

Response to Reviewers

Title: Disentangling Atmospheric, Hydrological, and Coupling Uncertainties in Compound Flood Modeling within a Coupled Earth System Model

Author Response 1st revision

Reviewer 1

Reviewer Comments:

This paper uses an earth systems modeling framework (1) to simulate compound flooding from Hurricane Irene with different processes represented (e.g., components activated) (2) to investigate the uncertainties that meteorological outputs from an atmospheric model introduce on flooding, and (3) to study the sensitivity of Hurricane Irene flooding under different hydrologic conditions. The authors find that coupling different model components (river, land, ocean) improve the model's ability to generate flooded area from Hurricane Irene (compared to satellite imagery). However, the E3M tends to overestimate flooding overall and requires the use of a computationally expensive local ocean model to resolve the surge inundation. The authors show that the contribution of the uncertainty of the meteorological inputs (via an atmospheric model) is large and that antecedent soil moisture conditions impact the river discharge.

I commend the authors on their analysis and think there are interesting results that inform complexities that arise with modeling hurricane flooding. I particularly enjoyed the analysis of the hydrologic conditions/drivers. The article covers topics relevant to NHESS and could be considered for publication after major revisions. I have included some general and specific comments for the authors.

Author Response:

We appreciate the reviewer for the critical assessment of our work. In the following we address your comments point by point. Our responses and all changes in the revised paper are marked in blue. If you have further questions or concerns, please let us know. Although NHESS does not allow us to share our revised manuscript at this stage of the review process, we provide excerpts throughout the response to help illustrate the changes. We will provide the revised manuscript when invited.

Major concerns:

R1C1:

The gaps are not clearly stated so it is hard to understand the novelty/contribution of this work in the context of existing literature. Overall, this paper seems to focus on how E3Ms new components can be used to study different things including model coupling uncertainty, meteorologic uncertainty, and the impact of different hydrologic conditions on river discharge. These analyses have been done more in-depth in previous studies which the authors could discuss and cite (e.g., Munoz et al. (2022); Gori et al. (2020); Eilander et al. (2023); Bermudez et al. (2021)). Is the focus of the study on improving modeling of compound riverine and coastal flooding compared to regional models (Line 11, 54)? Is it on uncertainties

– model coupling, meteorological, and initial hydrologic conditions (Line 59, 91)? The layout of the introduction could be improved with more relevant citations for the issues the paper is addressing.

Author Response:

Thank you for the insightful comment. We recognize the extensive studies on compound flooding (CF) uncertainty analysis in the literature (Abbaszadeh et al., 2022; Xu et al., 2023; Muñoz et al., 2022; Muñoz et al., 2024), typically based on single or dual-coupled regional/local river and coastal models. While these efforts are highly valued, our research explores a different venue by implementing global-scale regional-refined Earth System Models (ESMs) to simulate complex, multivariate extreme events.

Although ESMs are likely less accurate than regional/local models, they can offer distinct insights about CF due to several prominent features. E3SM, for example, leverages its exascale computing capabilities, provides a tightly coupled framework of multiple Earth system components (i.e., atmosphere, land, river, and ocean), and resolves inherent physical processes, and also features interactive coupling among all model components.

Although a recent preprint (Zhang & Yu, 2024) has also demonstrated the importance of using a fully coupled ESM framework for CF modeling, our recent research (Feng et al., 2024) has performed the first integration of new features in E3SM, including two-way land-river-ocean coupling, coastal refined unstructured meshes in atmosphere, land, and river models, and 2-dimensional barotropic ocean models with wetting and drying. It is important to note that some of these features are relatively new in the domain of ESMs. Alongside model development efforts, it is crucial to evaluate the capability and uncertainty of ESMs in the context of CF modeling. Our developed framework is essential for a comprehensive evaluation of atmospheric, hydrological, and coupling uncertainties.

Uncertainty propagation in large-scale ESMs remains an important topic in Earth system modeling community. This complexity is further exacerbated by the two-way land-river-ocean coupling, where downstream models can affect upstream models at high temporal frequencies. Such uncertainty propagation has not been adequately addressed, particularly for extreme events, partly due to the lack of a fully coupled framework that integrates all model components at high resolutions and the high computational cost of running ESM ensembles. Even the study of Muñoz et al. (2024), which inspires our research in its preprint stage, does not comprehensively address atmospheric uncertainty. Although their uncertainty analysis includes a scenario that turned off rainfall, their study relies on a single set of external atmospheric forcing (i.e. ERA5) and does not account for backpropagating uncertainty from two-way coupling. Specifically, we have conducted ensemble atmosphere model simulations to assess the propagation of atmospheric uncertainty through the inherent systems.

Moreover, our study compares atmospheric uncertainty with the coupling uncertainties at both land/river and river/ocean interfaces. While it is true that such coupling techniques have been developed elsewhere, as detailed in R2C2, it remains unclear whether coupling uncertainty is significant compared to atmospheric uncertainty, due to a lack of direct comparisons.

Additionally, we address hydrological uncertainty by running a large ensemble of simulations with perturbed atmospheric forcing and other hydrological parameters, including Antecedent Soil Moisture Conditions (AMC) and two runoff generation parameters. This large ensemble enables the application of our newly proposed machine learning method and the related SHAP method to understand the relative

contributions of hydrological factors. While the importance of AMC has been documented by Deb et al. (2023), none of these studies have demonstrated the role of underlying hydrological factors under different AMC conditions. Bilskie et al. (2021) and Santiago-Collazo et al. (2024) highlighted that the CF impact may be exacerbated if the CF event is preceded by antecedent rainfall, but their studies implemented the rain-on-grid method and did not account for the complexity hydrological processes, such as soil moisture, infiltration, or evapotranspiration. We included this discussion in the revision.

With our ensemble run, we discovered that AMC plays a significant role, resulting in more than twice the inundated area if it co-occurs with a wetter TC. The methodologies and findings together underscore the novelty of our study, which would not have been possible without the new E3SM framework. More importantly, our research sheds light on future directions and supports the ongoing development of more comprehensive model coupling methods.

In response to this comment, we have restructured the introduction, which begin with a more in-depth literature review on the CF studies, followed by a discussion of recent advancements in ESM model implementation and its associated uncertainties. We elaborated the existing gaps and research questions that need to be addressed. Additionally, we have more explicitly elaborated on the scope, objectives, and novelty of this study. Please see below the updated introduction starting from the second paragraph.

“Modeling CF is inherently challenging because it is triggered and governed by the interactions of processes within multiple Earth system components, including atmosphere, land, river, and ocean, as well as the associated uncertainties (Xu et al., 2023). Traditional CF modelling is typically based on one or multiple coupled hydrological models (Feng et al., 2022; Ikeuchi et al., 2017), hydraulic models (Bakhtyar et al., 2020; Bermúdez et al, 2021; Gori et al., 2020b) and hydrodynamic coastal/ocean models (Bennet et al., 2023; Kerns & Chen, 2023; Xiao et al., 2021; Ye et al, 2020) at local, regional and global scales. Over time, more sophisticated methodologies have been developed to enhance CF modeling. These include combined statistical-numerical modeling approaches (Olbert et al., 2023), deep learning (Feng et al., 2023a; Muñoz et al., 2021), data assimilation (Muñoz et al., 2022), reduced-physics ocean models (Eilander et al., 2023; Leijnse et al., 2021), new compound inundation models (Santiago-Collazo et al., 2024), and two-way river-ocean model coupling (Bao et al., 2022; Bao et al., 2024; Feng et al., 2024; Zhang et al., 2024). The CF modeling uncertainties can be sourced from model structures, parameters, input data, boundary and initial conditions (Abbaszadeh et al., 2022; Beven et al., 2018; Fan et al., 2021). These uncertainties may also cascade through the system (Meresa et al., 2021; Hasan Tanim & Goharian, 2021) and their contributions change dynamically over time (Muñoz et al., 2024).

Emerged recently is the use of fully coupled Earth System Models (ESMs) to simulate compound flooding (Feng et al., 2024; Zhang & Yu, 2024). One significant advantage of ESMs is the integration of multiple earth system components in a single, tightly coupled framework that does not necessarily requires external forcing data. Such a framework allows for predictive understanding of multi-scale flow processes and their interactions with other relevant processes involving heat, energy, biogeochemical and sediment transport, as well as their impacts on Earth’s climate (Ward et al., 2020). Feng et al. (2024) performed the first fully-coupled ESM simulation for CF using the Energy Exascale Earth System Model (E3SM), by integrating recent advancements in E3SM including regionally refined unstructured meshes for atmosphere, land/river and ocean components in the global domain (Deb et al., 2024; Feng et al., 2022),

two-way online land-river-ocean coupling (Xu et al., 2022b; Feng et al., 2024), and a 2-dimensional (2D) barotropic ocean model (Lilly et al., 2023).

While state-of-the-art ESMs are being implemented to simulate local extremes, this advancement can inevitably introduce additional uncertainties. Compared with regional simulations using prescribed atmospheric forcing derived from observation or bias-corrected reanalysis datasets, the model simulated atmospheric forcings are more uncertain (Hersbach et al., 2020). Atmospheric forcing has critical impacts on the flood simulation (Cloke and Pappenberger, 2009; Hjelmstad et al., 2021; Xu et al., 2025). The river discharge intensity, storm surge levels, and CF inundation extents are directly influenced by the TC's track and intensity, as well as the rainfall rate and timing (Gori et al., 2020a; Pappenberger et al., 2005; Zhong et al., 2010). These factors are the primary drivers of the riverine and coastal flooding dynamics. The uncertainty originated from atmospheric forcings would propagate to land, river, and ocean components through the multi-component framework (Deb et al., 2023; Blanton et al., 2020; Joyce et al., 2018). Likewise, the hydrological uncertainties in the land and river components (Giuntoli et al., 2018; Feng et al., 2023b) and the model coupling schemes can also propagate and even amplify. Typically, the cascading meteorological uncertainty is handled by the ensemble approach (Hamill et al., 2011; Villarini et al., 2019). Multiple realizations of a TC event with perturbed initial conditions and/or model physics represent a range of scenarios that evolve differently based on the dynamics of the models (Blanton et al., 2020). However, the cascading meteorological uncertainty has not been systematically estimated for CF modeling (Abbaszadeh et al., 2022; Xu et al., 2023). It remains unclear whether such uncertainty will amplify or diminish when constrained by the physical processes inherent in ESMs.

The cascading uncertainty becomes even more complex with two-way interactive model coupling. In online two-way coupling, a downstream model, while receiving data from its upstream component, sends back real-time computed information at predefined time intervals, enabling bi-directional data exchange. For instance, a river model may send floodplain inundated water flux to the land surface model (Xu et al., 2022b). Similarly, an ocean model can provide its modeled water levels (Bao et al., 2022; Bao et al., 2024; Feng et al., 2024) or velocities (Zhang et al., 2024) to the river model. While other uncertainties have been extensively discussed (Camacho et al., 2015; Feng et al., 2019; Muñoz et al., 2024; Willis et al., 2019), the uncertainty relevant to the two-way model coupling has rarely been explored because the coupling capabilities have only recently been developed. Questions are raised regarding the role of model coupling and the magnitude of related uncertainty compared to meteorological uncertainty, especially given the characteristic spatiotemporal scales invoked in land-river and river-ocean coupling. Addressing these questions is critical to refining the performance of interactively coupled ESMs, which is essential for achieving a more comprehensive understanding of the complex interactions and uncertainties associated with CF simulations. Moreover, assessing the enhancements provided by the two-way coupling schemes sheds light on the application of these couplings in future scenarios.

Furthermore, the cascading uncertainty changes with the variability and complexity of hydrological drivers represented in models, because these factors are critical for determining how precipitation is partitioned into runoff and infiltration. As rainfall initially infiltrates the soil, subsurface runoff moves slowly through the soil layers. When the rainfall intensity exceeds the soil's absorption capacity, saturation-excess water leads to surface runoff. The rate of infiltration, which determines the balance between surface and subsurface runoff, is influenced by soil properties, antecedent moisture conditions (AMC) (Ivancic and Shaw, 2015), and land cover types. The runoffs are then routed through river networks, resulting in high river discharge (Fig. 1) (Bevacqua et al., 2020). Understanding the hydrological drivers, including the

sensitivity of flood responses to various conditions such as [different](#) AMC and rainfall scenarios, is crucial (Tramblay et al., 2010). These factors provide key insights for predicting different flood scenarios (Miguez-Macho and Fan, 2012; Schrapffer et al., 2020). In particular, AMC plays a critical role in the generation of peak runoff and modulating riverine flooding characteristics during heavy precipitation events (Berghuijs et al., 2019; Nanditha and Mishra, 2022). A saturated AMC can significantly amplify flood impacts compared to drier conditions. The relative importance of rainfall and AMC varies depending on the watershed area. Soil moisture becomes a more dominant factor in larger watersheds (Ran et al., 2022). However, the role of these hydrologic drivers in cascading uncertainties sourced from atmospheric forcing has not been thoroughly investigated in the context of CF, partly due to the absence of a tightly coupled modeling system (Jalili Pirani and Najafi, 2020) [or insufficient investigation into hydrological processes](#) (Lin et al., 2024). [Although Bilskie et al. \(2021\) and Santiago-Collazo et al. \(2024\) highlighted the critical consequence if CF is preceded by an antecedent rainfall event, their implementation of the rain-on-grid method does not account for the hydrological processes, such as runoff generation. Addressing these processes would require a detailed hydrological model or land surface component of ESMs. The full coupled E3SM provides a feasible framework for quantifying the hydrological uncertainties in the CF modelling.](#)

The above-mentioned uncertainties are complicated but must be carefully evaluated for ESMs as they will be more frequently applied for CF simulations in the context of climate change. This study focuses on exploring and disentangling the atmospheric, hydrological, and coupling uncertainties of coastal CF modeling within the coupled E3SM framework. We first provide a comprehensive description of the physical processes during a TC-induced CF event. We then evaluate the model coupling uncertainties and the cascading meteorological uncertainty using a simulation ensemble of a specific TC event. Using the atmospheric ensemble as a basis, we generated an expanded ensemble [and proposed a new machine learning approach](#) to analyze the relative contributions of different hydrological drivers to CF and how these contributions affect the accuracy and reliability of CF simulations [over time](#). Finally, various hydrological and meteorological scenarios are used to delineate a spectrum of plausible CF outcomes in the designated region.”

