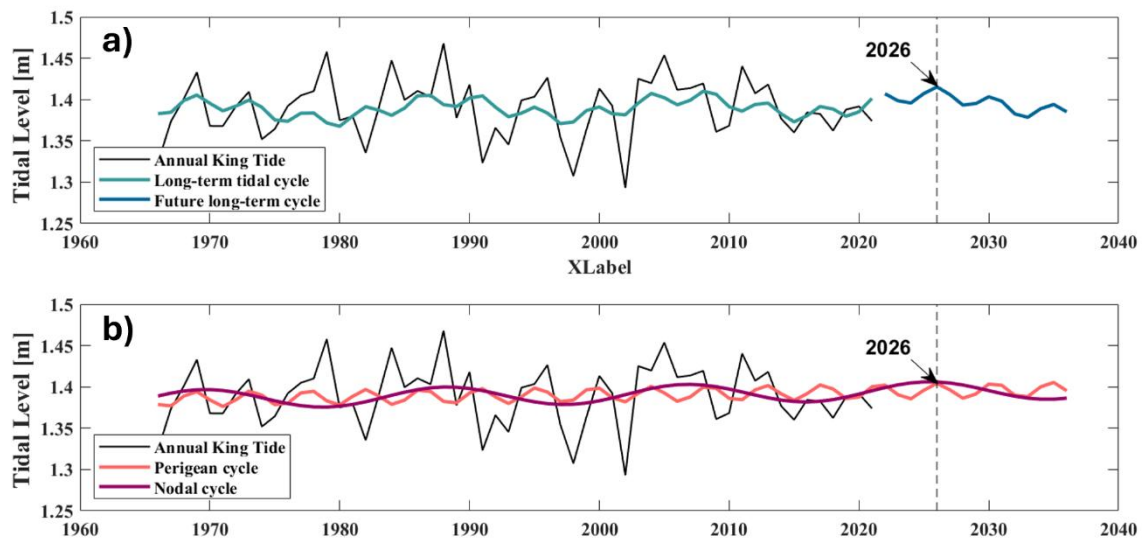



Figure 5 Annual king tide levels (as the maximum tidal level) at the tide-gauge of Philadelphia used in this study. Estimated long-term variability of king tides (a) from the combined nodal (18.6-y) and perigean (4.4-y) cycles including prediction of future

400   **combined peak of these long-term cycles expected for 2026; (b) separated nodal (18.6-y) and perigean (4.4-y) cycles for the historical and future period.**

These results show that the variability of the NTR hydrograph, together with the variability of the tidal curve, have very large effects on the resulting flooding since events with almost equal NTR peaks can produce very different flooding. The

405   topography also plays an important role (Fig. S1); when the water level at the Delaware River boundary exceeds the elevation of the coastline, the large low-lying region behind it floods. Thus, small increases in water levels along the hydrograph can cause large changes in flood extents when certain thresholds are exceeded.

We also compare the flooding arising from all ~~1% AEP~~ events with a 1% chance of occurring in any given year with the "most likely" event (i.e. "design event") in order to assess the uncertainties related to the use of a single design event when assessing flood hazards. This is commonly done when following the event-based approach for compound flood hazard modeling. In terms of flood extent and volumes (Fig. 6), most of the ~~1% AEP~~ events with a 1% chance of happening in any given year (17 and 19 of the 25, respectively) produce larger flooding than the "most likely" design event. However, there are substantial spatial variations between events, as some can produce larger flooding in some areas and smaller flooding in other areas (not shown). We calculate the total flood extent and volume of all 5,000 events to estimate the empirical return periods of these two flood metrics (Fig. 7). Based on the empirical distribution, the "most likely" event has a return period of 38 years in terms of extent and 33 years in terms of total flood volume. 20 of the ~~1% AEP~~ events with 1% chance of happening in any given year have return periods <100 years (> 1% AEP, >1% chance of happening in any given year) in terms of total flood extent, while only 12 events have return periods <100 years in terms of total flood volumes. This shows that using a single design event when assessing compound flood hazards can lead to large uncertainties in both flood extent and depth, often resulting in an underestimation of the extent in our case study.

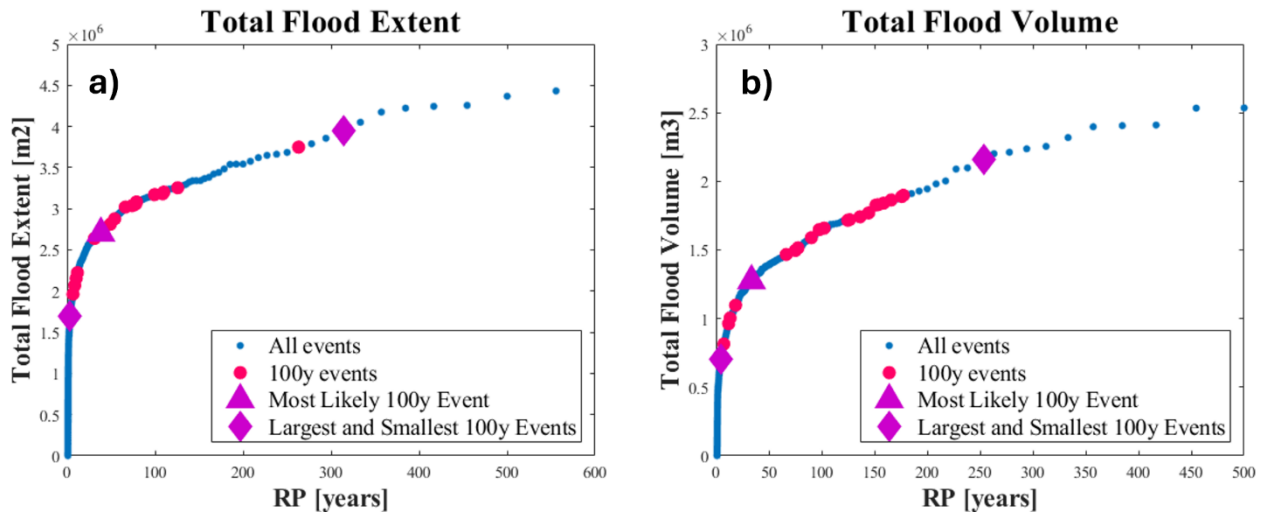



Figure 6. (a) Empirical distribution of total flood extent from all events (blue points), 1% AEP events (pink dots) including the "most likely" (purple triangle) and largest and smallest flood (purple diamonds); (b) same as (a) but in terms of total flood volumes.

