Nat. Hazards Earth Syst. Sci., 25, 1353–1375, 2025
https://doi.org/10.5194/nhess-25-1353-2025

Natural Hazards and Earth System Sciences
Open Access
EGU

Accelerating compound flood risk assessments through active learning: A case study of Charleston County (USA)

Lucas Terlinden-Ruhl[1,2], Anaïs Couasnon[2,3], Dirk Eilander[2,3], Gijs G. Hendrickx[1], Patricia Mares-Nasarre[1], and José A. Á. Antolínez[1]

[1]Department of Hydraulic Engineering, Delft University of Technology, Delft, the Netherlands
[2]Department of Inland and Water Systems, Deltares, Delft, the Netherlands
[3]Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

Correspondence: Lucas Terlinden-Ruhl (lucas.terlindenruhl@gmail.com)

Received: 10 October 2024 – Discussion started: 14 October 2024
Revised: 26 January 2025 – Accepted: 10 February 2025 – Published: 9 April 2025

Abstract. Flooding is the natural hazard most likely to affect individuals and can be driven by rainfall, river discharge, storm surge, tides, and waves. Compound floods result from their co-occurrence and can generate a larger flood hazard when compared to the synthetic flood hazard generated by the respective flood drivers occurring in isolation from one another. Current state-of-the-art stochastic compound flood risk assessments are based on statistical, hydrodynamic, and impact simulations. However, the stochastic nature of some key variables in the flooding process is often not accounted for as adding stochastic variables exponentially increases the computational costs (i.e., the curse of dimensionality). These simplifications (e.g., a constant flood driver duration or a constant time lag between flood drivers) may lead to a misquantification of the flood risk. This study develops a conceptual framework that allows for a better representation of compound flood risk while limiting the increase in the overall computational time. After generating synthetic events from a statistical model fitted to the selected flood drivers, the proposed framework applies a treed Gaussian process (TGP). A TGP uses active learning to explore the uncertainty associated with the response of damages to synthetic events. Thereby, it informs regarding the best choice of hydrodynamic and impact simulations to run to reduce uncertainty in the damages. Once the TGP predicts the damage of all synthetic events within a tolerated uncertainty range, the flood risk is calculated. As a proof of concept, the proposed framework was applied to the case study of Charleston County (South Carolina, USA) and compared with a state-of-the-art stochastic compound flood risk model, which used equidistant sampling with linear scatter interpolation. The proposed framework decreased the overall computational time by a factor of 4 and decreased the root mean square error in damages by a factor of 8. With a reduction in overall computational time and errors, additional stochastic variables such as the drivers' duration and time lag were included in the compound flood risk assessment. Not accounting for these resulted in an underestimation of 11.6 % (USD 25.47 million) in the expected annual damage (EAD). Thus, by accelerating compound flood risk assessments with active learning, the framework presented here allows for more comprehensive assessments as it loosens constraints imposed by the curse of dimensionality.

1 Introduction

Flooding has been identified as the natural hazard most likely to affect individuals (UNDRR, 2020). Moreover, climate change is expected to increase the magnitude and frequency of extreme water levels (e.g., Hirabayashi et al., 2013; Blöschl, 2022), which are driven by precipitation, river discharge, surge, tide, and waves (e.g., Couasnon et al., 2020; Hendry et al., 2019; Parker et al., 2023; Ward et al., 2018). Low-lying coastal areas are especially susceptible to compound events of these drivers, which enhance the flood hazard (Wahl et al., 2015). In addition, migration patterns are causing an increase in assets and people in these areas (e.g., Swain et al., 2020; Neumann et al., 2015). Governing bodies tackle the challenge of compound flooding using the con-

cept of stochastic flood risk (e.g., Klijn et al., 2015; Muis et al., 2015). This concept takes into account the three following metrics: (1) flood hazard, which is the intensity and frequency of a flood event; (2) exposure, which is the assets and/or people susceptible to flooding; and (3) vulnerability, which is the economic and/or social consequences of exposed elements as a result of a flood hazard (e.g., Klijn et al., 2015; Koks et al., 2015). The risk associated with flooding can be diminished through the development of resilient infrastructure (e.g., Jongman, 2018; Woodward et al., 2014).

To develop such infrastructure under unknown future scenarios, quantifying the risk from compound flooding requires an accurate quantification from the interactions between the many flood drivers impacting the flood hazard (e.g., Bates et al., 2023; Woodward et al., 2013; Barnard et al., 2019). Currently, state-of-the-art stochastic compound flood risk assessments usually apply the following four steps to quantify risk (e.g., Wyncoll and Gouldby, 2013; Couasnon et al., 2022; Rueda et al., 2015). Firstly, the joint probability distribution between the selected flood drivers is modeled based on the observed dependence and is used to generate synthetic events. Secondly, the flood hazard is modeled using a hydrodynamic model to account for the non-linear interactions of the flood drivers. Thirdly, the damage is modeled by combining the flood hazard with information on exposure and vulnerability. Lastly, the risk is modeled by accounting for the probability and damages associated with the flood hazard. To obtain an accurate probability distribution of damages, many events must be generated in the first step. To this end, a brute force Monte Carlo simulation (MCS) of the selected flood drivers can be applied (e.g., Wu et al., 2021; Winter et al., 2020). This can model the joint probability distribution of the different stochastic variables that define the time record of the flood drivers. While this minimizes the number of simplifications, it requires a large number of simulations from hydrodynamic models to quantify the flood hazard of each event, which can be computationally infeasible (e.g., Eilander et al., 2023b; Rueda et al., 2015).

Consequently, compound flood risk assessments have focused on reducing the computational time of performing hydrodynamic simulations while ensuring the risk estimate is accurate. Examples of strategies in the literature include the following: (1) improving computational resources (e.g., Apel et al., 2016), (2) using faster reduced-physics hydrodynamic models (e.g., Bates et al., 2010; Leijnse et al., 2021), (3) reducing the number of hydrodynamic simulations through various sampling techniques (e.g., Moftakhari et al., 2019; Barnard et al., 2019; Diermanse et al., 2014; Bakker et al., 2022), and (4) replacing hydrodynamic simulations with data-driven (i.e., regression) models (e.g., Moradian et al., 2024; Fraehr et al., 2024). The above examples are not mutually exclusive (e.g., Eilander et al., 2023b; Gouldby et al., 2017). Nonetheless, the largest reduction in computational time can be expected by focusing on the last two strategies (e.g., Rueda et al., 2015), which generate a surrogate model by combining sampling and regression techniques to obtain an estimate of the damages from all events in the MCS.

State-of-the-art surrogate models select simulations a priori by performing equidistant/factorial sampling for the events contained in the MCS (e.g., Jane et al., 2022; Eilander et al., 2023b; Gouldby et al., 2017; Rueda et al., 2015). However, compound flood events are characterized by many stochastic variables, such as the flood driver magnitude and duration, and the time lag between drivers. Therefore, a priori sampling requires a large number of hydrodynamic and impact simulations to provide a robust quantification of the damages of non-simulated events in the MCS. This often results in simplifications (e.g., Diermanse et al., 2023; Eilander et al., 2023b; Couasnon et al., 2022; Jane et al., 2022; Rueda et al., 2015). Examples of these include the following: (1) flood drivers are omitted; (2) stochastic variables such as the duration and the time lag are taken as constants; and (3) interpolation techniques are used for regression, as they minimize the computational time involved in training a data-driven model, for which ensuring it generalizes well to unseen locations or forcing conditions can be computationally expensive (e.g., Fraehr et al., 2024; Moradian et al., 2024). These series of simplifications may impact the distribution of flood damages and result in a mis-quantification of compound flood risk.

This mis-quantification can be minimized by accelerating the process of obtaining a robust surrogate model. This can be potentially achieved by using active learning, which guides the sampling technique by optimizing towards a goal. In the context of compound flood risk, a suitable goal could be to minimize the uncertainty related to the damages from the previous hydrodynamic and impact simulations. This would allow for hydrodynamic and impact simulations to be selected a posteriori. Treed Gaussian process (TGP) models can make use of active learning and have been shown to reduce the number of hydrodynamic simulations associated with high-dimensional datasets (e.g., Hendrickx et al., 2023, did so for modeling salt intrusion). Moreover, TGP models can also provide a reasonable regression model with limited training data (Gramacy and Lee, 2009). To the authors' knowledge, active learning is not used in state-of-the-art stochastic compound flood risk frameworks but has successfully been used in other types of risk assessments (Tomar and Burton, 2021).

Therefore, this study aims to explore active learning to improve the quantification of compound flood risk assessments while limiting the increase in overall computational time. To this end, a new conceptual framework based on the TGP model is proposed, which (1) leverages the uncertainty in the response of damages to flood drivers to minimize the number of required hydrodynamic and impact simulations and (2) can account for more stochastic variables in compound flood risk assessments. Therefore, this framework results in a more robust and comprehensive characterization of com-

pound flood risk. As a proof of concept, the framework is applied to the case study of Charleston County in South Carolina (USA).

2 Methods

Our framework for compound flood risk assessments that utilize active learning uses the five following general steps, which can be visualized as follows (Fig. 1; e.g., Eilander et al., 2023b; Rueda et al., 2015).

1. Based on the characteristics of the case study, parameterize the selected flood drivers (Sect. 2.1).
2. Infer the natural variability of compound flood drivers to generate stochastic event sets (Sect. 2.2).
3. Simulate the damages associated with a synthetic event (Sect. 2.3).
4. Use a surrogate model to select simulations (synthetic events) with active learning and model the input-to-output (i.e., flood driver parameters to damages associated with a geographic location) relationship associated with a stochastic event set (Sect. 2.4).
5. Model the risk by combining information on the probability and the damages of synthetic events to obtain an estimate of the expected annual damage (EAD) (Sect. 2.5).

The main difference between the proposed framework and the current state-of-the-art lies in the fourth step (Sect. 2.4). Therefore, an experiment was designed in Sect. 2.6 to compare the active learning-based framework with a state-of-the-art one based on equidistant sampling and a linear scatter interpolation (e.g., Eilander et al., 2023b; Jane et al., 2022).

2.1 Case study, flood drivers, and data

Charleston County, located in South Carolina (USA), is on the coast of the Atlantic Ocean. Figure 2 shows Charleston County subdivided by sub-county (United States Census Bureau, 2024). According to their distance to the open coast, the sub-counties can be classed as either "inland" or "coastal" (Fig. 2).

Charleston County is prone to compound flooding caused by hurricanes and extratropical events (e.g., Parker et al., 2023; Nederhoff et al., 2024). These events cause precipitation, and wind shear and pressure effects offshore, which can result in the co-occurrence of astronomical tides, runoff, storm surges, and waves (e.g., Barnard et al., 2023), which regularly causes damages (e.g., Samadi and Lunt, 2023). Moreover, sea level rise (SLR) is expected to worsen the flood hazards in the future (e.g., Morris and Renken, 2020). This has caused the city to create a plan to manage its infrastructure (City of Charleston, 2015).

Since validating a hydrodynamic model is not the purpose of this article, the hydrodynamic SFINCS model (Leijnse et al., 2021) validated by Diermanse et al. (2023) for Charleston County was used (see Sect. 2.3). The small offshore model domain meant waves could not be taken into account (Diermanse et al., 2023). Moreover, waves are not a dominant driver for Charleston County (e.g., Parker et al., 2023). Thus, the same flood drivers as Diermanse et al. (2023) were investigated here. These are the storm surge, tides, and precipitation.

Data for the still water level and the tides in Charleston Harbor were obtained at an hourly resolution from the National Oceanic and Atmospheric Administration (NOAA), at the tide gauge location shown in Fig. 2 (station ID: 8665530). The still water level time record contained the tidal, sea level rise, and non-tidal residual components. The tidal time record was obtained from NOAA's stationary harmonic analysis. The non-tidal residual component was assumed to be equivalent to the storm surge. The storm surge time record was calculated by subtracting the tidal time record from the still water level time record. Data for precipitation were obtained at an hourly resolution from the ERA5 reanalysis dataset (Hersbach et al., 2020) at the grid location of 32.75° N, 79.75° W (see Fig. 2). The ERA5 dataset has a spatial resolution of 0.25°, roughly equivalent to 30 km. The time record for the storm surge, tides, and precipitation had an overlapping time record of 24 years and 4 months.

Increased still water levels from storm surges induce a phase shift in the tidal signal, creating spurious peaks in the storm surge time record (Williams et al., 2016). Therefore, the skew surge was considered. It is the difference between the highest still water level and the high tide within a tidal period (Williams et al., 2016; Couasnon et al., 2022; Diermanse et al., 2023). A tidal period was taken as the time between two consecutive low tides. To recreate a time record with an hourly resolution, the skew surge values were assumed to be constant over their tidal period, which was approximately 12 h for Charleston Harbor. Furthermore, to remove the sea level rise component, the 1-year moving average of the skew surge time record was subtracted from the skew surge time record (Arns et al., 2013). The sea level rise component was also used to identify the current sea level as 0.2 m above mean sea level (a.m.s.l.).

