Investigating coastal backwater effects and flooding in the coastal zone using a global river transport model on an unstructured mesh

Dongyu Feng[1], Zeli Tan[1], Darren Engwirda[2], Chang Liao[1], Donghui Xu[1], Gautam Bisht[1], Tian Zhou[1], Hong-Yi Li[3], L. Ruby Leung[1]

[1]Pacific Northwest National Laboratory, Richland, WA, 99354, USA
[2]T-3 Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, Los Alamos, NM, 87545, USA
[3]Department of Civil and Environmental Engineering, University of Houston, TX, 77204, US

Correspondence to: Zeli Tan (zeli.tan@pnnl.gov)

Abstract. Coastal backwater effects are caused by the downstream water level increase as the result of elevated sea level, high river discharge and their compounding influence. Such effects have crucial impacts on floods in densely populated regions but have not been well represented in large-scale river models used in Earth System Models (ESMs), partly due to model mesh deficiency and oversimplifications of river hydrodynamics. Using two mid-Atlantic river basins as a testbed, we perform the first attempt to simulate the backwater effects comprehensively over a coastal region using the MOSART river transport model under an Earth system model framework i.e., Energy Exascale Earth System Model (E3SM) configured on a regionally-refined unstructured mesh, with a focus on understanding the backwater drivers and their long-term variations. By including sea level variations at the river downstream boundary, the model performance in capturing backwaters is greatly improved. We also propose a new flood event selection scheme to facilitate the decomposition of backwater drivers into different components. Our results show that while storm surge is a key driver, the influence of extreme discharge cannot be neglected, particularly when the river drains to a narrow river-like estuary. Compound flooding, while not necessarily increasing the flood peaks, exacerbates the flood risk by extending the duration of multiple coastal and fluvial processes. Furthermore, our simulations and analysis highlight the increasing strength of backwater effects due to sea level rise and more frequent storm surge during 1990-2019. Thus, backwaters need to be properly represented in ESMs for improving predictive understanding of coastal flooding.

Keywords: Earth System Model, hydrologic modeling, backwater effects, compound flood, unstructured mesh, sea level rise

1 Introduction

Backwater zones are regions at the river downstream sections, a fluvial-marine transition area between upstream flow and
estuary/coastal river plume, where the river flow is affected by the coastal processes, such as sea level changes, tides and storm surge (Lamb et al., 2012), and can extend hundreds of kilometers upstream in low-lying watersheds (e.g., up to 500 km in Mississippi River). Coastal backwaters are created by elevated sea level that can cause upstream propagation of flood waves and the attenuation of the spatial and temporal water stage fluctuations (Luo et al., 2017). These effects play a critical role in floodplain storage and river discharge (Paiva et al., 2013) and also have a key influence on the biogeochemistry and geomorphology at the terrestrial-aquatic interface (Dykstra & Dzwonkowski, 2020; Lamb et al., 2012; Ward et al., 2020). With population growth near coastal regions (Tellman et al., 2021), coastal backwaters are expected to exert a greater impact on human and natural systems.

The backwater zones usually face severe flood risks as a result of tide, storm surge, rainfall runoff and their combined effects. During a landfalling hurricane with strong winds and heavy rainfall, storm surge drives coastal waters to propagate into the river network and interact with high river discharge (Bilskie & Hagen, 2018). When multiple drivers occur simultaneously or in close successions, the flood event is referred as compound flooding (Santiago-Collazo et al., 2019). Coastal backwater induced floods have strong temporal and spatial variabilities (Hendry et al., 2019), depending largely on the local topography and storm characteristics (Gori et al., 2020). Due to climate warming, the frequency and intensity of such compound flooding have exhibited an increasing trend (Bates et al., 2021; Rahmstorf, 2017), as a result of intensified storm surge (Camelo et al., 2020; Marsooli et al., 2019), more frequent extreme precipitation (Alfieri et al., 2016) and accelerated sea level rise (SLR) (Kulp & Strauss, 2019; Orton et al., 2019). Although SLR and storm intensification are considered the most influential flooding drivers (Hwang et al., 2020), projected changes in river discharge also play an important role in modulating the flood potentials (Bermúdez et al., 2021).

Understanding the backwater drivers is prerequisite to mitigating the related flood risks. However, the interactions among the backwater drivers and their respective contributions through fluvial processes, storm and climate are not well understood (Dykstra & Dzwonkowski, 2021). Although there are extensive literatures that address the storm surge induced coastal inundation (or flooding on land) and the related impacts on flood risks in coastal cities (Hinkel et al., 2014; Ye et al., 2020), limited efforts are made towards understanding the extreme surge that propagates into the river network (Ikeuchi et al., 2017). The latter is more critical in low-lying mega-delta regions that reside over 0.5 billion people globally (Syvitski & Saito, 2007). Streamflow in the backwater zones is affected by river topology, upstream discharge, sea level variations and their interactive effects (Castelltort et al., 2020; Hellmers & Fröhle, 2022). Specifically, river topology is characterized by the river channel geometry, riverbed elevation and river's receiving water body. Among these factors, riverbed elevation, due to its control on the backwater propagation extent, has been widely recognized in previous studies (Gori et al., 2020). In contrast, the river's receiving water body has not yet drawn much attention. As rivers contribute to a variety of water bodies including deltaic floodplains, estuaries and coastal oceans (Mikhailov & Gorin, 2012), the interactive effects of river

discharge and sea level vary substantially. For example, the impact of river discharge on the local sea level is much more intense in a narrow tidal-river estuary than in an open sea (Rayson et al., 2015; Chegini et al., 2022). In a narrow estuary, flood risks are further exacerbated because high discharge increases the local sea level, high sea level induced by storm surge impedes river discharge to the ocean, and the interaction of these two mechanisms intensifies the backwater effects (Eilander et al., 2020).

Large-scale river models are one of the major components of Earth System Models (ESMs) that couple the atmosphere, land, river, and ocean models to simulate the global water cycle (e.g., Golaz et al., 2019; Leung et al., 2020) and assess flood risks (Hirabayashi et al., 2013; Towner et al., 2019). Although hydraulic or hydrodynamic models were used more often in previous studies to simulate storm surge induced coastal inundation (Bakhtyar et al., 2020; Muñoz et al., 2020), there have been growing applications of large-scale river models to assess the compound fluvial and coastal flooding at basin (Chen et al., 2013), regional (Ikeuchi et al., 2017; Yamazaki et al., 2012) and global scales (Eilander et al., 2020; Mao et al., 2019) because they are more computationally efficient for a large spatiotemporal assessment. The long-term evolution of flood drivers and risks can be quantified in the context of climate change. Moreover, such models, when directly coupled with other components of ESMs, can also provide estimations of energy, biogeochemical and sediment processes that are often neglected in pure flood inundation models (Li et al., 2022). However, several limitations in the current generation of ESMs impair the realistic representation of coastal backwaters and human–land–river–ocean interactions at the terrestrial-aquatic interface (Ward et al., 2020). First, most ESMs are configured with one-way coupled river and ocean models, in which water only flows from rivers to oceans and the impact of elevated sea levels on upstream river stage is ignored. Second, the meshes used in most ESMs are too coarse to represent backwater effects. For example, in the high-resolution configuration of Energy Exascale Earth System Model (E3SM), a uniform resolution of 12.5 km is used for the river model (Caldwell et al., 2019). The resolutions of other widely-used large-scale river models also only range from 5 km to 25 km. Much higher spatial resolutions (~ km) are required to resolve the smaller-scale topology near the coastline (Bates et al., 2021; Trigg et al., 2016) for coastal backwaters. Last, while most existing ESMs apply structured meshes in their river components, unstructured meshes are needed to achieve more flexible variable resolutions within areas of interest, such as high resolutions along the coastline, as well as to accommodate the high spatial variation of coastal processes.

Motivated by the increasing flood risks in a warming climate, this study is part of a larger effort to develop capabilities in representing land-river-ocean interactions in E3SM for modeling the changing compound flood risks in coastal regions and the potential implications for the regional and global water and biogeochemical cycles. More specifically, the objectives of this study are to (a) assess the capability of two-way coupled river and ocean models on a regionally-refined mesh to capture coastal backwater effects; and (b) understand the major and interactive backwater drivers and their long-term variations under climate change in two contrasting coastal river basins. The backwater drivers are decomposed using a novel extreme flood event selection scheme. Each selected event is identified by the dominant flood drivers. In Section 2, we provide an overview of the study domain, the river model, the unstructured mesh, and the methods of extreme event selection and

drivers decomposition. Model evaluation and analyses are provided in Sections 3, 4 and 5. In Section 6, we discuss the
95   findings and limitations. Finally, the conclusions are provided in Section 7.

2 Methods and Data

This section describes the study domain of two mid-Atlantic river basins. A global river routing model on a regionally-refined unstructured mesh and with two-way river-ocean coupling physics is introduced. We also develop a method to select extreme events and decompose the flood drivers of the selected events. Two hurricane events are selected to demonstrate the
100   applicability of the proposed methods.

2.1 Study domain

The mid-Atlantic region of the US is exposed to frequent tropical cyclones that bring intense precipitation and storm surge (Sun et al., 2021). In this study, we define the mid-Atlantic region as Susquehanna River Basin (SRB) and Delaware River Basin (DRB) (Fig. 1). Susquehanna River drains 71,228 km^2 to the northern end of Chesapeake Bay, contributing ~50% of
105   freshwater inflow to the estuary (Leathers et al., 2008). Chesapeake Bay is the largest estuary in the US with a surface area of 11,601 km^2 and a shoreline extending over 7000 km. Chesapeake Bay has varied tidal characteristics across the estuary, e.g., mixed tide in the northern portion and semidiurnal tide near the bay mouth. In addition to Susquehanna River, several other large rivers also drain to this estuary. For Chesapeake Bay, the amount of the freshwater outflow is approximately the same as the seawater inflow from the mid-Atlantic coastal waters (Valle-Levinson, 1995). Delaware River drains 35,070 km^2
110   to Delaware Bay and contributes 58% of freshwater to the estuary (Whitney & Garvine, 2006). Delaware Bay has 2,030 km^2 in surface area and is dominated by semidiurnal tide. The tidal range is 1.5 m at the bay mouth and increases towards Trenton (USGS gauge 01463500 in Figure 1). Trenton is referred as the downstream limit of freshwater (Sharp, 1983), as there is a hydraulic jump at 2.7 km downstream of the station due to an abrupt decrease in channel bathymetry (Zhang et al., 2020). Together, SRB and DRB have over 4 million residents and DRB provides drinking water to 6% of the US population.
115   The locations of in-situ observations are provided in Figure 1. Among over 100 USGS gauges in the mid-Atlantic region, we selected all USGS gauges in the mainstem of Susquehanna River and Delaware River, respectively, for simulated streamflow validation. The water level data at 6 NOAA tidal gauges are also selected for data analysis and/or model validation. While the coastal tidal gauge (8534720) is used for identifying storm surge, the two tidal gauges near the river mouths (8573364 and 8545240) provide the downstream boundary condition (BC) of the river model for Susquehanna River and Delaware
120   River, respectively. The four tidal gauges at the downstream section of Delaware River (8545240, 8546252, 8539094 and 8548989) are used in model evaluation.

2.2 Global River Routing Model

The Model for Scale Adaptive River Transport (MOSART) (Li et al., 2013; Li et al., 2015), the river component in E3SM (Golaz et al., 2019) is used for river modeling. MOSART is a river routing model applicable across local, regional, and global scales. The model is driven by runoff from a land surface model and simulates water flow from hillslopes to tributary subnetworks and to main channels. The routing schemes in MOSART include kinematic wave and diffusion wave equations, two simplified forms of the 1-dimensional Saint Venant equations. The routing of surface runoff in hillslopes and tributary is represented using the kinematic wave method. The flow in the main channel is represented by the diffusive wave method. The momentum equation in the diffusive wave method is (Chow, 1988)

$$\frac{\partial h}{\partial x} - S_0 + S_f = 0, \tag{1}$$

where h is the water depth in the channel, S_0 is riverbed slope and S_f is the friction slope. Compared to the diffusive wave method, the kinematic wave method neglects the first term of Eq. 1 in its momentum equation. The flow velocity (v) is estimated using the Chezy-Manning equation

$$v = \frac{|S_f|}{S_f} n^{-1} R^{\frac{2}{3}} |S_f|^{\frac{1}{2}}, \tag{2}$$

in which Manning's n is used as the frictional coefficient and R is the hydraulic radius. The backwater effect can be represented in the diffusive wave method, as the flow velocity is determined by both the riverbed slope (S_0) and the water level variations (h) along the river channels (Luo et al., 2017). In extreme conditions, when the downstream water stage is higher than that of the current channel, S_f becomes negative, resulting in a backwater. The extreme reverse flow is recently observed in Mississippi River during Hurricane Ida (Miller, 2021). In our study domain, the backwater processes in Susquehanna River and Delaware River were reported by USGS (U.S. Geological Survey, 2016) and showed significant impacts during Hurricane Irene (Zhang et al., 2020).

The cross-section of the main channel is specified as rectangular in MOSART when channel water depth is no more than the bankfull depth (H). The channel width (W) and bankfull depth (H) are estimated from the total upstream drainage area (A_{total}) using empirical formulations (Bent & Waite, 2013):

$$W = a(A_{total})^b, \tag{3}$$

$$H = a(A_{total})^b, \tag{4}$$

where a and b are empirical parameters. When channel water depth exceeds H, an elevation profile is invoked to capture the elevation variation in the floodplain (Luo et al., 2017).