4.2 Response-based flood hazard

We estimate the response-based flood hazard map by calculating the empirical distributions of water depths at each model cell (of 1 m²) and show the water depth with a 1% chance of happening in any giver year ~~corresponding to the~~ (1% AEP~~;~~ ~~(~~Fig. 6). This 1% AEP response flood hazard can thus be produced by different events in different regions. Comparing the response flood hazard to the flood hazard of the different ~~1% AEP~~ events with a 1% chance of happening in any given year (1% AEP), the response flood hazard has generally larger flood extents and water depths, with a few exceptions. In the 1% AEP response flood hazard map, there is a larger area in the south of Gloucester City facing the Delaware River (Fig. 6) with water depths

up to 1 m. However, this region is only flooded by a few ~~1% AEP~~ of the events with a 1% chance of happening in any given year (1% AEP; Fig. 4). In contrast, the flood hazard in the northern Delaware region from the response-based flood hazard map is similar to the region flooded by all ~~1% AEP~~ the events with a 1% chance of happening in any given year (1% AEP). In this region, the ~~1% AEP~~ event with a 1% chance of happening in any given year that causes the largest overall flooding (Fig.

435   2c) also produces more extensive flooding. This might be caused by the relatively longer hydrograph of that event combined with a larger-than-average king tide. Comparing the differences between the response-based and event-based flood hazard in the pluvial hotspots (areas not hydrologically connected to the Delaware River), only one ~~1% AEP~~ of the event with a 1% chance of happening in any given year causes larger flooding than the response-based approach. In the northeast region of the domain, ten of the ~~1% AEP~~ events with a 1% chance of happening in any given year (1% AEP) produce larger flooding than

440   the response-based 1% AEP floodplain. This can point to effects of the spatial variability of rainfall fields between events, which are masked in the joint probabilities because these are based on the 18h accumulated average rainfall in the catchment.