R1C2:

The methods section was unclear at times. This work refers to previous papers that describe the modeling framework, validation, and application. However, there is not enough information provided so that the paper is readable by someone in the field. For example, in Section 2.1, are these components 1D or 2D? What are the spatial and temporal resolutions? How have the different model components been used in previous studies and/or validated? By improving the description of the different model components in the paper, it would help readers understand the limitations of the different analysis. For example, if the river model uses a macroscale inundation scheme (Line 284) that should be stated earlier on.

[Author Response:](#)

[We appreciate the reviewer’s suggestion. In the revised manuscript, we have moved the paragraph in the previous introduction \(L49-58\) that describes E3SM to the methodology section and add more details of the new E3SM framework and methods developed in our previous studies. The macroscale inundation scheme is also mentioned in the Methodology section. Please see below the updated Section 2.1.](#)

“This study uses a recently developed configuration in the Energy Exascale Earth System Model (E3SMv2) (Feng et al., 2024). E3SM represents a significant advancement in Earth System modeling (Golaz et al., 2019, 2022). As a state-of-the-science model, E3SM features a tightly integrated multi-component framework that supports dynamic exchanges and propagation of information across its different components. Additionally, several other developments have been recently implemented to further improve the modeling of coastal extremes, including the introduction of high-resolution regional-refined unstructured meshes in global river models (Feng et al., 2022; Liao et al., 2022, 2023a, b), the implementation of interactively coupled land-river-ocean models (Xu et al., 2022b; Feng et al., 2024), and the global tide model with wetting and drying schemes in the ocean component (Barton et al., 2022; Pal et al., 2023). Compared with regional models that may provide more detailed inundation at the street level (Costabile et al., 2023; Ivanov et al., 2021), E3SM excels at coupling processes across various earth system components. This capability is crucial for capturing the complex responses of earth systems to climate change and projecting climate-driven flood hazards.

The new E3SM configuration (hereafter “E3SM coastal configuration”) integrates the global three-dimensional (3D) E3SM atmospheric model (EAM), one-dimensional (1D) E3SM land model (ELM), 1D E3SM river model MOSART (MOSART: Model for Scale Adaptive River Transport), and two-dimensional (2D) barotropic version of E3SM ocean model MPAS-O (MPAS-O: Model for Prediction Across Scales ocean model) (Fig. 2a). This configuration uses three different variable-resolution meshes to improve the E3SM’s capability in modeling coastal processes. The EAM mesh features a global resolution of 100 km, with enhanced refinement to about 25 km over the North Atlantic Ocean. Both ELM and MOSART use a land mesh with a coarse resolution of 60 km globally, which is further refined to 30 km across the contiguous US and to 3 km within the Mid-Atlantic watersheds. The MPAS-O mesh offers the highest resolution of 250 m along the US East Coast, specifically designed to capture estuary dynamics, with a broader global resolution of around 1 km everywhere else. The global bathymetry data of MOSART and MPAS-O are sampled from the 90-m HydroSHEDS digital elevation model (DEM) (Lehner et al., 2008) and the 450-m GEBCO dataset (IOC and IHO, 2020), respectively. The river networks and flow directions are derived using HexWatershed which performs hybrid depression filling and breaching stream burning for river routing in unstructured meshes (Liao et al., 2022, 2023a, b).

The novel two-way hydrological coupling between land and river components enables E3SM to capture the infiltration of inundated river water in floodplains and, subsequently, the enhancement of subsurface runoff and evapotranspiration from saturated floodplain soils (Xu et al., 2022b). The two-way river-ocean coupling was developed for E3SM to better represent the dynamic interaction between rivers and oceans, especially during CF events (Feng et al., 2024). This new approach allows for an accurate representation of coastal backwater effects and the mutual influences of river discharge and ocean sea surface height (SSH), providing a more realistic assessment of CF hazards (Feng et al., 2022).

Using the E3SM coastal configuration, we simulated Hurricane Irene, a TC event that occurred in August 2011 and had large flooding impacts across the Mid-Atlantic region (Fig. 2b). Irene led to significant riverine and coastal flooding in the Delaware River Basin (DRB) and Delaware River Estuary (DBE) due to concurrent intense precipitation and storm surge. Following Feng et al. (2024), an ensemble of 25 EAM simulations with perturbed model parameters were performed to reproduce Irene and associated meteorological outcomes. EAM is initialized from ECMWF Reanalysis v5 (Hersbach et al., 2020) in the

midnight of 26 August 2011. Atmospheric nudging is not applied. The EAM ensemble can reproduce the TC characteristics, including the storm track and intensity (see Appendix A in Deb et al. (2024)). These “prerun” EAM simulations were then prescribed within E3SM to drive the land, river, and ocean components. EAM (in data mode) provides atmospheric forcing to ELM and MPAS-O at a 15-min frequency. MOSART is interactively coupled with ELM and MPAS-O at the 1-hour interval via the E3SM coupler (Craig et al., 2012). The model outputs are archived at 15 minutes for EAM and hourly for ELM, MOSART and MPAS-O. We spun up ELM and MOSART from a 10-year historical simulation forced by Global Soil Wetness Projects version 3 (GSWPv3; Kim, 2017), and MPAS-O from 1-month simulation with the global tide model. MOSART was validated against the measured streamflow at 6 USGS gauges along the Delaware River main channel with averaged coefficient of determination (r^2) of 0.79 and Kling–Gupta efficiency (KGE; Gupta et al., 2009) of 0.84. MPAS-O was assessed for water level at 6 NOAA tidal gauges across the DBE, showing an averaged r^2 of 0.72 and root mean squared error (RMSE) of 0.41 m. Please refer to Feng et al., (2024) for a more detailed description of the E3SM configuration and the model evaluation.

Fluvial and coastal inundations are simulated in MOSART and MPAS-O, respectively. The riverine inundation in MOSART is simulated using a macroscale inundation scheme that assumes no between-cell flux exchange and estimates the inundation fraction using the elevation profiles within the grid cell (Luo et al., 2017; Yamazaki et al., 2011). Coastal inundation simulated on the MPAS-O inland mesh is aggregated onto the coarser MOSART mesh in the DRB. Within each MOSART grid cell, the inundation fraction is determined by the percentage of MPAS-O cells with a simulated water depth over 1 m. This value represents adjustments made to the MPAS-O inland bottom elevation data near the DRB coastline during the upscale sampling. Whenever there is a discrepancy between the inundation area from MOSART and MPAS-O in their overlapped cells near the coastline, the MPAS-O inundation is considered more accurate and will be used.

Here, the total simulated inundation extent of Irene is benchmarked against a 250-m resolution inundation extent dataset based on satellite imagery (Tellman et al., 2021). The dataset is aggregated onto the MOSART mesh for comparison. Within each MOSART cell, we compute the fraction of the observed inundation. The model performance is evaluated using flood metrics defined by Wing et al. (2017), including hit rate (HR), false rate (FR) and success index (SI)

$$HR = \frac{M_1 B_1}{M_1 B_1 + M_0 B_1}, \quad (1)$$

$$FR = \frac{M_1 B_0}{M_1 B_0 + M_1 B_1}, \quad (2)$$

$$SI = \frac{M_1 B_1}{M_1 B_1 + M_0 B_1 + M_1 B_0}, \quad (3)$$

where M and B are the pixels (or grid cells) from model simulations and benchmark data, respectively. The subscripts 1 and 0 represent wet (inundated) and dry cells, respectively. For all the three metrics, a score of 0 indicates poor performance, while a score of 1 represents perfect model performance. In our simulations, a wet cell is identified if the simulated inundation fraction is above a small unitless threshold of 0.02. This threshold minimizes the influence of cells that may only be marginally inundated—likely due to data and model uncertainties—thus ensuring a more reliable assessment of flood extent. The predicted flooded area (FA) is calculated by multiplying the flooded fraction by the corresponding cell area.”

R1C3:

The clarity of the manuscript could be improved. The authors should define terms and use them consistently throughout the paper including forcing, drivers, processes, components, responses, conditions, scenarios/experiments (for example Lines 78-81, Lines 210-220, Line 355, Line 371-380, Line 387). The authors should state in the abstract and early in the introduction that E3SM is global. It wasn't until Line 109 that this is mentioned making it unclear in the introduction how/why "regional models" are mentioned in comparison (Line 55-57, Line 391).

Author Response:

Thanks for the comment. We made sure the terminology is defined consistently in the revision. Hydrological drivers represent the surface runoff, subsurface runoff, infiltration and soil water storage, which are defined when they are first mentioned in the methodology section (Section 2.4): "The hydrological drivers we selected for uncertainty analysis include surface runoff, subsurface runoff, infiltration, and soil water storage."

We rephrased "hydrological processes" in previous L355 and "hydrological parameters" in previous L383 to "hydrological drivers".

"Forcing" refers to atmospheric forcing in the global E3SM and boundary forcing in regional models.

"Processes" is only used when referring to the general physical processes, such as "atmospheric, land and river processes" in L387.

"components" is typically used for atmosphere, land, river and ocean components in Earth System models (ESMs).

"Response" refers to the change in one variable as a result of a change in another variable. We rephrased "hydrological responses" in previous L197 and L459 to "hydrological drivers".

"Conditions" refer to a specific state or set of circumstances. For example, "meteorological condition" denotes the current state of the atmosphere, while "soil moisture condition" indicates the current level of moisture content in the soil. We rephrased "various hydrological conditions such as AMC and rainfall scenarios" to "various conditions such as different AMCs and rainfall scenarios".

"Scenario" refers to a set of physical conditions or a simulated situation used to explore the possible outcomes of various conditions. "Experiment" specifically refers to "numerical experiment". We double checked throughout the revised manuscript and made sure they represent their correct definitions.

In the revised introduction, we have added a new third paragraph that includes a brief description of the new E3SM framework. In this paragraph, we emphasized the simulations conducted on a global domain:

"Feng et al. (2024) performed the first fully-coupled ESM simulation for CF using the Energy Exascale Earth System Model (E3SM), by integrating recent advancements in E3SM including regionally refined unstructured meshes respectively in atmosphere, land/river and ocean components in the global domain".

R1C4:

The framing of the results and conclusions could be improved. I wonder if it would be more effective to narrow the scope of the manuscript, especially because it addresses so many issues and methods (models, structural equation, ANN). Some of my confusion with the results and discussion relate back to the lack of clarity in the gap/objective of the study. For example, Figure 4. results show that the hit rate is very low when the ocean is not resolved. The large-scale river model is not reliable for event-scale riverine flooding (Line 285). However, discharge and surge are key processes in compound flooding from hurricanes (Lines 35-40) and this is the event the model is being evaluated against.

Your results seem to support the need for resolving flood processes at finer resolutions (e.g., regional or smaller) (again Line 340-342) which E3M is not able to do efficiently due to computational constraints (Line 427). Why did you choose a hurricane to study? Would it be more interesting to look at the impacts of coupling, inputs, conditions, etc. on long-term simulations which the E3Ms seems more suited for (as the authors suggest in Line 435-438)?

Author Response:

We appreciate the suggestion and have revised the introduction to more specifically address the atmospheric, hydrological, and coupling uncertainties within E3SM. We emphasize the novel approach of using global ESMs, which effectively address atmospheric and hydrological uncertainties through a fully coupled framework using simulation ensembles and a detailed land surface model. Additionally, we highlight the importance of understanding how these uncertainties propagate and the related coupling uncertainties. Please refer to our response to R1C1 for the updated introduction.

The use of structural equation modeling (SEM) and ANN depends on the ensemble size. The latter requires a much larger ensemble, which is not feasible for global MPAS-O simulations at the current resolution. We have clarified this point in the revision: “structural equation modeling (SEM) is applied as a path analysis method (Wright, 1921) to trace the flow of data and uncertainty from the 25 ensemble members.” In the uncertainty analysis of the hydrological drivers, the expanded ensemble simulation uses the MPAS-O simulated water level as MOSART’s coastal boundary condition to reproduce the backwater effects, thereby allowing for the more accurate ANN and SHAP methods. This is mentioned in Section 2.4.