In this study, the runoff inputs for MOSART are obtained from Global Reach-level Flood Reanalysis (GRFR) (Yang et al., 2021), an offline simulation from a high-resolution land surface model that has been calibrated and bias corrected. The original configuration of the MOSART diffusive wave method applies a static coastal boundary condition (CBC), i.e., either normal depth or fixed mean sea level at the river mouth (Luo et al., 2017). The normal depth boundary assumes that the friction slope (S_f) equals to the riverbed slope (S_0) at the river outlet cell. This simplification, while reasonable for global

simulations in which the influence of coastal processes is limited, can be problematic in low-lying coastal regions. To represent the backwater effects induced by the dynamic sea level variation, this study introduces a new, dynamic CBC option in MOSART to read in time-varying water level data at a time interval consistent with the land-river coupling frequency in E3SM. The coupling time interval is set as 1 hour in this study. This dynamic CBC is only configured for the rivers of interest, while the static CBC is used in all other outlet boundaries due to the limited data availability.

2.3 Coastal refined global unstructured mesh and flow direction map

A global unstructured mesh with a resolution of ~100 km has been developed using the JIGSAW mesh library (Engwirda & Ivers, 2016; Engwirda, 2017), which enables: (a) the flexibility of embedding high-resolution subdomain within the ESM's global configuration; (b) oversampled geometrical features (O(<100m)), e.g., river network and coastline, to be simplified to coarser ESM length scales on orders of 2-60km; (c) close alignment of the complex geography of coastline, watershed boundaries and river networks (Engwirda & Chang, 2021). The global mesh is developed to allow for more seamless coupling of the land, river and ocean components in E3SM (Fig. 2) for more consistent modeling of global surface processes. In the mid-Atlantic region, the mesh resolution is refined to ~3 km to better resolve local coastal and watershed processes (Fig. 2b). Significant effort is also made to ensure that the cells in the high-resolution mesh match the prescribed dam locations and the orientations of the edges conforming to the flow direction along the main channel.

The river networks and flow directions are modeled using HexWatershed (Liao et al., 2020; Liao et al., 2022; Liao, 2022), a watershed and flow direction model that supports both structured and unstructured meshes for river routing models. HexWatershed uses a topological relationship-based approach to define river networks in the mid-Atlantic region (Lehner et al., 2008). To generate the flow direction for the entire domain, HexWatershed uses a hybrid depression filling and breaching stream burning algorithm to remove local depressions while minimizing modifications to surface elevation and produces flow routing parameters including the flow direction map, channel slope, and drainage area, which are critical for accurately representing coupled land-river-ocean processes.

2.4 Extreme event selection

The extreme events of fluvial flood (FF) and storm surge (SS) are separately selected based on their corresponding time-series observations. The selection of FF follows the strategy proposed in (Zhang et al., 2021). In the time-series of discharge data, a flood event is identified by first selecting the flood peaks using the peaks-over-threshold (POT) approach (Lang et al., 1999). The threshold is determined based on automatic threshold selection and the independence of the peak series is examined with a declustering method (Zhang et al., 2021). The start and end dates are specified using the empirical formulation:

$$Q_S \leq aQ_P;\ T_P - T_S \leq b(5 + \ln(\frac{A}{1.609^2})), \tag{5}$$

$$Q_E \leq aQ_P;\ T_E - T_S \leq b(5 + \ln(\frac{A}{1.609^2})), \tag{6}$$

where T_P, T_S and T_E are the peak date, start date and end date, Q_P, Q_S and Q_E are the discharge on the corresponding date, and A is the basin drainage area. The empirical parameters a and b are specified as 0.5 and 1.5, respectively.

The selection of storm surge (SS) is performed in three major steps. The 1st major step is to extract the SS component from the hourly total water level (TWL) data at the NOAA tidal gauge (i.e. 8534720 in Figure 1): (i) the TWL time series is detrended by extracting the annual mean sea level, which removes the effects of SLR; (ii) the predicted astronomical tides are then derived from the harmonic tidal analysis performed on the detrended TWL on a year-by-year basis using 8 major tidal constituents; (iii) SS is the non-tidal residual obtained by extracting the tides from the detrended TWL.

The 2nd major step is to filter the extreme SS events using a peak detection algorithm (Brakel, 2014). When a data point is m-fold standard deviations away from the moving mean, an event peak is identified. The start and end times of the corresponding event are the nearest data points that change signs. In addition to m, the other input parameters of this algorithm are *lag* and *influence*: *lag* is the number of observations to smooth the data, or the length of the moving window; *influence* represents the influence of new signals on the threshold. Here we performed the SS event selection by setting m = 5, *lag* = 30 and *influence* = 0. These parameters are determined to ensure that the selected SS events include all documented hurricane events in the mid-Atlantic region.

The 3rd major step is to select the extreme SS events of interest by extracting the events with SS peaks larger than the 99.5th percentile. When there is an overlap between the FF and SS events, a compound flood event is identified, for which the duration is defined as the combined period of the two events. The applied event selection method can be more accurate than those used in continental and global applications, where the event period is simply determined using a predefined time-window (e.g. ±1 days of a peak event) (Nasr et al., 2021; Ward et al., 2018; Wu et al., 2021).

2.5 Decomposition of backwater drivers

The backwater drivers are decomposed into three different levels (Fig. 3). The first level considers the river topology and the forcing that affects the river flow directly. The direct forcing in the backwater zone is considered as the upstream discharge and the TWL at the river mouth. To address the impact of topology, we compare the backwater effects in Susquehanna River and Delaware River, which differ significantly in terms of the riverbed elevation along the downstream section and the receiving water body. As previous studies have shown the crucial impact of sea level variations on the backwater effects (Yamazaki et al., 2012), we further decompose the TWL into low-frequency surge (LFS) and tide in the second level. The predicted tide is estimated using harmonic tidal analysis and the LFS is obtained by subtracting the tide from the TWL. It should be noted that this level of decomposition must be applied to a tidal gauge as the harmonic tidal analysis requires measurable tidal effects. The third level decomposes the LFS to high discharge, SS and their compound effects. It is worth noting that the LFS differs from the SS in that the LFS is extracted from the TWL at the NOAA gauge nearest to the river mouth (e.g., 8545240 in Figure 1) where the river discharge can have a significant influence on the LFS. In particular, the LFS is dominated by both river discharge and storm surge in narrow tidal rivers, such as Delaware Bay. In contrast, the river discharge influence is negligible in larger estuaries or coastal oceans, such as Chesapeake Bay. Previous studies mainly

focused on the drivers of river discharge and SS (Nasr et al., 2021; Ward et al., 2018), assuming they are distinct mechanisms that have little mutual interactions. Such assumption is valid when the river drains directly to a large receiving water body. However, cases with a river ending within a small estuary have not been explored. In such cases, the high discharge during fluvial flood increases the water level within the estuary and attenuates the spatial variation of water stage along the river (Luo et al., 2017), creating backwater effects. Thus, in the third level, we attempt to understand the respective role of high discharge, SS and their interactive impacts on the LFS induced backwater effects. These drivers are separated based on the selected flood events in Section 2.4. Events dominated by the drivers of high discharge, SS and their compound effects are denoted as $LFS \cap FF \cap \overline{SS}$, $LFS \cap SS \cap \overline{FF}$ and $LFS \cap compound$, respectively. The symbol $\cap$ means the intersection period between two events and the overline means to exclude the different drivers if there is an intersection. Each event is measured in terms of the event duration and the peak water level during the event.

The trend analysis was performed to assess the annual trends of the backwater drivers. For each event type, we calculated the annual duration, occurrence and the maximum peak value on an annual basis. The impact of SLR is addressed by selecting the SS event without detrending the TWL data, while the other steps are the same as those in Section 2.4. The corresponding case is referred as SS+SLR. The nonparametric Mann-Kendall (MK) test (Tosunoglu & Kisi, 2017) was employed to statistically assess if there is a monotonic trend. The null hypothesis (H_0) and the alternative (H_a) are: no monotonic trend is present; and monotonic trend is present, respectively. For each MK test, we set the significance level at 0.1 and calculated the standard MK statistics (Z) and p-value. The positive (negative) value of Z corresponds to the increase (decrease) trend.

2.6 Numerical experiments

MOSART simulations were performed from 1990 to 2019 with the first year excluded from analysis as the spin-up time. This period has sufficient data coverage of both runoff and the water level at NOAA gauges. We configured five simulations based on the aforementioned downstream CBCs: (a) normal depth; (b) total water level (TWL); (c) mean sea level (MSL); (d) low-frequency surge (LFS); (e) tide. The backwater effects are quantified by comparing the TWL and MSL simulations in terms of two quantification metrics along the main channel: water depth change Δh and water volume change ΔV:

$$\Delta h(t,i) = h_{exp}(t,i) - h_{MSL}(t,i), \qquad (7)$$

$$\Delta V(t) = \sum_i^N \left(h_{exp}(t,i) - h_{MSL}(t,i) \right) L(i)W(i), \qquad (8)$$

where the subscript exp represents TWL, LFS or tide, h is the main channel water depth, t is the model output time step and i is the grid cell index, L and W are the length and width of the main channel within the ith cell, respectively, and N is the number of cells with nonzero Δh. The two metrics measure the backwater-induced changes within the river channel that are created from the variation in the downstream water level. The model performance is assessed using coefficient of determination (r^2), root mean square error (RMSE) and Kling–Gupta efficiency (KGE) (Gupta et al., 2009)

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\delta_m}{\delta_o} - 1\right)^2 + \left(\frac{\sum X_m}{\sum X_o} - 1\right)^2}, \qquad (9)$$

where X_m and X_o represents the model simulation and the observation, δ_m and δ_o are the corresponding standard deviations, and r is their linear correlation.

2.7 Extreme Events

The flood events in Delaware River during Hurricane Irene (followed by Tropical Storm Lee) and Hurricane Sandy are used to demonstrate the selection of extreme events and the quantification of the backwater effects. Hurricane Irene, one of the most destructive tropical cyclones in the US history, made its first landfall on the coast of North Carolina on August 27, 2011 as a Category 1 hurricane, followed by another landfall in the southeastern New Jersey on August 27 and a third landfall in New York City. The storm surge peaked at 1.8 m along the coast of New Jersey and the wind speed was up to 105 km/h. The maximum rainfall is 10 inches in DRB. Tropical Storm Lee is the subsequent storm event that formed over the gulf coast and swept the east coast. Lee brought 10~12 inches of precipitation to the mid-Atlantic region, resulting in mainly fluvial processes rather than coastal surges (Ye et al., 2020). The combined events caused two consecutive flow peaks in Delaware River from Aug 28 to Sep 10, 2011 (Fig. 4). Irene is affected by the interactive storm surge and precipitation-induced fluvial flood and has been studied extensively as an example of compound flood (Xiao et al., 2021; Zhang et al., 2020). Hurricane Sandy, the largest Atlantic hurricane on the US record, made landfall on the coast of New Jersey on October 29, 2012 with a sustained wind speed of 130 km/h. Sandy caused a maximum storm surge of 4 m near New York city and 1.5 m near the Delaware coast. The observed peak flow during Sandy is ~800 m^3/s at Trenton.

3 Model evaluation

In this section, the MOSART performance for simulating river discharge and water level is compared between the experiments using the downstream CBC of normal depth, TWL and MSL to demonstrate the importance of imposing an appropriate downstream CBC. The model performance in reproducing the observed water level in the downstream section of Delaware River is significantly improved when the TWL is enforced at the boundary.

3.1 River discharge

The MOSART simulated daily discharge is compared with the USGS observations over the simulation period (1991–2019) at the gauges along the main stem of Susquehanna River and Delaware River (Fig. 4). The coefficient of determination r^2 and KGE are calculated for each gauge. The MOSART simulation compares reasonably well with r^2 and KGE with both over 0.5 across all gauges (Towner et al., 2019). The model performance is generally higher in Susquehanna River and decreases towards the upstream regions. The highest r^2 and KGE (≥ 0.75) are found at the downstream gauges of the two rivers, as the forcing may not capture the runoff accurately for smaller drainage areas. A closer look at the time series of discharge at these two gauges from 2011 to 2012 shows that the model can capture the hydrograph and smaller peaks of river discharge well, and there is no significant difference of model performance among the different CBC configurations because

9

even the USGS gauges closest to the river outlets are still too far upstream to capture effects of dynamic CBCs. The model has a large bias in Susquehanna River during Hurricane Irene (August 21 to 30, 2011) with the extreme flood peak significantly underestimated. While the observed peak flow is about 20000 m^3/s, the simulated flow is about 5000 m^3/s. This bias is likely caused by uncertainty in the GRFR runoff forcing (Yang et al., 2021). The global runoff forcing, although is bias corrected, may underestimate the extremes over specific regions. This is a known challenge in global runoff generation schemes. Additionally, there also exists uncertainty in USGS discharge data estimated from rating curve during extreme events (Di Baldassarre et al., 2012), because these measurements occur less frequently and usually do not cover extreme events (Turnipseed & Sauer, 2010). Overall, the evaluation indicates that the MOSART simulated river discharge reasonably captures the spatial and temporal variability of the observed discharge in the mid-Atlantic region.