Based on skew surge, tides, and precipitation, six stochastic variables were selected to parameterize compound flood events in Charleston County: skew surge magnitude (*S.Mag*), precipitation magnitude (*P.Mag*), tidal magnitude (*T.Mag*), precipitation duration (*P.Dur*), skew surge duration (*S.Dur*), and precipitation lag (*P.Lag*).

2.2 Inference of natural variability

Describing and inferring natural variability is key in risk management. Due to the relatively short time span of measured and/or reanalysis data, statistical models are often used

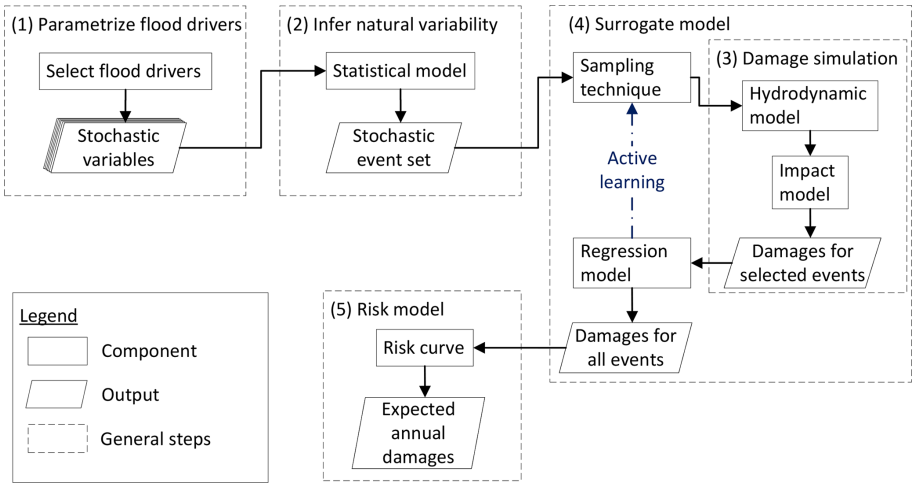

Figure 1. Typical framework required to characterize the flood risk associated with compound floods when using a surrogate model and our implementation for this study.

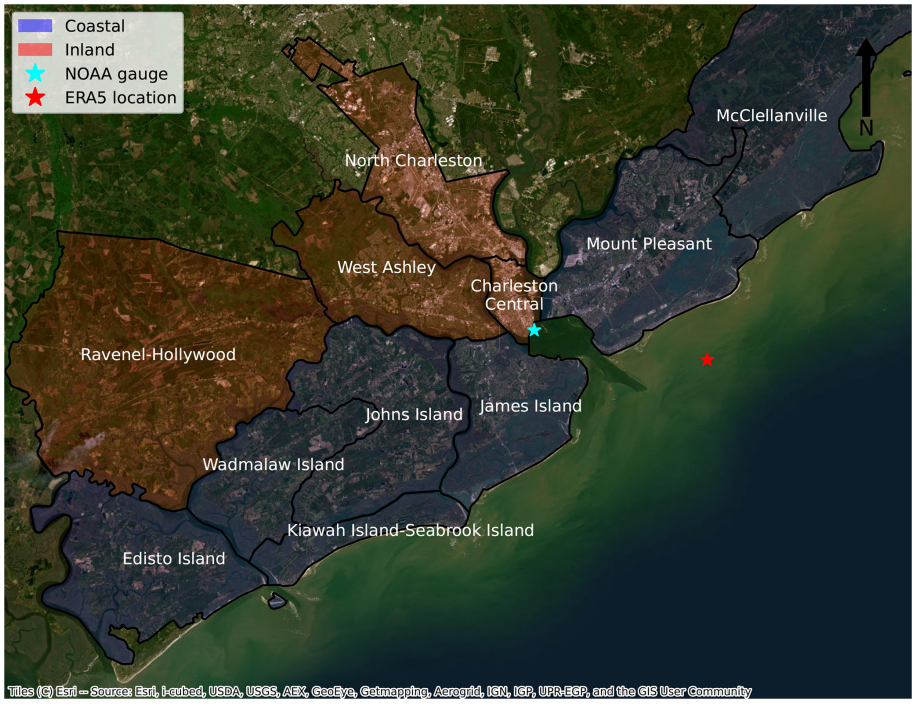

Figure 2. Charleston County with sub-county divisions that are color-coded according to their proximity to the open coast. Markers indicate the location of where data on the drivers were collected.

to stochastically generate a large number of synthetic compound events (e.g., Couasnon et al., 2022; Bevacqua et al., 2017; Bates et al., 2023). The statistical model presented in the proposed framework follows the following four steps.

1. Identify extreme high-water events (Sect. 2.2.1).
2. Quantify flood driver parameters (Sect. 2.2.2).
3. Model the joint probability distribution (Sect. 2.2.3).
4. Generate stochastic event sets (Sect. 2.2.4).

2.2.1 Identify extreme high-water events

Charleston Harbor's tidal record showed daily inequalities larger than semi-diurnal differences in *S.Mag*s. Therefore, applying peak over threshold (POT) to all *S.Mag*s could have identified high-water events that were not extreme as they could co-occur with lower high *T.Mag*s. Therefore, to only identify extreme high-water events, POT was only applied to

S.Mag when it co-occurred with a higher high (HH) *T.Mag*. POT identifies extremes as exceedances over a threshold and uses a declustering time window to ensure all extremes are independent and identically distributed. Therefore, the number of identified extremes is dependent on the threshold and declustering time window. A threshold of 0.32 m relative to mean sea level (MSL) and a declustering time window of 14 d between each extreme *S.Mag* were chosen. The latter was based on the longest *S.Dur* before restricting the duration of events (see Sect. 2.2.2). This ensured consecutive extreme events were not embedded in events longer than the declustering time window. This resulted in 2.91 extreme high-water events per year.

2.2.2 Quantify flood driver parameters

All six flood driver parameters had to be quantified for each extreme high-water event. POT was applied on *S.Mag* in Sect. 2.2.1. For *P.Mag*, the largest value that co-occurred within ±3 d of all identified *S.Mag* extremes was used. For *T.Mag*, the co-occurring HH tide with all identified *S.Mag* extremes were used. *P.Dur* and *S.Dur* were taken as the duration of *P.Mag* and *S.Mag* to continuously remain above a critical value within the ±3 d window used to quantify *P.Mag*. For precipitation and skew surge, the values used to define the duration were 0.3 $mm\,h^{-1}$ and 0.2 m, respectively. *P.Lag* was defined as the difference in hours between *S.Mag* and *P.Mag* for each extreme high-water event.

2.2.3 Model the joint probability distribution

The joint probability distribution between the different stochastic variables was modeled using a vine copula because of the method's flexibility in high dimensional datasets (Czado and Nagler, 2022; Bedford and Cooke, 2002) and successful applications in other compound flooding studies (e.g., Bevacqua et al., 2017; Eilander et al., 2023b). A vine copula is defined by three components (e.g., Czado, 2019): (1) bivariate copulas; (2) a graph, named regular vine, which composed of a series of nested trees; and (3) marginal cumulative distribution functions (CDFs). A vine copula constructs a multivariate distribution using bivariate copulas. A bivariate copula models the dependence between two stochastic variables in the normalized ranked space. A vine copula organizes the bivariate copulas into a series of trees. The first tree represents the unconditional dependence between the stochastic variables. The following trees add a layer of conditional dependence. To transform the observations and generated data to and from the normalized ranked space where the vine copula is defined, marginal CDFs are required.

To select and fit a regular vine for a given problem, two options are possible: (1) brute force (Morales-Nápoles et al., 2023) or (2) heuristic algorithms. It has been shown that the number of regular vines grows extremely fast with the number of stochastic variables (Morales-Nápoles, 2010). Therefore, in this study the regular vine was chosen using Dißmann's algorithm (Dißmann et al., 2013) as implemented in the *pyvinecopulib* Python package (Nagler and Vatter, 2023) and fitted to minimize the Bayesian information criterion (BIC; Schwarz, 1978).

To mimic different levels of simplifications typically used in compound flood risk assessments, multiple (vine) copula models were fitted, each one considering different numbers of stochastic variables. The starting point was the copula model fitted between *S.Mag* and *P.Mag*. Then, *T.Mag*, *P.Dur*, *S.Dur*, and *P.Lag* were added one at a time. This resulted in one copula and four vine copulas fitted to two, three, four, five, and six stochastic variables, respectively. To simplify the models, if the stochastic variable that was added had independent copulas between all the pairs, it was removed from the vine copula. This simplified the number of models from five to three. Table A1 located in Appendix A shows an overview of the models.

Marginal CDFs were defined for each stochastic variable. For *S.Mag*, both the exponential and generalized Pareto distribution (GPD) were fitted to the data using the L-moments method and the best fit was selected based on the BIC using the *HydroMT* Python package (Eilander et al., 2023a). For *T.Mag*, the empirical CDF of all HH tides was used as *T.Mag* is expected to be independent of *S.Mag* (e.g., Williams et al., 2016). For *P.Mag*, 80 continuous distributions available in the *scipy* Python package (Virtanen et al., 2020) were fitted using maximum likelihood estimates, and the best fit was selected based on the BIC. For *P.Dur*, *S.Dur*, and *P.Lag*, extrapolation is not desired, so only the truncated distributions available in the *scipy* Python package (Virtanen et al., 2020) were considered. The smallest sum squared error (SSE) was used to choose the models. For *P.Mag*, *P.Dur*, *S.Dur*, and *P.Lag*, the *fitter* Python package was used (Cokelaer et al., 2024) to apply the exposed methodology. Table A2 located in Appendix A summarizes the CDFs and respective parameters chosen for the different stochastic variables.

2.2.4 Generate stochastic event sets

The fitted vine copula models were used to generate stochastic event sets to represent the different extents of simplifications in compound flood risk assessments. Stochastic event sets were generated with the inverse Rosenblatt transform in the *pyvinecopulib* Python package (Nagler and Vatter, 2023). When a variable was not yet stochastic, a constant was defined. This constant was assumed to be the median value from the empirical distribution. Moreover, in some cases, the incremental addition of a stochastic variable resulted in this variable only being contained in statistically independent bivariate copulas. To simplify, this variable was removed from the vine copula model, which resulted in the vine copula model only modeling statistically dependent variables. Thus, to obtain a stochastic event set that included a statistically in-

dependent variable, the stochastic event set generated by the simplified vine copula model was combined with data generated from the marginal of the independent variable.

Two different types of stochastic event sets were generated. For the first type, two "benchmark" event sets with 500 synthetic events with two and six stochastic variables were generated. For these event sets, the damage of each event was simulated (Sect. 2.3). For the second type, five "testing" event sets with 10 000 synthetic events with two, three, four, five, and six stochastic variables were generated.

2.3 Damage simulation

The damage related to synthetic events is required to train a surrogate model. This requires: (1) the simulation of the flood hazard and (2) the simulation of the damages associated with a flood hazard. This study uses a hydrodynamic and impact model.

Many fully physics-solving and reduced-physics hydrodynamic models are available (e.g., Delft-3D, Deltares (2022); MIKE 21, DHI (2017); HEC-RAS, USACE Hydrologic Engineering Center (2025); LISFLOOD-FP, Bates et al. (2010)). The SFINCS (Leijnse et al., 2021) model was used to estimate the flood hazard map associated with the boundary conditions of a synthetic event. SFINCS was chosen because (1) a model was validated for Charleston County; (2) the model uses high-resolution local datasets; and (3) the computational grid is small, reducing the computational time (Diermanse et al., 2023). SFINCS is a reduced-physics hydrodynamic model optimized for the fast calculation of the flood hazard (Leijnse et al., 2021). The governing equations are based on the local inertia equations (Bates et al., 2010), tuned for coastal and compound flooding by including additional terms, such as wind stress and advection. These equations are solved at the resolution of the computational grid using subgrid information about the topography and conveyance capacity (van Ormondt et al., 2025). For more details on SFINCS, see Leijnse et al. (2021) and van Ormondt et al. (2025). For Charleston County, the SFINCS model had a 200×200 m grid resolution. The native 1×1 m resolution information for the topo bathymetry and land roughness were included with a subgrid lookup table. The topo bathymetry data were based on the Coastal National Elevation Database (CoNED; Danielson et al., 2016; Cushing et al., 2022). For the spatially varying land roughness, the National Land Cover Database (NCLD; Homer et al., 2020) was used and reclassified to Manning roughness values following Nederhoff et al. (2024). Drainage was handled with (1) pumps located in the Charleston Central sub-county (Diermanse et al., 2023), and (2) the curve number infiltration scheme, which was based on the United States General Soil Map (STATSGO2; U.S. Department of Agriculture, 2020) following Nederhoff et al. (2024). For our application, only two boundary conditions were required: (1) the still water level at the coast and (2) the precipitation. The time series for the still water level was reconstructed by linearly superposing three components: the constant MSL equal to 0.2 m (Sect. 2.1), a historical tidal time series from the HH tide empirical distribution associated with a given *T.Mag*, and the skew surge time series. A Gaussian distribution was used to reconstruct the time series for skew surge and precipitation. For skew surge, *S.Mag* and *S.Dur* were used. For precipitation, *P.Mag*, *P.Dur*, and *P.Lag* were used. Figure A1 in Appendix A shows how these variables were combined to create the time series of boundary conditions for the downstream water level and precipitation. In terms of spatial distribution, both the downstream water level and precipitation were spatially uniform.

