Title: Integrated Coastal-River-Urban Total Water Level Forecast System for Tidal Rivers: Calibration, Validation, and Operational Evaluation

Author names and affiliations:

Arslaan Khalid[1], Celso Ferreira[1], Jason Elliott[2]

[1]Department of Civil, Environmental and Infrastructure Engineering, George Mason University

[2]NOAA/National Weather Service, Baltimore/Washington Forecast Office

Corresponding author:

Arslaan Khalid[1,*]

[1] Department of Civil, Environmental and Infrastructure Engineering,

George Mason University, 4400 University Drive, Fairfax, VA 22030, USA

*Corresponding Author: akhalid6@gmu.edu; Tel.: +1-703-678-9528

Key Points:

1. Ocean scale surge guidance systems underrepresent total water level in upstream tidal rivers in the Chesapeake Bay
2. Lack of river flows, urban runoff and local winds leads to inaccurate total water level forecasting
3. Ensemble-based total water level forecasting improved predictions for upstream tidal rivers

Integrated Coastal-River-Urban Total Water Level Forecast System for Tidal Rivers: Calibration, Validation, and Operational Evaluation

Arslaan Khalid[1], Celso Ferreira[1], Jason Elliott[2]

[1]Department of Civil, Environmental and Infrastructure Engineering, George Mason University

[2]NOAA/National Weather Service, Baltimore/Washington Forecast Office

Abstract

Existing real-time coastal flooding guidance systems in the US tend to underestimate total water level (TWL) predictions in upstream tidal areas of the Chesapeake Bay rivers, impacting flood forecasts for highly vulnerable areas, such as the National Capital Region. These under-predictions are mostly due to missing physical processes, lack of integration between hydrological and hydrodynamic models, and an oversimplification of the model setups used to predict TWL. In this study, an integrated TWL forecast system was introduced, where a high-resolution two-dimensional coastal storm surge model (ADCIRC) was implemented to simulate the combined influence of various flood drivers (storm tide, river flows, urban runoff, and local wind forcing) in the Potomac River. In this framework, the downstream boundaries of storm tide predictions are provided by existing coastal guidance systems, whereas, streamflow forecasts at upstream rivers and local urban runoff are provided by the National Weather Service and the National Water Model. Additionally, high-resolution wind fields from the North American Mesoscale and the National Blend of Models are added to account for local wind effects on TWL. This model setup was successfully validated with a range of historical events and it also demonstrated improved forecast performance against the existing large-scale coastal guidance systems in a reforecast evaluation during 2020. Unlike other studies, we provided a comprehensive evaluation on the influence of individual flood drivers on TWL modeling and clearly demonstrated that the absence of one or more flood drivers in the model framework can underestimate simulated TWL in the National Capital Region.

Keywords: *ADCIRC; Flood Forecasts; Total Water Level; Washington, DC; Tidal Potomac.*

1. Introduction

Coastal metropolitan cities are dynamic and mostly located at the interface of multiple flood hazards (Depietri et al., 2018). These coastal communities are frequently hit by hurricanes and tropical storms in the US Atlantic and Gulf of Mexico coasts; therefore, subject to both heavy inland rainfall and coastal storm surge (Ray et al., 2011). The total global exposure from river and coastal flooding was estimated to be 46 trillion USD in 2010, with a probable increase of up to 158 trillion USD by 2050 (Jongman et al., 2012). However, these floods are becoming increasingly

common in the tidal areas away from the open coast. Co-occurrence of different flood drivers in tidal areas like storm surges, river flow, rainfall-runoff, sea-level rise, and wind can result in increased water levels leading to compound floods (Herdman et al., 2018). The number of compound flooding events have increased significantly over the past century in many coastal cities (Wahl et al., 2015). Given an increase in flood events due to sea level rise (SLR) change (Atkinson et al., 2012; Zhong et al., 2008), as well as their intrinsic complexity in urban areas surrounding tidal rivers, accurate predictions of combined coastal-river-urban flooding are essential for cost-effective storm mitigation, emergency management plans, flood insurance, and planning. In 2019, for instance, one of the National Oceanographic and Atmospheric Administration (NOAA) monitoring station in the National Capital Region (Washington, DC) recorded water levels higher than the National Weather Service (NWS) "Action" flood level during portions of 195 days, compared to 150 days in 2017.