3.2 Water level

The simulated water level is compared with the observations at four NOAA tidal gauges along the downstream section of Delaware River. The model performance is quantified in terms of r^2 and RMSE (Fig. 5). The TWL simulation results in the best performance, in which r^2 is over 0.5 among all the gauges, much higher than the other two configurations. The lowest RMSE is also obtained from the TWL simulation. By setting the TWL as the downstream CBC, the model's capability at reproducing the water level variation is greatly enhanced. The same conclusion can also be drawn in the time-series comparison from 2011 to 2012 (Fig. 6). The TWL simulation accurately captures the small variations in the observed water level, which are missing in the simulations with normal depth and MSL boundary conditions. The extreme peaks are overestimated in the TWL simulation, as well as the normal depth simulation in which no data are enforced at the downstream boundary. The MSL simulation tends to produce smaller variations and lower peaks as the downstream boundary is forced by a constant water level. The overestimation in water level peaks by the TWL and normal depth simulations is likely a result of the uncertainties in the channel topology in MOSART. Moreover, the diffusive wave equation (Eq. 1) simplifies the momentum transport by neglecting the inertia terms (local and convective). In the diffusive wave method, because the flood wave is considered as subcritical and diffusive (Trigg et al., 2009), the water level is mainly controlled by the upstream discharge. In low-lying rivers, while gravity and friction may not be the dominant forcing, the inertial force related with velocity changes in space and time dominates the flow momentum. As such, the flood wave propagation from the downstream boundary is underestimated in the backwater zone. As shown in Fig. 5, the improvement of the TWL simulation in predicting water level is reduced towards upstream. This is not unexpected in a reduced-physics river model (Hodges, 2013) because implied in the diffusive wave equation (Eq. 1), the energy head at the downstream CBC is created by the pressure gradient as a result of the variation in water surface but it is lost rapidly upstream due to the increase in riverbed elevation.

The model evaluation results illustrate that a river model on a regionally-refined global mesh can represent backwater effects at the basin scale when properly specified downstream CBC is used. Thus, the model can be used to further examine the contribution of the backwater drivers. It should be noted that the 1D river models are by no means comparable to 3D

hydrodynamic models (e.g. models in (Gori et al., 2020) and (Zhang et al., 2020)) at reproducing coastal flood events or resolving the complex flow dynamics (Neal et al., 2012). Therefore, our analysis focuses on a larger temporal scale by extracting the extreme events from a long period, which are then used to quantify the backwater drivers.

4 Flood event simulation

The event selection method (Section 2.4) is used to select the extreme SS, LFS and FF events during Hurricane Irene, Tropical Storm Lee and Hurricane Sandy from long-term observations at the NOAA coastal gauge (8534720), the NOAA gauge closest to the river mouth (8545240) and the USGS gauge at Trenton (01463500), respectively (Fig. 7). Gauge 8545240, despite located at the upstream reach of Delaware Bay, is dominated by semidiurnal tides even at high-flow conditions during extreme storm events (Xiao et al., 2021). A lag time of 4 hours is added to the water level data at the coastal gauge to compensate for the phase lag between this NOAA tidal gauge and that at the river mouth. As there is an overlap between the SS and FF events during Irene (Fig. 7a), a compound flood is identified over the combined period. The obvious difference is observed between the LFS and SS events, which were obtained using the same method but at different locations. At the Delaware River mouth, the LFS event can be attributed to both SS and river discharge, resulting in a duration period comparable to an FF event and much longer than a SS event. This highlights the importance of considering the influence of high river discharge on compound flooding, particularly for rivers contributing to a small receiving water body. Sandy did not induce significant fluvial processes, and the LFS event was primarily caused by SS. Thus, no compound flood is identified over this period.

Because the TWL simulation shows a reasonable performance at reproducing both river discharge and water level during the hurricane periods (Section 3), the TWL simulation is used to estimate the water depth change (Δh) and water volume change (ΔV) in Eq (7) and (8) to assess the backwater effects. Δh shows the backwater propagation extent, which is roughly 60 km upstream from the river mouth (Figure 7). This extent is determined by the riverbed elevation. Increasing the elevation to ~5 m, Δh implies a large increase in the downstream water level. As a spatially aggregated quantity Δh, ΔV is also consistent with LFS, with ΔV following the LFS variation and peaking on the same dates. By comparing Δh and ΔV over the two hurricanes, it is not difficult to conclude that the compound flood caused by two consecutive events of Hurricane Irene and Tropical Storm Lee has a much larger impact on the backwater effects than that of Hurricane Sandy, even though their LFS peaks are at similar levels. This is probably because the duration of the SS event during Sandy is much shorter than the combined SS and FF events during Irene.

11

5 Backwater drivers

The interactive effects of water level and discharge inspire us to further decompose the backwater drivers. This section provides the analysis on the drivers based on the three-level decomposition introduced in Section 2.5. We examine the contribution of each driver to the backwater effects and their corresponding long-term trend under climate change.

5.1 Decomposition of backwater drivers

5.1.1 Discharge, TWL and topology

The first decomposition level assesses the impacts of river discharge (Q), TWL and topology by comparing ΔV between Susquehanna River and Delaware River over the entire simulation period (Fig. 8). While ΔV is computed from the TWL simulation, Q and TWL are obtained from the paired streamflow and tidal gauges nearest to the river outlets, that is, gauge 01578310 and 8573364 for Susquehanna River and gauge 01463500 and 8545240 for Delaware River. The result shows the key role of river topology and TWL in affecting backwaters. The maximum ΔV in Susquehanna River is roughly 3 orders of magnitude smaller than that in Delaware River. This is the result of a larger gradient in riverbed elevation profile of Susquehanna River that impedes the backwater propagation. Over the 5-km downstream section, the elevation increases from 0 to ~20 m in Susquehanna River but by less than 1 m in Delaware River.

In both rivers, ΔV is dominated by TWL, with the corresponding correlation coefficient (r) over 0.9 (Fig. 9). However, the influence of Q differs significantly between the two rivers. In Susquehanna River, Q is negatively correlated with ΔV ($r =$ −0.11). The increase in Q reduces the TWL impact on ΔV. For instance, at the same TWL, a smaller Q could result in a higher ΔV (Fig. 8a). This behavior is also evident in the slopes of the fitted linear regression lines: the fitted slope for $Q \geq$ 1000 m³/s is smaller than those for low discharge (Fig. 8a). This result is expected because high upstream discharge can attenuate the propagation of downstream backwaters. In Delaware River, ΔV increases with Q and r between ΔV and Q is 0.36. The regression slopes are similar at different discharge conditions. These contrasting results between the two rivers imply that the impact of Q on the backwater effects depends on the river's receiving water body. Because Delaware River contributes to Delaware Bay, a much narrower estuary than Chesapeake Bay, the effect of its discharge on the water level variation of the estuary is much stronger than that of Susquehanna River. Consistently, there is a much higher r between Q and the TWL (0.27) in Delaware River than that (-0.03) in Susquehanna River (Fig. 9). In addition, the channel constriction in Delaware River might also facilitate the formation of backwaters (Castelltort et al., 2020).

Between the two draining estuaries, Chesapeake Bay behaves more like an ocean as the river impact is limited and the coastal and fluvial processes are distinct. Such situation is usually taken as the general case for compound effects and has been addressed extensively using statistical models (Nasr et al., 2021; Ward et al., 2018) and large-scale river models (Ikeuchi et al., 2017). In contrast, cases like Delaware River have rarely been documented in any global-scale studies, even though they are ubiquitous and may witness more coastal backwaters. A possible reason for why such cases were overlooked

370   is that previous global meshes have large deficiency in resolving narrow estuaries properly. Thus, we focus the following analysis on Delaware River.

5.1.2 LFS and tide

Given the key impact of TWL, the second decomposition level examines the respective role of the LFS and tide. The simulations configured with the downstream BCs of the TWL, LFS and tide (Section 2.6) are used to derive the maximum

375   water depth change (Δh_{max}) for each grid cell:

$$\Delta h_{max}(i) = \max \{\Delta h(t,i): t = 1,2,...,T\}, \tag{9}$$

where T is the simulation period.

The Δh_{max} comparison shows the dominance of LFS over tide in increasing the maximum water depth (Fig. 10). In the along-channel profile (Fig. 10a), the value of Δh_{max} in the LFS simulation is close to that in the TWL simulation: ≥ 1 m

380   within the 40 km upstream range of the outlet and then gradually reduced to 0 with a sharp increase of the riverbed elevation. In contrast, the tide simulation produces a much smaller Δh_{max} with the value never exceeding 0.5 m. Among the simulation cases, the highest Δh_{max} occurs at roughly 25 km upstream from the mouth. This along-channel profile reveals the interaction of the discharge and the upstream propagation of tide and surge momentum. It is also observed that Δh_{max} is slightly higher in the LFS simulation than in the TWL simulation, which is likely the result of the negative impact of low

385   tide on TWL. The spatial variation of Δh_{max} is shown in Figure 10b. The backwater effects are limited to the low-lying section of Delaware River, i.e. the downstream of Trenton. Backwater propagation occurs along the main channel as well as some small contributing tributaries, for which the extent is determined by the corresponding elevation. The propagation extent is similar between the TWL and LFS simulations and is much smaller in the tide simulation.

5.1.3 High discharge, SS and compound effect

390   The LFS impact is further decomposed into high discharge, SS and their compound effect using the LFS simulation. We compared the variation of the event accumulated ΔV with respect to the event duration and peak water level among the drivers (Fig. 11). Regardless of the drivers, ΔV is mainly determined by the event duration. Its value linearly increases with the duration with high correlations. The peak water level provides the secondary effect. Higher peaks generally increase ΔV, resulting in values above the fitted regression line.

395   Our result indicates that the influence of each driver on ΔV is more dependent on the frequency and duration of the corresponding events rather than their extremes. For example, the FF events are more influential on ΔV because they are more frequent than the SS and compound flood events. Also, the ΔV is higher in the FF and compound flood events than in the SS events because the latter lasts much shorter. A remarkable case occurred during Hurricane Irene when the highest ΔV in our study period was produced by the combined long-lived FF and compound flood events (Fig. 7). Noticeably, the

400   slope and correlation between ΔV and duration is the largest in the compound events, which means that the strength of the

compound events increases more rapidly with duration. In all, our driver comparison indicates that high discharge is the key driver of the backwaters in Delaware River due to the higher frequency and longer duration of the corresponding FF events.

5.2 Trend analysis

This section performs an analysis of annual trend on the LFS and decomposed drivers, as well as their influences on the
405   backwater effects. The result reveals the impacts of SLR and increasing frequency of SS events during 1990-2019 in exacerbating the backwater effects.

5.2.1 Trend in the backwater drivers

The annual trend of the backwater drivers shows an increasing trend of SS due to both SLR and increasing SS frequency (Fig. 12), with p-value as 0.039 and 0.031, respectively, for annual duration and occurrence in the SS+SLR case (Table 1).
410   When SLR is considered, the number of SS occurrence increases from ~1 to 2~5 times per year over the study period (Fig. 12b). Accordingly, the annual duration of the SS event increases from ≤15 days (1990-2005) to >25 days over multiple years (2005-2019) and is up to over 50 days in 2016 (Fig. 12a). The SS peaks are also increased by SLR but do not present any trend (Fig. 12c). Neither do we notice clear trends from the FF events nor the compound events (Table 1). The annual characteristics of the FF events vary significantly between wet and dry years. In the very wet year of 1996, the duration of
415   the FF events reaches 100 days with up to 6 events per year. The frequency of compound flood events is low, only occurring 1~3 times per year between 2003 and 2012.

Affected by the drivers of high discharge and SS, the LFS events show an increasing trend in both duration and frequency with p-value at 0.094 and 0.042, respectively. No clear trend is found in the LFS peaks (Table 1). The LFS trend is basically consistent with that of the SS events with a few exceptions in flood years (i.e., year 1996 and 1998) when the LFS events
420   were caused by high discharge. Except for the flood years, the LFS occurrence is as low as one time per year in the 1990s and increases to 2~6 times per year since 2003. Correspondingly, the LFS duration increases from ≤5 to 20~40 days since 2004. The increased trend in LFS is likely a result of SLR that leads to more frequent occurrences of the SS events.

5.2.2 Trend in the backwater effects

The annual trend of the backwater effects is analyzed for the different drivers (Table 1) in terms of the annually accumulated
425   ΔV (Fig. 13). The trend of ΔV is consistent with the event duration and frequency trends of the corresponding drivers. The SS and LFS induced backwater effects are increasing but no clear trends can be observed for the FF and compound flood induced backwater effects. The ΔV trends are significant in the SS+SLR, SS and LFS cases and insignificant in the FF and compound cases (Table 1). Our result also demonstrates the critical impact from high discharge. The resulting ΔV by FF over the flood years (e.g., year 1996) can be several times higher than the ΔV caused by the other drivers. This is likely due to the
430   long duration of the FF events. Because SLR and intensified SS increased coastal backwaters in river channels, our analyses

14

call for better representations of the related processes in ESMs for predictive understanding of associated flood risks under climate change and effects on the water and biogeochemical cycles through land-river-ocean interactions and possible impacts on atmospheric processes. However, we caution attributing changes based on modeling and analysis of a 30-year period to climate change as internal climate variability and other anthropogenic effects likely also play important roles in the increasing sea level and storm surge frequency.