We apologize for any confusion. In L285, we did not mean that the large-scale river model is not reliable for event-scale riverine flooding. Instead, we aim to explain the slightly reduced performance of the two-way land-river coupling in Figure 5. This statement is rephrased to improve clarity: “While it is crucial for capturing hydrological processes at larger spatiotemporal scales, the two-way land-river coupling, building upon the macroscale inundation scheme, potentially makes the coupling less reliable for event-scale riverine flooding”.

The low hit rate when the ocean model is excluded is because the river model does not resolve coastal flooding. This was mentioned as “although MOSART is capable of simulating extensive riverine inundation in coastal regions, it cannot simulate the finer details of inundation immediately adjacent to the coastline (Fig. 4a), where coastal tide and storm surge play a significant role.”

In L340-342, we are trying to explain the different behaviors of river discharge Q and riverine inundation extent A_{river} . While Q is dominated by surface and subsurface runoff and A_{river} is affected by localized processes such as infiltration. Although a regional or smaller model at higher resolutions may improve the

flood simulation, our discussion here is not to compare models, which as the reviewer pointed out, is only mentioned in L427. To respond to the reviewer's concern about using a coarser ESM rather than a finer local model to simulate the coastal extremes, we highlighted the motivation in the revised introduction "One significant advantage of ESMs is the integration of multiple earth system components in a single, tightly coupled framework that does not necessarily requires external forcing data. Such framework allows for predictive understanding of multi-scale flow processes and their interactions with other relevant processes involving heat, energy, biogeochemical and sediment transport, as well as their impacts on Earth's climate (Ward et al., 2020)"

and in the methodology section

"Compared with regional models that may provide more detailed inundation at the street level (Costabile et al., 2023; Ivanov et al., 2021), E3SM excels at coupling processes across various earth system components. This capability is crucial for capturing the complex responses of earth systems to climate change and projecting climate-driven flood hazards."

Regarding the last point of this comment, ESMs are indeed suited for long-term simulations. The MOSART-alone simulation using the same mesh presented good performance in the Mid-Atlantic region over 1991~2019 (Feng et al., 2022). This study, on the other hand, particularly explores the E3SM's capability and uncertainties of simulating coastal extremes using the new configuration as its very first step. The computational constraint posed by the atmosphere model and the fine-resolution global ocean model is another reason why we cannot pursue a long-term simulation in the current stage. The 10-day simulation of MPAS-O alone takes over 6 hours on 640 CPUs. Thus, more computationally efficient schemes, such as local time stepping method (Capodaglio & Petersen, 2022; Lilly et al., 2023), are required to reduce the computational cost. We have added further elaboration to Section 4.4:

"Future work may focus on enhancing the computational efficiency and feasibility of MPAS-O, for instance, by implement advanced schemes such as local time stepping (Capodaglio & Petersen, 2022; Lilly et al., 2023) and/or developing a regional ocean model within the E3SM framework."

R1C5:

Line 19: The findings in the abstract are not very specific and it would be beneficial to add more detail. For example, what is the measure of the "effectiveness" of the model framework? Is it reproducing the flood extent from Irene?

Author Response:

We appreciate this comment and rephrased "demonstrate E3SM's effectiveness in handling multivariate flooding on the coast" to "demonstrate E3SM's capability in capturing compound flood extent near the coast, with a hit rate over 0.75".

R1C6:

Line 23: The authors state that flood damage could be tripled if Hurricane Irene was preceded by extreme AMC. I do not see this supported anywhere in the results. It would be more appropriate to state the results of changes in the flood hazard (e.g., discharge, flooded area) since these are modeled, not damages.

Author Response:

By “tripled flood damage”, we referred to Figure 11b, where the flooded areas in the worst-case scenario could increase to more than twice (~ 2.4) of the flooded areas in the best scenario and the actual Irene event. We apologize for the confusion and have rephrased this statement to

“For instance, the flooded area could increase by more than twice (~ 2.4) if Hurricane Irene was preceded by an extreme antecedent soil moisture condition (AMC).”

R1C7:

Line 61: What are the “critical” impacts?

Author Response:

Thanks for the comment. We briefly described the atmospheric forcing impact in the next sentence of the original manuscript (L62): “The water movement in terrestrial and aquatic environments during a TC is strongly influenced by the TC’s track and intensity (Pappenberger et al., 2005; Zhong et al., 2010).” In the revision, we rephrased this sentence to improve clarity:

“The river discharge intensity, storm surge levels, and CF inundation extents are directly influenced by the TC’s track and intensity, as well as the rainfall rate and timing (Gori et al., 2020a; Pappenberger et al., 2005; Zhong et al., 2010). These factors are the primary drivers of the riverine and coastal flooding dynamics.”

R1C8:

Line 110: What is the scale of the coastal refined meshes? Are they comparable size and resolution to “regional models”?

Author Response:

Thanks for the comment. The resolutions of the global EAM, ELM/MOSART and MPAS-O meshes have been added to section 2.1: “The EAM mesh features a global resolution of 100 km, with enhanced refinement to about 25 km over the North Atlantic Ocean. Both ELM and MOSART use a land mesh with a coarse resolution of 60 km globally, which is further refined to 30 km across the contiguous US and to 3 km within the Mid-Atlantic watersheds. The MPAS-O mesh offers the highest resolution of 250 m along the US East Coast, specifically designed to capture estuary dynamics, with a broader global resolution of around 1 km.” We note that the resolutions of the global meshes, even regionally refined, could still be lower than regional and local scale models.

R1C9:

Lines 130-135: It would be helpful to include what a perfect score is for each metric (e.g., 0 or 1).

Author Response:

Thank you for your suggestion. In the revised manuscript, We have included a brief explanation for the flood metrics.

“For all metrics, a score of 0 indicates poor performance, while a score of 1 represents perfect model performance.”

R1C10:

Line 138: Why did you choose a small threshold of 0.02? I’m assuming meters is the units.

Author Response:

This threshold is unitless, representing the minimum fraction of wet cell. Because we use the binary flood metrics to evaluate the model performance, a cell could be wet or dry even with a very small inundation fraction, especially considering the grid resolution of 5 km. To account for the uncertainty within the satellite data (see our response to R1C21) and the macroscale inundation scheme in the 5-km mesh (see the updated Section 2.1 under the response to R1C2, and our discussion about this uncertainty: “The macroscale inundation scheme may not be sufficient to represent water depth heterogeneity within each cell and the subgrid flow dynamics given the grid resolution of 5 km.”), this small threshold is able to eliminate the cell with a very small fraction of inundation potentially caused by the uncertainty.

This threshold is unitless and represents the minimum fraction of a cell considered to be inundated. We have chosen this minimal threshold in order to refine our binary flood metrics that categorize cells as either wet or dry. This approach is particularly pertinent and helps mitigate potential misclassifications in cell status due to the inherent uncertainties presented in the satellite data (see our response to R1C21), the macroscale inundation scheme and the 5-km mesh. The macroscale inundation scheme is now provided in the updated Section 2.1 under the response to R1C2. And its corresponding uncertainty is added to the result discussion: “The macroscale inundation scheme may not be sufficient to represent water depth heterogeneity within each cell and the subgrid flow dynamics given the grid resolution of 5 km.” We also elaborated the selection of this threshold in the revision.

“In our simulations, a wet cell is identified if the simulated inundation fraction is above a small unitless threshold of 0.02. This threshold minimizes the influence of cells that may only be marginally inundated—likely due to data and model uncertainties—thus ensuring a more reliable assessment of flood extent.”

R1C11:

Line 144-146: Why did you choose a depth of 1m to calculate the inundation fraction?

Author Response:

Thanks for the comment. As we explained in the following sentence, “this number is used to reflect adjustments for the MPAS-O inland bottom elevation bias relative to the actual data”. The bathymetry

data of the global MPAS-O is sampled from a global gridded bathymetric data set, the General Bathymetric Chart of the Oceans (GEBCO) (IOC and IHO, 2020). This upsampling results in roughly 1-m bias of the inland elevation near the DRB coastline. The inland elevation has to be adjusted to obtain the coastal inundation extent.

In the revision, we made further clarifications here:

“This value represents adjustments made to the MPAS-O inland bottom elevation data near the DRB coastline during the upscale sampling.”

In the discussions of the limitation, we highlighted the need for a more accurate global ocean model:

“Currently, MPAS-O is geared towards global simulations, but adapting it for regional use [with the merging of high-resolution local bathymetry data](#) could allow for more accurate and locally relevant flood simulations, integrating two-way land-ocean coupling to account for ocean water intrusion (Xu et al., 2024) and its effect on soil moisture along coastlines.”

R1C12:

Line 190: I think this should be flood hazards, not flood risks. I would check the use of risk throughout the paper as this is typically defined as the combination of hazard, exposure, and vulnerability.

Author Response:

We acknowledge that “risk” is broadly defined, encompassing flood hazard, exposure, and vulnerability (Kron, 2005). Specifically, flood hazard and exposure risks represent the frequency of flooding events and the extent of human exposure to these events, respectively (Feng et al., 2023). In the revised manuscript, the term “risk” is only used in accordance with its general definition. We checked throughout the manuscript to ensure that the use of “hazard” or “exposure” is consistently accurate and precise. We made the revisions below following this comment.

L119: “CF hazard risks” to “CF hazards”

L191: “flood risks” to “flood exposure”

L246: “flood hazard risk assessment” to “flood exposure risk assessment”

L380 and L407: “flood risks” to “flood exposure”

L414: “increasing risks to coastal residents” to “increasing exposure risks to coastal residents”

R1C13:

Line195: How did you select the 5 ensemble members? How do you define a reasonable spread?

Author Response:

The five ensemble members were chosen from the full set of 25, selected to represent a roughly even distribution of the accumulated discharge and precipitation during Hurricane Irene, which should also

result in a wide spread of peak discharge along Delaware River and riverine flood extent (previous Figure S2, now Figure S3). In the revised manuscript, we clarified our selection in the main text:

“selecting 5 ensemble members that represent a roughly even distribution of river discharge and precipitation during Hurricane Irene (Figure S1 and S2)”.

Additionally, we have included a new supplementary figure (Figure S2) to further clarify the basis of our selection.

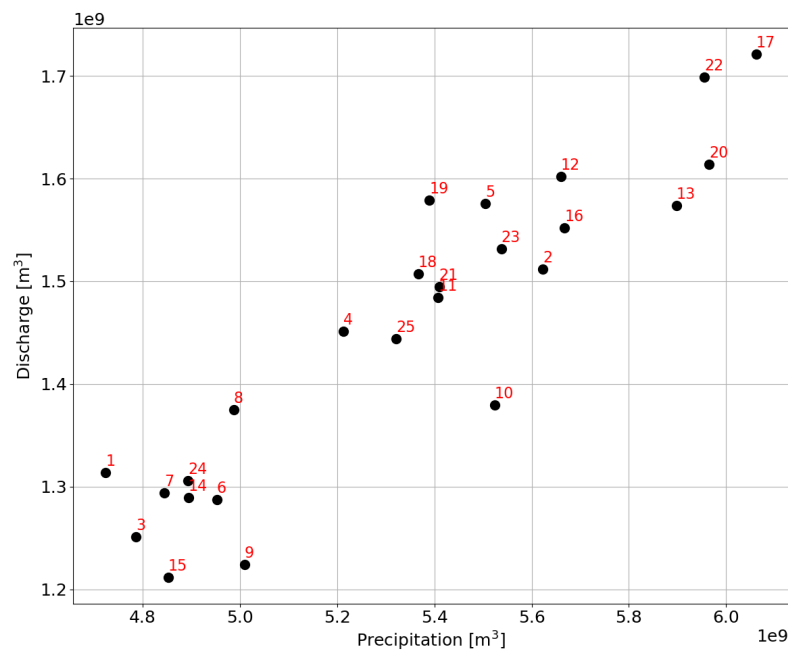


Figure S2 The accumulated river discharge against precipitation during Irene for the 25 ensemble simulations.

R1C14:

Lines195-200: What data did you use to get trends in historical soil moisture? To better help the reader understand the range you are looking at, can you state the basin averaged soil moisture range you are using in the simulations? From the supplemental, it looks like the range is 0.065 to 0.087.

Author Response:

Thanks for the comment. The historical soil moisture data is obtained from a historical ELM simulation, as stated at L198: “basin-averaged soil moisture during hurricane seasons from 2005 to 2011 as modeled in a historical ELM simulation (Fig. S1).”

We have added more clarification in the caption of Figure S3 (previous Figure S1):

“Time series of soil moisture averaged in Delaware River Basin (DRB) obtained from a historical ELM simulation forced by the Global Soil Wetness Projects version 3 (GSWPv3) climate forcing data set (Kim, 2017) running on the same mesh as Irene simulations.”

In the revised main text, we stated the basin averaged soil moisture range as $0.067 \sim 0.087 \text{ kg m}^{-2}$.

R1C15:

Line 200: Since these two parameters are reported in the results, you should provide more detail on what they represent in the text.

Author Response:

Agree. In the revision, we have clarified:

“Two parameters in ELM (f_{over} and f_{drain} , that determine the runoff generation are considered, where f_{over} determines the saturation fraction, i.e. how much surface runoff is generated from precipitation, and f_{drain} controls the amount of subsurface runoff (see Appendix A for a detailed definition).”

R1C16:

Line 205: What do you mean by widely distributed impacts?