Impact models combine hazard, exposure, and vulnerability to quantify the economic and/or social consequences of a flooding event (e.g., Bates et al., 2023). The Delft-FIAT (Deltares, 2024) model was used to compute the damages associated with a synthetic event. Delft-FIAT was chosen because (1) a model was validated for Charleston County and (2) it uses data from the United States Army Corps of Engineers and Federal Emergency Management Agency (Diermanse et al., 2023). Delft-FIAT combines the hazard map obtained from the SFINCS model, with the exposure (maximum damages of a building footprint), and the vulnerability (depth–damage fraction curves for each building footprint). This allowed for the computation of the damages associated with each building footprint included in the model for a synthetic event. Summing all building footprints in the model results in economic damage for Charleston County. Here, the damages associated with different geographic locations (referred to as outputs hereinafter) were also investigated. This was done per sub-county and according to the classifications (i.e., coastal vs. inland) made in Sect. 2.1. These different spatial scales are referred to as the complete, sub-county, and classified models, respectively.

2.4 Surrogate model

A surrogate model approximates the behavior of a more complex and computationally expensive model. By being computationally faster, they are an asset in compound flood risk assessments for various reasons (e.g., quantifying uncertainty and testing risk reduction measures in Eilander et al., 2023b). The development of a surrogate model requires (1) a sampling technique that selects a subset of events and configurations out of a large multivariate space, (2) running this subset in the reference numerical model, and (3) training a regression model to obtain the outcome for any other possible event or configuration. Previous compound flood risk studies only use surrogate models that (1) sample based on input parameters (e.g., Jane et al., 2022) and (2) are commonly combined with linear interpolation (e.g., Couasnon et al., 2022), although more complex regression techniques are available (e.g., neural networks in Hendrickx et al., 2023, and radial basis functions in Antolínez et al., 2019). However, active

learning can be used to minimize the number of simulations by leveraging the uncertainty in the output (e.g., Tomar and Burton, 2021; Hendrickx et al., 2023). Here, the same sampling technique as in Hendrickx et al. (2023) was applied, which builds upon the work of Gramacy and Lee (2009). The tool used to perform active learning was also used as a regression model. Section 2.6 shows how these surrogate models were used and compared.

A treed Gaussian process, limiting linear model (TGP-LLM; Gramacy and Lee, 2009) was used. It is a generative model able to provide different estimates for the damages. These estimates are used to calculate a mean and a confidence interval for the damages from each possible synthetic event from a stochastic event set. The mean is used as the regression model (Gramacy and Lee, 2009), while the confidence interval can be used as a metric to drive a sampling technique (e.g., Hendrickx et al., 2023). This is done by choosing the simulation with the largest standard deviation as it is expected to bring the largest gain in information (MacKay, 1992). From here on, the standard deviation will be referred to as the active learning Mackay (ALM; MacKay, 1992) statistic.

A TGP-LLM uses Bayesian tree regression to partition the input into different subdomains, allowing different Gaussian processes or linear models to fit different regions in the input space. This allows the TGP-LLM to be non-stationary and account for heteroskedasticity (Gramacy and Lee, 2009), preventing the magnitude of the ALM statistic from propagating from one subdomain to another. This could result in large ALM statistics in uninteresting areas (Gramacy and Lee, 2009). Therefore, a TGP-LLM enables the selection of simulations to only occur in feature-rich sub-domains.

To minimize the number of simulations, performing active learning with the TGP-LLM model is beneficial. The conceptual framework, therefore, repeats three steps: (1) fit a TGP-LLM to the damages associated with the subset; (2) based on the highest ALM statistic, select a simulation to perform from the stochastic event set; and (3) perform the simulation to obtain the damages and append to the subset. The TGP-LLM has a computational cost that is proportional to the number of simulations currently in the subset (N): $O(N^3)$ (Gramacy and Lee, 2009).

When provided with a small subset of simulations, the TGP-LLM may sample randomly, providing a small improvement in information for the computational cost (Hendrickx et al., 2023). To this end, we initially used a maximum dissimilarity algorithm (MDA; Kennard and Stone, 1969) as it selects simulations at the outskirts of a stochastic event set (Camus et al., 2011). Given an initial subset of simulated events, and a dissimilarity measure, an MDA repeats the following two steps until the final subset contains a predetermined number of simulated events. Firstly, given the current subset of simulated events, it assigns non-simulated events to the simulated event that it is the least dissimilar to. Secondly, based on this assignment, the non-simulated event with the largest remaining dissimilarity to a simulated event is added to the subset and is simulated.

For our implementation, to prevent bias towards a variable, the stochastic event set was first normalized using a min–max scaler from the *scikit-learn* Python package (Pedregosa et al., 2011). The MDA was initialized by providing the synthetic event associated with the largest *S.Mag*. Euclidean distance was used as a metric of dissimilarity. We used the MDA until 2^d (where d represents the number of dimensions) events were contained in the subset, as it was proportional to the number of vertices. After the MDA simulations, active learning with the TGP-LLM was performed, as proposed by Hendrickx et al. (2023).

To further minimize the computational cost, we developed our own stopping criterion as this is still an active area of research (e.g., Ishibashi and Hino, 2021; Tomar and Burton, 2021; Hendrickx et al., 2023). The learning curve of the TGP-LLM was assessed by comparing its mean with the benchmark (Sect. 2.2.4) after each TGP-LLM iteration. To this end, different metrics were computed: the root mean square error (RMSE) of the simulated and non-simulated events, the mean and maximum ALM statistic of the non-simulated events, and the EAD (Sect. 2.5). A two-sample Kolmogorov–Smirnov (KS) test (Sect. 2.6) was applied to compare the empirical CDFs of the benchmark and the TGP-LLM.

The stopping criterion was defined as an ALM mean smaller than 0.1 for two consecutive TGP-LLM models for an output. This is based on our findings that (1) the mean ALM is correlated with the RMSE of the simulated events, (2) the EAD shows more stability and certainty as the number of simulations increases, (3) the change in the RMSE per simulation is below USD 1 million per simulation when the mean ALM is below 0.1, and (4) the two-sample KS test always has a significant p value for six stochastic variables (Fig. A2).

The implementation of a stopping criterion meant it was unknown how well these simulations would perform for an unseen output. Therefore, a round-robin schedule was used. After each simulation, the TGP-LLM was fit to a different output in a predetermined loop. This process was repeated until all outputs for a model reached the stopping criterion. To reduce the computational costs, outputs that reached the stopping criterion were incrementally removed from the round-robin schedule.

2.5 Risk modeling

To model the risk, the method used by Couasnon et al. (2022) was followed. Each event in the stochastic event set was assumed to occur at a constant frequency, equivalent to the reciprocal of the average number of floods per year identified by the POT. Then, all events were ranked according to the magnitude of their economic damage, creating an empirical CDF. The risk curve was obtained by converting the rank of

each event to a return period (RP; Gumbel, 1941). The EAD is an important metric in flood risk assessments (e.g., Olsen et al., 2015) as it allows one to perform a cost–benefit analysis (e.g., Haer et al., 2017). To obtain the EAD, the empirical CDF of damages was integrated.

2.6 Experiment

An experiment was designed to validate, test, and assess the proposed framework under different scenarios. Firstly, a comparison was made with the state-of-the-art (Sect. 2.6.1). Secondly, the scalability was tested (Sect. 2.6.2). Finally, the effect of simplifications on flood risk was assessed (Sect. 2.6.3).

To compare the different risk estimates, statistical tests were performed. A two-sample Kolmogorov–Smirnov (KS) test was applied to compare two empirical CDFs (Hodges, 1958). The null hypothesis assumes that both empirical CDFs are drawn from the same parent distribution. To compare two EADs, a Mann–Whitney U (MWU) rank test was used (Mann and Whitney, 1947). The null hypothesis assumes both empirical CDFs have the same EAD. For the latter, an empirical bootstrap (Efron, 1979) was repeated 500 times for each empirical CDF to obtain a distribution of EADs. For the implementation of both tests, the *scipy* Python package (Virtanen et al., 2020) was used. Two risk estimates were considered significantly different if the p value was smaller than 0.05.

When comparing the computational cost of different surrogate models, the computational time was assessed on an AMD Ryzen 5 5600X 6-core-processor 3.70 GHz CPU.

2.6.1 State-of-the-art vs. active learning

To validate the proposed framework, the active learning approach was compared with a current state-of-the-art equidistant sampling approach based on the methods used by Jane et al. (2022) and Eilander et al. (2023b). Both approaches were only compared for the models with two stochastic variables.

For the current state-of-the-art approach, MDA sampling was combined with linear scatter interpolation based on the recommendation of Jane et al. (2022). In a linear scatter interpolation, simulated events are triangulated (for two dimensions), and linear interpolation occurs within each triangle. To prevent extrapolation, 2^2 simulations representing the vertices of the stochastic event set were simulated. Then, an MDA was applied to the stochastic event set using the same implementation as the active learning approach (Sect. 2.4). In total 64 events were simulated, which is comparable to the number of events used by Eilander et al. (2023b) based on factorial sampling with 8^d simulations. The *gridddata* function from the *scipy* Python package was used to implement the linear scatter interpolation (Virtanen et al., 2020).

The testing and benchmark event sets in two dimensions were used to compare the approaches. The testing event set was used to select the simulations to run as it provided a larger diversity in the synthetic events. Then, the regression model was used to calculate the damages of all synthetic events in the benchmark event set, which allowed for the quantification of the RMSE.

2.6.2 Scalability of the proposed framework

To test the scalability of the proposed framework, the active learning approach was deployed on the testing event sets. Since compound flood drivers respond differently based on the geographic location (e.g., Gori et al., 2020), the active learning approach was not only deployed on the complete (all sub-counties combined) model but also on the classified (coastal and inland sub-counties separated) model for all testing events. It was also deployed on the testing event for two stochastic variables for the sub-county model.

2.6.3 Effect of simplifications on flood risk

To assess the effect of simplifications on current compound flood risk assessments, the flood risk associated with the damage of the complete model to the five testing event sets was modeled.

3 Results and discussion

3.1 State-of-the-art vs. active learning

This section presents and discusses the accuracy and computational time of equidistant sampling and active learning surrogate models. Figure 3 shows the RMSE of both surrogate models compared to the benchmark events set. The active learning approach outperforms the equidistant sampling approach in two ways: (1) for any number of simulations, the RMSE is smaller; and (2) the smallest RMSE is reached after fewer simulations (14 vs. 52). This results in the active learning approach improving the accuracy by a factor of 8 (USD 90.8 vs. 11.2 million), while reducing the number of numerical simulations that would normally be performed by a factor of 4 (64 vs. 14). This increase in accuracy provides a better estimate of the EAD for the active learning approach (37.3 % vs. 1.68 % error) when compared to the benchmark EAD (Fig. A3).

The difference between the accuracy and number of numerical simulations of the approaches is caused by their respective sampling and regression techniques. On the one hand, equidistant sampling cannot explore uncertainty and thus selects a large number of numerical simulations. When combined with linear scatter interpolation, it is not flexible enough to correctly represent the non-linear response of damages to the flood drivers. To achieve a similar accuracy to the active learning approach, a large number of numeri-

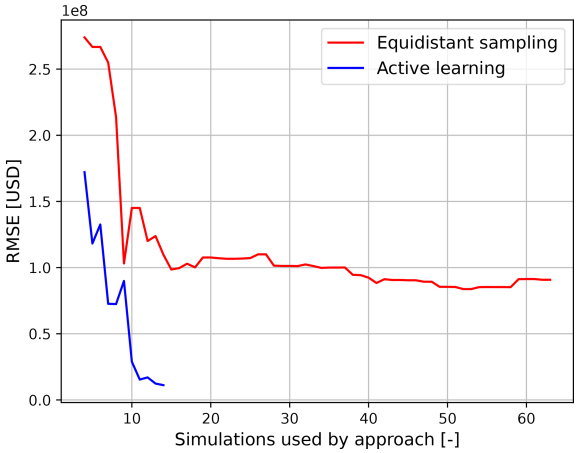

Figure 3. RMSE of both approaches as a function of the number of simulations from the testing event set with two stochastic variables, measured based on the difference with benchmark events.