In some areas, the storm surge and wave components play the most important role in certain coastal flooding events; for others, the predominant component is the freshwater from the upstream river flow (Dresback, Fleming, Blanton, Kaiser, Gourley, Tromble, Luettich, Kolar, Hong, Cooten, et al., 2013). Co-occurrence of rainfall and coastal flooding can be more destructive, in contrast to if they occur separately (Ikeuchi et al., 2017). Exceptionally high tides, also called "King tide" in combination with slight storm surges can significantly increase the flooding potential in vulnerable areas (Loftis & Forrest, 2018). Atmospheric forcing (especially winds at 10m height) also play an important role in generating the local changes to water levels based on the station location and prevailing winds (Möller et al., 2001). Poor handling of storm water runoff by the community also results in increased volume of water entering the rivers (Walsh et al., 2012). Since these processes (surge, river flow, runoff and local winds) have a direct effect on the total water levels, absence of and variability in one or more physical processes directly leads to inaccurate estimates of total water levels (Lyddon et al., 2018). A comprehensive review of the studies of extreme flood events emphasized the importance of including multiple flood drivers (surge, river flow, runoff) in modeling tools for more accurate flood forecasts (Santiago-Collazo et al., 2019).

The lack of physical processes and its integration in the numerical modeling poses challenges for accurate total water level forecasting in real-time (Tshimanga et al., 2016). Current storm surge prediction methods are limited in their precision as they may regard tide and surge as separate processes, while completely ignoring the significance of a coupled tide-surge interaction (Bobanović et al., 2006). Additionally, fewer forecast systems extend the computational domain into shallower and narrower reaches of upstream rivers, leading to inaccurate hydrodynamic modeling in these areas. This simplification in the total water level forecasting is often adopted to avoid large computational costs related to high-resolution modeling in real-time environments (Ikeuchi et al., 2017). In addition to the simplification within the models, most of the models guidance and the NWS TWL forecasts are deterministic single-value outputs. In order to provide a better estimation of uncertainty and improving TWL accuracy, forecasters have also started considering ensemble forecasts. Furthermore, various studies investigated the influence of only

storm surge, or stream flows, or a combination of both on the water level modeling (J. Garzon & Ferreira, 2016; Herdman et al., 2018; Mashriqui et al., 2014; Svensson & Jones, 2004; Wahl et al., 2015; Wu et al., 2018; Zheng et al., 2013), therefore, a better understanding of the impact of flood drivers in TWL prediction needs to be developed for upstream tidal areas in complex estuaries.

The National Capital Region (NCR) of the US, located at the confluence of the Anacostia and the Potomac River, both of which are under the tidal influence of the Chesapeake Bay, is a perfect example of a region of national relevance that is impacted by compound flooding events (Sumi & Ferreira, 2019). Furthermore, existing flood guidance systems in operation underestimate the total water levels for the NCR area (Khalid & Ferreira, 2020). Even though the NWS issues watches and warnings for floods to protect life and property damage, the threat from flooding stays considerably high. The majority of forecast models currently in use do not integrate a range of forcing, such as river flows, urban runoff and local winds, to reflect compound events (Herdman et al., 2018). For instance, the existing two-dimensional (2D) or three-dimensional (3D) estuary-ocean models running at NOAA in an operational environment are one way coupled (Kourafalou et al., 2015) and do not account for observed or forecast freshwater in their calculations (Mashriqui et al., 2014). Additionally, the current unsteady flow models used by the NWS for riverine flood forecasting do not include the effect of urban runoff and local wind forcings (Mashriqui et al., 2014). While existing guidance systems provide the water level forecasts and the downstream boundary conditions for river scale forecasting, the simplification and lack of physical processes in the modeling framework results in significant underestimations (Herdman et al., 2018).

The objective of this study is to evaluate the interaction of coastal-river-urban processes that play an important role in real-time forecasting of compound flooding in tidal areas and demonstrate how the integration of these physical processes leads to an improved real-time TWL prediction in tidal rivers. Furthermore, the study aims to better understand the influences of storm surge, river discharge, urban runoff, and local winds on total water level predictions and provide recommendations for developing ensemble-based integrated total water level forecast systems for tidal rivers in large estuaries.