Author Response:

By “widely distributed impacts”, we refer to the selection of each parameter covering a broad range of possible scenarios to quantify their effects. We apologize for the confusion and rephrase this sentence to

“The peak discharge observed in the main channel of the Delaware River and the extent of riverine flooding indicate a broad range of possible scenarios with the perturbed parameters of atmospheric conditions, AMC, and the parameters f_{over} and f_{drain} (Fig. S2 and S4~S6).”

R1C17:

Line 226-237: Some of the statements about the model performance (e.g., “effectively” and “successfully” simulates compound riverine and coastal inundation) might be better suited in the discussion or should be given after you have stated the results that can support these findings.

Author Response:

Agree. We modified the first two sentences to present the capabilities and limitations of the models in a balanced way, focusing on what the results show rather than preemptively describing the models as successful or effective.

“In Experiment 3, the coastal configuration of E3SM employs the coupled MOSART and MPAS-O models to simulate compound riverine and coastal inundation (Fig. 4). The results indicate that MOSART can predict riverine flooding along the lower Delaware River and its upstream tributaries.”

R1C18:

Line230-231: Are these inundated cells in the upstream parts of the basin fluvial or pluvial flooding? It is hard to tell if they are located along the river network since that is not shown on Figure 4.

Author Response:

Thanks for the comment. In this study, we do not consider pluvial flooding due to the coarse resolution of the large-scale MOSART model, which is noted in the caption of Figure 4 as 'MOSART simulated riverine inundation.' We prefer to use “riverine flooding” over “fluvial flooding”, typically because the latter is usually discussed alongside pluvial flooding. In response to this comment, we have updated Figure 4 to include major river channels in the Delaware River Basin (DRB), which are used for MOSART’s river network delineation. This addition shows that the inundated areas are primarily along the main channel of the Delaware River.

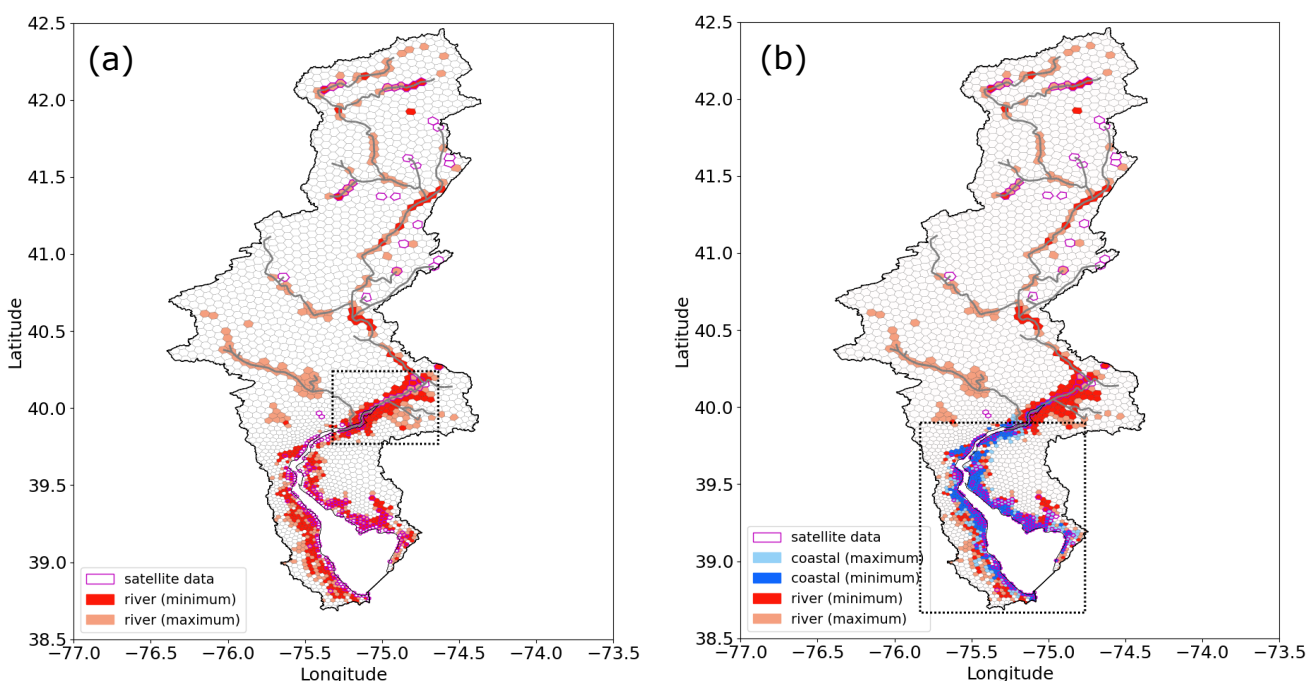


Figure 4: (a) MOSART simulated riverine inundation (red) against satellite-measured inundation (magenta box). The black dashed box highlights the lower Delaware River reach. (b) E3SM simulated riverine (red) and coastal (blue) total inundation against satellite data (magenta box). The black dashed box represents the coastline of DBE where extensive coastal inundation occurred. In both panels, dark and light colors represent the minimum and maximum inundated extent from the ensemble simulations, respectively. The gray lines are the major river channels.

R1C19:

Line 240: Is the predicted flooded area just an indicator of flooding from the model output or is this measurement used in the comparison with the satellite data?

Author Response:

Thanks for the comment. The predicted flooded area is an indicator of the flood hazard, computed directly from the model’s output, which definition has been added to the methodology section (Section 2.1): “The

predicted flooded area (*FA*) is calculated by multiplying the flooded fraction by the corresponding cell area.” We do not use flooded area for comparison with the satellite data. Due to the uncertainty in satellite data (see our response to R1C21), these data are only used to benchmark the model in terms of the binary metrics.

R1C20:

Figure 4. It would be helpful to spell out hit rate, false rate, and success index in the caption. It would also be helpful to include 1 sentence that lists what the Experiments are again based on the components land, river, ocean.

Author Response:

The caption of Figure 5 has been updated:

“Figure 5: Flood metrics of hit rate (*HR*), false rate (*FR*), success index (*SI*) and flooded area (*FA*) used to compare riverine flooding in Experiments 1~3 and the combined riverine and coastal flooding in Experiment 3. Experiment 1 and 2 include land and river components, while Experiment 3 runs all land, river and ocean components (Table 1). Whiskers extend to 1.5 times the interquartile range from the quartile boundaries.”

R1C21:

Line 245: What is the error/uncertainty in the satellite data you are comparing to? Does it perform well in capturing pluvial, fluvial, and coastal flooding from Hurricanes which generally have significant cloud cover? It would be helpful to discuss this in the context of your model results.

Author Response:

We appreciate the reviewer’s suggestion and have added more discussions on the uncertainty of the satellite data in the revised section 3.1:

“The overestimation in *FR* is likely due to the bias in the MODIS satellite data, the macroscale inundation scheme in MOSART and the MOSART mesh resolution. The flood extent dataset (Tellman et al. 2021) is known to result in an underestimation of the actual flooding area due to the cloud cover removing technique (Zhang & Yu, 2024). Its fidelity further decreases towards upstream due to the vegetation covers (Sexton et al., 2013).”

R1C22:

Figure 276: Do you mean the influence of the two-way coupling on flooded area (*FA*) is more significant... not discharge?

Author Response:

We apologize for the confusion. This sentence is rewritten:

“The impact of the new two-way coupling schemes on accurately capturing the flood extent (Fig. 5) is less significant compared to their effect on modulating the discharge near the river outlet (Fig. 6a),”

R1C23:

Line 293-294: Are you saying that the small variations in the soil water storage is because it can't exceed the maximum soil storage capacity? What are the ranges of capacity of the soil?

Author Response:

Yes, the small variation in soil water storage is a result of the soil saturation capability, especially when the rainfall intensity during TCs exceeds the soil's infiltration rate, resulting in water inputs surpass what the soil can absorb. This process depends on a few factors, including the antecedent moisture content of the soil, soil type and rain intensity and duration. In the Delaware River Basin, the soil storage capacity varies significantly across the basin due to a diversity of soil types and their associated physical properties. Soil types range from sandy loams to clay loams. The available water capacity of these soils generally ranges from about 100 mm to 250 mm according to USDA soil surveys. In the response, we elaborate our statement to improve the clarity:

“The minimal variation in soil water during a TC event is likely because the soil reaching its saturation capacity, especially when rainfall intensity exceeds the soil's infiltration rate.”

R1C24:

Line 297: Do you consider sea level rise in this modeling framework?

Author Response:

In the current modeling framework, sea level rise (SLR) is not considered in MPAS-O, which functionality is still under development. In the revised discussion, we acknowledged this limitation:

“The current version MPAS-O does not account for sea level rise (SLR) (Kulp and Strauss, 2019). As a key driver of coastal flooding, SLR can also interact with future AMC scenarios (Deb et al, 2023) and further amplify the flood hazard risks, which should be considered for more accurate CF risk assessment.”

R1C25:

Figure 8: redefine all variables so caption can be understood standalone

Author Response:

The caption of Figure 8 has been updated: “Figure 8: CV (light bars) and NMAD (dark bars with black margins) computed from the simulation ensembles for the variables selected in Section 2.3, including precipitation (*precip*), surface runoff (Q_{sur}), subsurface runoff (Q_{sub}), infiltration (Q_{infl}) and soil water storage (Q_{soil}), river discharge (Q), riverine inundation area (A_{river}), coastal inundation area (A_{ocean}),

sea surface height (SSH), air pressure (P_{air}) and wind speed (U_{wind}). Red and blue bars indicate riverine and coastal flood drivers, respectively.”

R1C26:

Figure 10: It would be helpful to have hyetograph and hydrographs of Hurricane Irene, even in the supplemental, to provide context on the results shown in Figure 10.

Author Response:

Thanks for the suggestion. The hyetograph and hydrograph have been added to Supplementary Figure S11.

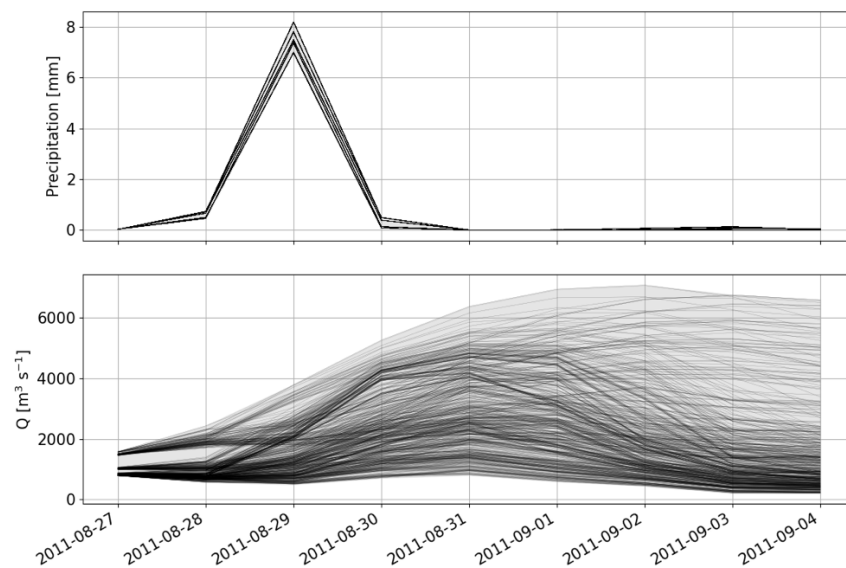


Figure S11: Time series of DRB-averaged precipitation and river discharge (Q) at the outlet corresponding to Figure 10.

R1C27:

Does flooded area shown in Figure 5 included the ocean model and then in Figure 9 it is only the river? I would explicitly state this because it is unclear.

Author Response:

Yes, in the expanded ensemble run, the ocean model is excluded as we focus on the hydrological uncertainties specifically and the running the global high-resolution MPAS-O for a large ensemble is extremely computationally expensive. Thus, as shown in Figure 9, we selected riverine inundation river (A_{river}) as the response variable, while in Figure 5, the flooded area (FA) of MOSART and MPAS-O in Experiment 3 included both river and ocean modelled flooding.

In the revised Figure 5 caption, we explicitly specified that: “MOSART only considers riverine inundation, while MOSART+MPASO accounts for both riverine and coastal inundation.”

In the revised Figure 9 caption, we clarified that: “The coastal inundation area (A_{ocean}) is not considered from this analysis as MPAS-O is excluded from the expanded ensemble simulations.”

R1C28:

Line 355: Did you look at how the time-varying impacts of hydrological processes on compound flooding, not just river discharge?

Author Response:

Thanks for the comment. We exclusively focus on the impact of hydrological processes on riverine flooding in terms of both river discharge (Q) and inundation area (A_{river}). This is noted in Section 2.4 (P7L190), “we focus exclusively on riverine flooding in this analysis”. The decision was made to “avoid the substantial computational burden associated with MPAS-O”. However, to account for the backwater effects on riverine flooding during compound flooding, we did impose the “MPAS-O simulated water level as the coastal boundary condition of MOSART” in the large ensemble simulations (Feng et al., 2022), and in the discussions of limitations (Section 4.4), we acknowledge that “One limitation of this study is the exclusion of the ocean model in the expanded ensemble simulations, primarily due to the high computational costs associated with running the global MPAS-O.”

In response to this comment, we further elaborated this point to increase the clarity:

“One limitation of this study is the exclusion of the ocean model in the expanded ensemble simulations, primarily due to the high computational costs associated with running the global MPAS-O. As a result, the influence of hydrological processes on compound riverine-coastal flooding could not be evaluated using the proposed ANN approach.”