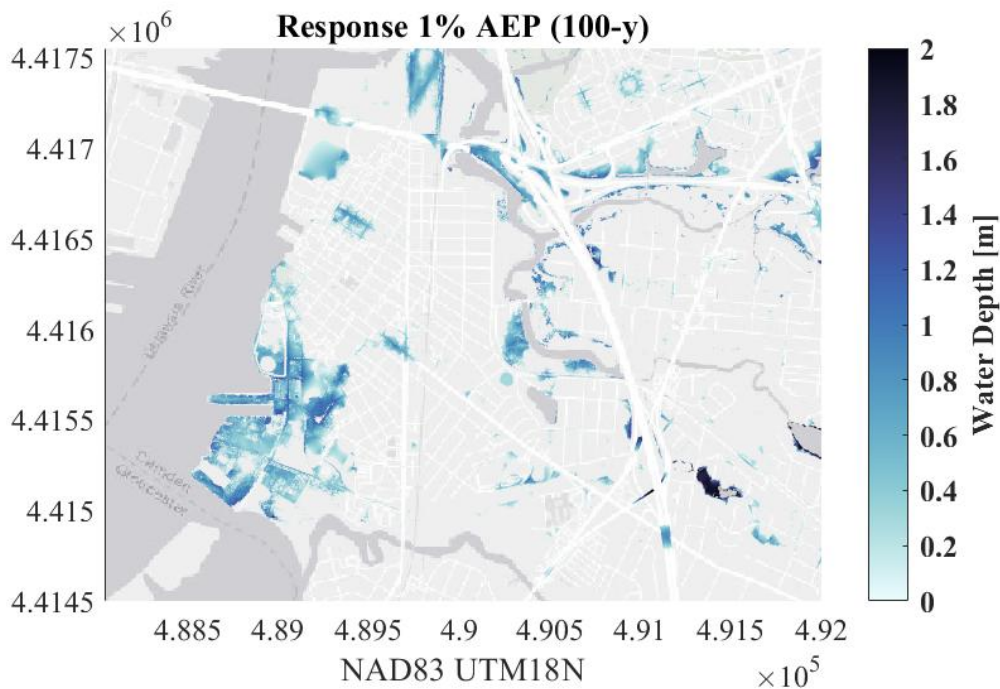



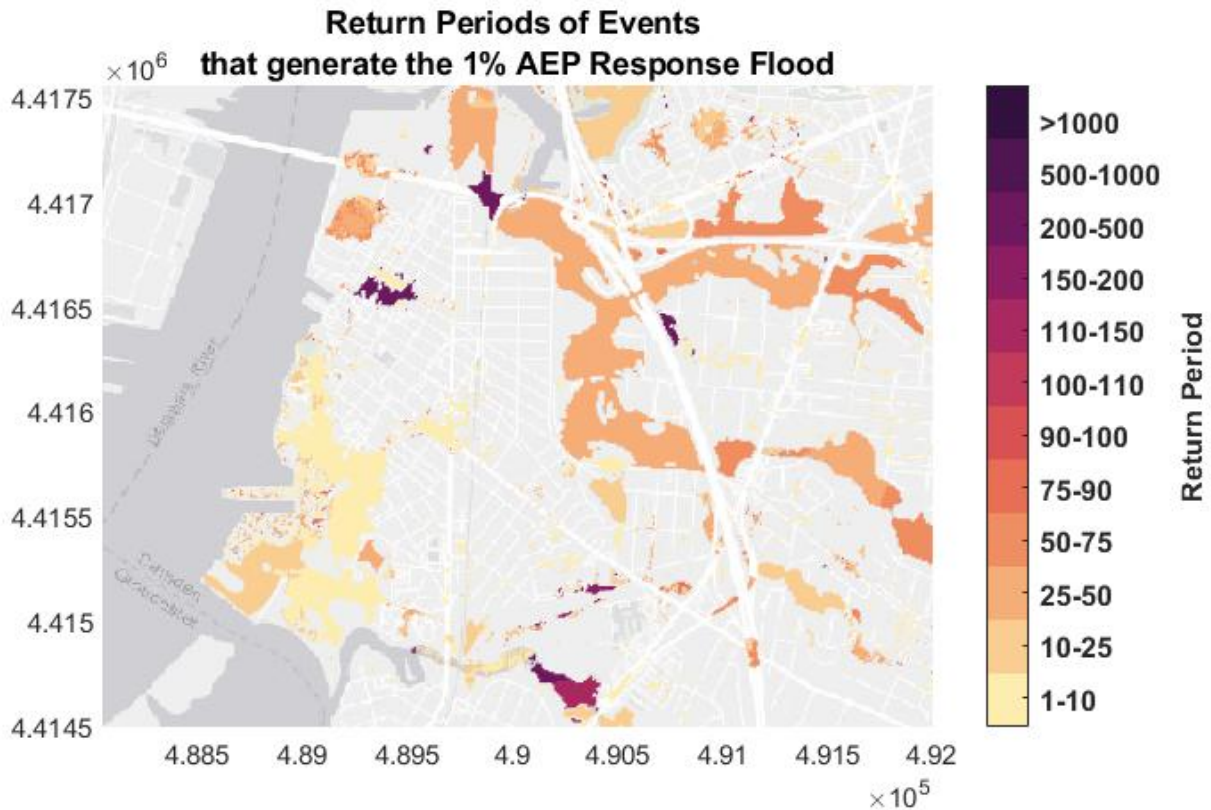
Figure 7. 1% AEP response flood hazard map calculated by the empirical distribution of water depths from 5,000 simulations; we show the 1% AEP water depth at each model cell (1m^2) [NAD83/UTM18N. ©Esri]

445   We trace the events that produce the ~~1% AEP~~ response-based flood hazard of 1% chance of happening in any given year (1% AEP) and group them by their corresponding return periods based on the joint probabilities of flood drivers (Fig. 8). This helps identifying the types of events causing the 1% AEP water depth in different areas of our study site. Across most of the study

domain, and especially in the urbanized region in the centre of the city, events with joint AEPs much higher than 1% can lead to the 1% AEP water depth as identified from the response-based approach. Most of the ~~1% AEP~~ water depths with a 1% chance of happening in any given year along the south coastal region of the city are produced by a single compound event with a~~n AEP of~~ ~50% chance of happening in any given year (50% AEP or 2-year return period; yellow) and another event with a~~n AEP of~~ ~7% chance of happening in any given year (7% AEP or 14-year return period; light orange) based on the joint probability distribution of NTR and 18h accumulated rainfall. Although the NTR peak of these events is around 1 m, and thus much smaller than other events, these two events have long hydrographs with sustained water levels for several hours and combined with tidal levels of around 1 m can produce the 1% AEP water depths in that area (Fig. S5~~4~~ of Supporting Material). The tidal levels of these two events (#605 and #3354 in Fig. 3) exceed the MHHW but remain below the mean spring tides, making them likely to occur on a fortnightly basis. The 1% AEP water depths in regions that are only affected by pluvial flood events are generally also caused by events with ~~AEPs~~ >2% chance of happening in any given year (AEPs). Notably, there are two pluvial hotspots in the city region produced by events with ~~AEPs~~ >10% chance of happening in any given year (less than 10-year return period or >10% AEPs). These are produced by two different events, both with AEPs of ~12% based on the joint probability distribution. More detailed assessment of the rainfall fields of these events reveals that they have larger rainfall over that area of the model domain, which gets masked when averaging the rainfall over the entire catchment (Fig. S6~~5~~-S6~~7~~).

450

455

460

465 **Figure 8. Return periods (based on the bivariate extreme model) of the events that generate the 100-year water depth from the response approach. [NAD83/UTM18N. ©Esri]**

5 Discussion

~~Lots~~ Much of the research has been dedicated to improving extreme statistics of compound events and to quantifying the uncertainties of extreme value analysis of flood drivers (e.g., Lucey and Gallien, 2024), assuming that the probability of the
470 event approximates the probability of the resulting flooding. However, little research has been focused on analyzing the latter, specifically for compound flooding, in which more than one driver is involved, and thus different combinations of flood driver magnitudes have the same joint probabilities. Here, we have assessed how ~~well the 1% AEP event approximates the 1% AEP~~linear is the relationship between the probability of the event and the probability of flooding for a case study in Gloucester City (NJ, U.S.) by comparing the ~~1% AEP~~ flood hazard with a 1% chance of happening in any given year (1% AEP) based on
475 the event- and response-based approaches. We find that the 1% AEP water depth can be produced by different events in different parts of the city and that the AEPs of these events are often much larger than 1%. This means that the relationship between the probability of the event and the probability of flooding does not follow a one-to-one relationship. These results are in line with previous studies that addressed the same question for rainfall-driven flooding (Baer, 2025; Perez et al., 2024; Winter et al., 2020). ~~Specifically, t~~We find that the region of Gloucester City with the largest 1% AEP flood hazard is the

480   coastal zone and it is caused by events with ~~AEPs~~ >5% chance of happening in any given year (>5% AEP and less than 20-year return period) based on the joint probability distribution of 18h rainfall accumulation and NTR peaks. However, these events (Fig. S~~4~~5) exhibit sustained high NTR levels, which, when combined with tidal levels larger than MHHW, can result in greater water depths than most of the analyzed ~~1% AEP~~ events with 1% chance of happening in any given year (1% AEP). Similarly, the regions impacted mainly by pluvial flooding also tend to experience the 1% AEP water depths from events with

485   ~~AEPs~~ >10% chance of happening in any given year (>10% AEP or less than 10-year return period). In this case, the events producing the 1% AEP water depths in the city show spatial variations in their rainfall fields, with larger precipitation rates over the urbanized city region. However, the spatial variability of the rainfall fields is smoothed when calculating the average rainfall in the catchment to perform the extreme value analysis, and together with small NTR peaks, these events get assigned high AEP values (or short return periods).