6 Discussion

Our study shows that using the diffusive wave method, large-scale river models configured on a coastal refined mesh are capable of reproducing backwater effects in low-lying river channels with appropriate downstream boundaries. The global unstructured mesh alleviates the computational burden in ESMs by relaxing the resolution in the inland and offshore regions to ~100km while embedding regions of "high" resolution of O(1km) near the river-ocean interface. Although the resolution is still not comparable to that used in local-scale models, the mesh is able to resolve the complex river networks near the coastline without having to merge multiple outlets into a single cell. The downstream boundary condition is critical for connecting the coastal and fluvial processes, transferring the water head energy upstream and thus simulating the backwater effects. The widely used diffusive wave method that uses the normal depth boundary and the more simplified kinematic wave method may be only applicable in high-gradient regions, as these methods do not incorporate any downstream information.

As an important finding, this study revealed the crucial difference in flood drivers between two distinct coastal rivers (i.e., Susquehanna River vs. Delaware River), with the former connected to a wide ocean-like estuary and the latter connected to a narrow river-like estuary, which is usually ignored by ESMs. The difference is mainly caused by the effects of river geometry and estuary size on LFS which is a direct driver of backwaters. For an ocean-like estuary, such as Chesapeake Bay, river discharge from its drainage basins hardly affects the water level fluctuations of the estuary. But when a coastal river drains to a narrow estuary, its LFS would be driven by not only storm surge but also river discharge and their compounding. Further, the backwater effects created by high discharge and storm surge are different. While storm surge generates an upstream-propagated energy head, high discharge gradually builds up the water level of the receiving water body. The increased water level would slowly move upstream, attenuating the river stage fluctuation and flood waves. High discharge that presents a higher frequency and a longer duration can occur in close successions with storm surge during compound flooding, e.g., Hurricane Irene and Tropical Storm Lee, creating extended backwater effects. In all, we show that in addition to the conventional flood drivers, such as riverbed elevation and sea level, ESMs need to properly represent the flood drivers for small estuaries, such as river discharge.

We demonstrated that the backwater effects are significant in the low-lying watersheds and have an increased trend over the recent 30 years. In Delaware River, the propagated backwaters account for up to 1.2 m increase in water depth and 1×10^8 m^3 increase in water volume per day during an extreme event. The effects could be several orders higher for larger river basins.

15

This increased flood risk will otherwise be underestimated if the backwater effects are not properly represented in ESMs. Furthermore, our simulation also shows the increased influence of climate change on backwaters, with SLR and more frequent storm surge increasing the strength of backwaters in the mid-Atlantic region.

Noticeably, there are still a few limitations in the river model used in this study, which may introduce uncertainties to our simulations. First, as a large-scale river transport model, MOSART simplifies the channel cross-section as rectangular and trapezoidal when the water depth is below and above the river's bankfull depth, respectively, partially due to lack of large-scale river cross-section data (Li et al., 2013). This simplification may affect the accuracy of simulated water depth for rivers with very irregular channel cross-section. The river channel width and bankfull depth estimated from empirical formulations may introduce uncertainties. Even though such estimation achieves reasonable accuracy at local basins, more reliable river geometry data should be considered at least for regions wherever the data is available. While global river width datasets have been developed for rivers with width >90 m (Allen & Pavelsky, 2018; Yamazaki et al., 2014), the river bankfull depth may also be derived from high-resolution remote sensing data. However, it remains challenging to upscale the observed river geometry to model resolution given the river is resolution free (Liao et al., 2022). Second, the global runoff forcing may underestimate the event extremes, such as the discharge peak during Hurricane Irene, affecting the reliability of backwater quantifications during the corresponding event. As we target at generalizing the analyses to the global scale, the bias-corrected global forcing that can capture extremes is desired. Third, we demonstrated that the direct impact of tide on the backwater effects is limited compared to LFS, but the quantification assumes that LFS and tide can be separated by removing tide from the total water level, ignoring the nonlinear interaction between tide and surge. In reality, as a dynamic component in Delaware Bay, the interaction between high tide and coastal surge may stimulate a further increase in the water level (Krien et al., 2017). Last but not least, because our simulation takes the in-situ observation as the boundary data, it does not account for the interaction between fluvial and coastal processes. The backwater effects are constrained by the prescribed boundary sea level. However, the river-ocean interface is a dynamic region that features complex multiscale processes. In the context of an intense storm, the upstream propagation of storm surge impedes river discharge, which in turn modulates the water level (Dykstra & Dzwonkowski, 2020). It remains unclear if this mutual interaction is critical in flood modeling and how it responds to sea level rise (Kulp & Strauss, 2019) and enhanced tidal dynamics (Talke & Jay, 2020) due to climate change. Thus, this study provides a basis for modeling coastal induced flooding in river basins. We aim to couple MOSART with the E3SM ocean model interactively in our next step for resolving the complex interactions at the river-ocean interface and eventually couple all land-river-ocean processes within E3SM to improve predictive understanding of compound flooding through pluvial, fluvial and coastal processes (Xu et al., 2022).

7 Conclusion

This research assesses the capability of the global-scale MOSART river transport model to simulate the coastal backwater effects at the basin scale by imposing the observed water level at the downstream boundary and using a coastal refined

495 unstructured mesh. The simulation is evaluated at two major river basins of the densely populated mid-Atlantic region. MOSART shows a reasonable agreement with the observed river discharge in both rivers and captures the water level variations in the downstream section of Delaware River, indicating the model's capability in representing the backwater effects. We performed numerical experiments and extracted extreme flood events to examine the contribution of various backwater drivers. Our analyses revealed the dependence of the backwater drivers on the river geometry and the river's

500 receiving water body. While storm surge is considered a dominant forcing, the impact of high discharge can be significant in a narrow river-like estuary, such as Delaware River, as the discharge modulates the low-frequency water level variations within the estuary. In addition, high discharge when occurring simultaneously with storm surge, creates strong compound flooding. The extreme impact of a compound event should be mainly attributed to the extended duration of the combined coastal and fluvial processes rather than extreme flood peaks. The trend analysis of the backwater drivers shows that the

505 strength of backwaters in Delaware River has been increasing in the recent decades due to SLR and more frequent storm surge. In the future, we plan to extend the current work from the mid-Atlantic region to the global domain and refine the coastal mesh globally. A framework of two-way coupled river and ocean models will be established to understand the complex river-ocean interactions.

Code and data availability

510 The water level observation can be assessed through Center for Operational Oceanographic Products and Services (CO-OPS, https://tidesandcurrents.noaa.gov/). The river discharge data are obtained from USGS National Water Information System (https://waterdata.usgs.gov/nwis). The developed MOSART code is available at https://github.com/fdongyu/E3SM/tree/fdongyu/mosart_tide.

Author contributions

515 DF and ZT designed the experiments. DE, CL and DX prepared the model mesh and parameters. DF carried out the analysis. DF and ZT wrote the initial draft of the manuscript. All authors contributed to the editing of the paper.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

collaborative Integrated Coastal Modeling (ICoM) project. PNNL is operated for DOE by Battelle Memorial Institute, United States under contract DE-AC05-76RL01830. All model simulations were performed using resources available through Research Computing at Pacific Northwest National Laboratory.

525 **References**

Alfieri, L., L. Feyen and G. Di Baldassarre (2016). Increasing flood risk under climate change: a pan-European assessment of the benefits of four adaptation strategies. Climatic Change 136(3): 507-521.

Allen, G. H. and T. M. Pavelsky (2018). Global extent of rivers and streams. Science 361(6402): 585-588.

Bakhtyar, R., K. Maitaria, P. Velissariou, B. Trimble, H. Mashriqui, S. Moghimi, A. Abdolali, A. Van der Westhuysen, Z.
530 Ma and E. Clark (2020). 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(9): e2019JC015822.

Bates, P. D., N. Quinn, C. Sampson, A. Smith, O. Wing, J. Sosa, J. Savage, G. Olcese, J. Neal and G. Schumann (2021). Combined modeling of US fluvial, pluvial, and coastal flood hazard under current and future climates. Water Resources Research 57(2): e2020WR028673.

535 Bent, G. C. and A. M. Waite (2013). Equations for estimating bankfull channel geometry and discharge for streams in Massachusetts, US Geological Survey.

Bermúdez, M., J. Farfán, P. Willems and L. Cea (2021). Assessing the effects of climate change on compound flooding in coastal river areas. Water Resources Research 57(10): e2020WR029321.

Bevacqua, E., M. I. Vousdoukas, T. G. Shepherd and M. Vrac (2020). Brief communication: The role of using precipitation
540 or river discharge data when assessing global coastal compound flooding. Natural Hazards and Earth System Sciences 20(6): 1765-1782.

Bilskie, M. and S. Hagen (2018). Defining flood zone transitions in low-gradient coastal regions. Geophysical Research Letters 45(6): 2761-2770.

Brakel, v. J. P. G. (2014). Robust peak detection algorithm using z-scores. Retrieved 2022-3-23, from
545 https://stackoverflow.com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data/22640362#22640362.

Caldwell, P. M., A. Mametjanov, Q. Tang, L. P. Van Roekel, J. C. Golaz, W. Lin, D. C. Bader, N. D. Keen, Y. Feng and R. Jacob (2019). The DOE E3SM coupled model version 1: Description and results at high resolution. Journal of Advances in Modeling Earth Systems 11(12): 4095-4146.

Camelo, J., T. L. Mayo and E. D. Gutmann (2020). Projected Climate Change Impacts on Hurricane Storm Surge Inundation
550 in the Coastal United States. Frontiers in Built Environment: 207.

Castelltort, F. X., E. Bladé, J. Balasch and M. Ribé (2020). The backwater effect as a tool to assess formative long-term flood regimes. Quaternary international 538: 29-43.

Chegini, T., G. Almeida Coelho, J. Ratcliff, C.M. Ferreira, K. Mandli, P. Burke and H.-Y. Li (2022). A Novel Framework for Parametric Analysis of Coastal Transition Zone Modeling. JAWRA Journal of the American Water Resources Association.

Chen, W.-B., W.-C. Liu and C.-Y. Wu (2013). Coupling of a one-dimensional river routing model and a three-dimensional ocean model to predict overbank flows in a complex river–ocean system. Applied Mathematical Modelling 37(9): 6163-6176.

Chow, V. T., D. R. Maidment and L. W. Mays. Applied Hydrology. McGraw-Hill Series in Water Resources and Environmental Engineering (1988).

Di Baldassarre, G., F. Laio and A. Montanari (2012). Effect of observation errors on the uncertainty of design floods. Physics and Chemistry of the Earth, Parts A/B/C 42: 85-90.

Dykstra, S. and B. Dzwonkowski (2020). The propagation of fluvial flood waves through a backwater-estuarine environment. Water Resources Research 56(2): e2019WR025743.

Dykstra, S. and B. Dzwonkowski (2021). The Role of Intensifying Precipitation on Coastal River Flooding and Compound River-Storm Surge Events, Northeast Gulf of Mexico. Water Resources Research 57(11): e2020WR029363.

Eilander, D., A. Couasnon, H. Ikeuchi, S. Muis, D. Yamazaki, H. C. Winsemius and P. J. Ward (2020). The effect of surge on riverine flood hazard and impact in deltas globally. Environmental Research Letters 15(10): 104007.

Engwirda, D. and D. Ivers (2016). Off-centre Steiner points for Delaunay-refinement on curved surfaces. Computer-Aided Design 72: 157-171.

Engwirda, D. (2017). JIGSAW-GEO (1.0): locally orthogonal staggered unstructured grid generation for general circulation modelling on the sphere. Geoscientific Model Development 10(6): 2117-2140.

Engwirda, D. and C. Liao (2021). 'UNIFIED' LAGUERRE-POWER MESHES FOR COUPLEDEARTH SYSTEM MODELLING. Zenodo. https://doi.org/10.5281/zenodo.5558988.

Esri (2012). "Topographic" [basemap]. Scale Not Given. "World Topographic Map". http://www.arcgis.com/home/item.html?id=30e5fe3149c34df1ba922e6f5bbf808f. (July 22, 2022).

Golaz, J. C., P. M. Caldwell, L. P. Van Roekel, M. R. Petersen, Q. Tang, J. D. Wolfe, G. Abeshu, V. Anantharaj, X. S. Asay-Davis and D. C. Bader (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. Journal of Advances in Modeling Earth Systems 11(7): 2089-2129.

Gori, A., N. Lin and J. Smith (2020). Assessing compound flooding from landfalling tropical cyclones on the North Carolina coast. Water Resources Research 56(4): e2019WR026788.

Gupta, H. V., H. Kling, K. K. Yilmaz and G. F. Martinez (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of hydrology 377(1-2): 80-91.

Hellmers, S. and P. Fröhle (2022). Computation of backwater effects in surface waters of lowland catchments including control structures–an efficient and re-usable method implemented in the hydrological open-source model Kalypso-NA (4.0). Geoscientific Model Development 15(3): 1061-1077.

Hendry, A., I. D. Haigh, R. J. Nicholls, H. Winter, R. Neal, T. Wahl, A. Joly-Laugel and S. E. Darby (2019). Assessing the characteristics and drivers of compound flooding events around the UK coast. Hydrology and Earth System Sciences 23(7): 3117-3139.