R1C29:

Figure S6: Where does the 625 number come from?

Author Response:

As described in Section 2.4.1, we selected 5 ensemble members from the atmosphere model runs, 5 AMC scenarios, 5 values for f_{over} and f_{drain} , respectively, resulting in a total of 625 ensemble runs. We apologize for being unclear at this point, which has been clarified in the revision in Section 2.4.1: “Using the selected 5 ensemble members, 5 AMC scenarios, 5 values each for f_{over} and f_{drain} , we performed a total of 625 ($625 = 5 \times 5 \times 5 \times 5$) ensemble simulations.”

R1C30:

Figure S7: What does the Irene color mean?

Author Response:

Sorry for the typo. The caption of Figure S8 (previous Figure S7) has been corrected to:

“The plausible outcomes of inundated extent in DRB with the five colors representing the minimum, 25th percentile, median, 75th percentile and maximum inundation from the 625 ensemble runs. The magenta boxes represent the flood extent of the satellite data.”

Reviewer 2

Reviewer Comments:

This article presents a compound flood assessment of Hurricane Irene in the Delaware River Basin. The authors used a new two-way/tightly-coupled approach to quantify the uncertainty of different aspects without the modeling approach, such as the environmental forcing and coupling technique.

Author Response:

We appreciate the reviewer for the critical assessment of our work. We have carefully addressed the reviewer's suggestions as follows. Excerpts of the revised manuscript are provided to explain our responses to the review comments, as NHESS does not allow us to share our revised manuscript at this stage of the review process.

General Comments:

R2C1:

First, the authors need to highlight better the novelty of the paper since its current version is difficult to identify clearly. For example, they say that their main goal is to assess the uncertainties in the compound flood model, but this has been done before by Muñoz et al. (2024). However, the authors mention this gap (L69-72) but fail to mention the work by Muñoz despite citing them for another reason in the discussion. Another example is their claim that the compound flood needs to be redefined to include AMC and not just TC impacts. However, Bilske et al. (2021) and Santiago-Collazo et al. (2023) simulated CF from a TC but were preceded by an antecedent rainfall event, which affects the AMC. My point here is that the manuscript needs to go into an intensive literature review process so it can better define the knowledge gaps and cite the appropriate references. In my judgment, presenting the manuscript novelty as an uncertainty assessment of this specific model they used is not enough to impact the science community, as we expect with any manuscript in top-tier journals like this one.

Author Response:

We appreciate the thoughtful comment, which made us realize that the structure of our introduction may have led to the reviewer's assertion that the manuscript's novelty in assessing the uncertainty of this specific modeling framework has limited impact on the scientific community. We recognize the extensive studies on compound flooding (CF) uncertainty analysis in the literature (Abbaszadeh et al., 2022; Xu et al., 2023; Muñoz et al., 2022; Muñoz et al., 2024), typically based on single or dual-coupled regional/local river and coastal models. While these efforts are highly valued, our research explores a different venue by implementing global-scale regional-refined Earth System Models (ESMs) to simulate complex, multivariate extreme events (Fig. R1).

Although ESMs are likely less accurate than regional/local models, they can offer distinct insights about CF due to several prominent features. E3SM, for example, leverages its exascale computing capabilities, provides a tightly coupled framework of multiple Earth system components (i.e., atmosphere, land, river,

and ocean), and resolves inherent physical processes, and also features interactive coupling among all model components.

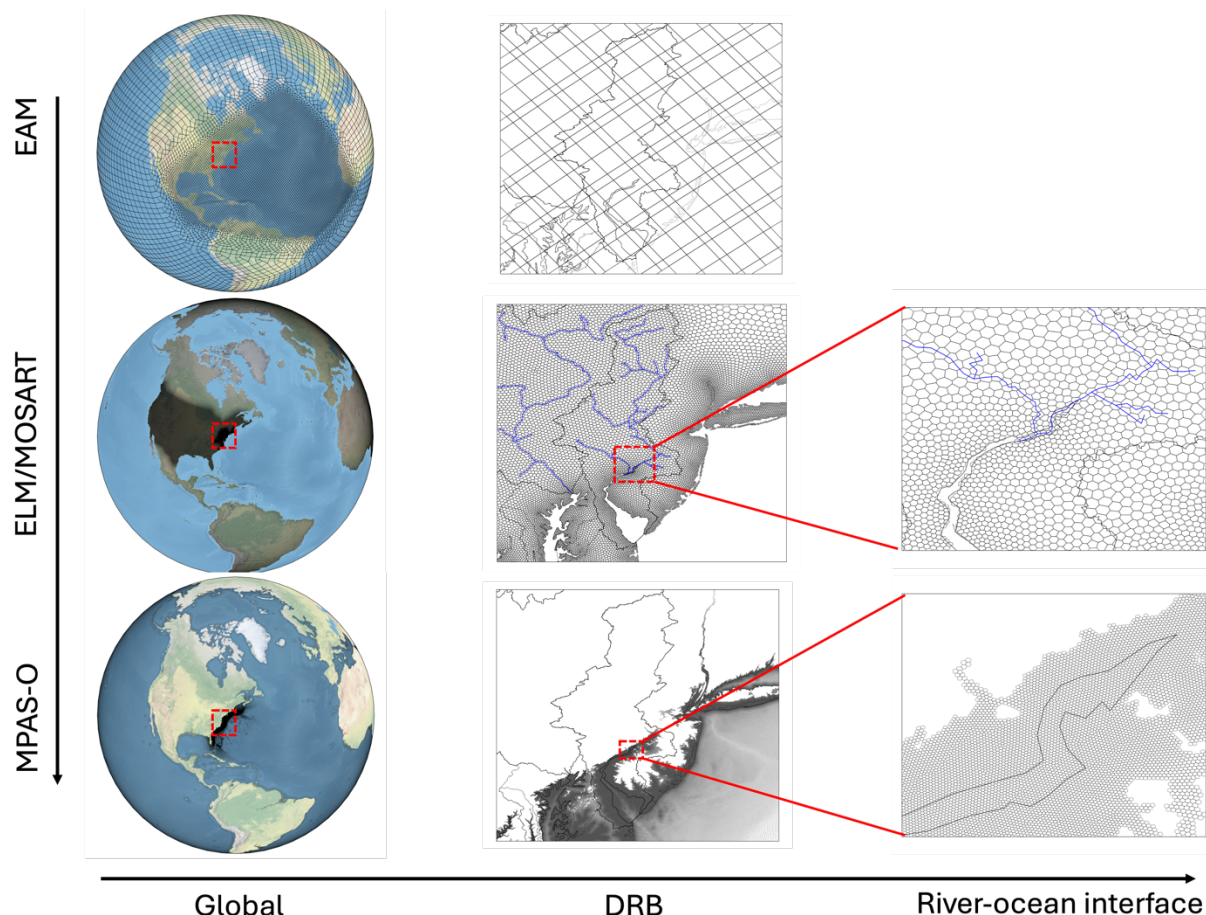


Figure R1. The fully coupled E3SM that features multiple variable-resolution global meshes designed for its atmosphere (EAM), land and river (ELM/MOSART) and ocean (MPAS-O) components, with magnified views at the Delaware River Basin, and the river-ocean interface where the river-ocean two-way coupling is implemented. A similar figure is presented in Feng et al. (2024); to avoid redundancy, it is included only in our response to reviewers. However, this figure illustrates the broader concept of advancing global Earth System modeling to simulate extreme coastal processes.

Although a recent preprint (Zhang & Yu, 2024) has also demonstrated the importance of using a fully coupled ESM framework for CF modeling, our recent research (Feng et al., 2024) has performed the first integration of new features in E3SM, including two-way land-river-ocean coupling, coastal refined unstructured meshes in atmosphere, land, and river models, and 2-dimensional barotropic ocean models with wetting and drying. It is important to note that some of these features are relatively new in the domain of ESMs. Alongside model development efforts, it is crucial to evaluate the capability and uncertainty of ESMs in the context of CF modeling.

Uncertainty propagation in large-scale ESMs remains an important topic in Earth system modeling community. This complexity is further exacerbated by the two-way land-river-ocean coupling, where downstream models can affect upstream models at high temporal frequencies. Such uncertainty propagation has not been adequately addressed, particularly for extreme events, partly due to the lack of a fully coupled framework that integrates all model components at high resolutions and the high

computational cost of running ESM ensembles. Even the study of Muñoz et al. (2024), which inspires our research in its preprint stage, does not comprehensively address atmospheric uncertainty. Although their uncertainty analysis includes a scenario that turned off rainfall, their study relies on a single set of external atmospheric forcing (i.e. ERA5) and does not account for backpropagating uncertainty from two-way coupling. Our developed framework is thus essential for a comprehensive evaluation of atmospheric, hydrological, and coupling uncertainties. Specifically, we have conducted ensemble atmosphere model simulations to assess the propagation of atmospheric uncertainty through the inherent systems.

Moreover, our study compares atmospheric uncertainty with the coupling uncertainties at both land/river and river/ocean interfaces. While it is true that such coupling techniques have been developed elsewhere, as detailed in R2C2, it remains unclear whether coupling uncertainty is significant compared to atmospheric uncertainty, due to a lack of direct comparisons.

Additionally, we address hydrological uncertainty by running a large ensemble of simulations with perturbed atmospheric forcing and other hydrological parameters, including Antecedent Soil Moisture Conditions (AMC) and two runoff generation parameters. This large ensemble enables the application of our newly proposed machine learning method and the related SHAP method to understand the relative contributions of hydrological factors. While the importance of AMC has been documented by Deb et al. (2023), none of these studies have demonstrated the role of underlying hydrological factors under different AMC conditions. Bilskie et al. (2021) and Santiago-Collazo et al. (2024) highlighted that the CF impact may be exacerbated if the CF event is preceded by antecedent rainfall but their studies implemented the rain-on-grid method and did not account for the complexity hydrological processes, such as soil moisture, infiltration, or evapotranspiration. We included this discussion in the revision.

With our ensemble run, we discovered that AMC plays a significant role, resulting in more than twice the inundated area if it co-occurs with a wetter TC. The methodologies and findings together underscore the novelty of our study, which would not have been possible without the new E3SM framework. More importantly, our research sheds light on future directions and supports the ongoing development of more comprehensive model coupling methods.

In response to this comment, we have restructured the introduction, which begin with a more in-depth literature review on the CF studies, followed by a discussion of recent advancements in ESM model implementation and its associated uncertainties. We elaborated the existing gaps and research questions that need to be addressed. Additionally, we have more explicitly elaborated on the scope, objectives, and novelty of this study. Please see below the updated introduction starting from the second paragraph.

“Modeling CF is inherently challenging because it is triggered and governed by the interactions of processes within multiple Earth system components, including atmosphere, land, river, and ocean, as well as the associated uncertainties (Xu et al., 2023). Traditional CF modelling is typically based on one or multiple coupled hydrological models (Feng et al., 2022; Ikeuchi et al., 2017), hydraulic models (Bakhtyar et al., 2020; Bermúdez et al, 2021; Gori et al., 2020b) and hydrodynamic coastal/ocean models (Bennet et al., 2023; Kerns & Chen, 2023; Xiao et al., 2021; Ye et al, 2020) at local, regional and global scales. Over time, more sophisticated methodologies have been developed to enhance CF modeling. These include combined statistical-numerical modeling approaches (Olbert et al., 2023), deep learning (Feng et al., 2023a; Muñoz et al., 2021), data assimilation (Muñoz et al., 2022), reduced-physics ocean models

(Eilander et al., 2023; Leijnse et al., 2021), new compound inundation models (Santiago-Collazo et al., 2024), and two-way river-ocean model coupling (Bao et al., 2022; Bao et al., 2024; Feng et al., 2024; Zhang et al., 2024). The CF modeling uncertainties can be sourced from model structures, parameters, input data, boundary and initial conditions (Abbaszadeh et al., 2022; Beven et al., 2018; Fan et al., 2021). These uncertainties may also cascade through the system (Meresa et al., 2021; Hasan Tanim & Goharian, 2021) and their contributions change dynamically over time (Muñoz et al., 2024).

Emerged recently is the use of fully coupled Earth System Models (ESMs) to simulate compound flooding (Feng et al., 2024; Zhang & Yu, 2024). One significant advantage of ESMs is the integration of multiple earth system components in a single, tightly coupled framework that does not necessarily requires external forcing data. Such a framework allows for predictive understanding of multi-scale flow processes and their interactions with other relevant processes involving heat, energy, biogeochemical and sediment transport, as well as their impacts on Earth's climate (Ward et al., 2020). Feng et al. (2024) performed the first fully-coupled ESM simulation for CF using the Energy Exascale Earth System Model (E3SM), by integrating recent advancements in E3SM including regionally refined unstructured meshes for atmosphere, land/river and ocean components in the global domain (Deb et al., 2024; Feng et al., 2022), two-way online land-river-ocean coupling (Xu et al., 2022b; Feng et al., 2024), and a 2-dimensional (2D) barotropic ocean model (Lilly et al., 2023).