490   These results show that the response-based approach leads to better representation of flood hazard at the household level. It accounts for the temporal variability of NTR hydrographs, combination with tides and mean sea level, and the spatial variability of rainfall fields. All of those are not explicitly accounted for in extreme value models to derive joint AEPs or joint return periods for the event-based approach. Nevertheless, some applications, such as emergency management, might need event-based flood hazard maps.

495   Event-based flood assessments commonly use a single design event with specific temporal and spatial structure, thus neglecting ~~temporal and spatial variations~~ the variability in the temporal and spatial evolution of the flood drivers between different events ~~along the duration of the event~~. We have shown that using a single design ~~1% AEP~~ event with a 1% chance of happening in any given year (1% AEP) can introduce large uncertainties in both flood extents and water depths that arise from the different combinations of the drivers' magnitude but are mostly due to ~~the~~ differences in temporal and spatial ~~variability~~

500   evolution ~~of~~ between events. Events of almost equal magnitude but different spatial rainfall fields and temporal ~~variability~~ distribution of the water level hydrographs can produce very different flood extents and water depths. The disparities in resulting flooding are more pronounced in the coastal areas of our study domain, where both flood drivers interact and are further influenced by changes in tidal variability. Considering the variability of the tide, rather than relying on a single MHHW level is also crucial, as tidal fluctuations over longer time scales (such as spring and king tides) can influence coastal flooding.

505   This is especially relevant now, as tides are in the ascending phase of their long-term cycles, which are projected to reach their peaks within this decade, with the first peak expected as soon as 2026 at our study site. This finding highlights the necessity of taking into account the variability in the tidal levels.

    Ignoring the variability in the spatio-temporal ~~variability~~ structure between~~of~~ extreme events by relying on a single design event can lead to significant uncertainties in flood exposure, which in turn can result in substantial uncertainties in flood risks.

510   One way to address this limitation when using the event-based approach is to employ ensembles of events that account for variations in the spatial and temporal ~~variability~~ structure of the flood drivers. From that one can produce an ensemble of flood hazard maps for a desired return period, similar to probabilistic flood maps but for a given AEP (or return period).

Our study has several limitations that highlight areas for further research. We focused on a small study site with a particular topography that is affected by two flood drivers with associated variabilities. Thus, the results cannot be extrapolated to other

515 coastal regions. However, we expect that our general conclusions are transferable to other regions. For example, the importance of temporal and spatial variations of the flood drivers has been pointed out by other studies in Germany (Kupfer et al., 2024; Santamaria-Aguilar et al., 2017) and in the UK (Quinn et al., 2014), showing that changes in water level hydrographs can produce large changes in flood hazards. Likewise, differences in rainfall-induced flooding between the event-based approach and the use of synthetic storms that capture the temporal and spatial variability of rainfall fields between events have been

520 shown to significantly influence flood hazard estimates in the East and Gulf coasts of US (Baer, 2025; Perez et al., 2024) and Austria (Winter et al., 2020). Another limitation of our study is that we use a synthetic event set developed using a data-driven statistical framework, which is limited to observed events. Although the statistical framework used to generate the synthetic events account for more dependencies between parameters that characterize the events (e.g. time lags) than other previous frameworks (Couasnon et al., 2018; Moftakhari et al., 2019), ~~Consequently,~~ it may not fully capture the full range of the

525 potential spatio-temporal variability of flood drivers. Tropical cyclones might also be underrepresented in the historical sample since their frequency of occurrence is very low. This limitation can be overcome by using synthetic tropical cyclones that are dynamically downscaled to the study site (e.g., Gori et al., 2020) ~~This is computationally expensive but has been done for other regions . This~~ Methods such as the JPM, which expand the storm climatology, enable the generation of a larger set of tropical cyclones, and capture greater variability in their spatio-temporal characteristics compared to historical records.

530 However, these methods are computationally demanding, as flood drivers must be generated in advance of the flood assessment using hydrodynamic models. Further research is needed to evaluate how different synthetic event generation approaches affect flood hazard estimates. Given the high computational demands of JPM, its application across large coastal areas may be impractical, making data-driven approaches like the one used in this study a more efficient alternative. Similarly, other data-driven techniques, such as stochastic storm transposition, are increasingly being adopted to generate synthetic rainfall fields

535 for assessing rainfall-driven flood hazards ~~st~~ (Baer, 2025; Perez et al., 2024; Winter et al., 2020). However, further investigation is needed to ensure that this method adequately preserves the interdependencies between coastal and rainfall processes when generating synthetic compound events for coastal flood assessments. ~~allows for increasing the number of extreme events in the sample but also for changes in the spatio-temporal variability of these events compared to the historical record.~~ A potential source of uncertainty in the variability captured by our synthetic event set arises from not disaggregating river- and coastal-

540 driven components of the NTR. In our mid-estuarine study area, both processes contribute to the NTR, along with their nonlinear interactions. Separating these contributions would introduce considerable complexity due to their tightly coupled dynamics. Our approach is supported by recent work from McKeon and Piecuch, (2025), who investigated the relative influence of coastal and fluvial drivers in the Delaware Estuary above flood thresholds. They found that most events observed at the Philadelphia tide gauge were primarily driven by coastal processes (e.g., tides and storm surge), but others resulted from

545 river discharge alone or a combination of both mechanisms.