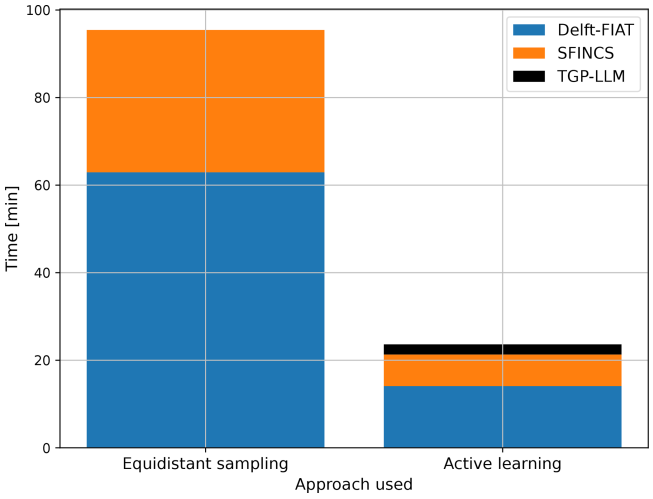

Figure 4. Computational time associated with both approaches for the testing event set with two stochastic variables.

cal simulations would be required, increasing the computational cost. This is similar to observations made by Gouldby et al. (2017). If the number of stochastic variables were to increase, the curse of dimensionality would cause the number of simulations to increase exponentially, making the problem infeasible. On the other hand, the active learning approach explores the uncertainty related to the output (i.e., damages related to a geographic location) of simulations already performed. This allows for the sampling of simulations that are expected to bring the largest gain in expected information. This is combined with a generative model that can capture some of the non-linear relationships between the flood driver parameters and the damages. However, the active learning approach is unable to provide a perfect fit as the RMSE never reaches 0 (Fig. A2), showing that a stopping criterion is necessary, as marginal gains in accuracy may bring high computational costs. This causes the active learning approach to show significant differences with the benchmark when using the two-sample KS test (Fig. A3).

Figure 4 shows the computational time associated with both approaches when they sample from the testing event set. The overall computational time can be split into three components: (1) Delft-FIAT, (2) SFINCS, and (3) TGP-LLM. For both approaches, Delft-FIAT is the component that requires the largest computational time. This is caused by the Delft-FIAT model being poorly optimized for Charleston County as it preprocesses the exposure data before each simulation. The added computational time (2.3 min) for the active learning approach due to the TGP-LLM is relatively small as the number of simulations is small. Therefore, the number of simulations is the main factor influencing the overall computational time. For our experiment, the overall computational time was reduced by a factor of 4 (95.4 vs. 23.6 min).

3.2 Scalability of the proposed framework

In this section, we present and discuss the computational time of the active learning approach under different extents of simplifications. These simplifications can take various forms, but only two of these are investigated: (1) the inclusion of additional stochastic variables and (2) increasing the number of outputs (i.e., damages related to different geographic locations). Figure 5 shows the computational time of the active learning approach for the testing event sets with different numbers of stochastic variables (or dimensions) and the number of outputs. For each, the overall computational time is subdivided into three components: (1) Delft-FIAT, (2) SFINCS, and (3) TGP-LLM. A horizontal red line is provided as a reference, showing the overall computational time of the equidistant sampling approach with two dimensions. Henceforth, this will be referred to as either the equidistant sampling computational time or the reference computational time. For a singular output (complete), the reference computational time is only exceeded with six dimensions. The difference in overall computational time between two and six dimensions is a factor of 5 (100 min). The overall computational time for two outputs (classified), is larger for all test event sets when compared to a singular output (complete). For one and two outputs, the overall computational time does not always increase as the number of dimensions increases. Moreover, the TGP-LLM dominates the overall computational time of the active learning approach for the classified model in six dimensions, where the TGP-LLM amounts to 75 % (1163.9 out of 1553.3 min) of the overall computational time. This caused it to exceed the reference computational time by a factor of 16. Finally, when increasing the number of outputs to 11 (sub-county), the active learning approach is below the reference computational time but requires substan-

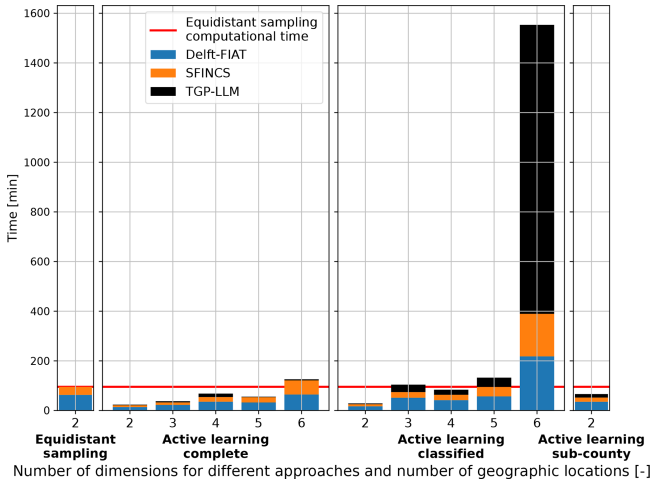

Figure 5. Computational time required for different approaches as a function of outputs (i.e., damages related to different geographic locations) and number of stochastic variables (dimensions).

tially more overall computational time compared to 1 and 2 outputs.

For a singular output, the cost associated with the TGP-LLM is limited as the stopping criterion is often met quickly after the MDA initialization (Table A3). Thus, the TGP-LLM only has to select a small number of simulations, which dominates the reduction in the overall computational time. The use of an initialization results in five dimensions having a smaller overall computational time than four dimensions. This is because the TGP-LLM was used a smaller number of times.

When increasing the number of outputs to two, the number of simulations also increases. This is because the response of economic damages to the flood drivers is different for both locations. This can be seen in Fig. A4 and is a known phenomenon in compound flood risk (e.g., Gori et al., 2020). The number of additional simulations will depend on how complex and diverse the response surfaces are. Here, the more complex response surface of the inland location requires a larger number of numerical simulations than the coastal location (Table A3). These additional simulations cause the overall computational time for the classified model to increase when compared to the complete model. This increase in simulations also causes the relative cost of the TGP-LLM to increase as its cost is proportional to the number of simulations cubed (Gramacy and Lee, 2009).

For 11 outputs, the overall computational time significantly increases when compared to one and two outputs. This is not only caused by the response of the flood drivers to each location but also by the stopping criterion. The round-robin schedule requires a minimum number of simulations, which is proportional to the number of outputs (Sect. 2.4).

3.3 Effect of simplifications on flood risk

This section presents and discusses the outcomes of incrementally adding stochastic variables to the estimate of economic risk. This is tested for a singular output using the active learning approach. Figure 6 shows the risk curves associated with the different numbers of stochastic variables (or dimensions). The samples represent the mean response of the TGP-LLM to a test event set in d dimensions. Additionally, since the TGP-LLM is a generative model (i.e., it captures and models the distribution of the output), the 5 % and 95 % confidence interval associated with the TGP-LLM damage response to the d-dimensional space is used to show the 5 % and 95 % confidence interval associated with each RP. The EAD associated with each risk curve is also shown. The logarithmic behavior of the x axis makes it easier to see the differences in the economic damages at large RPs. The uncertainty bands of the TGP-LLM can be used to show statistically significant differences between the risk curve in two dimensions and the risk curves in higher dimensions when the RP > 10 years. In addition, based on the two-sample KS test, all combinations show significant differences (Fig. A5). When considering the EADs, all but one of the combinations show a significant difference (Fig. A6). When comparing models with two and three dimensions (all driver magnitudes) and the three and six dimensions (adding driver durations and time lag), the EAD difference is 11.3 % (USD 22.01 million) and 11.6 % (USD 25.47 million), respectively. The return values are also directly compared. For an RP of 1, 10, and 100 years the differences between (1) two- and three-dimension models are USD 2.98, 115.08, and 279.57 million, respectively; and (2) six- and three-dimension models are USD 3.03, 31.03, and 66.08 million, respectively.

The uncertainty bands in Fig. 6 show the uncertainty driven by the confidence of the TGP-LLM in modeling the damage response to a d-dimensional space. The uncertainty-based stopping criterion ensures the mean ALM for non-simulated events is at least a certain value for all dimensions. However, differences in the width of the confidence interval at different RPs can be expected because of the following three reasons. Firstly, the uncertainty associated with simulated events will be smaller. This is most noticeable at large RPs, where the density of events is lower in the d-dimensional space, making them more likely to be simulated. Moreover, the uncertainty and damage of an RP are likely to represent the same synthetic event, as the difference in damages between consecutive RPs is large. At smaller RPs, this is less noticeable as (1) a large proportion of these events have not been simulated, (2) the logarithmic scale on this x axis makes it difficult to visualize the uncertainty associated with specific RPs, and (3) the uncertainty and damage of an RP may not represent the same synthetic event because of small differences in damages between consecutive RPs. Secondly, the TGP-LLM can partition the d-dimensional space, which

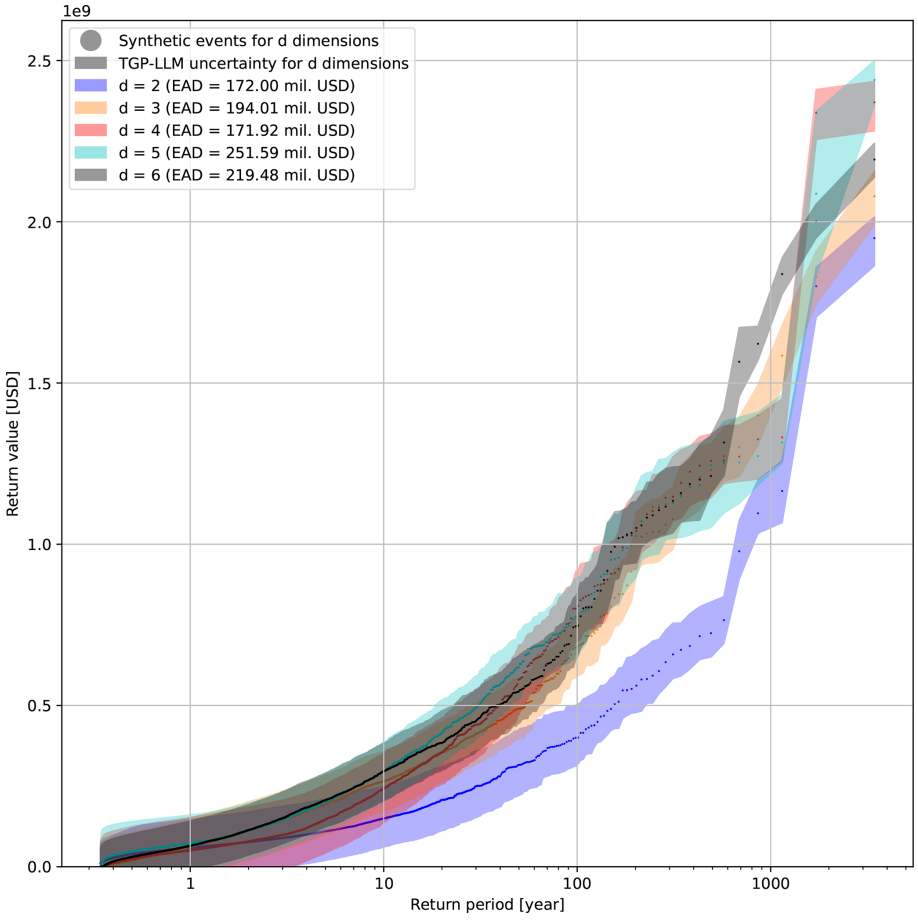

Figure 6. Risk curves for models with a different number of stochastic variables (dimensions). The uncertainty bands represent the 5th and 95th percentiles. The legend includes the EAD estimate for each model.

allows the TGP-LLM to model heteroskedasticity. However, this did not affect the uncertainty bands in Fig. 6, as the TGP-LLM did not have any partitions for all the d-dimensional spaces when the stopping criterion was met. Thirdly, the predictability of the damage response by the TGP-LLM changes depending on the magnitude of the synthetic events and the number of dimensions considered.

A stricter stopping criterion will lead to a larger confidence in the risk curve, which could show a larger number of significant differences between the risk curves. However, this will come at the expense of a larger computational cost.

The main reason for the risk curve in two dimensions showing large differences with other risk curves in higher dimensions is the omission of *T.Mag* as a stochastic variable. This greatly influences the quantification of flood risk as Charleston County is located close to the open coast. This means that the tide and the surge drivers will have a large effect on the economic damage. When *T.Mag* is included, the more extreme coastal water levels have a larger magnitude. Combined with non-linear vulnerability curves (Diermanse et al., 2023), this causes a non-linear increase in damages.

Significant differences between almost all risk curves and EAD show that neglecting drivers' duration and time lag will lead to a mis-quantification in flood risk estimates. In general, the effect of an added stochastic variable on the flood risk depends on two things: (1) how the economic damages respond to the stochastic variable and (2) the assumed constant value used in a model with fewer dimensions and how well that represents its probability distribution and dependence on other stochastic variables. If a constant is used to represent a stochastic variable, it will be unable to provide the same risk curve but may provide a similar estimate in the EAD or return value. For the EAD, this is shown in Figs. A5 and A6 where two and four dimensions show significantly different risk curves but have similar estimates in the EAD. This is because the EAD is an integral of the empirical CDF. Thus, differences can be offset, showing that the EAD is not the best metric to use when justifying simplifications.