While state-of-the-art ESMs are being implemented to simulate local extremes, this advancement can inevitably introduce additional uncertainties. Compared with regional simulations using prescribed atmospheric forcing derived from observation or bias-corrected reanalysis datasets, the model simulated atmospheric forcings are more uncertain (Hersbach et al., 2020). Atmospheric forcing has critical impacts on the flood simulation (Cloke and Pappenberger, 2009; Hjelmstad et al., 2021; Xu et al., 2025). The river discharge intensity, storm surge levels, and CF inundation extents are directly influenced by the TC's track and intensity, as well as the rainfall rate and timing (Gori et al., 2020a; Pappenberger et al., 2005; Zhong et al., 2010). These factors are the primary drivers of the riverine and coastal flooding dynamics. The uncertainty originated from atmospheric forcings would propagate to land, river, and ocean components through the multi-component framework (Deb et al., 2023; Blanton et al., 2020; Joyce et al., 2018). Likewise, the hydrological uncertainties in the land and river components (Giuntoli et al., 2018; Feng et al., 2023b) and the model coupling schemes can also propagate and even amplify. Typically, the cascading meteorological uncertainty is handled by the ensemble approach (Hamill et al., 2011; Villarini et al., 2019). Multiple realizations of a TC event with perturbed initial conditions and/or model physics represent a range of scenarios that evolve differently based on the dynamics of the models (Blanton et al., 2020). However, the cascading meteorological uncertainty has not been systematically estimated for CF modeling (Abbaszadeh et al., 2022; Xu et al., 2023). It remains unclear whether such uncertainty will amplify or diminish when constrained by the physical processes inherent in ESMs.

The cascading uncertainty becomes even more complex with two-way interactive model coupling. In online two-way coupling, a downstream model, while receiving data from its upstream component, sends back real-time computed information at predefined time intervals, enabling bi-directional data exchange. For instance, a river model may send floodplain inundated water flux to the land surface model (Xu et al., 2022b). Similarly, an ocean model can provide its modeled water levels (Bao et al., 2022; Bao et al., 2024; Feng et al., 2024) or velocities (Zhang et al., 2024) to the river model. While other uncertainties have been extensively discussed (Camacho et al., 2015; Feng et al., 2019; Muñoz et al., 2024; Willis et al., 2019), the uncertainty relevant to the two-way model coupling has rarely been explored because the coupling

capabilities have only recently been developed. Questions are raised regarding the role of model coupling and the magnitude of related uncertainty compared to meteorological uncertainty, especially given the characteristic spatiotemporal scales invoked in land-river and river-ocean coupling. Addressing these questions is critical to refining the performance of interactively coupled ESMs, which is essential for achieving a more comprehensive understanding of the complex interactions and uncertainties associated with CF simulations. Moreover, assessing the enhancements provided by the two-way coupling schemes sheds light on the application of these couplings in future scenarios.

Furthermore, the cascading uncertainty changes with the variability and complexity of hydrological drivers represented in models, because these factors are critical for determining how precipitation is partitioned into runoff and infiltration. As rainfall initially infiltrates the soil, subsurface runoff moves slowly through the soil layers. When the rainfall intensity exceeds the soil's absorption capacity, saturation-excess water leads to surface runoff. The rate of infiltration, which determines the balance between surface and subsurface runoff, is influenced by soil properties, antecedent moisture conditions (AMC) (Ivancic and Shaw, 2015), and land cover types. The runoffs are then routed through river networks, resulting in high river discharge (Fig. 1) (Bevacqua et al., 2020). Understanding the hydrological drivers, including the sensitivity of flood responses to various conditions such as [different](#) AMC and rainfall scenarios, is crucial (Tramblay et al., 2010). These factors provide key insights for predicting different flood scenarios (Miguez-Macho and Fan, 2012; Schrapffer et al., 2020). In particular, AMC plays a critical role in the generation of peak runoff and modulating riverine flooding characteristics during heavy precipitation events (Berghuijs et al., 2019; Nanditha and Mishra, 2022). A saturated AMC can significantly amplify flood impacts compared to drier conditions. The relative importance of rainfall and AMC varies depending on the watershed area. Soil moisture becomes a more dominant factor in larger watersheds (Ran et al., 2022). However, the role of these hydrologic drivers in cascading uncertainties sourced from atmospheric forcing has not been thoroughly investigated in the context of CF, partly due to the absence of a tightly coupled modeling system (Jalili Pirani and Najafi, 2020) [or insufficient investigation into hydrological processes](#) (Lin et al., 2024). [Although Bilskie et al. \(2021\) and Santiago-Collazo et al. \(2024\) highlighted the critical consequence if CF is preceded by an antecedent rainfall event, their implementation of the rain-on-grid method does not account for the hydrological processes, such as runoff generation. Addressing these processes would require a detailed hydrological model or land surface component of ESMs. The full coupled E3SM provides a feasible framework for quantifying the hydrological uncertainties in the CF modelling.](#)

The above-mentioned uncertainties are complicated but must be carefully evaluated for ESMs as they will be more frequently applied for CF simulations in the context of climate change. This study focuses on exploring and disentangling the atmospheric, hydrological, and coupling uncertainties of coastal CF modeling within the coupled E3SM framework. We first provide a comprehensive description of the physical processes during a TC-induced CF event. We then evaluate the model coupling uncertainties and the cascading meteorological uncertainty using a simulation ensemble of a specific TC event. Using the atmospheric ensemble as a basis, we generated an expanded ensemble [and proposed a new machine learning approach](#) to analyze the relative contributions of different hydrological drivers to CF and how these contributions affect the accuracy and reliability of CF simulations [over time](#). Finally, various hydrological and meteorological scenarios are used to delineate a spectrum of plausible CF outcomes in the designated region.”

R2C2:

Second, the authors need to improve their introduction section so it can be more comprehensive than its current state. For example, one of the main points of discussion in the manuscript is the coupling technique that their model uses, such as the twoway/tightly coupling approach, but fails to explain to the reader (at least briefly) how this technique works and compares it with others. At a minimum, the authors should briefly define and explain this and cite a reference such as Santiago-Collazo et al. (2019). Similarly, the paragraph (L49-58) that describes the E3Sm model should be moved to methods and include more details of the model itself. Without this, the authors are expecting the reader to go into another reference to read basic details and model configuration, which is not appropriate. At a minimum, they should talk/show their numerical modeling domain, validation results and simulation set, and model assumptions. The authors only cite a paper from Feng et al. (2024) to point the reader to all the important details about the model for this study, which it should not. Furthermore, the author makes various unsupported claims that are not true. For example, (L90) said that model coupling uncertainties have not been studied before (which is not true; see Muñoz et al., 2024) since this coupling technique has been recently developed. However, the authors fail to comment on the many available models that can do this, such as SFICNS, SCHISIM, and the work done by George XU from LSU.

Author Response:

Thank you for the feedback. We are aware of the comprehensive discussion on the different model coupling approaches reviewed by Santiago-Collazo et al. (2019). We cited the reference and only briefly mentioned the limitation of the one-way coupling method on compound flooding: “Traditional modeling approaches that rely on one-way coupling between any two model components thus have a limited ability to capture CF (Santiago-Collazo et al., 2019)”, as we consider their classification between two-way loosely coupling and tightly coupling is slightly different than the definition in Earth System Models (ESMs) that consider multiple components as a single modeling system. In the revision, we have further elaborated on the difference between one-way and two-way coupling:

“The cascading uncertainty becomes even more complex with two-way interactive model coupling. In online two-way coupling, a downstream model, while receiving data from its upstream component, sends back real-time computed information at predefined time intervals, enabling bi-directional data exchange. For instance, a river model may send floodplain inundated water flux to the land surface model (Xu et al., 2022b). Similarly, an ocean model can provide its modeled water levels (Bao et al., 2022; Bao et al., 2024; Feng et al., 2024) or velocities (Zhang et al., 2024) to the river model.”

We apologize for any confusion caused, leading the reviewer to believe we made unsupported false claims about model coupling and the associated uncertainty. While some of our ideas on addressing uncertainty propagation in a coupled modeling system were inspired by Muñoz et al., 2024, their focused uncertainties using one-way coupled regional models differ from our study’s focus and objectives. Specifically, our aim is to evaluate the capability and uncertainty of ESMs in simulating complex CF events. We concentrate on a different set of uncertainties inherent to the Earth modeling system, which are assessed through ensemble atmospheric simulations. In contrast, regional models are more susceptible to boundary forcing uncertainties, which are not applicable to our framework, as the coupled framework in the global domain does not require external forcing.

Additionally, other two-way coupled regional modeling studies (Bao et al., 2022; Bao et al., 2024; Zhang et al., 2024) primarily couple regional river and ocean models, such as WRF-Hydro, SCHISM, and ROMS. These model development studies lack the in-depth uncertainty analysis as did by Muñoz et al. (2024). The question of whether coupling uncertainty is significant, given the dominance of atmospheric uncertainty in such cases, remains unaddressed. We appreciate the reviewer's valuable recommendations and have expanded our literature review accordingly. Please see the updated introduction in our response to R1C1.

Following the reviewer's suggestion, we have relocated the detailed description of E3SM to the methodology section. We have intentionally not duplicated the mesh and result presentation in Feng et al., (2024). However, in the revised manuscript, we carefully describe the methodology being employed here, including the modeling schemes, mesh resolution, spatial and temporal scales and validation metrics. Please see below the updated Section 2.1.

“This study uses a recently developed configuration in the Energy Exascale Earth System Model (E3SMv2) (Feng et al., 2024). E3SM represents a significant advancement in Earth System modeling (Golaz et al., 2019, 2022). As a state-of-the-science model, E3SM features a tightly integrated multi-component framework that supports dynamic exchanges and propagation of information across its different components. Additionally, several other developments have been recently implemented to further improve the modeling of coastal extremes, including the introduction of high-resolution regional-refined unstructured meshes in global river models (Feng et al., 2022; Liao et al., 2022, 2023a, b), the implementation of interactively coupled land-river-ocean models (Xu et al., 2022b; Feng et al., 2024), and the global tide model with wetting and drying schemes in the ocean component (Barton et al., 2022; Pal et al., 2023). Compared with regional models that may provide more detailed inundation at the street level (Costabile et al., 2023; Ivanov et al., 2021), E3SM excels at coupling processes across various earth system components. This capability is crucial for capturing the complex responses of earth systems to climate change and projecting climate-driven flood hazards.

The new E3SM configuration (hereafter “E3SM coastal configuration”) integrates the global three-dimensional (3D) E3SM atmospheric model (EAM), one-dimensional (1D) E3SM land model (ELM), 1D E3SM river model MOSART (MOSART: Model for Scale Adaptive River Transport), and two-dimensional (2D) barotropic version of E3SM ocean model MPAS-O (MPAS-O: Model for Prediction Across Scales ocean model) (Fig. 2a). This configuration uses three different variable-resolution meshes to improve the E3SM's capability in modeling coastal processes. The EAM mesh features a global resolution of 100 km, with enhanced refinement to about 25 km over the North Atlantic Ocean. Both ELM and MOSART use a land mesh with a coarse resolution of 60 km globally, which is further refined to 30 km across the contiguous US and to 3 km within the Mid-Atlantic watersheds. The MPAS-O mesh offers the highest resolution of 250 m along the US East Coast, specifically designed to capture estuary dynamics, with a broader global resolution of around 1 km everywhere else. The global bathymetry data of MOSART and MPAS-O are sampled from the 90-m HydroSHEDS digital elevation model (DEM) (Lehner et al., 2008) and the 450-m GEBCO dataset (IOC and IHO, 2020), respectively. The river networks and flow directions are derived using HexWatershed which performs hybrid depression filling and breaching stream burning for river routing in unstructured meshes (Liao et al., 2022, 2023a, b).

The novel two-way hydrological coupling between land and river components enables E3SM to capture the infiltration of inundated river water in floodplains and, subsequently, the enhancement of subsurface runoff and evapotranspiration from saturated floodplain soils (Xu et al., 2022b). The two-way river-ocean coupling was developed for E3SM to better represent the dynamic interaction between rivers and oceans, especially during CF events (Feng et al., 2024). This new approach allows for an accurate representation of coastal backwater effects and the mutual influences of river discharge and ocean sea surface height (SSH), providing a more realistic assessment of CF hazards (Feng et al., 2022).

Using the E3SM coastal configuration, we simulated Hurricane Irene, a TC event that occurred in August 2011 and had large flooding impacts across the Mid-Atlantic region (Fig. 2b). Irene led to significant riverine and coastal flooding in the Delaware River Basin (DRB) and Delaware River Estuary (DBE) due to concurrent intense precipitation and storm surge. Following Feng et al. (2024), an ensemble of 25 EAM simulations with perturbed model parameters were performed to reproduce Irene and associated meteorological outcomes. EAM is initialized from ECMWF Reanalysis v5 (Hersbach et al., 2020) in the midnight of 26 August 2011. Atmospheric nudging is not applied. The EAM ensemble can reproduce the TC characteristics, including the storm track and intensity (see Appendix A in Deb et al. (2024)). These “prerun” EAM simulations were then prescribed within E3SM to drive the land, river, and ocean components. EAM (in data mode) provides atmospheric forcing to ELM and MPAS-O at a 15-min frequency. MOSART is interactively coupled with ELM and MPAS-O at the 1-hour interval via the E3SM coupler (Craig et al., 2012). The model outputs are archived at 15 minutes for EAM and hourly for ELM, MOSART and MPAS-O. We spun up ELM and MOSART from a 10-year historical simulation forced by Global Soil Wetness Projects version 3 (GSWPv3; Kim, 2017), and MPAS-O from 1-month simulation with the global tide model. MOSART was validated against the measured streamflow at 6 USGS gauges along the Delaware River main channel with averaged coefficient of determination (r^2) of 0.79 and Kling–Gupta efficiency (KGE; Gupta et al., 2009) of 0.84. MPAS-O was assessed for water level at 6 NOAA tidal gauges across the DBE, showing an averaged r^2 of 0.72 and root mean squared error (RMSE) of 0.41 m. Please refer to Feng et al., (2024) for a more detailed description of the E3SM configuration and the model evaluation.