590 Hinkel, J., D. Lincke, A. T. Vafeidis, M. Perrette, R. J. Nicholls, R. S. Tol, B. Marzeion, X. Fettweis, C. Ionescu and A. Levermann (2014). Coastal flood damage and adaptation costs under 21st century sea-level rise. Proceedings of the National Academy of Sciences 111(9): 3292-3297.

Hirabayashi, Y., R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim and S. Kanae (2013). Global flood risk under climate change. Nature climate change 3(9): 816-821.

595 Hodges, B. R. (2013). Challenges in continental river dynamics. Environmental modelling & software 50: 16-20.

Hwang, S., S. Son, C. Lee and H.-D. Yoon (2020). Quantitative assessment of inundation risks from physical contributors associated with future storm surges: a case study of Typhoon Maemi (2003). Natural Hazards 104(2): 1389-1411.

Ikeuchi, H., Y. Hirabayashi, D. Yamazaki, S. Muis, P. J. Ward, H. C. Winsemius, M. Verlaan and S. Kanae (2017). Compound simulation of fluvial floods and storm surges in a global coupled river-coast flood model: Model development

600 and its application to 2007 Cyclone Sidr in B angladesh. Journal of Advances in Modeling Earth Systems 9(4): 1847-1862.

Krien, Y., L. Testut, A. Islam, X. Bertin, F. Durand, C. Mayet, A. Tazkia, M. Becker, S. Calmant and F. Papa (2017). Towards improved storm surge models in the northern Bay of Bengal. Continental Shelf Research 135: 58-73.

Kulp, S. A. and B. H. Strauss (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. Nature communications 10(1): 1-12.

605 Lamb, M. P., J. A. Nittrouer, D. Mohrig and J. Shaw (2012). Backwater and river plume controls on scour upstream of river mouths: Implications for fluvio-deltaic morphodynamics. Journal of Geophysical Research: Earth Surface 117(F1).

Lang, M., T. B. Ouarda and B. Bobée (1999). Towards operational guidelines for over-threshold modeling. Journal of hydrology 225(3-4): 103-117.

Leathers, D. J., M. L. Malin, D. B. Kluver, G. R. Henderson and T. A. Bogart (2008). Hydroclimatic variability across the

610 Susquehanna River Basin, USA, since the 17th century. International Journal of Climatology: A Journal of the Royal Meteorological Society 28(12): 1615-1626.

Lehner, B., K. Verdin and A. Jarvis (2008). New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union 89(10): 93-94.

Leung, L. R., D. C. Bader, M. A. Taylor, and R. B. McCoy (2020). An introduction to the Energy Exascale Earth System

615 Model (E3SM) Special Collection: Goals, Science Drivers, Development, and Analysis. Journal of Advancing Modeling of Earth Systems, 12, https://doi.org/10.1029/2019MS001821.

Li, H., M. S. Wigmosta, H. Wu, M. Huang, Y. Ke, A. M. Coleman and L. R. Leung (2013). A physically based runoff routing model for land surface and earth system models. Journal of Hydrometeorology 14(3): 808-828.

Li, H.-Y., L. R. Leung, A. Getirana, M. Huang, H. Wu, Y. Xu, J. Guo and N. Voisin (2015). Evaluating global streamflow simulations by a physically based routing model coupled with the community land model. Journal of Hydrometeorology 16(2): 948-971.

Li, H.-Y., Z. Tan, H. Ma, Z. Zhu, G. W. Abeshu, S. Zhu, S. Cohen, T. Zhou, D. Xu and L. R. Leung (2022). A new large-scale suspended sediment model and its application over the United States. Hydrology and Earth System Sciences 26(3): 665-688.

Liao, C., T. Tesfa, Z. Duan and L. R. Leung (2020). Watershed delineation on a hexagonal mesh grid. Environmental Modelling & Software 128: 104702.

Liao, C., T. Zhou, D. Xu, R. Barnes, G. Bisht, H.-Y. Li, Z. Tan, T. Tesfa, Z. Duan and D. Engwirda (2022). Advances in hexagon mesh-based flow direction modeling. Advances in Water Resources: 104099.

Liao. C. (2022). HexWatershed: a mesh independent flow direction model for hydrologic models (0.1.1). Zenodo. https://doi.org/10.5281/zenodo.6425881.

Luo, X., H.-Y. Li, L. R. Leung, T. K. Tesfa, A. Getirana, F. Papa and L. L. Hess (2017). Modeling surface water dynamics in the Amazon Basin using MOSART-Inundation v1. 0: impacts of geomorphological parameters and river flow representation. Geoscientific Model Development 10(3): 1233-1259.

Mao, Y., T. Zhou, L. R.Leung, T. K. Tesfa, H. Y. Li, K. Wang, Z. Tan and A. Getirana (2019). Flood Inundation Generation Mechanisms and Their Changes in 1953–2004 in Global Major River Basins. Journal of Geophysical Research: Atmospheres. 124(22), pp.11672-11692.

Marsooli, R., N. Lin, K. Emanuel and K. Feng (2019). Climate change exacerbates hurricane flood hazards along US Atlantic and Gulf Coasts in spatially varying patterns. Nature communications 10(1): 1-9.

Mikhailov, V. and S. Gorin (2012). New definitions, regionalization, and typification of river mouth areas and estuaries as their parts. Water Resources 39(3): 247-260.

Miller, B. (2021). Hurricane ida forces mississippi river to reverse flow. Retrieved 2022-03-03, from https://www.cnn.com/2021/08/29/weather/mississippi-river-hurricane-ida/index.html.

Muis, S., M. Verlaan, R. J. Nicholls, S. Brown, J. Hinkel, D. Lincke, A. T. Vafeidis, P. Scussolini, H. C. Winsemius and P. J. Ward (2017). A comparison of two global datasets of extreme sea levels and resulting flood exposure. Earth's Future 5(4): 379-392.

Muñoz, D., H. Moftakhari and H. Moradkhani (2020). Compound effects of flood drivers and wetland elevation correction on coastal flood hazard assessment. Water Resources Research 56(7): e2020WR027544.

Nasr, A. A., T. Wahl, M. M. Rashid, P. Camus and I. D. Haigh (2021). Assessing the dependence structure between oceanographic, fluvial, and pluvial flooding drivers along the United States coastline. Hydrology and Earth System Sciences 25(12): 6203-6222.

Neal, J., I. Villanueva, N. Wright, T. Willis, T. Fewtrell and P. Bates (2012). How much physical complexity is needed to model flood inundation? Hydrological Processes 26(15): 2264-2282.

21

Orton, P., N. Lin, V. Gornitz, B. Colle, J. Booth, K. Feng, M. Buchanan, M. Oppenheimer and L. Patrick (2019). New York City panel on climate change 2019 report chapter 4: coastal flooding. Annals of the New York Academy of Sciences 1439(1): 95-114.

Paiva, R. C., W. Collischonn and D. C. Buarque (2013). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon. Hydrological Processes 27(3): 333-346.

Rahmstorf, S. (2017). Rising hazard of storm-surge flooding. Proceedings of the National Academy of Sciences 114(45): 11806-11808.

Rayson, M. D., E. S. Gross and O. B. Fringer (2015). Modeling the tidal and sub-tidal hydrodynamics in a shallow, micro-tidal estuary. Ocean Modelling 89: 29-44.

Santiago-Collazo, F. L., M. V. Bilskie and S. C. Hagen (2019). A comprehensive review of compound inundation models in low-gradient coastal watersheds. Environmental Modelling & Software 119: 166-181.

Sharp, J. H. (1983). The Delaware estuary: Research as background for estuarine management and development. Univ. Delaware and New Jersey Mar. Sci. Consort.

Sun, N., M. S. Wigmosta, D. Judi, Z. Yang, Z. Xiao and T. Wang (2021). Climatological analysis of tropical cyclone impacts on hydrological extremes in the Mid-Atlantic region of the United States. Environmental Research Letters 16(12): 124009.

Syvitski, J. P. and Y. Saito (2007). Morphodynamics of deltas under the influence of humans. Global and Planetary Change 57(3-4): 261-282.

Talke, S. A. and D. A. Jay (2020). Changing tides: The role of natural and anthropogenic factors. Annual review of marine science 12: 121-151.

Tellman, B., J. Sullivan, C. Kuhn, A. Kettner, C. Doyle, G. Brakenridge, T. Erickson and D. Slayback (2021). Satellite imaging reveals increased proportion of population exposed to floods. Nature 596(7870): 80-86.

Tosunoglu, F. and O. Kisi (2017). Trend analysis of maximum hydrologic drought variables using Mann–Kendall and Şen's innovative trend method. River Research and Applications 33(4): 597-610.

Towner, J., H. L. Cloke, E. Zsoter, Z. Flamig, J. M. Hoch, J. Bazo, E. Coughlan de Perez and E. M. Stephens (2019). Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin. Hydrology and Earth System Sciences 23(7): 3057-3080.

Trigg, M. A., M. D. Wilson, P. D. Bates, M. S. Horritt, D. E. Alsdorf, B. R. Forsberg and M. C. Vega (2009). Amazon flood wave hydraulics. Journal of Hydrology 374(1-2): 92-105.

Trigg, M., C. Birch, J. Neal, P. Bates, A. Smith, C. Sampson, D. Yamazaki, Y. Hirabayashi, F. Pappenberger and E. Dutra (2016). The credibility challenge for global fluvial flood risk analysis. Environmental Research Letters 11(9): 094014.

Turnipseed, D. P. and V. B. Sauer (2010). Discharge measurements at gaging stations, US Geological Survey.

U.S. Geological Survey, 2016, National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), accessed [Sep 20, 2022], at URL [http://waterdata.usgs.gov/nwis/].

22

Valle-Levinson, A. (1995). Observations of barotropic and baroclinic exchanges in the lower Chesapeake Bay. Continental Shelf Research 15(13): 1631-1647.

Ward, P. J., A. Couasnon, D. Eilander, I. D. Haigh, A. Hendry, S. Muis, T. I. Veldkamp, H. C. Winsemius and T. Wahl (2018). Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. Environmental Research Letters 13(8): 084012.

Ward, N. D., J. P. Megonigal, B. Bond-Lamberty, V. L. Bailey, D. Butman, E. A. Canuel, H. Diefenderfer, N. K. Ganju, M. A. Goñi and E. B. Graham (2020). Representing the function and sensitivity of coastal interfaces in Earth system models. Nature Communications 11(1): 1-14.

Whitney, M. M. and R. W. Garvine (2006). Simulating the Delaware Bay buoyant outflow: Comparison with observations. Journal of Physical Oceanography 36(1): 3-21.

Wu, W., S. Westra and M. Leonard (2021). Estimating the probability of compound floods in estuarine regions. Hydrology and Earth System Sciences 25(5): 2821-2841.

Xiao, Z., Z. Yang, T. Wang, N. Sun, M. Wigmosta and D. Judi (2021). Characterizing the non-linear interactions between tide, storm surge, and river flow in the delaware bay estuary, United States. Front. Mar. Sci. 8: 715557.

Xu, D., G. Bisht, T. Zhou, L. R. Leung and M. Pan (2022). Development of Land-River Two-Way Hydrologic Coupling for Floodplain Inundation in the Energy Exascale Earth System Model. Journal of Advances in Modeling Earth Systems: e2021MS002772.

Yamazaki, D., H. Lee, D. E. Alsdorf, E. Dutra, H. Kim, S. Kanae and T. Oki (2012). Analysis of the water level dynamics simulated by a global river model: A case study in the Amazon River. Water Resources Research 48(9).

Yamazaki, D., F. O'Loughlin, M. A. Trigg, Z. F. Miller, T. M. Pavelsky and P. D. Bates (2014). Development of the global width database for large rivers. Water Resources Research 50(4): 3467-3480.

Yang, Y., M. Pan, P. Lin, H. E. Beck, Z. Zeng, D. Yamazaki, C. H. David, H. Lu, K. Yang and Y. Hong (2021). Global Reach-Level 3-Hourly River Flood Reanalysis (1980–2019). Bulletin of the American Meteorological Society 102(11): E2086-E2105.

Zhang, Y. J., F. Ye, H. Yu, W. Sun, S. Moghimi, E. Myers, K. Nunez, R. Zhang, H. Wang and A. Roland (2020). Simulating compound flooding events in a hurricane. Ocean Dynamics 70(5): 621-640.

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

Zhang, Q., L. Zhang, D. She, S. Wang, G. Wang and S. Zeng (2021). Automatic procedure for selecting flood events and identifying flood characteristics from daily streamflow data. Environmental Modelling & Software 145: 105180.