3.4 Limitations

In this study, the surrogate models were assessed for the case study of Charleston County. Other case study locations will

show a different response to economic damages as these will have (1) a different combination of flood drivers (e.g., Eilander et al., 2023b), (2) different physical properties that determine the response of the flood hazard to the drivers, and (3) different spatial distributions of exposure and vulnerability (e.g., Koks et al., 2015). While the findings about the exact reduction in overall computational time and increased accuracy are case specific, we expect similar results in other locations. Active learning allows for the number of numerical simulations to be minimized given an input-to-output relationship. The number of simulations will depend on how complex this relationship is, but as shown in Sect. 3.2, the TGP-LLM is still able to reach the stopping criterion for different outputs that have a different response to the flood drivers. A stricter stopping criterion can be used to achieve a higher confidence in the results. However, this will increase the computational cost. Furthermore, the stopping criterion may never be reached as the TGP-LLM is regularized. The results of this study only investigate up to six stochastic variables, but certain case study locations can be affected by more than six drivers (e.g., California Bay-Delta; Cloern et al., 2011), leading to a stronger curse of dimensionality if their duration, time lag, and spatial distribution are also included. We expect that the conceptual framework can still be applied to these case studies, but the overall computational time will largely depend on the interaction of the flood drivers and the number of outputs to be modeled. The overall computational time of the conceptual framework can be further minimized by reducing the number of times the TGP-LLM is applied (e.g., each x simulation rather than after each single simulation).

The choice of boundary condition will not affect the conclusions drawn from the conceptual framework as long as the damage response to the d-dimensional space is consistent for all synthetic events in the stochastic event set. Nonetheless, the current representation of the boundary conditions for the flood drivers is not accurate compared to the state-of-the-art (e.g., Apel et al., 2016; Bakker et al., 2022; Anderson et al., 2019; Marra et al., 2023). This is because (1) Gaussian distributions force the time series to be symmetrical, rigid, and monotonically increasing/decreasing before/after the peak magnitude of the event, and (2) spatially homogeneous boundary conditions do not represent historical events if the model domain is large and are based on the data from a point source. These boundary conditions were used because they minimize the number of parameters, ensure the results are interpretable, and can be applied to all flood drivers. Moreover, they facilitate the use of vine copulas, which offer more flexibility when inferring the natural variability.

Choosing the number of flood drivers and associated stochastic variables to include is currently based on knowledge from previous studies, expert knowledge, or a preliminary sensitivity analysis. We expect the TGP-LLM to understand if a stochastic variable is not contributing to economic damages as it can associate a linear model with this dimension. However, this will require additional numerical simulations. Moreover, it may become difficult to generate representative synthetic events from a robust statistical model in higher dimensions (Morales-Nápoles et al., 2023).

SFINCS is used here as a hydrodynamic model as it has been validated during a previous study (Diermanse et al., 2023). SFINCS significantly decreases the computational cost associated with the hydrodynamic simulation of an event. However, it neglects certain physical processes (e.g., morphodynamics) (Leijnse et al., 2021), potentially increasing the modeling bias associated with the flood hazard. The use of the conceptual framework proposed in this study reduces the number of numerical simulations. This allows for the use of more computationally demanding numerical models, which are expected to represent the flood hazard better in more complex cases. Nonetheless, the majority of the modeling bias is caused by the input datasets used to calibrate and validate the hydrodynamic model, which would also be present in more complex hydrodynamic models (Bates et al., 2021).

When fitting the statistical model to the observed extreme high-water events, we grouped extratropical and tropical events together because of the short time record for the available data (see Sect. 2.1). Ideally, these should be taken into account using separate distributions when modeling risk (e.g., Nederhoff et al., 2024).

During this study, only the uncertainty associated with economic consequences was explored. However, risk can also be social. Authors who use the equidistant/factorial sampling approach make a necessary simplification that the simulations chosen to represent economic risk also represent social risk (e.g., Diermanse et al., 2023; Eilander et al., 2023b). This simplification is not required for the active learning approach as the social consequences can be included as a different type of output in the round-robin schedule of the TGP-LLM.

4 Conclusions and recommendations

A conceptual framework that uses active learning to leverage the input-to-output uncertainty was applied to the case study of Charleston County. The proposed framework uses uncertainty related to the economic damages caused by flood driver parameters. This framework reduces the overall computational time of performing a compound flood risk assessment and/or allows for reducing the number of simplifications usually taken in compound flood risk assessments.

When comparing the state-of-the-art equidistant sampling surrogate model with the proposed active learning surrogate model, the RMSE in damages was reduced by a factor of 8 (USD 90.8 vs. 11.2 million), while reducing the overall computational time by a factor of 4 (95.4 vs. 23.6 min), resulting in a win–win scenario. This reduction in error results in higher accuracy risk estimates. The reduction in overall com-

putational time makes it possible to include more stochastic variables (i.e., to reduce the number of simplifications) to improve risk estimates. For a singular output (one geographic location), the overall computational time increased by a factor of 5 (100 min) between two and six stochastic variables. Not accounting for all the simplifications resulted in an underestimation of the EAD of USD 47.48 million, and an underestimation of the 100-year return period (RP) of USD 345.65 million. Exploring the uncertainty associated with different outputs (i.e., geographic locations) increased the overall computational time as the economic damages responded differently to the flood drivers, requiring additional numerical simulations. In the case of six dimensions and two outputs, this caused the overall computational time to increase by a factor of 16 when compared with the current state-of-the-art. While the proposed active learning surrogate model was only assessed for a single case study, there is a high expectation to see similar benefits to other case studies.

Based on these findings, we have three recommendations for future studies. Firstly, different strategies could limit the computational time associated with training TGP-LLMs, for instance, by simulating the synthetic events with the x highest active learning Mackay (ALM) statistic in each iteration or increasing the number of simulations in the initialization. Secondly, some case studies will have stochastic variables associated with processes that do not dominate the compound flood risk. To limit the increase in overall computational time, future research should propose simplifications that do not affect the quantification of compound flood risk. However, our results show that similar EAD values can result from different risk curves, and hence, care should be taken when using EAD to validate these simplifications. Finally, to propose optimal risk reduction measures, the framework would have to be deployed once for each possible adaptation measure. Future research should investigate how to limit the computational cost of this operation by understanding and predicting how the input-to-output response surface will change with certain adaptation measures.

Appendix A: Supplementary information

Table A1. Regular vine copula models (e.g., Czado, 2019) for different numbers of stochastic variables with associated regular vine distribution and bivariate copula models.

Two stochastic variables			
Tree	Edge	Copula model	Parameters
1	*S.Mag*, *P.Mag*	Gaussian	0.348
Four stochastic variables			
Tree	Edge	Copula model	Parameters
1	*S.Mag*, *P.Mag*	Gaussian	0.348
	P.Mag, *P.Dur*	Joe 180°	2.44
	T.Mag, *P.Dur*	Frank	1.69
2	*T.Mag*, *P.Mag* | *P.Dur*	Independence	
	P.Dur, *S.Mag* | *P.Mag*	Independence	
3	*T.Mag*, *S.Mag* | *P.Dur*, *P.Mag*	Independence	
Six stochastic variables			
Tree	Edge	Copula model	Parameters
1	*T.Mag*, *P.Dur*	Frank	1.69
	P.Dur, *P.Mag*	Joe 180°	2.44
	P.Mag, *S.Mag*	Gaussian	0.348
	S.Dur, *S.Mag*	Independence	
	S.Mag, *P.Lag*	Frank	1.98
2	*T.Mag*, *P.Mag* | *P.Dur*	Independence	
	P.Dur, *S.Mag* | *P.Mag*	Independence	
	P.Mag, *S.Dur* | *S.Mag*	Independence	
	S.Dur, *P.Lag* | *S.Mag*	Joe 270°	1.35
3	*T.Mag*, *S.Mag* | *P.Dur*, *P.Mag*	Independence	
	P.Dur, *S.Dur* | *P.Mag*, *S.Mag*	Independence	
	P.Mag, *P.Lag* | *S.Mag*, *S.Dur*	Joe 180°	1.30
4	*T.Mag*, *S.Dur* | *P.Dur*, *P.Mag*, *S.Mag*	Independence	
	P.Dur, *P.Lag* | *P.Mag*, *S.Mag*, *S.Dur*	Independence	
5	*T.Mag*, *P.Lag* | *P.Dur*, *P.Mag*, *S.Mag*, *S.Dur*	Independence	

Table A2. Summary of the marginal CDFs for the different stochastic variables.

Stochastic variable	Distribution	Parameters
S.Mag	exponential	{mean: 0.322; scale: 0.102}
P.Mag	exponential	{mean: 0.0796; scale: 5.60}
T.Mag	empirical: HH tides	
P.Dur	truncated normal	{a: 0; b: 406; mean: 0.158; scale: 18.9}
S.Dur	truncated Gumbel	{c: 0.726; mean: 0.880; scale: 5.26}
P.Lag	truncated normal	{a: −72.0; b: 72.0; mean: −10.7.; scale: 30.6}

Table A3. Total number of simulations required to reach stopping criterion for different dimensions and number of outputs. For multiple outputs, these are ordered top-down in the order they appear in the round-robin schedule. Bold numbers represent the output dictating the total number of simulations required to reach the final stopping criterion.

		Number of dimensions [–]				
Model	Output	2	3	4	5	6
Complete	total	**14**	**24**	**36**	**33**	**65**
Classified	inland	**17**	**38**	**43**	**59**	**224**
	coastal	13	30	37	35	67
Sub-county	Charleston Central	**36**				
	Edisto Island	27				
	James Island	28				
	Johns Island	29				
	Kiawah Island Seabrook Island	30				
	McClellanville	31				
	Mount Pleasant	32				
	North Charleston	33				
	Ravenel-Hollywood	34				
	Wadmalaw Island	24				
	West Ashley	35				

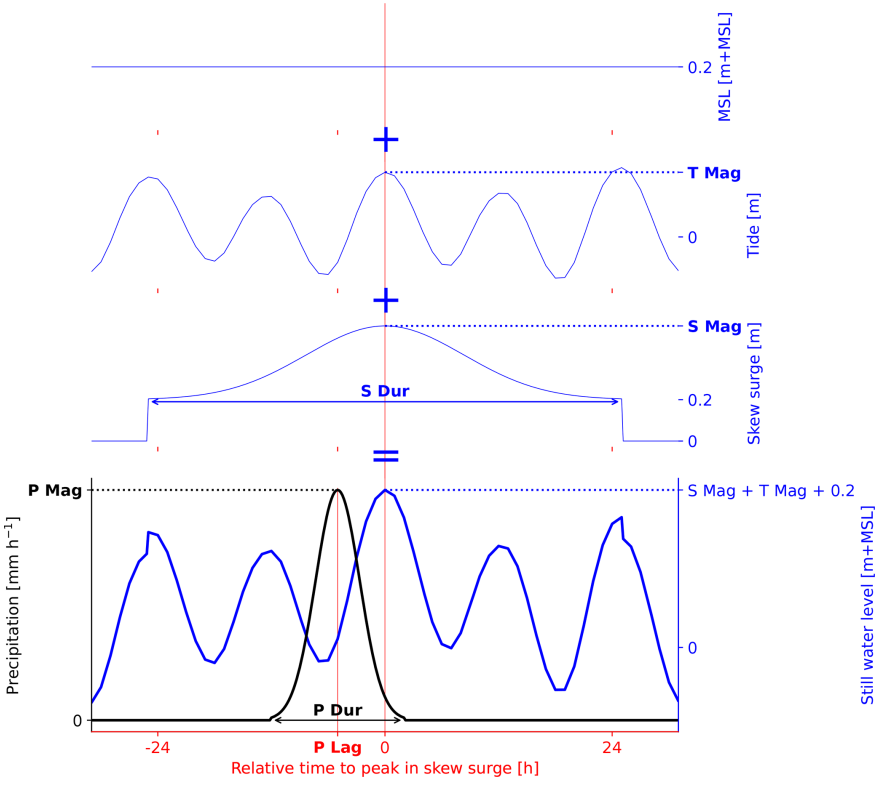

Figure A1. Schematization of SFINCS model boundary conditions.