Fluvial and coastal inundations are simulated in MOSART and MPAS-O, respectively. The riverine inundation in MOSART is simulated using a macroscale inundation scheme that assumes no between-cell flux exchange and estimates the inundation fraction using the elevation profiles within the grid cell (Luo et al., 2017; Yamazaki et al., 2011). Coastal inundation simulated on the MPAS-O inland mesh is aggregated onto the coarser MOSART mesh in the DRB. Within each MOSART grid cell, the inundation fraction is determined by the percentage of MPAS-O cells with a simulated water depth over 1 m. This value represents adjustments made to the MPAS-O inland bottom elevation data near the DRB coastline during the upscale sampling. Whenever there is a discrepancy between the inundation area from MOSART and MPAS-O in their overlapped cells near the coastline, the MPAS-O inundation is considered more accurate and will be used.

Here, the total simulated inundation extent of Irene is benchmarked against a 250-m resolution inundation extent dataset based on satellite imagery (Tellman et al., 2021). The dataset is aggregated onto the MOSART mesh for comparison. Within each MOSART cell, we compute the fraction of the observed inundation. The model performance is evaluated using flood metrics defined by Wing et al. (2017), including hit rate (HR), false rate (FR) and success index (SI)

$$HR = \frac{M_1 B_1}{M_1 B_1 + M_0 B_1}, \quad (1)$$

$$FR = \frac{M_1 B_0}{M_1 B_0 + M_1 B_1}, \quad (2)$$

$$SI = \frac{M_1 B_1}{M_1 B_1 + M_0 B_1 + M_1 B_0}, \quad (3)$$

where M and B are the pixels (or grid cells) from model simulations and benchmark data, respectively. The subscripts 1 and 0 represent wet (inundated) and dry cells, respectively. For all metrics, a score of 0 indicates poor performance, while a score of 1 represents perfect model performance. In our simulations, a wet cell is identified if the simulated inundation fraction is above a small unitless threshold of 0.02. This threshold minimizes the influence of cells that may only be marginally inundated—likely due to data and model uncertainties—thus ensuring a more reliable assessment of flood extent. The predicted flooded area (FA) is calculated by multiplying the flooded fraction by the corresponding cell area.”

R2C3:

Third, the study limitations should be shown before the results section so the reader is aware of all of them before “believing” the results. For example, I was wondering about the lack of the coastal flood model (deep ocean, for example) since it was not explained in the methods, but then in the limitation section, I learned that the model configuration used in this study does not have one. It is not clear how they model the coastal flooding using their proposed framework since they do not have the coastal flood model (L443) nor use boundary conditions. Thus, I am not even sure if they model coastal processes. I also understand that the authors had a different purpose than to provide a high-resolution model, but the coarse resolution inland is producing an overestimation of the fluvial flood inland.

However, the authors said that it is prudent to overestimate the potential flooding from a flood hazard risk assessment perspective (L244-245). As a civil engineer myself, I am in complete disagreement with such a statement. It is not the same to overestimate the flood depth by a couple of inches, but to estimate areas being severely flooded, where there should not be any flood, is just a “bad model.” Thus, the author’s use of this type of statement represents a lousy justification for their poor validation/calibration processes. I strongly recommend going back to the MOSART model component and calibrating it to a better fit before moving forward with the manuscript. This has a big implication in their study findings since if the simulated storm was a pluvial/fluvial-dominated CF event, then the results would be overestimated greatly, including the uncertainty, since Irene was a coastal flood-dominated event.

Author Response:

Our framework indeed has the global ocean model MPAS-O integrated and run in the 25 ensemble simulations, which is described in Section 2.1. The ocean is only excluded from the expanded ensemble simulations as explained in Section 2.4, in which case MPAS-O simulated water level provides the coastal boundary condition to MOSART (L190-193). In the revision, we complemented more details about the MPAS-O setup and validation. Please see our response to R2C2 for the updated Section 2.1.

MOSART has been validated against the measured streamflow at USGS gauges along the Delaware River main channel with averaged coefficient of determination (r^2) of 0.79 and Kling–Gupta efficiency (KGE)

(Gupta et al., 2009) of 0.84, which is comparable to that of a high-resolution regional model using the same atmospheric forcing (Deb et al., 2024). The bias in the streamflow simulation is due to the EAM simulation of Hurricane Irene, which does not have nudging nor is its output bias-corrected in the ensemble simulation. The overestimation in *FR* is likely due to the bias in the MODIS satellite data, the macroscale inundation scheme in MOSART, and the 5-km resolution of the MOSART mesh. The flood extent dataset (Tellman et al. 2021) is known to likely underestimate the actual flooding area due to the uncertainty in the cloud cover removing technique, which is also reported in Zhang & Yu, 2024. Its fidelity further decreases towards upstream due to the vegetation covers (Sexton et al., 2013). Such data, unlike high-resolution lidar data, should be only used to “benchmark” (rather than “validate”) the modeled results. This was mentioned in the previous L128. The macroscale inundation scheme typically used in large-scale river models assumes no between-cell flux exchange and estimates the inundation fraction using the elevation profiles within the grid cell (Luo et al., 2017; Yamazaki et al., 2011). Such scheme may not be sufficient to represent water depth heterogeneity within each cell and the subgrid flow dynamics given the grid resolution of 5 km. We agree with the reviewer that high-resolution local-scale models, which is able to provide street level flood prediction, consider the bias of a few inches as poor performance. However, global-scale Earth System Models (ESMs) offer their unique advantages in simulating coastal extremes, which is further elaborated in the updated introduction: “One significant advantage of ESMs is the integration of multiple earth system components in a single, tightly coupled framework that does not necessarily requires external forcing data. Such framework allows for predictive understanding of multi-scale flow processes and their interactions with other relevant processes involving heat, energy, biogeochemical and sediment transport, as well as their impacts on Earth’s climate (Ward et al., 2020).” See our response to R1C1 and R2C1 for the updated introduction.

We agree with the reviewer that it is not appropriate to state the prudence to overestimate the potential flooding. In response, we first clarified the objective and advantage of the global E3SM in the introduction. We then complemented more details about MOSART macroscale inundation scheme and the validation result in Section 2.1. Finally, in the revised Section 3.1, we provided more elaboration on this bias: “The overestimation in *FR* is likely due to the bias in the MODIS satellite data, the macroscale inundation scheme in MOSART, and the MOSART mesh resolution. The flood extent dataset (Tellman et al. 2021) is known to likely underestimate the actual flooding area due to the uncertainty in the cloud cover removing technique (Zhang & Yu, 2024). Its fidelity further decreases towards upstream due to the existence of vegetation covers (Sexton et al., 2013). Likewise, the macroscale inundation scheme may be not sufficiently skillful to represent water depth heterogeneity within each cell and the subgrid flow dynamics given the grid resolution of 5 km.”

Specific Comments:

R2C4:

Figure 1: remove the adjectives (e.g., large, heavy, high) for the flood drivers labels. This will make the figure more general and applicable to a broader audience.

Author Response:

Thanks. Figure 1 has been updated following the reviewer’s suggestion.

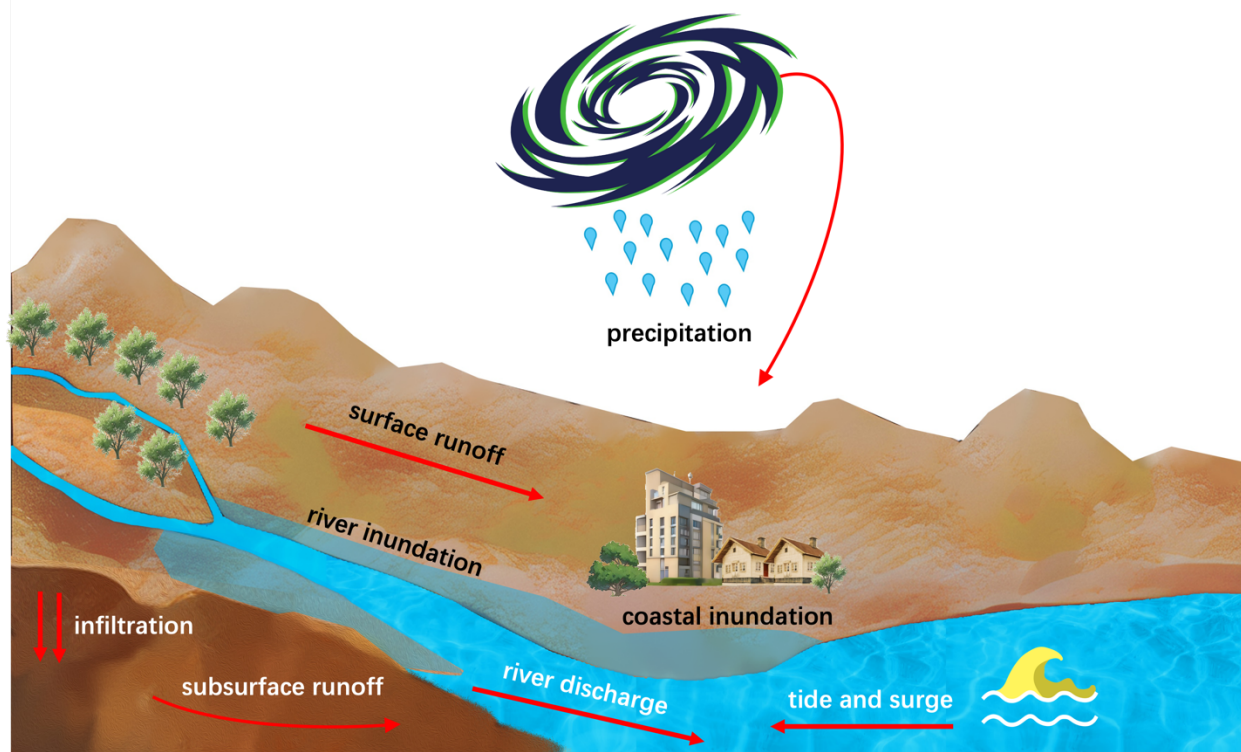


Figure 1: Compound flooding processes in a coastal river basin during a TC event. This conceptual diagram shows the key elements contributing to CF, resulting from combined riverine and coastal inundation along the river channels and adjacent coastal areas.

R2C5:

Figure 2: why are the MOSART network lines so different from the real stream network? The authors say that is a limitation but do not offer details on why.

Author Response:

We appreciate this comment. Upon checking the MOSART river network and the Delaware River streams in Figure 2b, we don't think they are very different at the 5-km resolution. In particular, we employed a rigorous way to generate the river network using HexWatershed (Liao et al., 2022, 2023a, b), a watershed and flow direction model that supports unstructured meshes for river routing models. HexWatershed employs a topology-based method to delineate river networks in the mid-Atlantic region (Lehner et al., 2008). This model generates flow direction across the domain using a combination of depression filling and stream burning algorithms. These algorithms are designed to eliminate local depressions with minimal alterations to the surface elevation, thereby producing essential flow routing parameters such as flow direction maps, channel slopes, and drainage areas. This information is complemented in the updated Section 2.1. See our response to R2C2 for the new section. We have also provided the river

parameter data and the HexWatershed simulation output in Code and data availability: The HexWatershed simulation results is archived at the Zenodo repository (Liao, 2025).

We agree with the reviewer that the original depiction of Figure 2b might lead to confusion, particularly due to the prominently displayed subbasin (HUC8) boundaries in thick black lines. Additionally, it's important to note that the Delaware streams shown in the figure, sourced directly from the Delaware River Basin Commission, were intended solely to illustrate geographic conditions and not for generating the river network. In the updated Figure 2b, we have omitted the Delaware streams and the MOSART river network. Instead, we now directly showcase the major channels of the Delaware River as derived from the HexWatershed output. We also changed to the subbasin boundaries to dashed gray lines.