2. Methods

In order to assess the relevance of an integrated framework for total water level modeling in upstream tidal areas, a dedicated hydrodynamic numerical modeling domain (section 2.2.2) was set-up for Potomac River. This framework was calibrated for accurate tidal modeling in the upstream tidal areas (section 2.2.3), where existing ocean scale coastal guidance systems mostly under-estimate extreme water levels under certain conditions. This calibration also focused on finding an optimal numerical mesh resolution that can allow high resolution modeling while keeping computational costs to a moderate level. The calibrated modeling framework was then validated against nine historical flooding events (section 2.3) to ensure model accuracy when

multiple flood drivers act simultaneously. We then focused on quantifying the change in water levels as a result of individual flood drivers' boundary conditions based on several hypothetical scenarios (section 2.4). Using real-time forecasted outputs from existing coastal and hydrologic guidance systems, we performed a reforecast test of several flooding events of 2020 within our integrated modeling framework to find the best set of boundary forcing for total water level forecasting (section 2.5). Lastly, through one case study, we demonstrated the system's ability to use ensemble based total water level forecasting (section 2.6) and assessed its performance against deterministic forecasting.

2.1. Study Area

The Potomac River is the largest tributary of the Chesapeake Bay, whose length from Washington, DC to the Chesapeake Bay is about 166 km. The tidal Potomac River is influenced by the tidal signal due to its downstream connection with the Chesapeake Bay. On the other hand, given the steep slope and the large incoming discharge, the water levels between Little Falls Pump Station (LFMD) and Chain Bridge are not tidally influenced, marking the end of the tidal signal influence in the river. The average daily flow at the LFMD is approximately 334 m3/s, whereas the maximum discharge of 13,705 m3/s (slightly greater than 100-year return period flow) was observed during the Great Flood of 1936. The Anacostia River joins the Potomac River in Washington DC, which is much shallower and narrower compared to the Potomac River (McDowell, 2016). Additionally, the average daily flow at the Anacostia River is as low as 4 m3/s, while the maximum measured discharge did not exceed 350 m3/s (less than 100 year return period flow). A number of small streams also flow into the Potomac River from Chain Bridge to Occoquan, with drainage areas ranging from 104 to 1554 square kilometers (km). The mean tidal range at Washington, DC is approximately 0.9 m while the tidal phase lags 5 h behind Lewisetta and 11.5 h behind that at Hampton Roads at the mouth of the Chesapeake Bay (NOAA tides and currents).

2.2. Numerical Model Setup and Calibration

2.2.1. Advanced Circulation (ADCIRC) Hydrodynamic Model

The Advanced Circulation (ADCIRC) model (Luettich et al., 1992) is a finite-element hydrodynamic model based on the generalized wave continuity equation (GWCE). These equations are solved on an unstructured computational grid in space and time to simulate the behavior of open water bodies like ocean, lake and rivers, forced by astronomical tides, coastal storms, and incoming river flows. It has been used extensively for modeling historical storm surges and forecasting flooding (Blain et al., 2010; Dresback, Fleming, Blanton, Kaiser, Gourley, Tromble, Luettich, Kolar, Hong, Van Cooten, et al., 2013; Funakoshi, Feyen, Aikman, Tolman,

Van Der Westhuysen, et al., 2012; Juan L. Garzon et al., 2018; Hanson et al., 2013; Shen et al., 2006). The two-dimensional, depth-integrated version of ADCIRC (ADCIRC-2DDI) is used in the barotropic mode to simulate the combined influence of astronomical tides, river inflows and storm surges on total water levels. ADCIRC is a FORTRAN-based open source numerical model and well documented, both in the published scientific literature (Luettich et al., 1992) and on the ADCIRC web site (http://adcirc.org/).

2.2.2. Numerical Mesh and Model Setup

The dedicated ADCIRC model grid developed for this study was constructed with upstream boundaries at Little Falls (USGS Station 01646500; latitude 38°56′59.2″ longitude 77°07′39.5″), MD, and the junction of Northeast and Northwest branch of Anacostia River, and a downstream boundary at Lewisetta, VA, as shown in **Figure 1**. The unstructured computational grid was developed using the automated mesh generator, OceanMesh2D (Roberts et al., 2019) for the ADCIRC model. This program allows the construction of varying resolution, project specific, numerical meshes in the area of interest. The high-resolution coastline from Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) (Wessel & Smith, 1996) was manually updated in the Potomac River to accurately represent the coastline in the numerical grid. The open ocean boundary was kept at Lewisetta as it allows the inclusion of the coastal boundary from this NOAA station or other forecast systems. Ocean boundary conditions for all the numerical experiments were provided by the NOAA predicted tides at the Lewisetta station to minimize the tidal prediction error resulting from low resolution Global Tidal Models. The topography and bathymetry datasets in our localized model were extracted from the USGS topographic Digital Elevation Model (DEM) and a set of other bathymetry sources (NOAA nautical charts (Austin, 2005), NOAA National Centers for Environmental Information (NCEI) (Caldwell et al., 2015) and The Coastal National Elevation Database (CoNED) (Thatcher et al., 2016)) and a mosaic (Merged DEM) (detailed description on supplementary materials, section 1.2) with a vertical reference adjusted to North American Vertical Datum (NAVD88). In total five numerical meshes were developed using various nearshore and channel resolutions. Furthermore, four set of bathymetries (Modified NOAA nautical charts, NCEI, CONED and Merged DEM) were tested on ADCIRC numerical grids for model conveyance. The ADCIRC model was configured to run in explicit form of the barotropic mode. Wetting and drying, non-linear bottom friction, advection, finite amplitude terms, convective acceleration and the time derivative of convective acceleration were all included in the simulations. A detailed description of each recording station used in this study is provided in **Table A1** of appendix, while **Figure 1** shows the numerical modeling domain, several input boundary points and recording stations.