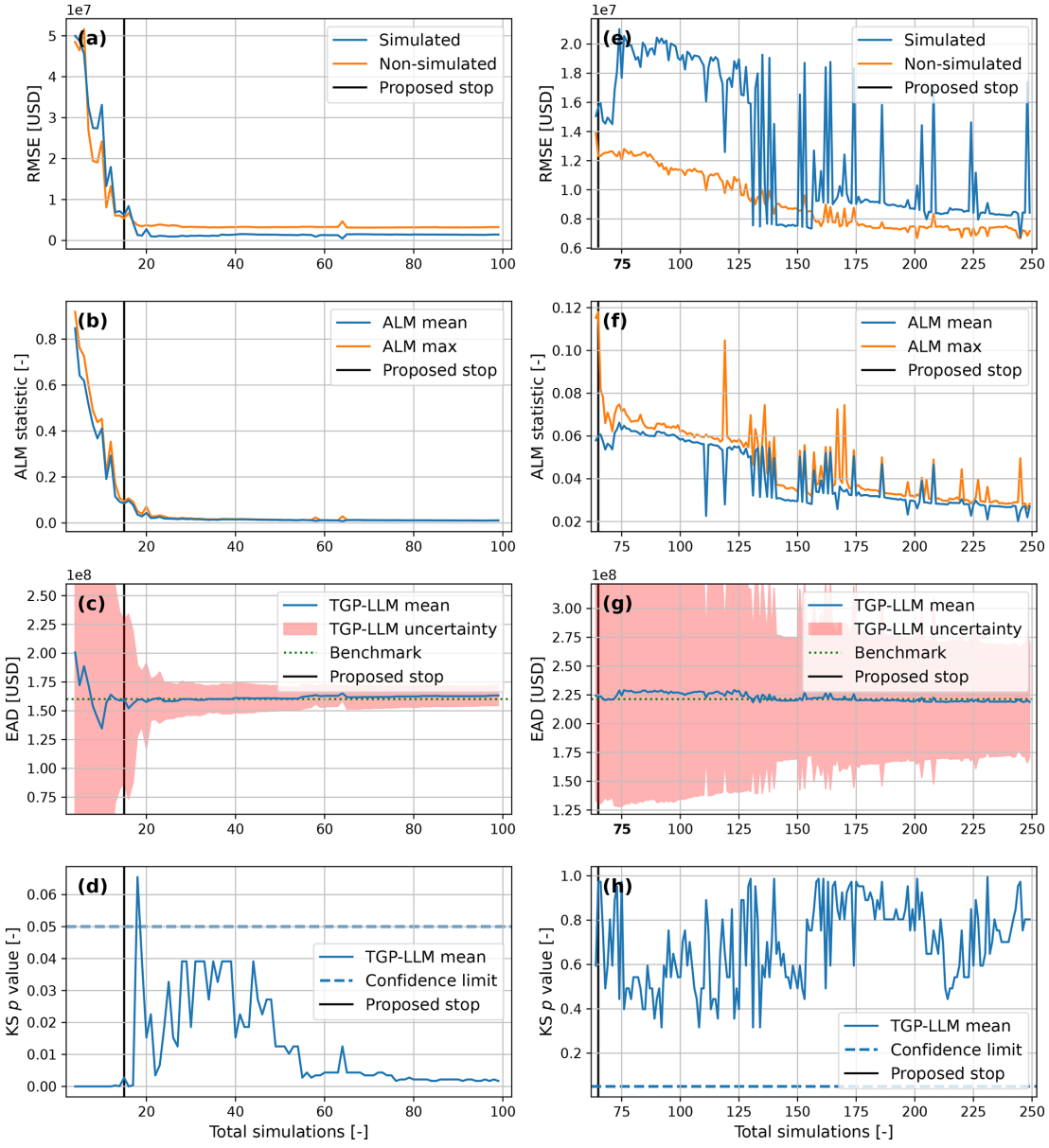

Figure A2. Developing a stopping criterion for compound flood risk. Panels **(a)**–**(d)** and **(e)**–**(h)** represent the benchmark datasets for two and six dimensions, respectively. Panels **(a)** and **(e)** show the error of the surrogate damage model. Panels **(b)** and **(f)** show information on the ALM statistic. Panels **(c)** and **(g)** compare the estimate of the EAD from the surrogate damage model with the benchmark. Panels **(d)** and **(h)** show the results of the two-sample KS test when comparing the empirical CDF of the surrogate damage model with the benchmark.

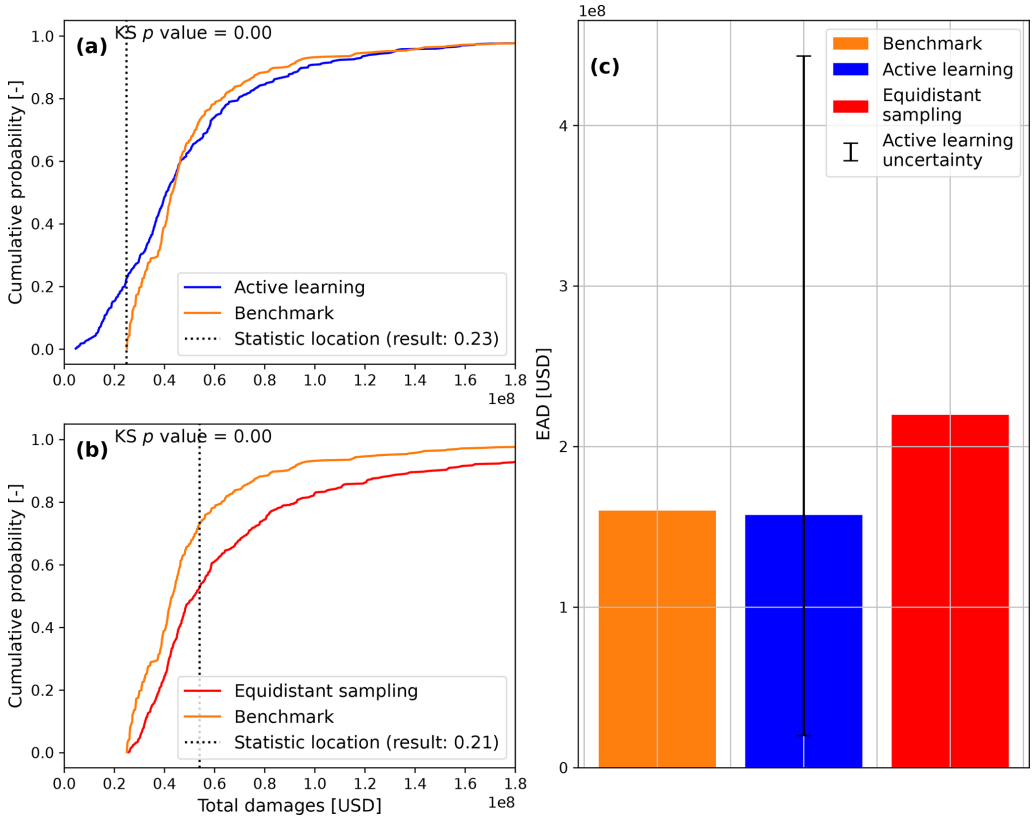

Figure A3. Risk estimates of both approaches when compared with the benchmark. Panels **(a)** and **(b)** show the outcomes of the two-sample KS test for the active learning and equidistant sampling approaches, respectively. Panel **(c)** shows the estimates of EAD from each approach.

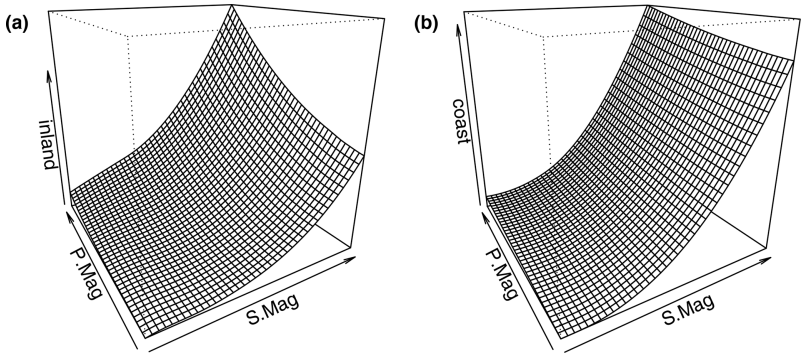

Figure A4. Response of economic damages to the testing event set in two dimensions for the different outputs of the classified model: panel **(a)** shows the response of the inland location and panel **(b)** shows the response of the coastal (coast) location. Plots are obtained from the *tgpllm* R package developed by Gramacy and Lee (2009).

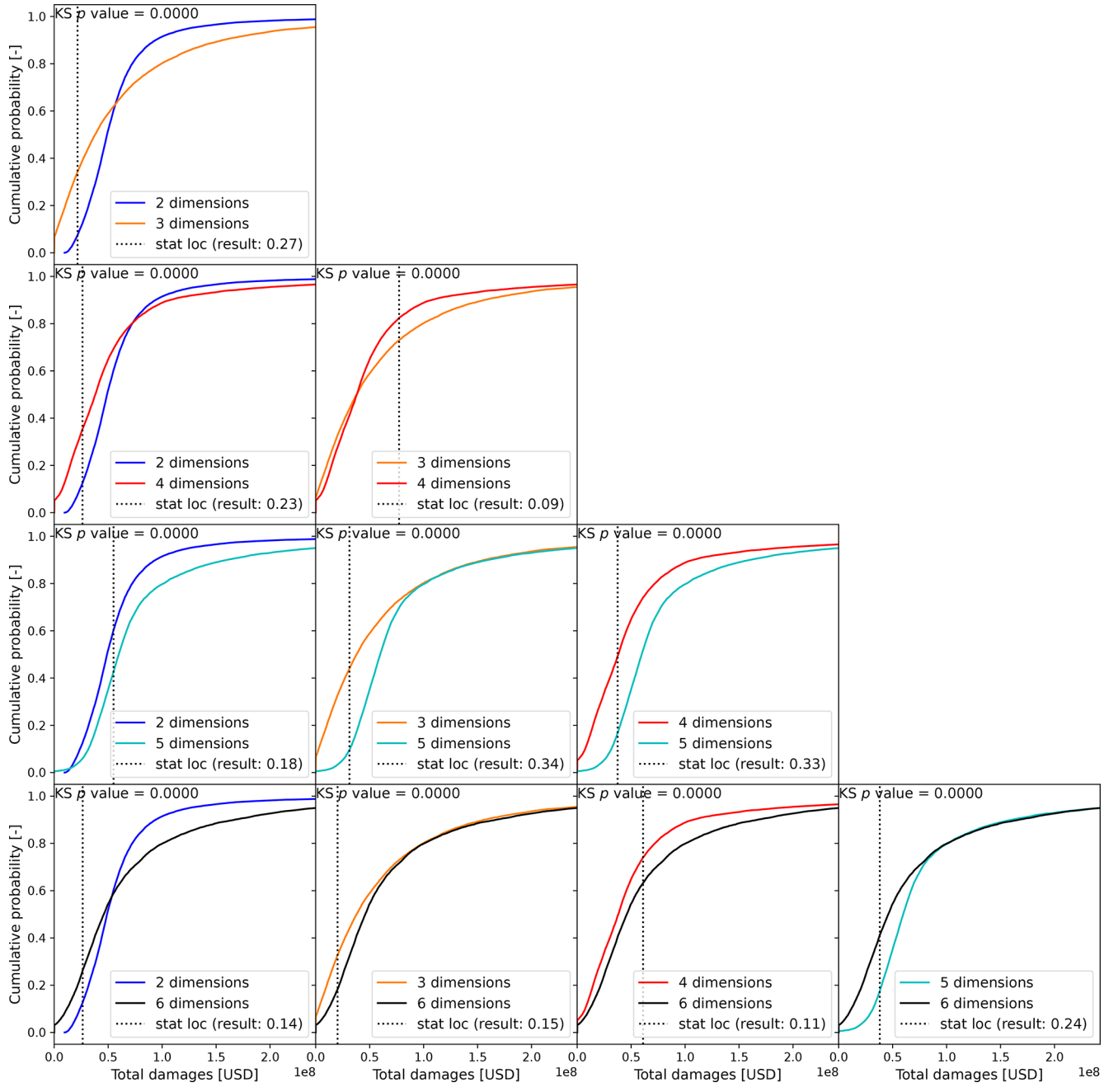

Figure A5. Test statistics and associated p values associated with two-sample KS test for the empirical CDFs of different dimensionalities when sampling from the total economic damages. If the p value is smaller than 0.05, the null hypothesis (empirical CDFs come from the same parent distribution) is rejected.