Another limitation of the synthetic event set used is the reliance on mathematically defined thresholds for event selection, rather than thresholds based on actual flood impacts. This approach may exclude relatively frequent, lower-magnitude events that fall outside the statistical tails of the drivers' distributions but are still capable of causing localized flooding, potentially influencing response-based flood estimates. In our study, we evaluated the flood response of events near the selected thresholds
550 and found that several produced no flooding, while others resulted in only minor inundation, with empirical return periods between 1 and 2.8 years. As a result, the selected thresholds did not affect our response-based flood estimates; however, this may not hold true in other regions with different hydrologic or exposure characteristics.

Our flood model approach also has some limitations. First, we are neglecting the stormwater system and thus we might overestimate flooding and neglect the fact that some areas might experience flooding due to backwater effects in the system.
555 Stormwater systems are typically designed for events with low to moderate return periods (often the 10-year event, or 10% AEP event) while we have focused on the 1% AEP (100-year) flood hazard. Such events would likely exceed the capacity of the stormwater system. Nevertheless, the exclusion of stormwater infrastructure may have a greater impact on the results for smaller, more frequent events, potentially leading to an overestimation of flooding in cases where the existing drainage system would likely manage the runoff. However, this should not affect the response-based estimates for the 1% AEP since the
560 empirical distribution will not change for rare large events. In addition, we neglected infiltration based on the validation of the model for one single event for which we have information on reported flooding at a single location. Although most of our study domain is urban and thus covered by impervious surfaces, we might underestimate infiltration in areas of the with larger amounts of vegetation such as areas around the creeks. The bathymetry of the creeks might also not be very accurately represented in the CoNED DEM used (Fig. S1). The limited depth representation of creek channels, combined with the
565 exclusion of infiltration processes, likely results in an overestimation of floodwater depths along the margins of the creeks. Both the validation of the flood model and calibration of parameters and processes such as infiltration can be improved if more observed flood data from past events is available. The lack of this data is a common problem worldwide (Merz et al., 2024; Molinari et al., 2019) and it can be overcome by systematically collecting flood data after flood events or making available datasets such as claims from the National Flood Insurance Program (Sebastian et al., 2021).

570 **6 Conclusions**

Coastal communities are experiencing growing flood hazards due to rising sea levels, more frequent extreme events, and an increase in population and assets in flood prone areas. Consequently, more robust flood hazard estimates are required to develop effective adaptation strategies to mitigate flood impacts. Although significant attention has been focused on reducing uncertainties in the estimation of probabilities of flood drivers, little is known about how well the probability of compound
575 events approximates the probability of flooding. Here, we addressed this issue by comparing flood hazard derived from the event- and response-based approaches for a case study in Gloucester City (NJ, U.S.), which is frequently affected by pluvial and coastal flooding.

Our findings reveal that the 1% AEP flood hazard derived from the response-based approach can be caused by different events in various parts of the city, with AEPs much larger than 1% (return periods <100-year). In the coastal area, events with ~~AEPs~~ >5% chance of occurring in any given year (>5% AEP or less than 20-year return period) can produce a 1% AEP water depth if the NTR hydrograph leads to prolonged high water levels when combined with tidal levels between the MHHW and average spring tides. In this context, our findings are in line with previous studies that highlighted that the long-term variability of tides can modulate both minor and extreme flooding (Enriquez et al., 2022; Thompson et al., 2021). We find that considering tidal variability is crucial, rather than relying on the assumption of a constant MHHW, as flooding from both low and high return period events can differ substantially depending on the tidal level considered. Tides are currently in the ascending phase of the nodal and perigean cycles, which are expected to peak in 2026 in our study region, making it more likely that storm surges coincide with high tide levels, thus increasing the probability of flooding. Similarly, not accounting for the variability in the spatial pattern of~~variability of~~ rainfall fields between events, which is masked when using catchment average values for extreme value analysis, can underestimate pluvial flood hazards. This study highlights the importance of considering the variability of the temporal and spatial ~~variability~~ structure of extreme events in flood hazard estimates. The traditional method of using a single design event in event-based assessments can lead to considerable uncertainties in flood extent and water depth, especially due to varying combinations of flood drivers. The response-based approach, which accounts for factors like tidal variations and ~~spatial variability in rainfall~~the full range of the variability of temporal and spatial distributions at event scale, provides a more robust representation of flood hazards. However, event-based maps remain essential for some applications such as emergency management. Using ensembles of events that account for these variations would enhance flood hazard estimates derived from the event-based approach.

While our results are not directly applicable to other regions, we expect similar conclusions elsewhere regarding the impacts on compound flood hazards from neglecting the temporal and spatial variability of flood drivers. Future work should focus on producing more robust flood hazard estimates by using many compound events including their temporal and spatial evolution rather than focusing on single design events for given AEPs or return periods. Similarly, future projections of flood hazards should also account for potential changes in the temporal and spatial evolution of events rather than focusing only on changes in their magnitude. Additionally, future research should aim to evaluate how different methods for generating synthetic events influence the resulting flood hazard estimates. Such comparisons can help inform best practices for generating more reliable flood hazard assessments under both current and future climate conditions.

Code availability

The SFINCS model is available at https://sfincs.readthedocs.io/en/latest/example.html#executable. The codes used for these analyses are available on GitHub (https://github.com/CoRE-Lab-UCF/MACH-Compound-Flooding/tree/main/Santamaria-Aguilar_et_al_2025_Event_Response) (The DOI and the final version of the codes will be available after addressing the reviewers' comments and suggestions.)

Data availability

The hydrologic units are available https://gisdata-njdep.opendata.arcgis.com/datasets/02599a9424254a4ea33e689941559e3c_17/explore. The DEM is available at https://www.usgs.gov/special-topics/coastal-national-elevation-database-applications-project/data, and land cover data is available at https://gisdata-njdep.opendata.arcgis.com/documents/njdep::land-use-land-cover-of-new-jersey-2015-download/about.

The SFINCS model files generated and used in this study are available at https://zenodo.org/records/14251309, and the 5,000 flood simulations are available at https://zenodo.org/records/15047845.