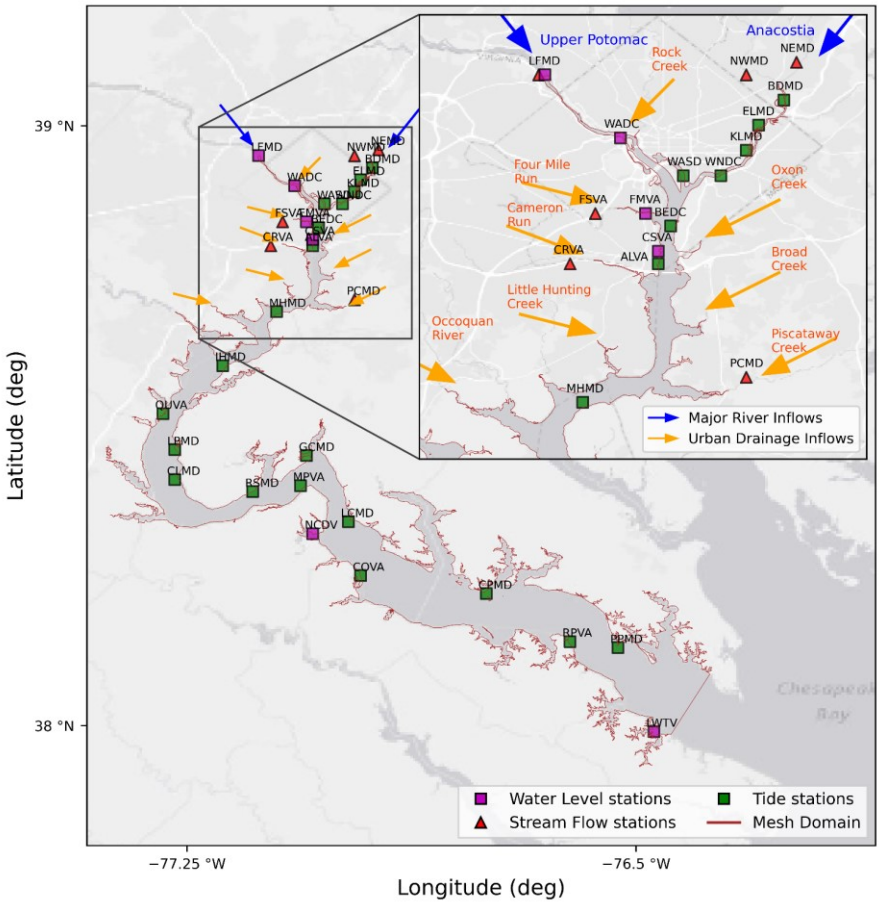


Figure 1. Location of study area

2.2.3. Model Calibration Parameters

A set of tests were performed with the ADCIRC model to determine the best range of calibration parameters (mesh resolution, bathymetry datasets, bottom friction, and eddy horizontal viscosity) for astronomical tidal modeling in the study area. Since several studies (Bacopoulos et al., 2017; Bakhtyar et al., 2020; J. Garzon & Ferreira, 2016; Kerr et al., 2013; Mied et al., 2006) have shown that these calibration parameters directly improve simulated tides and model stability, we selected our test range for each parameter (**Table 1**) based on recommended values provided by such literature. In these tests, we performed tides-only simulations, where model results from a baseline simulation were compared against simulated tests to assess changes in the total tidal amplitude at individual stations. Results suggested that: 1) high-resolution mesh with defined channels (*high_Ch*) exhibited the smallest Mean Absolute Error (MAE) in modeled tides for WASD; 2) Horizontal Eddy Viscosity (ESLM) value of 0.5 yielded the smallest MAE while keeping the model simulation stable; 3) Modified NOAA Nautical DEM provided bathymetry shows smallest MAE, however, it showed major

discrepancies (much shallower) compared to other regional bathymetry sources (NCEI and CoNED); and 4) *manning's n* roughness value of 0.01 in open water presented smallest MAE in modeled tides. The details on the methods and results of model calibration can be found in the supplementary materials (section 1 and 3.1).