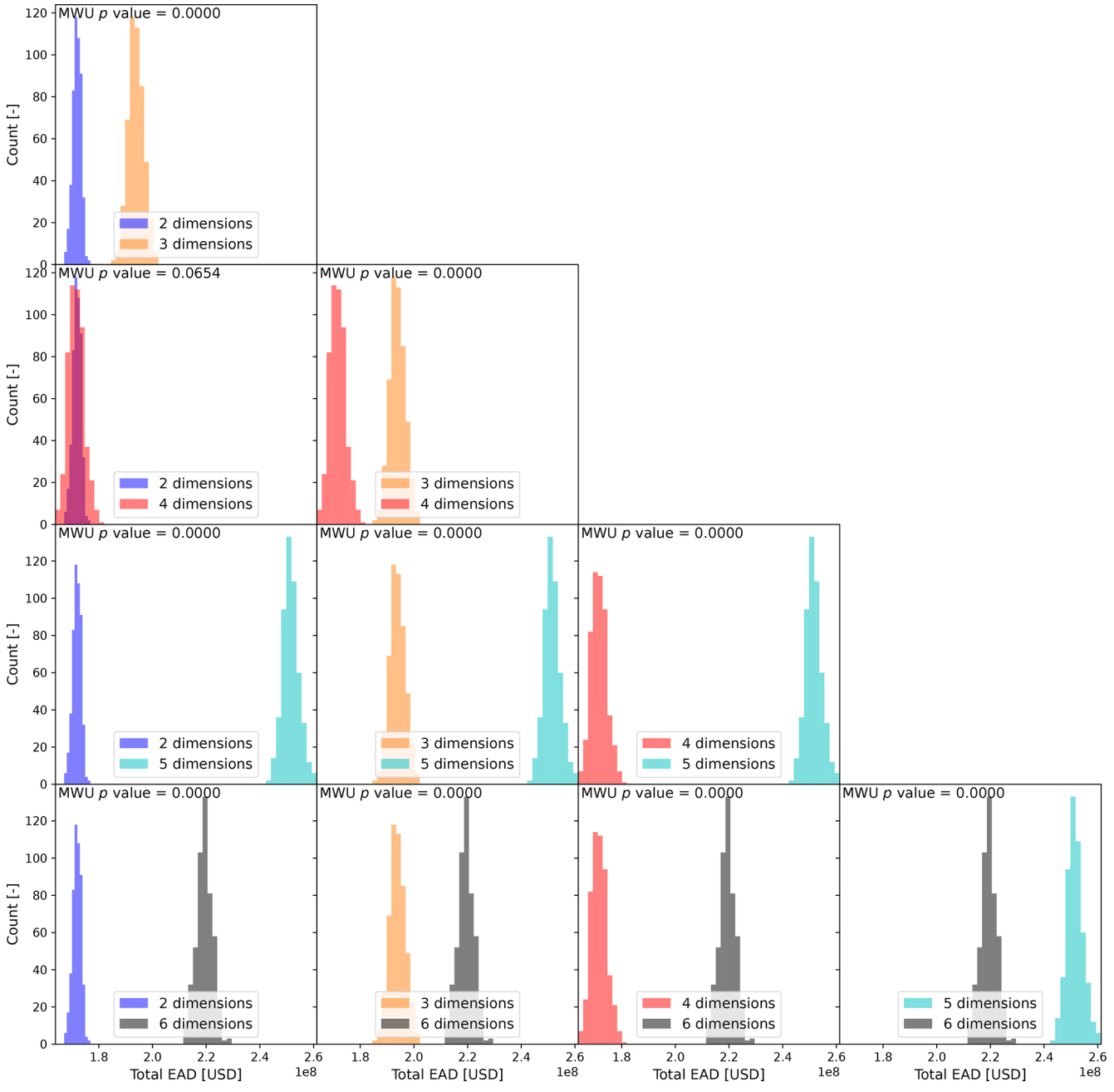

Figure A6. p values associated with MWU test for empirical CDFs of different dimensionalities. If the p values are smaller than 0.05, the null hypothesis (empirical CDFs have the same location (EAD)) is rejected.

Code and data availability. The scripts and data used to set up the experiments in this study are available from Zenodo at https://doi.org/10.5281/zenodo.13910108 (Terlinden-Ruhl, 2024).

Author contributions. Author contributions follow the CRediT authorship categories. LTR: conceptualization, methodology, software, formal analysis, investigation, writing – original draft, writing – review and editing and visualization. AC, DE, GGH: conceptualization, methodology, writing – original draft, writing – review and editing. PMN: writing – original draft, writing – review and editing. JAAA: writing – original draft, writing – review and editing, supervision.

Competing interests. At least one of the (co-)authors is a guest member of the editorial board of *Natural Hazards and Earth System Sciences* for the special issue "Methodological innovations for the analysis and management of compound risk and multi-risk, including climate-related and geophysical hazards (NHESS/ESD/ESSD/GC/HESS inter-journal SI)". The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Special issue statement. This article is part of the special issue "Methodological innovations for the analysis and management of compound risk and multi-risk, including climate-related and geophysical hazards (NHESS/ESD/ESSD/GC/HESS inter-journal SI)". It is not associated with a conference.

Acknowledgements. We would like to thank the FloodAdapt team for providing their support with the hydrodynamic and impact models.

Financial support. This research has been supported by the Deltares (Moonshot 2: Making the world safer from flooding).

Review statement. This paper was edited by Silvia De Angeli and reviewed by two anonymous referees.

References

Anderson, D., Rueda, A., Cagigal, L., Antolinez, J. A. A., Mendez, F. J., and Ruggiero, P.: Time-Varying Emulator for Short and Long-Term Analysis of Coastal Flood Hazard Potential, J. Geophys. Res.-Oceans, 124, 9209–9234, https://doi.org/10.1029/2019jc015312, 2019.

Antolínez, J. A. A., Méndez, F. J., Anderson, D., Ruggiero, P., and Kaminsky, G. M.: Predicting Climate-Driven Coastlines With a Simple and Efficient Multiscale Model, J. Geophys. Res.-Earth Surface, 124, 1596–1624, https://doi.org/10.1029/2018jf004790, 2019.

Apel, H., Martínez Trepat, O., Hung, N. N., Chinh, D. T., Merz, B., and Dung, N. V.: Combined fluvial and pluvial urban flood hazard analysis: concept development and application to Can Tho city, Mekong Delta, Vietnam, Nat. Hazards Earth Syst. Sci., 16, 941–961, https://doi.org/10.5194/nhess-16-941-2016, 2016.

Arns, A., Wahl, T., Haigh, I., Jensen, J., and Pattiaratchi, C.: Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise, Coast. Eng., 81, 51–66, https://doi.org/10.1016/j.coastaleng.2013.07.003, 2013.

Bakker, T. M., Antolínez, J. A., Leijnse, T. W., Pearson, S. G., and Giardino, A.: Estimating tropical cyclone-induced wind, waves, and surge: A general methodology based on representative tracks, Coast. Eng., 176, 104154, https://doi.org/10.1016/j.coastaleng.2022.104154, 2022.

Barnard, P., Befus, K. M., Danielson, J. J., Engelstad, A. C., Erikson, L., Foxgrover, A., Hardy, M. W., Hoover, D. J., Leijnse, T., Massey, C., McCall, R., Nadal-Caraballo, N., Nederhoff, K., Ohenhen, L., O'Neill, A., Parker, K. A., Shirzaei, M., Su, X., Thomas, J. A., van Ormondt, M., Vitousek, S. F., Vos, K., and Yawn, M. C.: Future coastal hazards along the U.S. North and South Carolina coasts, U.S. Geological Survey [data set], https://doi.org/10.5066/P9W91314, 2023.

Barnard, P. L., Erikson, L. H., Foxgrover, A. C., Hart, J. A. F., Limber, P., O'Neill, A. C., van Ormondt, M., Vitousek, S., Wood, N., Hayden, M. K., and Jones, J. M.: Dynamic flood modeling essential to assess the coastal impacts of climate change, Sci. Rep., 9, 4309, https://doi.org/10.1038/s41598-019-40742-z, 2019.

Bates, P. D., Horritt, M. S., and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, J. Hydrol., 387, 33–45, https://doi.org/10.1016/j.jhydrol.2010.03.027, 2010.

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., 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, e2020WR02867, https://doi.org/10.1029/2020wr028673, 2021.

Bates, P. D., Savage, J., Wing, O., Quinn, N., Sampson, C., Neal, J., and Smith, A.: A climate-conditioned catastrophe risk model for UK flooding, Nat. Hazards Earth Syst. Sci., 23, 891–908, https://doi.org/10.5194/nhess-23-891-2023, 2023.

Bedford, T. and Cooke, R. M.: Vines–a new graphical model for dependent random variables, Ann. Stat., 30, 1031–1068, https://doi.org/10.1214/aos/1031689016, 2002.

Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., and Vrac, M.: Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy), Hydrol. Earth Syst. Sci., 21, 2701–2723, https://doi.org/10.5194/hess-21-2701-2017, 2017.

Blöschl, G.: Three hypotheses on changing river flood hazards, Hydrol. Earth Syst. Sci., 26, 5015–5033, https://doi.org/10.5194/hess-26-5015-2022, 2022.

Camus, P., Mendez, F. J., Medina, R., and Cofiño, A. S.: Analysis of clustering and selection algorithms for the study of multivariate wave climate, Coast. Eng., 58, 453–462, https://doi.org/10.1016/j.coastaleng.2011.02.003, 2011.

City of Charleston: Sea Level Rise Strategy, City of Charleston, https://www.charleston-sc.gov/DocumentCenter/View/10089, (last access: 6 August 2024), 2015.

Cloern, J. E., Knowles, N., Brown, L. R., Cayan, D., Dettinger, M. D., Morgan, T. L., Schoellhamer, D. H., Stacey, M. T., van der Wegen, M., Wagner, R. W., and Jassby, A. D.: Projected

Evolution of California's San Francisco Bay-Delta-River System in a Century of Climate Change, PLoS ONE, 6, e24465, https://doi.org/10.1371/journal.pone.0024465, 2011.

Cokelaer, T., Kravchenko, A., Varma, A., L, B., Eadi Stringari, C., Brueffer, C., Broda, E., Pruesse, E., Singaravelan, K., Russo, S., and Li, Z.: cokelaer/fitter: v1.7.0, Zenodo [code], https://doi.org/10.5281/zenodo.10459943, 2024.

Couasnon, A., Eilander, D., Muis, S., Veldkamp, T. I. E., Haigh, I. D., Wahl, T., Winsemius, H. C., and Ward, P. J.: Measuring compound flood potential from river discharge and storm surge extremes at the global scale, Nat. Hazards Earth Syst. Sci., 20, 489–504, https://doi.org/10.5194/nhess-20-489-2020, 2020.

Couasnon, A., Scussolini, P., Tran, T. V. T., Eilander, D., Muis, S., Wang, H., Keesom, J., Dullaart, J., Xuan, Y., Nguyen, H. Q., Winsemius, H. C., and Ward, P. J.: A Flood Risk Framework Capturing the Seasonality of and Dependence Between Rainfall and Sea Levels – An Application to Ho Chi Minh City, Vietnam, Water Resour. Res., 58, e2021WR030002, https://doi.org/10.1029/2021wr030002, 2022.

Cushing, W. M., Taylor, D., Danielson, J. J., Poppenga, S., Beverly, S., and Shogib, R.: Topobathymetric Model of the Coastal Carolinas, 1851 to 2020, U.S. Geological Survey [data set], https://doi.org/10.5066/P9MPA8K0, 2022.

Czado, C.: Analyzing Dependent Data with Vine Copulas, Springer, ISBN 978-3-030-13785-4, https://doi.org/10.1007/978-3-030-13785-4, 2019.

Czado, C. and Nagler, T.: Vine Copula Based Modeling, Annu. Rev. Stat. Appl., 9, 453–477, https://doi.org/10.1146/annurev-statistics-040220-101153, 2022.

Danielson, J. J., Poppenga, S. K., Brock, J. C., Evans, G. A., Tyler, D. J., Gesch, D. B., Thatcher, C. A., and Barras, J. A.: Topobathymetric Elevation Model Development using a New Methodology: Coastal National Elevation Database, J. Coast. Res., 76, 75–89, https://doi.org/10.2112/si76-008, 2016.

Deltares: Delft3D-FM User Manual, Deltares, https://content.oss.deltares.nl/delft3d/D-Flow_FM_User_Manual.pdf (last access: 6 April 2025), 2022.

Deltares: Delft FIAT, Deltares, https://deltares.github.io/Delft-FIAT/stable/ (last access: 29 September 2024), 2024.

DHI: MIKE 21 Flow Model User Guide, DHI, https://manuals.mikepoweredbydhi.help/2017/Coast_and_Sea/M21HD.pdf (last access: 6 August 2024), 2017.

Diermanse, F., Roscoe, K., van Ormondt, M., Leijnse, T., Winter, G., and Athanasiou, P.: Probabilistic compound flood hazard analysis for coastal risk assessment: A case study in Charleston, South Carolina, Shore & Beach, 9–18, https://doi.org/10.34237/1009122, 2023.

Diermanse, F. L. M., De Bruijn, K. M., Beckers, J. V. L., and Kramer, N. L.: Importance sampling for efficient modelling of hydraulic loads in the Rhine–Meuse delta, Stoch. Env. Res. Risk A., 29, 637–652, https://doi.org/10.1007/s00477-014-0921-4, 2014.

Dißmann, J., Brechmann, E., Czado, C., and Kurowicka, D.: Selecting and estimating regular vine copulae and application to financial returns, Comput. Stat. Data An., 59, 52–69, https://doi.org/10.1016/j.csda.2012.08.010, 2013.