Authors contribution

The study was conceived by TW and SSA. SSA performed the simulations, undertook the analyses, and wrote the first draft of the paper under the guidance of TW, PW, and ARE. All authors co-wrote the final paper.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

S.S.A, P. M. and T.W. were supported by the National Science Foundation as part of the Megalopolitan Coastal Transformation Hub (MACH) under NSF award ICER-2103754. This is MACH contribution number 76. We would like to acknowledge Maarten van Ormondt for providing technical assistance with the SFINCS model. We would also like to thank Stefan Talke for his support in the analysis of tidal levels.

References

Alfonso, L., Mukolwe, M. M., and Di Baldassarre, G.: Probabilistic Flood Maps to support decision-making: Mapping the Value of Information, Water Resour Res, 52, 1026–1043, https://doi.org/10.1002/2015WR017378, 2016.

Baer, J. A.: Design storms underestimate flood hazard and risk derived from the stochastic storm transposition, Chapel Hill, 56 pp., 2025.

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

Bates, P. D.: Remote sensing and flood inundation modelling, Hydrol Process, 18, 2593–2597, https://doi.org/10.1002/hyp.5649, 2004.

Bates, P. D., Horritt, M. S., Aronica, G., and Beven, K.: Bayesian updating of flood inundation likelihoods conditioned on flood extent data, Hydrol Process, 18, 3347–3370, https://doi.org/10.1002/hyp.1499, 2004.

640 Bates, P. D., Dawson, R. J., Hall, J. W., Horritt, M. S., Nicholls, R. J., and Wicks, J.: Simplified two-dimensional numerical modelling of coastal flooding and example applications, Coastal Engineering, 52, 793–810, https://doi.org/10.1016/j.coastaleng.2005.06.001, 2005.

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., Savage, J., Olcese, G., Neal, J., Schumann, G., Giustarini, L., Coxon, G., Porter, J. R., Amodeo, M. F., Chu, Z., Lewis-Gruss, S., Freeman, N. B., Houser, T., Delgado, M., Hamidi, A.,

645 Bolliger, I., E. McCusker, K., Emanuel, K., Ferreira, C. M., Khalid, A., Haigh, I. D., Couasnon, A., E. Kopp, R., Hsiang, S., and Krajewski, W. F.: Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates, Water Resour Res, 57, 1–29, https://doi.org/10.1029/2020WR028673, 2021.

de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., and Aerts, J. C. J. H. J. H.: A global database of historic and real-time flood events based on social media, Sci Data, 6, 311, https://doi.org/10.1038/s41597-019-0326-9, 2019.

650 Codiga, D. L.: Unified Tidal Analysis and Prediction Using the UTide Matlab Functions, 59 pp., https://doi.org/10.13140/RG.2.1.3761.2008, 2011.

Cosby, A. G., Lebakula, V., Smith, C. N., Wanik, D. W., Bergene, K., Rose, A. N., Swanson, D., and Bloom, D. E.: Accelerating growth of human coastal populations at the global and continent levels: 2000–2018, Sci Rep, 14, 1–10, https://doi.org/10.1038/s41598-024-73287-x, 2024.

655 Couasnon, A., Sebastian, A., and Morales-Nápoles, O.: 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 (Switzerland), 10, https://doi.org/10.3390/w10091190, 2018.

Dawson, R. J., Hall, J. W., Bates, P. D., and Nicholls, R. J.: Quantified analysis of the probability of flooding in the Thames estuary under imaginable worst-case Sea Level Rise scenarios, Int J Water Resour Dev, 21, 577–591,

660 https://doi.org/10.1080/07900620500258380, 2005.

Dullaart, J. C. M., Muis, S., De Moel, H., Ward, P. J., Eilander, D., and Aerts, J. C. J. H.: Enabling dynamic modelling of coastal flooding by defining storm tide hydrographs, Natural Hazards and Earth System Sciences, 23, 1847–1862, https://doi.org/10.5194/nhess-23-1847-2023, 2023.

Enriquez, A. R., Wahl, T., Baranes, H. E., Talke, S. A., Orton, P., Booth, J. F., and Haigh, I. D.: Predictable changes in extreme

665 sea levels and coastal flood risk due to long-term tidal.pdf, J Geophys Res Oceans, 2022.

FEMA: Flood Risk Report Camden County Coastal Project Area, New Jersey, 2016.

FEMA: Guidance for Flood Risk Analysis and Mapping Superseded, 2022.

Flores, A. B., Collins, T. W., Grineski, S. E., Amodeo, M., Porter, J. R., Sampson, C. C., and Wing, O.: Federally Overlooked Flood Risk Inequities in Houston, Texas: Novel Insights Based on Dasymetric Mapping and State-of-the-Art Flood Modeling, Ann Am Assoc Geogr, 113, 240–260, https://doi.org/10.1080/24694452.2022.2085656, 2023.

Gori, A., Lin, N., and Xi, D.: Tropical Cyclone Compound Flood Hazard Assessment: From Investigating Drivers to Quantifying Extreme Water Levels, Earths Future, 8, https://doi.org/10.1029/2020EF001660, 2020.

Gouldby, B., Méndez, F. J. J., Guanche, Y., Rueda, a., and Mínguez, R.: A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis, Coastal Engineering, 88, 15–26, https://doi.org/10.1016/j.coastaleng.2014.01.012, 2014.

Haigh, I. D., Eliot, M., and Pattiaratchi, C.: Global influences of the 18.61 year nodal cycle and 8.85 year cycle of lunar perigee on high tidal levels, J Geophys Res Oceans, 116, 1–16, https://doi.org/10.1029/2010JC006645, 2011.

Helgeson, C., Auermuller, L., Bennett Gayle, D., Dangendorf, S., Gilmore, E. A., Keller, K., Kopp, R. E., Lorenzo-Trueba, J., Oppenheimer, M., Parrish, K., Ramenzoni, V., Tuana, N., and Wahl, T.: Exploratory scoping of place-based opportunities for convergence research, OSF Preprints, April 29, https://doi.org/10.31219/osf.io/z5ue4, 2024.