Zscheischler, J., O. Martius, S. Westra, E. Bevacqua, C. Raymond, R. M. Horton, B. van den Hurk, A. AghaKouchak, A. Jézéquel and M. D. Mahecha (2020). A typology of compound weather and climate events. Nature reviews earth & environment 1(7): 333-347.
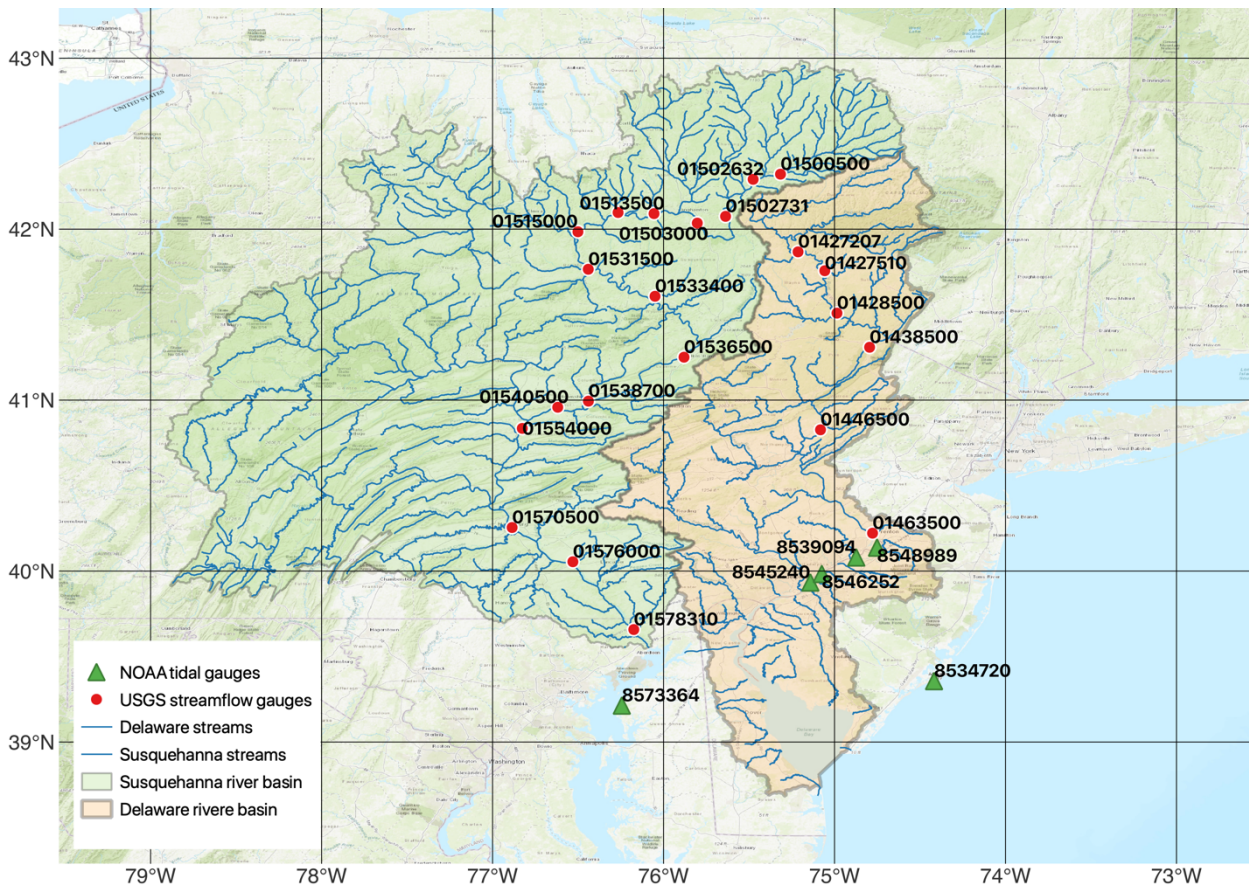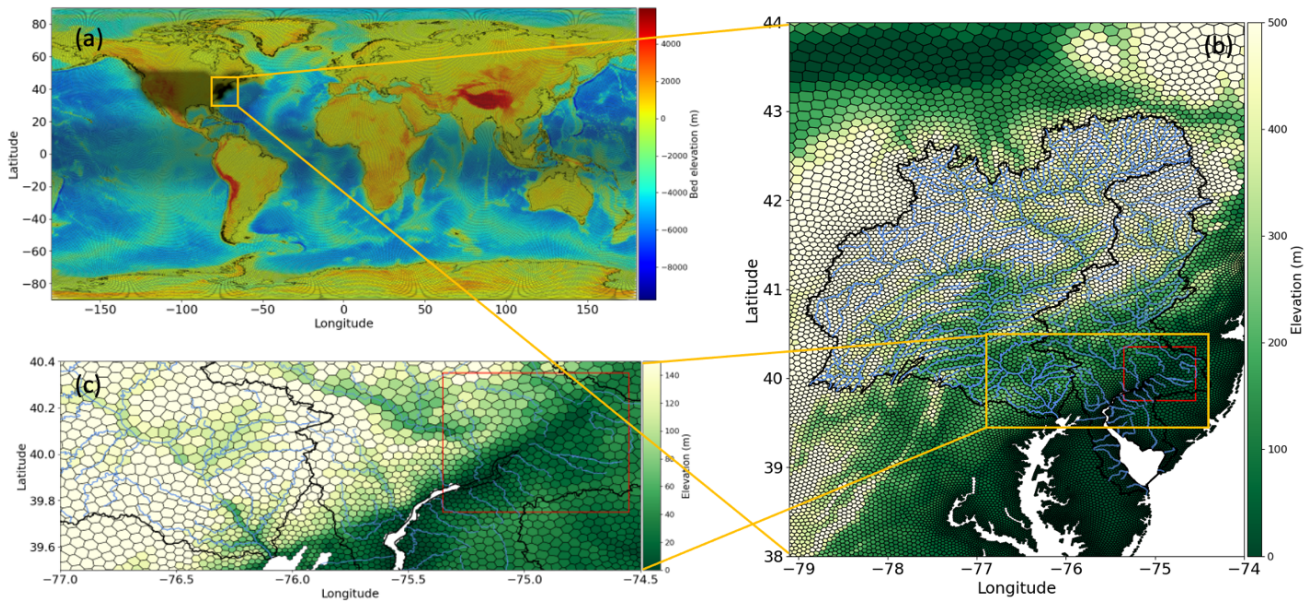
23

Figure 1: An overview of Susquehanna River Basin (SRB) and Delaware River Basin (DRB). This map is created using the free and open source QGIS on the world topographic map (ESRI, 2012).

725

Figure 2: (a) The global unstructured mesh of E3SM. (b) A magnified view of the mid-Atlantic region (orange rectangle in (a). (c) A magnified view near the mouths of Susquehanna River and Delaware River (orange rectangle in (b). The red rectangles in (b) and (c) represent the downstream section of Delaware River used for backwater propagation analysis.
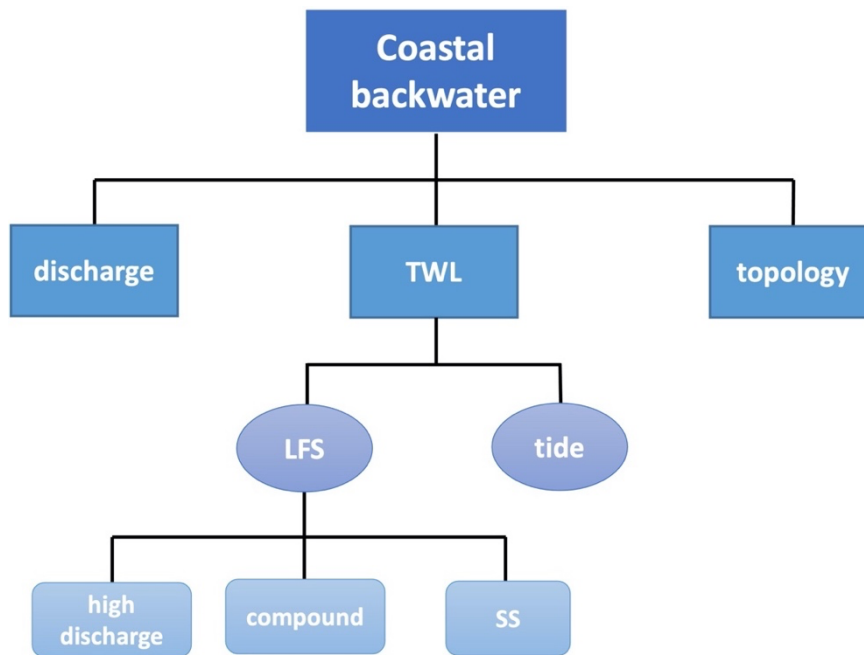
730

Figure 3: Decomposition of the drivers of backwater effects. TWL is total water level, LFS is low-frequency surge, and SS is storm surge.
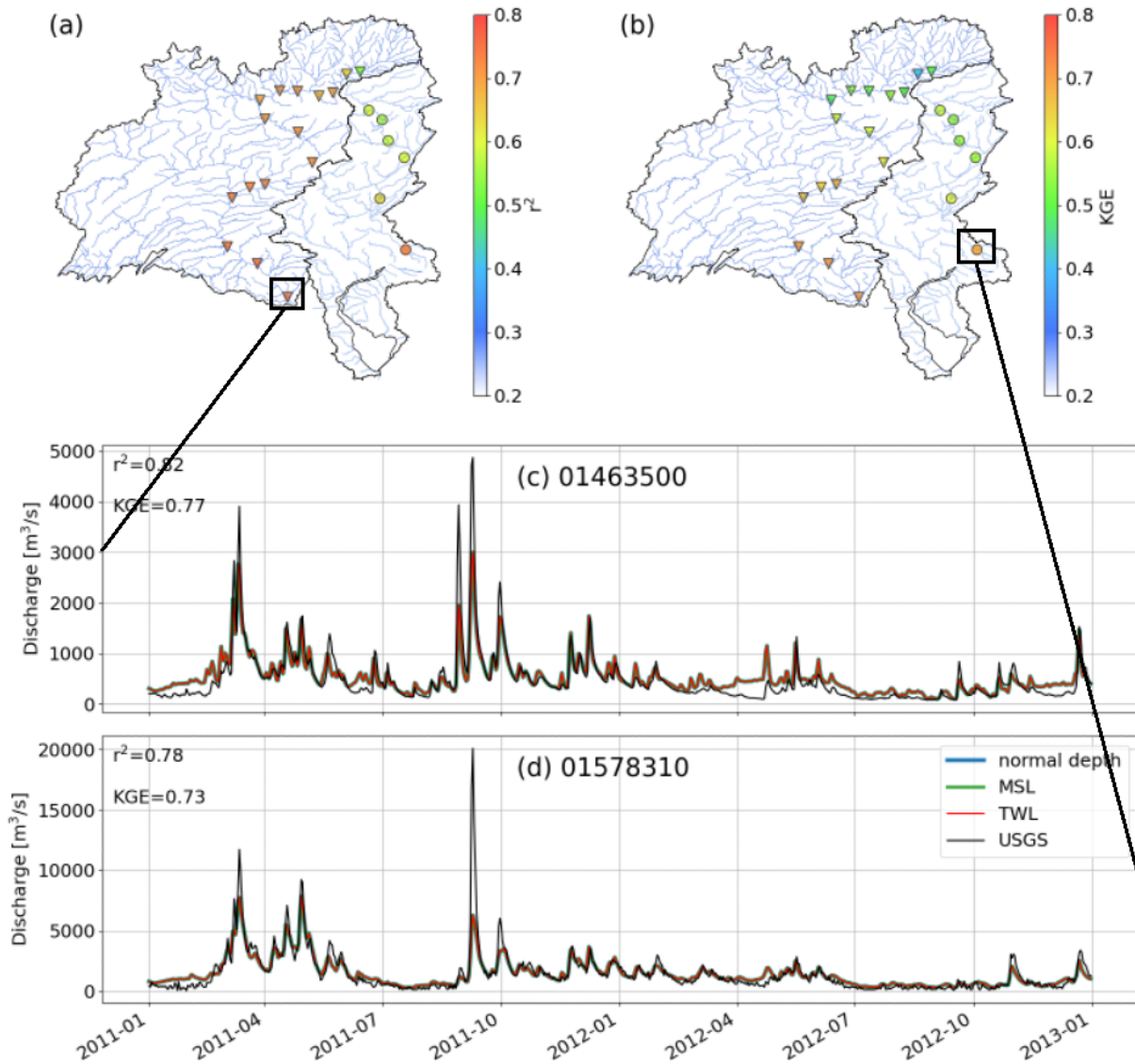
735 Figure 4: The river discharge evaluation in SRB and DRB: (a) r^2, (b) KGE, (c) hydrograph at USGS gauge 01463500, (d) hydrograph at USGS gauge 01578310. The triangles and the circles in (a) and (b) represent the USGS gauges in the main stem of Susquehanna River and Delaware River, respectively.