Efron, B.: Bootstrap Methods: Another Look at the Jackknife, Ann. Stat., 7, 1–26, https://doi.org/10.1214/aos/1176344552, 1979.

Eilander, D., Boisgontier, H., Bouaziz, L. J. E., Buitink, J., Couasnon, A., Dalmijn, B., Hegnauer, M., de Jong, T., Loos, S., Marth, I., and van Verseveld, W.: HydroMT: Automated and reproducible model building and analysis, Journal of Open Source Software, 8, 4897, https://doi.org/10.21105/joss.04897, 2023a.

Eilander, D., Couasnon, A., Sperna Weiland, F. C., Ligtvoet, W., Bouwman, A., Winsemius, H. C., and Ward, P. J.: Modeling compound flood risk and risk reduction using a globally applicable framework: a pilot in the Sofala province of Mozambique, Nat. Hazards Earth Syst. Sci., 23, 2251–2272, https://doi.org/10.5194/nhess-23-2251-2023, 2023b.

Fraehr, N., Wang, Q. J., Wu, W., and Nathan, R.: Assessment of surrogate models for flood inundation: The physics-guided LSG model vs. state-of-the-art machine learning models, Water Res., 252, 121202, https://doi.org/10.1016/j.watres.2024.121202, 2024.

Gori, A., Lin, N., and Xi, D.: Tropical Cyclone Compound Flood Hazard Assessment: From Investigating Drivers to Quantifying Extreme Water Levels, Earths Future, 8, e2020EF001660, https://doi.org/10.1029/2020ef001660, 2020.

Gouldby, B., Wyncoll, D., Panzeri, M., Franklin, M., Hunt, T., Hames, D., Tozer, N., Hawkes, P., Dornbusch, U., and Pullen, T.: Multivariate extreme value modelling of sea conditions around the coast of England, P. I. Civil Eng. Mar. En., 170, 3–20, https://doi.org/10.1680/jmaen.2016.16, 2017.

Gramacy, R. B. and Lee, H. K. H.: Bayesian treed Gaussian process models with an application to computer modeling, arXiv [preprint], https://doi.org/10.48550/arXiv.0710.4536, 17 March 2009.

Gumbel, E. J.: The Return Period of Flood Flows, Ann. Math. Stat., 12, 163–190, https://doi.org/10.1214/aoms/1177731747, 1941.

Haer, T., Botzen, W. W., Zavala-Hidalgo, J., Cusell, C., and Ward, P. J.: Economic evaluation of climate risk adaptation strategies: Cost-benefit analysis of flood protection in Tabasco, Mexico, Atmósfera, 30, 101–120, https://doi.org/10.20937/atm.2017.30.02.03, 2017.

Hendrickx, G. G., Antolínez, J. A., and Herman, P. M.: Predicting the response of complex systems for coastal management, Coast. Eng., 182, 104289, https://doi.org/10.1016/j.coastaleng.2023.104289, 2023.

Hendry, A., Haigh, I. D., Nicholls, R. J., Winter, H., Neal, R., Wahl, T., Joly-Laugel, A., and Darby, S. E.: Assessing the characteristics and drivers of compound flooding events around the UK coast, Hydrol. Earth Syst. Sci., 23, 3117–3139, https://doi.org/10.5194/hess-23-3117-2019, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S.: Global flood

risk under climate change, Nat. Clim. Change, 3, 816–821, https://doi.org/10.1038/nclimate1911, 2013.

Hodges, J. L.: The significance probability of the smirnov two-sample test, Ark. Mat., 3, 469–486, https://doi.org/10.1007/bf02589501, 1958.

Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., Gass, L., Funk, M., Wickham, J., Stehman, S., Auch, R., and Riitters, K.: Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database, ISPRS J. Photogramm., 162, 184–199, https://doi.org/10.1016/j.isprsjprs.2020.02.019, 2020.

Ishibashi, H. and Hino, H.: Stopping Criterion for Active Learning Based on Error Stability, arXiv [preprint], https://doi.org/10.48550/arXiv.2104.01836, 9 April 2021.

Jane, R. A., Malagón-Santos, V., Rashid, M. M., Doebele, L., Wahl, T., Timmers, S. R., Serafin, K. A., Schmied, L., and Lindemer, C.: A Hybrid Framework for Rapidly Locating Transition Zones: A Comparison of Event- and Response-Based Return Water Levels in the Suwannee River FL, Water Resour. Res., 58, e2022WR032481, https://doi.org/10.1029/2022wr032481, 2022.

Jongman, B.: Effective adaptation to rising flood risk, Nat. Commun., 9, 1986, https://doi.org/10.1038/s41467-018-04396-1, 2018.

Kennard, R. W. and Stone, L. A.: Computer Aided Design of Experiments, Technometrics, 11, 137–148, https://doi.org/10.1080/00401706.1969.10490666, 1969.

Klijn, F., Kreibich, H., de Moel, H., and Penning-Rowsell, E.: Adaptive flood risk management planning based on a comprehensive flood risk conceptualisation, Mitig. Adapt. Strat. Gl., 20, 845–864, https://doi.org/10.1007/s11027-015-9638-z, 2015.

Koks, E., Jongman, B., Husby, T., and Botzen, W.: Combining hazard, exposure and social vulnerability to provide lessons for flood risk management, Environ. Sci. Policy, 47, 42–52, https://doi.org/10.1016/j.envsci.2014.10.013, 2015.

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, Coast. Eng., 163, 103796, https://doi.org/10.1016/j.coastaleng.2020.103796, 2021.

MacKay, D. J. C.: Information-Based Objective Functions for Active Data Selection, Neural Comput., 4, 590–604, https://doi.org/10.1162/neco.1992.4.4.590, 1992.

Mann, H. B. and Whitney, D. R.: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other, Ann. Math. Stat., 18, 50–60, https://doi.org/10.1214/aoms/1177730491, 1947.

Marra, J., Sweet, W., Leuliette, E., Kruk, M., Genz, A., Storlazzi, C., Ruggiero, P., Leung, M., Anderson, D. L., Merrifield, M., Becker, J., Robertson, I., Widlansky, M. J., Thompson, P. R., Mendez, F., Rueda, A., Antolinez, J. A. A., Cagigal, L., Menendez, M., Lobeto, H., Obeysekera, J., and Chiesa, C.: Advancing best practices for the analysis of the vulnerability of military installations in the Pacific Basin to coastal flooding under a changing climate – RC-2644, U.S. Department of Defense Strategic Environmental Research and Development Program, https://pubs.usgs.gov/publication/70244064 (last access: 19 January 2025), 2023.

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.

Moradian, S., AghaKouchak, A., Gharbia, S., Broderick, C., and Olbert, A. I.: Forecasting of compound ocean-fluvial floods using machine learning, J. Environ. Manage., 364, 121295, https://doi.org/10.1016/j.jenvman.2024.121295, 2024.

Morales-Nápoles, O.: Counting Vines, World Scientific, ISBN 9789814299886, 189–218, https://doi.org/10.1142/9789814299886_0009, 2010.

Morales-Nápoles, O., Rajabi-Bahaabadi, M., Torres-Alves, G. A., and 't Hart, C. M. P.: Chimera: An atlas of regular vines on up to 8 nodes, Scientific Data, 10, 337, https://doi.org/10.1038/s41597-023-02252-6, 2023.

Morris, J. T. and Renken, K. A.: Past, present, and future nuisance flooding on the Charleston peninsula, PLOS ONE, 15, e0238770, https://doi.org/10.1371/journal.pone.0238770, 2020.

Muis, S., Güneralp, B., Jongman, B., Aerts, J. C., and Ward, P. J.: Flood risk and adaptation strategies under climate change and urban expansion: A probabilistic analysis using global data, Sci. Total Environ., 538, 445–457, https://doi.org/10.1016/j.scitotenv.2015.08.068, 2015.

Nagler, T. and Vatter, T.: pyvinecopulib, Zenodo [code], https://doi.org/10.5281/zenodo.10435751, 2023.

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?, Nat. Hazards, 120, 8779–8825, https://doi.org/10.1007/s11069-024-06552-x, 2024.

Neumann, B., Vafeidis, A. T., Zimmermann, J., and Nicholls, R. J.: Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding – A Global Assessment, PLOS ONE, 10, e0118571, https://doi.org/10.1371/journal.pone.0118571, 2015.

Olsen, A., Zhou, Q., Linde, J., and Arnbjerg-Nielsen, K.: Comparing Methods of Calculating Expected Annual Damage in Urban Pluvial Flood Risk Assessments, Water, 7, 255–270, https://doi.org/10.3390/w7010255, 2015.

Parker, K., Erikson, L., Thomas, J., Nederhoff, K., Barnard, P., and Muis, S.: Relative contributions of water-level components to extreme water levels along the US Southeast Atlantic Coast from a regional-scale water-level hindcast, Nat. Hazards, 117, 2219–2248, https://doi.org/10.1007/s11069-023-05939-6, 2023.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Rueda, A., Gouldby, B., Méndez, F., Tomás, A., Losada, I., Lara, J., and Díaz-Simal, P.: The use of wave propagation and reduced complexity inundation models and metamodels for coastal flood risk assessment, J. Flood Risk Manag., 9, 390–401, https://doi.org/10.1111/jfr3.12204, 2015.

Samadi, V. and Lunt, S.: Historical Floods of South Carolina, Clemson University, https://lgpress.clemson.edu/publication/historical-floods-of-south-carolina/ (last access: 15 July 2024), 2023.

Schwarz, G.: Estimating the Dimension of a Model, Ann. Stat., 6, 461–464, https://doi.org/10.1214/aos/1176344136, 1978.

Swain, D. L., Wing, O. E. J., Bates, P. D., Done, J. M., Johnson, K. A., and Cameron, D. R.: Increased Flood Exposure Due to Climate Change and Population Growth in the United States, Earths Future, 8, e2020EF001778, https://doi.org/10.1029/2020ef001778, 2020.

Terlinden-Ruhl, L.: Compound_TGP, Zenodo [code], https://doi.org/10.5281/zenodo.13910108, 2024.

Tomar, A. and Burton, H. V.: Active learning method for risk assessment of distributed infrastructure systems, Comput.-Aided Civ. Inf., 36, 438–452, https://doi.org/10.1111/mice.12665, 2021.

UNDRR: The human cost of disasters: an overview of the last 20 years (2000–2019), UNDRR, https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019 (last access: 12 July 2024), 2020.

United States Census Bureau: https://www2.census.gov/geo/maps/DC2020/DC20BLK/st45_sc/cousub/, last access: 6 May 2024.

USACE Hydrologic Engineering Center: HEC-RAS Documentation, USACE Hydrologic Engineering Center, https://www.hec.usace.army.mil/confluence/rasdocs (last access: 8 April 2025), 2025.

U.S. Department of Agriculture: U.S. General Soil Map (STATSGO2) for Florida, Georgia, South Carolina, North Carolina and Virginia, U.S. Department of Agriculture, https://gdg.sc.egov.usda.gov/ (last access: 8 January 2021), 2020.

van Ormondt, M., Leijnse, T., de Goede, R., Nederhoff, K., and van Dongeren, A.: Subgrid corrections for the linear inertial equations of a compound flood model – a case study using SFINCS 2.1.1 Dollerup release, Geosci. Model Dev., 18, 843–861, https://doi.org/10.5194/gmd-18-843-2025, 2025.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

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. Change, 5, 1093–1097, https://doi.org/10.1038/nclimate2736, 2015.

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., Veldkamp, T. I. E., Winsemius, H. C., and Wahl, T.: Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries, Environ. Res. Lett., 13, 084012, https://doi.org/10.1088/1748-9326/aad400, 2018.

Williams, J., Horsburgh, K. J., Williams, J. A., and Proctor, R. N. F.: Tide and skew surge independence: New insights for flood risk, Geophys. Res. Lett., 43, 6410–6417, https://doi.org/10.1002/2016gl069522, 2016.

Winter, B., 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, Nat. Hazards Earth Syst. Sci., 20, 1689–1703, https://doi.org/10.5194/nhess-20-1689-2020, 2020.

Woodward, M., Kapelan, Z., and Gouldby, B.: Adaptive Flood Risk Management Under Climate Change Uncertainty Using Real Options and Optimization, Risk Anal., 34, 75–92, https://doi.org/10.1111/risa.12088, 2013.

Woodward, M., Gouldby, B., Kapelan, Z., and Hames, D.: Multiobjective Optimization for Improved Management of Flood Risk, J. Water Res. Pl, 140, 201–215, https://doi.org/10.1061/(asce)wr.1943-5452.0000295, 2014.

Wu, W., Westra, S., and Leonard, M.: Estimating the probability of compound floods in estuarine regions, Hydrol. Earth Syst. Sci., 25, 2821–2841, https://doi.org/10.5194/hess-25-2821-2021, 2021.

Wyncoll, D. and Gouldby, B.: Integrating a multivariate extreme value method within a system flood risk analysis model, J. Flood Risk Manag., 8, 145–160, https://doi.org/10.1111/jfr3.12069, 2013.