Hino, M. and Nance, E.: Five ways to ensure flood-risk research helps the most vulnerable, Nature, 595, 27–29, https://doi.org/10.1038/d41586-021-01750-0, 2021.

IPCC: SPECIAL REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE MANAGING THE RISKS OF EXTREME EVENTS AND DISASTERS TO ADVANCE CLIMATE CHANGE ADAPTATION, NY, USA, 2012.

Jane, R., Cadavid, L., Obeysekera, J., and Wahl, T.: Multivariate statistical modelling of the drivers of compound flood events in south Florida, Natural Hazards and Earth System Sciences, 20, 2681–2699, https://doi.org/10.5194/nhess-20-2681-2020, 2020.

Jane, R., Wahl, T., Santos, V. M., Misra, S. K., and White, K. D.: Assessing the Potential for Compound Storm Surge and Extreme River Discharge Events at the Catchment Scale with Statistical Models: Sensitivity Analysis and Recommendations for Best Practice, J Hydrol Eng, 27, https://doi.org/10.1061/(ASCE)HE.1943-5584.0002154, 2022.

Kim, H., Villarini, G., Jane, R., Wahl, T., Misra, S., and Michalek, A.: On the generation of high-resolution probabilistic design events capturing the.pdf, International Journal of Climatology, 43, 761–771, https://doi.org/10.1002/joc.7825, 2023.

Kitzmiller, D. H. , W. W. , Z. Z. , P. N. , & T. X.: The Analysis of Record for Calibration: A High- Resolution Precipitation and Surface Weather Dataset for the United States, in: AGU Fall Meeting, 2018.

Kupfer, S., MacPherson, L. R., Hinkel, J., Arns, A., and Vafeidis, A. T.: A Comprehensive Probabilistic Flood Assessment Accounting for Hydrograph Variability of ESL Events, J Geophys Res Oceans, 129, https://doi.org/10.1029/2023JC019886, 2024.

Leijnse, T., van Ormondt, M., Nederhoff, K., and van Dongeren, A.: Modeling compound flooding in coastal systems using a computationally efficient reduced-physics solver: Including fluvial, pluvial, tidal, wind- and wave-driven processes, Coastal Engineering, 163, https://doi.org/10.1016/j.coastaleng.2020.103796, 2021.

Lucey, J. T. D. and Gallien, T. W.: Quantifying compound flood event uncertainties in a wave and tidally dominated coastal region: The impacts of copula selection, sampling, record length, and precipitation gauge selection, J Flood Risk Manag, 17, https://doi.org/10.1111/jfr3.12984, 2024.

705 Maduwantha, P., Wahl, T., Santamaria-Aguilar, S., Jane, R. A., Booth, J. F., Kim, H., and Villarini, G.: A multivariate statistical framework for mixed populations in compound flood analysis, Natural Hazards and Earth System Science, https://doi.org/10.5194/egusphere-2024-1122, 2024.

Maduwantha, P., Wahl, T., Santamaria-Aguilar, S., Jane, R., Dangendorf, S., Kim, H., and Villarini, G.: Generating Boundary Conditions for Compound Flood Modeling in a Probabilistic Framework, https://doi.org/10.5194/egusphere-2025-1557, 17

710 April 2025.

McKeon, K. and Piecuch, C. G.: Compound Minor Floods and the Role of Discharge in the Delaware River Estuary, J Geophys Res Oceans, 130, https://doi.org/10.1029/2024JC021716, 2025.

Merz, B., Blöschl, G., Jüpner, R., Kreibich, H., Schröter, K., and Vorogushyn, S.: Invited perspectives: Safeguarding the usability and credibility of flood hazard and risk assessments, EGUsphere, Preprint, 1–24, https://doi.org/10.5194/egusphere-

715 2024-856, 2024.

Moftakhari, H., Schubert, J. E., AghaKouchak, A., Matthew, R. A., and Sanders, B. F.: Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries, Adv Water Resour, 128, 28–38, https://doi.org/10.1016/j.advwatres.2019.04.009, 2019.

Molinari, D., De Bruijn, K. M., Castillo-Rodríguez, J. T., Aronica, G. T., and Bouwer, L. M.: Validation of flood risk models:

720 Current practice and possible improvements, International Journal of Disaster Risk Reduction, 33, 441–448, https://doi.org/10.1016/j.ijdrr.2018.10.022, 2019.

National Academies of Sciences, E. and M.: Framing the challenge of urban flooding in the United States, The National Academies Press, Washington DC, 89 pp., https://doi.org/10.17226/25381, 2019.

(NCEI), N. N. C. for E. I.: U.S. Billion-Dollar Weather and Climate Disasters, https://doi.org/10.25921/stkw-7w73, 2024.

725 Nederhoff, K., Leijnse, T. W. B., Parker, K., Thomas, J., O'Neill, A., van Ormondt, M., McCall, R., Erikson, L., Barnard, P. L., Foxgrover, A., Klessens, W., Nadal-Caraballo, N. C., and Massey, T. C.: Tropical or extratropical cyclones: what drives the compound flood hazard, impact, and risk for the United States Southeast Atlantic coast?, Natural Hazards, 120, 8779–8825, https://doi.org/10.1007/s11069-024-06552-x, 2024.

Peña, F., Obeysekera, J., Jane, R., Nardi, F., Maran, C., Cadogan, A., de Groen, F., and Melesse, A.: Investigating compound

730 flooding in a low elevation coastal karst environment using multivariate statistical and 2D hydrodynamic modeling, Weather Clim Extrem, 39, https://doi.org/10.1016/j.wace.2022.100534, 2023.

Perez, G., Coon, E. T., Rathore, S. S., and Le, P. V. V.: Advancing process-based flood frequency analysis for assessing flood hazard and population flood exposure, J Hydrol (Amst), 639, 131620, https://doi.org/10.1016/j.jhydrol.2024.131620, 2024.