26
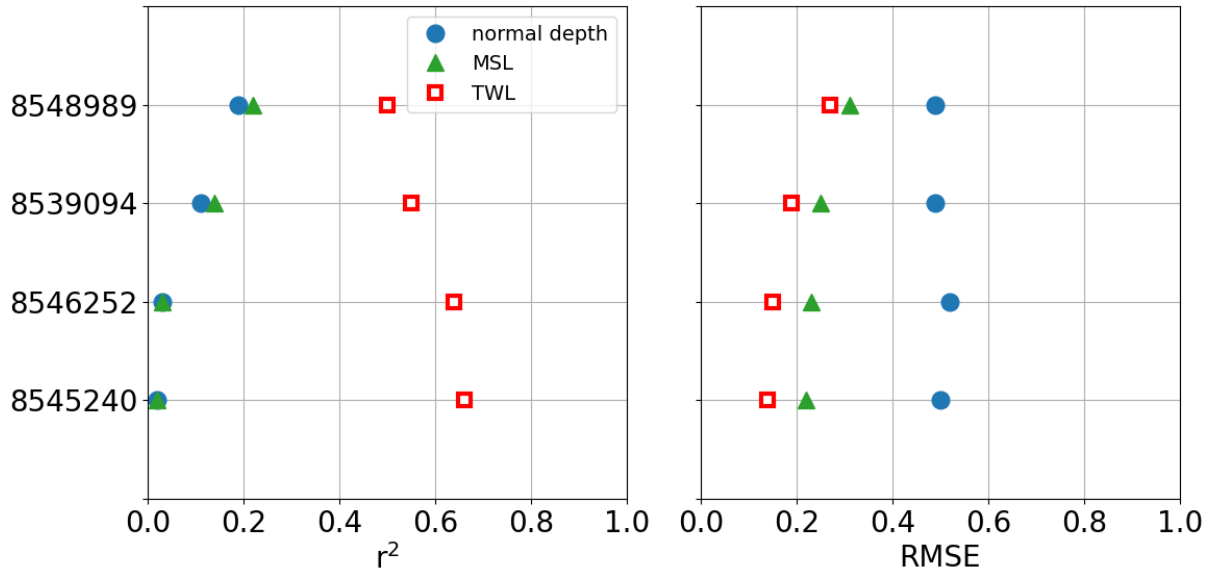
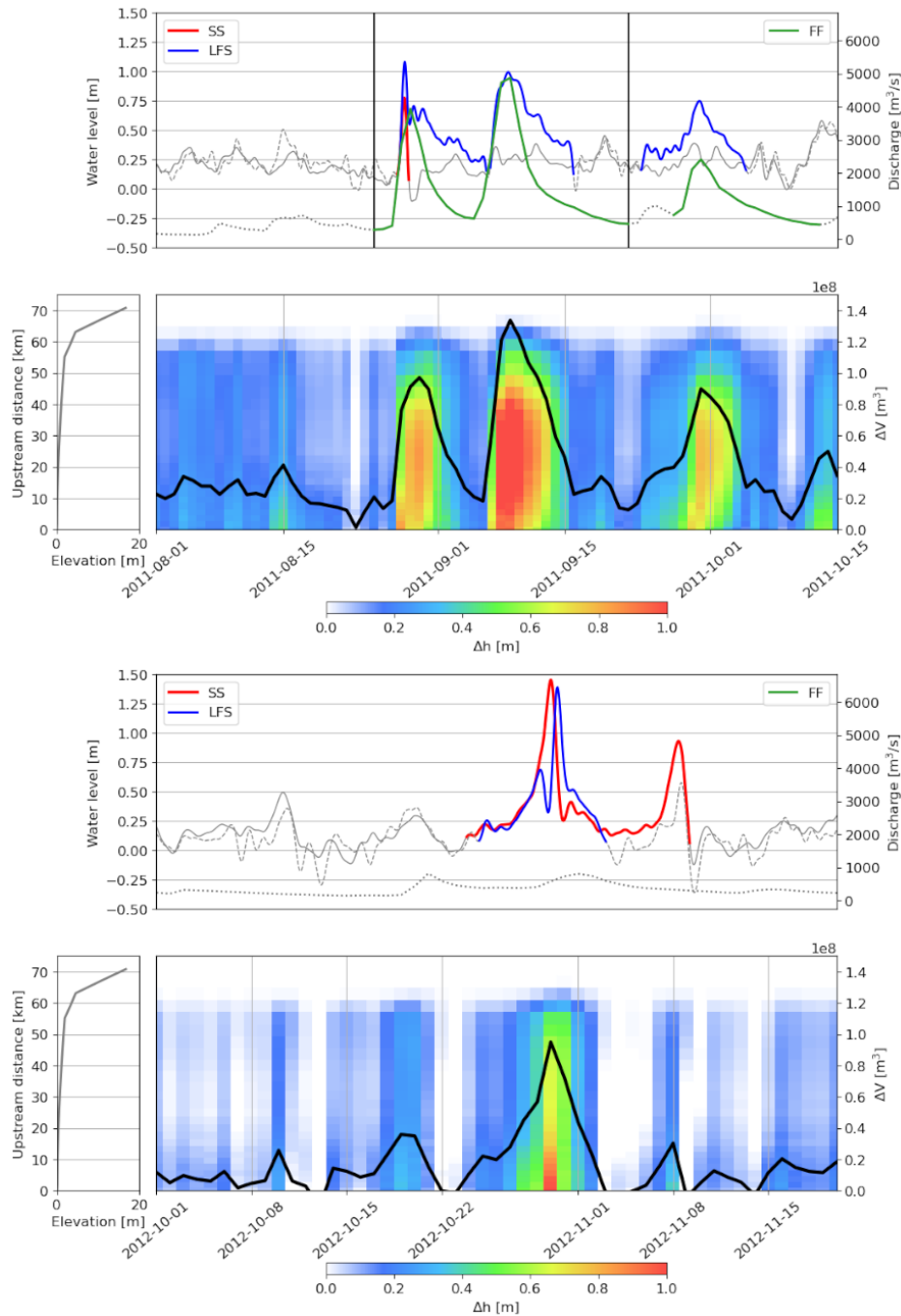**Figure 5: The r^2 and RMSE of the MOSART simulations at the 4 NOAA gauges.**
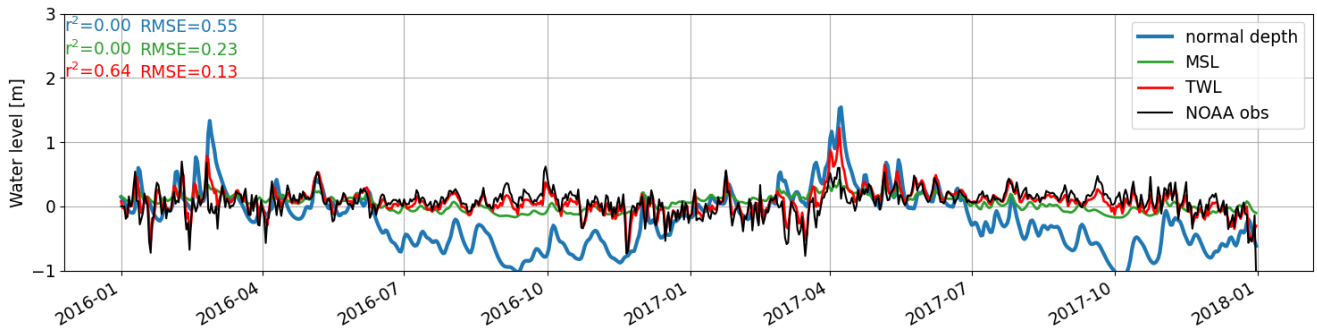
Figure 6: Comparison of daily-averaged simulated and observed water level at NOAA gauge 8546252.

Figure 7: The top panels are the three types of selected extreme events (SS, LFS, FF) overlaid on the corresponding time series of SS, LFS and discharge, represented by grey solid, dashed and dotted lines, respectively, for Hurricane Irene and Tropical Storm Lee (a) and Hurricane Sandy (b). The compound flood event is marked between the two black vertical lines. The bottom left panels are the riverbed elevation along the backwater propagation extent in Delaware River and the bottom right panels are Δh (color shading) along the upstream distance and ΔV (black curve) derived from the MOSART simulations.
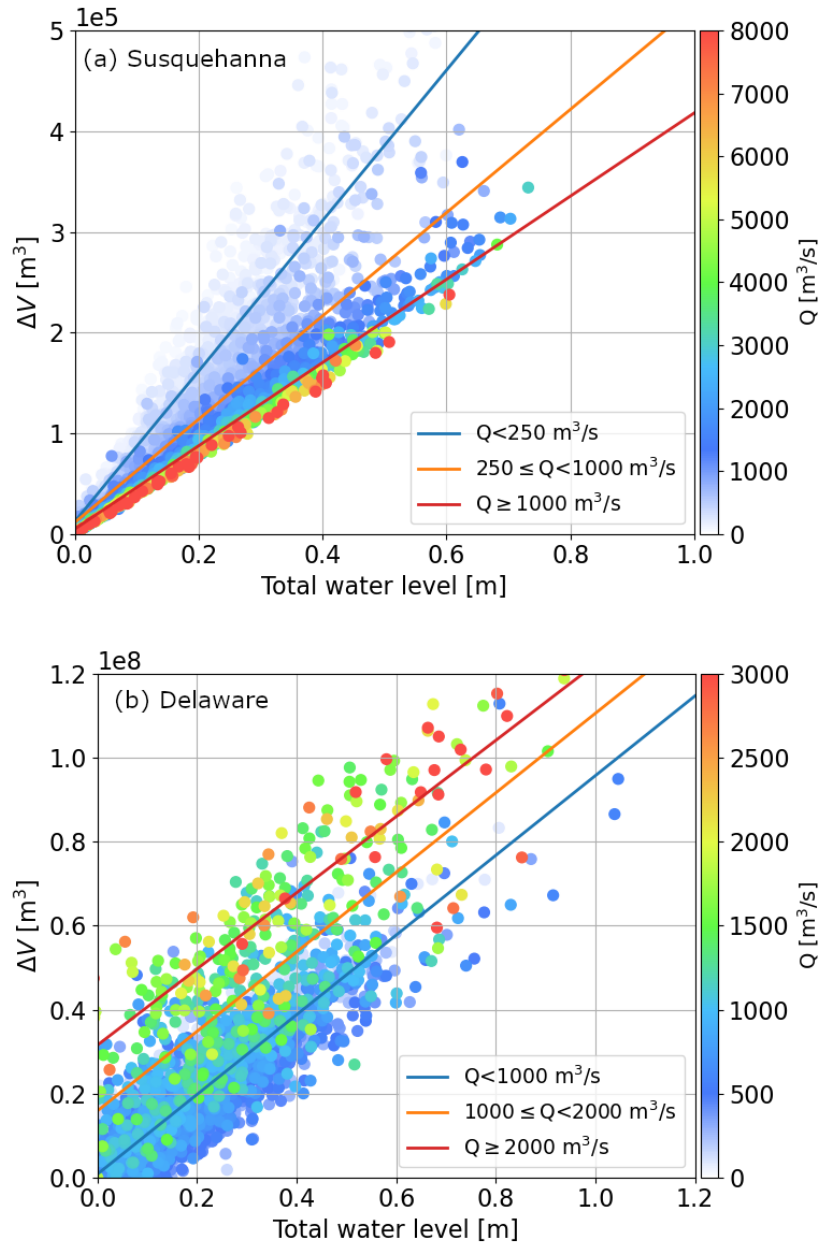
Figure 8: Scatter plot of ΔV against TWL in Susquehanna River (a) and Delaware River (b). Colored circles represent the corresponding river discharge (Q). The TWL and Q data are respectively obtained from gauges 8573364 and 01578310 for Susquehanna River, and from gauges 8546252 and 01463500 for Delaware River. The solid lines are the fitted linear regression under different discharge ranges.
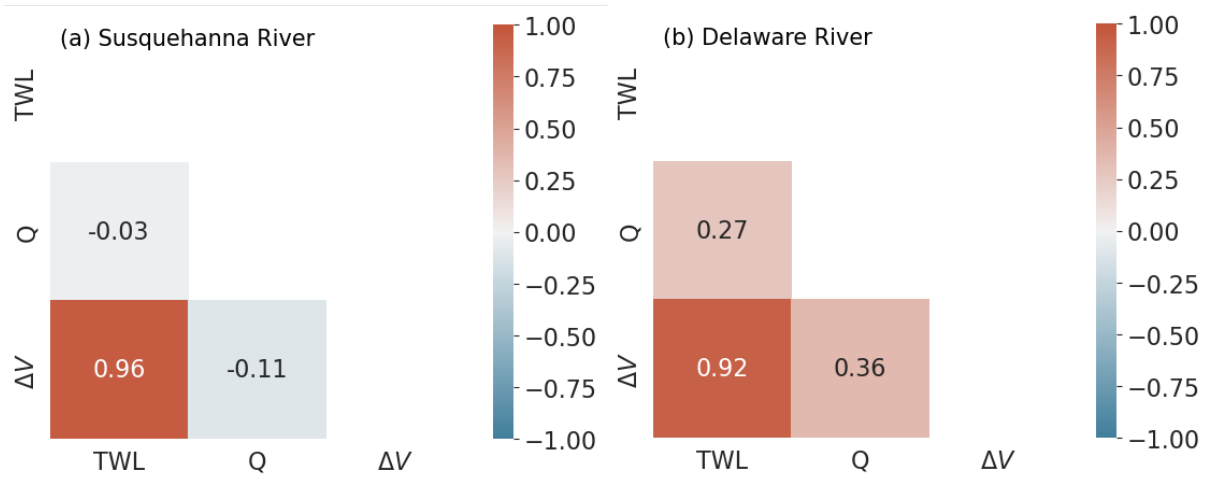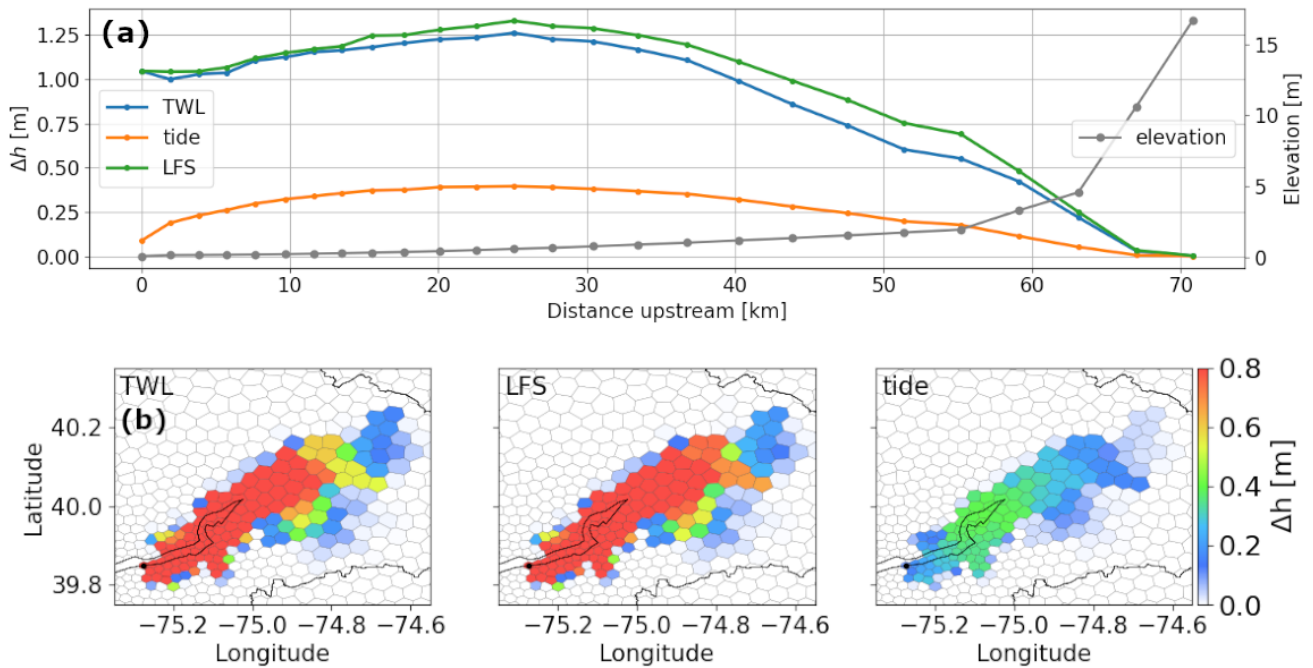
Figure 9: The correlation coefficient (r) matrix of ΔV, Q and TWL in Susquehanna River (a) and Delaware River (b).