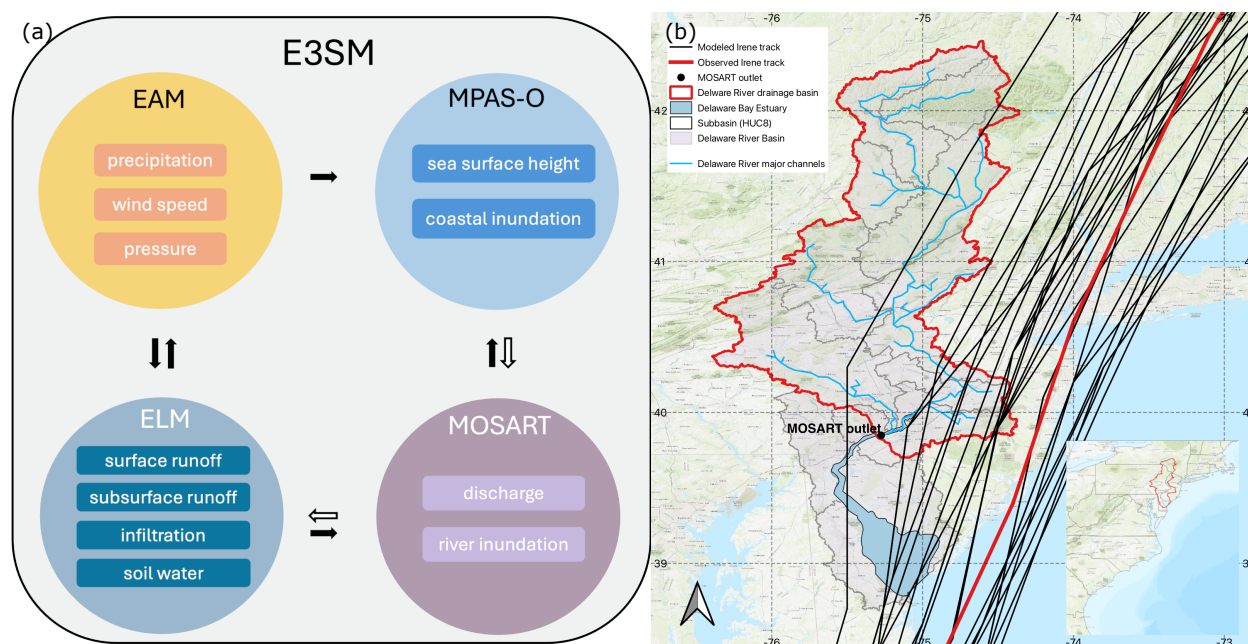


Figure 2: (a) The multi-component E3SM framework and drivers used for analyses within each model component. The black arrows represent the data flow via the one-way coupled framework. The white arrows are the new flow directions from the 2-way land-river and river-ocean models. (b) Map of Delaware river basin (DRB), Delaware bay estuary (DBE), and the observed (red) and modeled (black) Irene tracks. The topographic map in (b) is from the ESRI world topographic map (ESRI, 2012).

R2C6:

L227: The word “effectively” should not be used since the authors do not show at least a time series hydrograph, which they are multiple along the river, of its results and the observed to verify how effective the simulation was.

Author Response:

The validation metrics of MOSART and MPAS-O against observations are now provided in Section 2.1. We prefer not to show the hydrograph time series as it has already been presented in Figure 5 of Feng et al., (2024). Additionally, we have modified the first two sentences of this paragraph to present the capabilities and limitations of the models in a balanced way, focusing on what the results show rather than preemptively describing the models as successful or effective.

“In Experiment 3, the coastal configuration of E3SM employs the coupled MOSART and MPAS-O models to simulate compound riverine and coastal inundation (Fig. 4). The results indicate that MOSART can predict riverine flooding along the lower Delaware River and its upstream tributaries.”

R2C7:

L 231: does the stream network used in MOSART include the real bathymetry? If so, the data was available?

Author Response:

Yes. The delineation of the global river network and river parameters (e.g. riverbed slope) of MOSART uses the 90-m HydroSHEDS digital elevation model (DEM) (Lehner et al., 2008). This information has been added to Section 2.1.

R2C8:

L283: there was no discussion/mention of the temporal scale of any model in the manuscript.

Author Response:

Sorry for the confusion. In the revision, we have clarified the model’s output frequency and coupling frequency.

“EAM (in data mode) provides atmospheric forcing to ELM and MPAS-O at a 15-min frequency. MOSART is interactively coupled with ELM and MPAS-O at the 1-hour interval via the E3SM coupler (Craig et al., 2012). The model outputs are archived at 15 minutes for EAM and hourly for ELM, MOSART and MPAS-O.”

Below is the list of newly added references.

Bakhtyar, R., Maitaria, K., Velissariou, P., Trimble, B., Mashriqui, H., Moghimi, S., Abdolali, A., Van der Westhuysen, A., Ma, Z., and Clark, E.: A new 1D/2D coupled modeling approach for a riverine-estuarine system under storm events: Application to Delaware river basin, *Journal of Geophysical Research: Oceans*, 125, e2019JC015822, 10.1029/2019JC015822, 2020.

Bao, D., Xue, Z. G., Warner, J. C., Moulton, M., Yin, D., Hegermiller, C. A., Zambon, J. B., and He, R.: A numerical investigation of Hurricane Florence-induced compound flooding in the Cape Fear Estuary using a dynamically coupled hydrological-ocean model, *Journal of Advances in Modeling Earth Systems*, 14, e2022MS003131, 10.1029/2022MS003131, 2022.

Bao, D., Xue, Z. G., and Warner, J. C.: Quantifying compound and nonlinear effects of hurricane-induced flooding using a dynamically coupled hydrological-ocean model, *Water Resources Research*, 60, e2023WR036455, 10.1029/2023WR036455, 2024.

Bennett, W. G., Karunarathna, H., Xuan, Y., Kusuma, M. S., Farid, M., Kuntoro, A. A., Rahayu, H. P., Kombaitan, B., Septiadi, D., and Kesuma, T. N.: Modelling compound flooding: a case study from Jakarta, Indonesia, *Natural Hazards*, 118, 277-305, 10.1007/s11069-023-06001-1, 2023.

Bermúdez, M., Farfán, J., Willems, P., and Cea, L.: Assessing the effects of climate change on compound flooding in coastal river areas, *Water Resources Research*, 57, e2020WR029321, 10.1029/2020WR029321, 2021.

Beven, K. J., Almeida, S., Aspinall, W. P., Bates, P. D., Blazkova, S., Borgomeo, E., Freer, J., Goda, K., Hall, J. W., and Phillips, J. C.: Epistemic uncertainties and natural hazard risk assessment—Part 1: A review of different natural hazard areas, *Natural Hazards and Earth System Sciences*, 18, 2741-2768, 10.5194/nhess-18-2741-2018, 2018.

Bilskie, M. V., Zhao, H., Resio, D., Atkinson, J., Cobell, Z., and Hagen, S. C.: Enhancing flood hazard assessments in coastal Louisiana through coupled hydrologic and surge processes, *Frontiers in Water*, 3, 609231, 10.3389/frwa.2021.609231, 2021.

Capodaglio, G. and Petersen, M.: Local time stepping for the shallow water equations in MPAS, *Journal of Computational Physics*, 449, 110818, 10.1016/j.jcp.2021.110818, 2022.

Craig, A. P., Vertenstein, M., and Jacob, R.: A new flexible coupler for earth system modeling developed for CCSM4 and CESM1, *The International Journal of High Performance Computing Applications*, 26, 31-42, 10.1177/1094342011428141, 2012.

Deb, M., Sun, N., Yang, Z., Wang, T., Judi, D., Xiao, Z., and Wigmosta, M. S.: Interacting effects of watershed and coastal processes on the evolution of compound flooding during Hurricane Irene, *Earth's Future*, 11, e2022EF002947, 10.1029/2022EF002947, 2023.

Fan, Y., Yu, L., Shi, X., and Duan, Q.: Tracing Uncertainty Contributors in the Multi-Hazard Risk Analysis for Compound Extremes, *Earth's Future*, 9, e2021EF002280, 10.1029/2021EF002280, 2021.

Gori, A., Lin, N., and Xi, D.: Tropical cyclone compound flood hazard assessment: From investigating drivers to quantifying extreme water levels, *Earth's Future*, 8, e2020EF001660, 10.1029/2020EF001660, 2020a.

Gori, A., Lin, N., and Smith, J.: Assessing compound flooding from landfalling tropical cyclones on the North Carolina coast, *Water Resources Research*, 56, e2019WR026788, 10.1029/2019WR026788, 2020b.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of hydrology*, 377, 80-91, 10.1016/j.jhydrol.2009.08.003, 2009.

Ikeuchi, H., Hirabayashi, Y., Yamazaki, D., Muis, S., Ward, P. J., Winsemius, H. C., Verlaan, M., and Kanae, S.: Compound simulation of fluvial floods and storm surges in a global coupled river-coast flood model: Model development and its application to 2007 Cyclone Sidr in Bangladesh, *Journal of Advances in Modeling Earth Systems*, 9, 1847-1862, 10.1002/2017MS000943, 2017.

Joyce, J., Chang, N.-B., Harji, R., Ruppert, T., and Singhofen, P.: Cascade impact of hurricane movement, storm tidal surge, sea level rise and precipitation variability on flood assessment in a coastal urban watershed, *Climate Dynamics*, 51, 383-409, 10.1007/s00382-017-3930-4, 2018.

Kerns, B. W. and Chen, S. S.: Compound effects of rain, storm surge, and river discharge on coastal flooding during Hurricane Irene and Tropical Storm Lee (2011) in the Mid-Atlantic region: coupled atmosphere-wave-ocean model simulation and observations, *Natural Hazards*, 116, 693-726, 10.1007/s11069-022-05694-0, 2023.

Kim, H.: Global soil wetness project phase 3 atmospheric boundary conditions (experiment 1), (No Title), 2017.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, *Eos, Transactions American Geophysical Union*, 89, 93-94, 10.1029/2008EO100001, 2008.

Leijnse, T., van Ormondt, M., Nederhoff, K., and van Dongeren, A.: Modeling compound flooding in coastal systems using a computationally efficient reduced-physics solver: Including fluvial, pluvial, tidal, wind-and wave-driven processes, *Coastal Engineering*, 163, 103796, 10.1016/j.coastaleng.2020.103796, 2021.

Liao, C.: MPAS Mesh-Based Flow Routing Datasets for Hydrologic Modeling in Mid-Atlantic Coastal Watersheds [Data set], Zenodo, <https://doi.org/10.5281/zenodo.14664927>, 2025.

Kulp, S. A. and Strauss, B. H.: New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding, *Nature communications*, 10, 1-12, 10.1038/s41467-019-12808-z, 2019.

Meresa, H., Murphy, C., Fealy, R., and Golian, S.: Uncertainties and their interaction in flood hazard assessment with climate change, *Hydrology and Earth System Sciences*, 25, 5237-5257, 10.5194/hess-25-5237-2021, 2021.

Olbert, A. I., Moradian, S., Nash, S., Comer, J., Kazmierczak, B., Falconer, R. A., and Hartnett, M.: Combined statistical and hydrodynamic modelling of compound flooding in coastal areas-Methodology and application, *Journal of Hydrology*, 620, 129383, 10.1016/j.jhydrol.2023.129383, 2023.

Santiago-Collazo, F. L., Bilskie, M. V., Bacopoulos, P., and Hagen, S. C.: Compound inundation modeling of a 1-D idealized coastal watershed using a reduced-physics approach, *Water Resources Research*, 60, e2023WR035718, 10.1029/2023WR035718, 2024.

Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K. M., Channan, S., and DiMiceli, C.: Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error, *International Journal of Digital Earth*, 6, 427-448, 10.1080/17538947.2013.786146, 2013.

Tanim, A. H. and Goharian, E.: Developing a hybrid modeling and multivariate analysis framework for storm surge and runoff interactions in urban coastal flooding, *Journal of Hydrology*, 595, 125670, 10.1016/j.jhydrol.2020.125670, 2021.

Ward, N. D., Megonigal, J. P., Bond-Lamberty, B., Bailey, V. L., Butman, D., Canuel, E. A., Diefenderfer, H., Ganju, N. K., Goñi, M. A., and Graham, E. B.: Representing the function and sensitivity of coastal interfaces in Earth system models, *Nature communications*, 11, 2458, 10.1038/s41467-020-16236-2, 2020.

Xiao, Z., Yang, Z., Wang, T., Sun, N., Wigmosta, M., and Judi, D.: Characterizing the non-linear interactions between tide, storm surge, and river flow in the delaware bay estuary, United States, *Frontiers in Marine Science*, 8, 715557, 10.3389/fmars.2021.715557, 2021.

Xu, D., Bisht, G., Engwirda, D., Feng, D., Tan, Z., and Ivanov, V. Y.: Uncertainties in simulating flooding during Hurricane Harvey using 2D shallow water equations, *Water Resources Research*, 61, e2024WR038032, 10.1029/2024WR038032, 2025.

Yamazaki, D., Kanae, S., Kim, H., and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, *Water Resources Research*, 47, 10.1029/2010WR009726, 2011.

Ye, F., Zhang, Y. J., Yu, H., Sun, W., Moghimi, S., Myers, E., Nunez, K., Zhang, R., Wang, H. V., and Roland, A.: Simulating storm surge and compound flooding events with a creek-to-ocean model: Importance of baroclinic effects, *Ocean Modelling*, 145, 101526, 10.1016/j.ocemod.2019.101526, 2020.

Zhang, A. and Yu, X.: Development of A Land-River-Ocean Coupled Model for Compound Floods Jointly Caused by Heavy Rainfalls and Storm Surges in Large River Delta Regions, *EGUsphere*, 2024, 1-26, 10.5194/egusphere-2024-3217, 2024.

Zhang, H., Shen, D., Bao, S., and Len, P.: Two-Way Coupling of the National Water Model (NWM) and Semi-Implicit Cross-Scale Hydrosience Integrated System Model (SCHISM) for Enhanced Coastal Discharge Predictions, *Hydrology*, 11, 145, 10.3390/hydrology11090145, 2024.