Pollack, A. B., Santamaria-aguilar, S., Maduwantha, P., Helgeson, C., Wahl, T., and Keller, K.: Funding rules that promote

735 equity in climate adaptation outcomes, PNAS, 2024.

Ponte, R. M., Carson, M., Cirano, M., Domingues, C. M., Jevrejeva, S., Marcos, M., Mitchum, G., van de Wal, R. S. W., Woodworth, P. L., Ablain, M., Ardhuin, F., Ballu, V., Becker, M., Benveniste, J., Birol, F., Bradshaw, E., Cazenave, A., De Mey-Frémaux, P., Durand, F., Ezer, T., Fu, L. L., Fukumori, I., Gordon, K., Gravelle, M., Griffies, S. M., Han, W., Hibbert, A., Hughes, C. W., Idier, D., Kourafalou, V. H., Little, C. M., Matthews, A., Melet, A., Merrifield, M., Meyssignac, B., Minobe, S., Penduff, T., Picot, N., Piecuch, C., Ray, R. D., Rickards, L., Santamaría-Gómez, A., Stammer, D., Staneva, J., Testut, L., Thompson, K., Thompson, P., Vignudelli, S., Williams, J., Simon, S. D., Wöppelmann, G., Zanna, L., and Zhang, X.: Towards comprehensive observing and modeling systems for monitoring and predicting regional to coastal sea level, Front Mar Sci, 6, 1–25, https://doi.org/10.3389/fmars.2019.00437, 2019.

Quinn, N., Lewis, M., Wadey, M. P., and Haigh, I. D. D.: Assessing the temporal variability in extreme storm-tide time series for coastal flood risk assessment, J Geophys Res Oceans, 119, 4983–4998, https://doi.org/10.1002/2014JC010197, 2014.

Salvadori, G., De Michele, C., and Durante, F.: On the return period and design in a multivariate framework, Hydrol Earth Syst Sci, 15, 3293–3305, https://doi.org/10.5194/hess-15-3293-2011, 2011.

Santamaria-Aguilar, S., Arns, A., and Vafeidis, A. T. A. T.: Sea-level rise impacts on the temporal and spatial variability of extreme water levels: A case study for St. Peter-Ording, Germany, J Geophys Res Oceans, 122, 1–22, https://doi.org/10.1002/2016JC012579, 2017.

Sebastian, A., Bader, D. J., Nederhoff, C. M., Leijnse, T. W. B., Bricker, J. D., and Aarninkhof, S. G. J.: Hindcast of pluvial, fluvial, and coastal flood damage in Houston, Texas during Hurricane Harvey (2017) using SFINCS, Natural Hazards, 109, 2343–2362, https://doi.org/10.1007/s11069-021-04922-3, 2021.

Smith, C.: Regional Flooding Study. Phase 1 Final Report, 2023.

Taherkhani, M., Vitousek, S., Barnard, P. L., Frazer, N., Anderson, T. R., and Fletcher, C. H.: Sea-level rise exponentially increases coastal flood frequency, Sci Rep, 10, https://doi.org/10.1038/s41598-020-62188-4, 2020.

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., and Slayback, D. A.: Satellite imaging reveals increased proportion of population exposed to floods, Nature, 596, 80–86, https://doi.org/10.1038/s41586-021-03695-w, 2021.

Thompson, P. R., Widlansky, M. J., Hamlington, B. D., Merrifield, M. A., Marra, J. J., Mitchum, G. T., and Sweet, W.: Rapid increases and extreme months in projections of United States high-tide flooding, Nat Clim Chang, 11, https://doi.org/10.1038/s41558-021-01077-8, 2021.

USACE: HEC-RAS River Analysis System Creating Land Cover , Manning ' s N Values , And % Impervious, 2021.

Vousdoukas, M. I., Voukouvalas, E., Mentaschi, L., Dottori, F., Giardino, A., Bouziotas, D., Bianchi, A., Salamon, P., and Feyen, L.: Developments in large-scale coastal flood hazard mapping, Natural Hazards and Earth System Sciences, 16, 1841–1853, https://doi.org/10.5194/nhess-16-1841-2016, 2016.

Wadey, M. P., Cope, S. N., Nicholls, R. J., McHugh, K., Grewcock, G., and Mason, T.: Coastal flood analysis and visualisation for a small town, Ocean Coast Manag, 116, 237–247, https://doi.org/10.1016/j.ocecoaman.2015.07.028, 2015.

Wahl, T., Jain, S., Bender, J., Meyers, S. D., and Luther, M. E.: Increasing risk of compound flooding from storm surge and rainfall for major US cities, Nat Clim Chang, 5, 1093–1097, https://doi.org/10.1038/nclimate2736, 2015.

Wang, R. Q., Mao, H., Wang, Y., Rae, C., and Shaw, W.: Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data, Comput Geosci, 111, 139–147, https://doi.org/10.1016/j.cageo.2017.11.008, 2018.

Wing, O. E. J., Bates, P. D., Quinn, N. D., Savage, J. T. S., Uhe, P. F., Cooper, A., Collings, T. P., Addor, N., Lord, N. S., Hatchard, S., Hoch, J. M., Bates, J., Probyn, I., Himsworth, S., Rodríguez González, J., Brine, M. P., Wilkinson, H., Sampson, C. C., Smith, A. M., Neal, J. C., and Haigh, I. D.: A 30 m Global Flood Inundation Model for Any Climate Scenario, Water Resour Res, 60, https://doi.org/10.1029/2023WR036460, 2024.

Winter, B., Winter, B., Schneeberger, K., Schneeberger, K., Förster, K., and Vorogushyn, S.: Event generation for probabilistic flood risk modelling: Multi-site peak flow dependence model vs. weather-generator-based approach, Natural Hazards and Earth System Sciences, 20, 1689–1703, https://doi.org/10.5194/nhess-20-1689-2020, 2020.