Figure 10: (a) The along channel profiles of the maximum Δh obtained using the TWL, SS and tide configurations. The river outlet is at x = 0 km. (b) The corresponding spatial map of the maximum Δh over a downstream region of Trenton. This region is specified as the red rectangle in Figure 2. The black dots represent the Delaware River outlet.
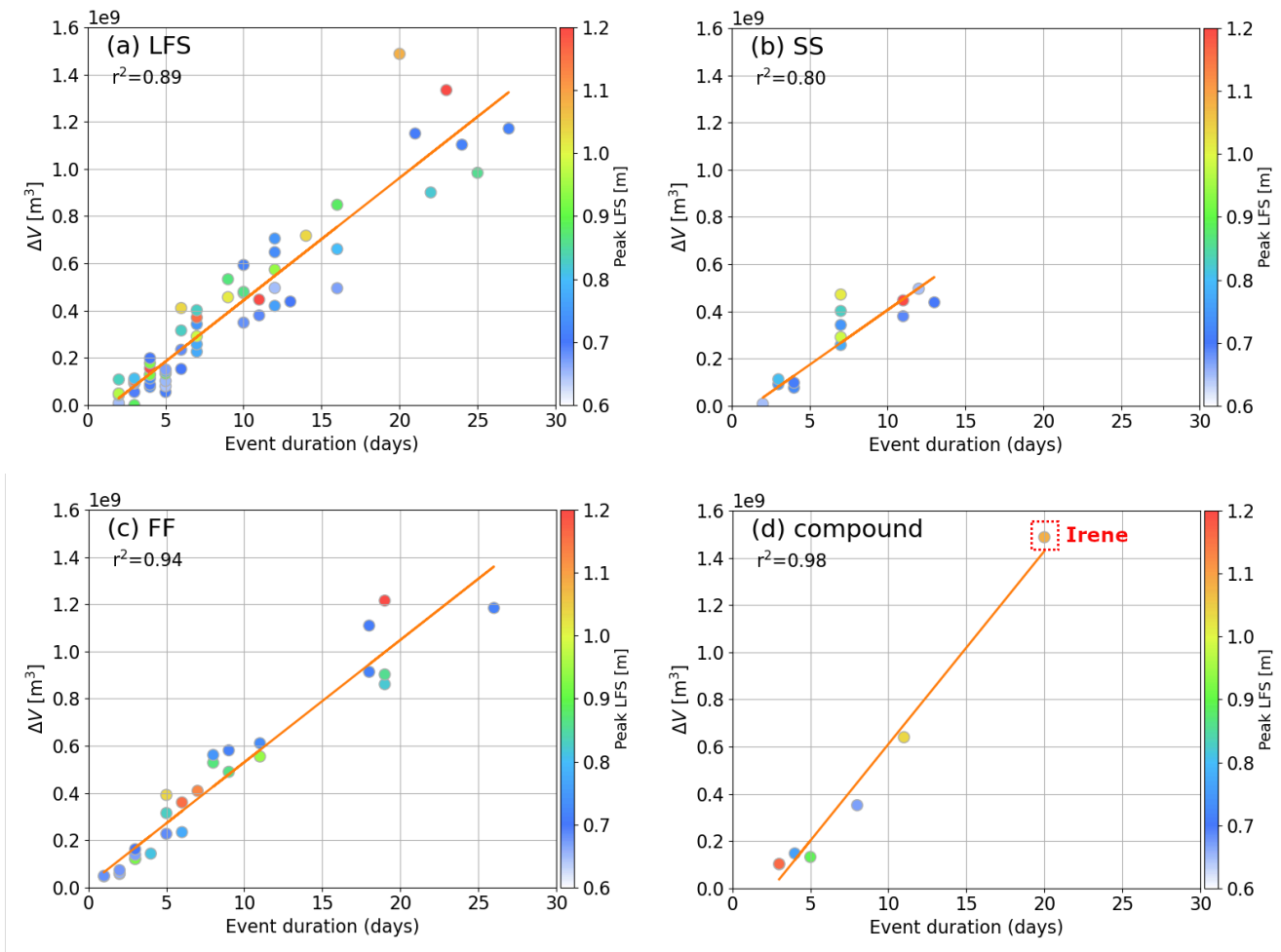
Figure 11: Scatter plot of the accumulated ΔV against the event duration in days for the LFS events (a), the SS events (b), the FF events (c) and the compound flood events (d). The color represents the corresponding peak LFS.
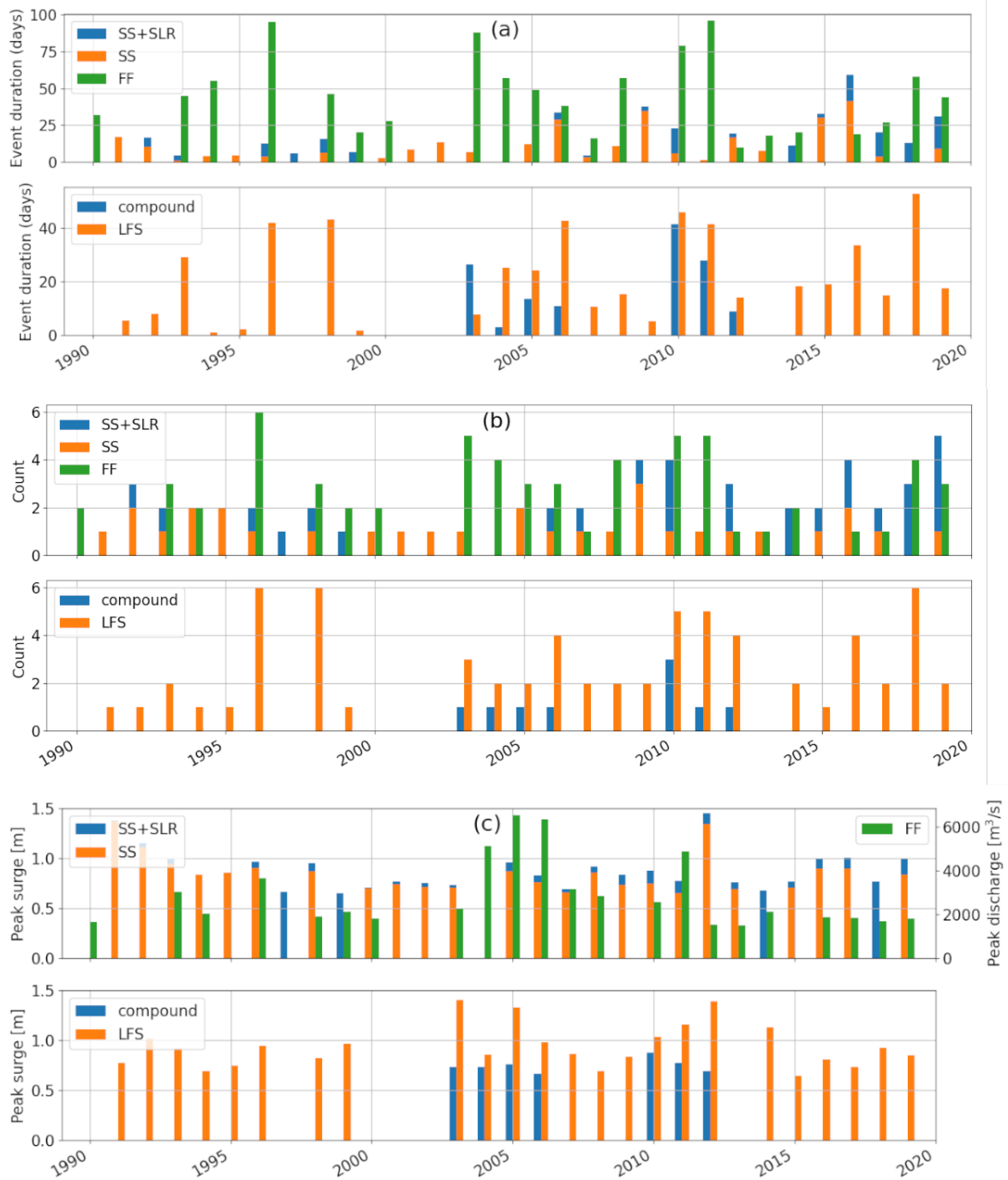
775

Figure 12: The interannual variability of (a) the event annual duration, (b) the event occurrence and (c) the maximum peak value for different types of flood events.
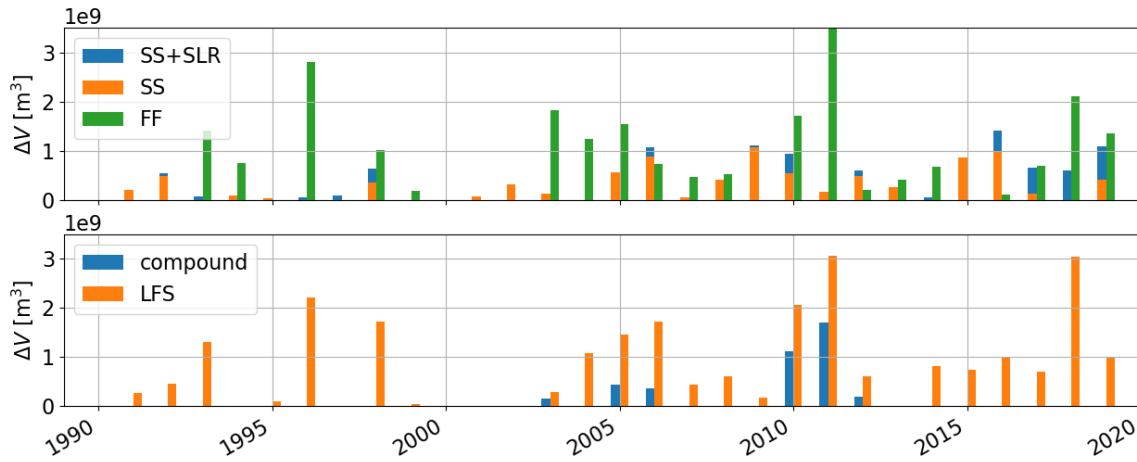
Figure 13: Similar to Figure 12 but for the interannual variability of ΔV.

Table 1: The MK statistics of long-term trends in annual duration, occurrence, peak value, and ΔV.

		SS+SLR	SS	FF	compound	LFS
Annual duration	Z	2.064	0.809	0.794	-0.091	1.674
	p-value	0.039	0.419	0.427	0.928	0.094
Occurrence	Z	2.157	-0.286	0.728	0.976	2.032
	p-value	0.031	0.775	0.467	0.329	0.042
Peak value	Z	0.446	-0.986	0.361	0.968	0.950
	p-value	0.656	0.324	0.718	0.333	0.342
ΔV	Z	2.720	1.900	1.286	0.742	1.900
	p-value	0.007	0.057	0.199	0.458	0.057