Table 1. Model calibration parameters used for astronomical tides modeling in ADCIRC

Parameters	Test1	Test2	Test3	Test4
Bathymetry Datasets	Modified NOAA nautical charts	NCEI	CONED	Merged DEM
Numerical Model Resolution	Low resolution with no defined channels	Low resolution with defined channels	High resolution with no defined channels	High resolution with defined channels
Bottom Friction (Manning's)	0.02	0.018	0.015	0.01
Horizontal Eddy Viscosity (m2/s)	0.5	1	5	10

2.3. Validation Cases

A number of historical case studies (9 in total) are used to validate the integrated numerical modeling framework using the calibrated parameters referenced in section 2.2.3. These case studies consist of three historical events for each flooding source type: River, Coastal, and Compound. A flood event is classified as a "River" driven flood if the increase in the water at WASD is majorly influenced by the upcoming discharge from the river (LFMD) leading to water levels above the NWS defined "Action" stage at WASD. A "Coastal" driven flood event is defined as the result from a strong storm surge signal from LWTV that travels upstream and a relatively low discharge from upstream (LFMD). A "Compound" flood is defined when a combination of both high river discharge and high coastal water levels occur at the same time. Three major input forcings are used based on the availability of the observed data: downstream water levels, upstream major river discharges, and WASD wind spread spatially over the modeling domain. The downstream boundary is retrieved from the observed water levels at LWTV, but in cases where observed water levels are not available for LWTV, data from Sewells Point (SWPV) is used and adjusted for the amplitude and timing of LWTV (-5 h). The upstream discharge boundary is provided by the daily-observed flow at Little Falls, which is interpolated to hourly flow using a spline interpolation. The wind data was not available for WASD station prior to 2008, and therefore, the observed wind data from LWTV was used for case studies after 1973 as an approximation. For model validation prior to 1973, no wind forcing was used. It must

be noted that due to data interpolation and using proxy data in the absence of observed data, some uncertainty can propagate in the model results. All these model validations were simulated for nearly 23 days, starting at least 10 days prior to observed maximum water levels at the WASD station to allow the model to warm up and increase stability. A detailed description for data availability of each case study is given in **Table 2** and the time series of observed data is shown in **Figure A2** of appendix. Additionally, the detailed description of each historical event used in the model validation is provided in the supplementary materials (section 2).

Table 2. Case studies for model validation

Sr. No	Events	Type	Year	Dates	DC max (m)	Observed Water			Observed Flow	Observed Winds
						WASD	LWTV	SWPV	LF	LWTV or WASD
1.	*Great flood of 1936*	River	1936	03/14-03/24	2.79	Y	N	Y	Y	N
2.	*Blizzard of 1996*	River	1996	01/14-01/25	2.04	Y	Y	Y	Y	Y
3.	*Hurricane Agnes*	River	1972	06/10-06/02	2.22	Y	N	Y	Y	N
4.	*Flood of 1937*	Compound	1937	04/22-05/02	2.18	Y	N	Y	Y	N
5.	*Hurricane Fran*	Compound	1996	09/04-09/15	2.04	Y	Y	Y	Y	Y
6.	*Hurricane Isabel*	Compound	2003	09/15-09/27	2.7	Y	Y	Y	Y	Y
7.	*TS Ernesto*	Coastal	2006	08/28-09/05	1.61	Y	Y	Y	Y	Y
8.	*Hurricane Sandy*	Coastal	2012	10/20-11/11	1.44	Y	Y	Y	Y	Y
9.	*Hurricane Florence*	Coastal	2018	09/05-09/20	1.48	Y	Y	Y	Y	Y

2.4. Influence of Input Boundary Conditions and Atmospheric Forcing

In order to investigate the effects of different flood drivers on the total water level in the Potomac River, a comprehensive sensitivity analysis was performed on scenarios representing a range of: 1) downstream boundary conditions (i.e., storm surges); 2) upstream boundary conditions (i.e., major river discharges); 3) additional lateral boundary conditions (local urban runoff); 4) combined upstream and lateral boundary conditions; and 5) surface wind forcing at the local scale. For each scenario, we ran the simulation for 15 days including 3 days for spin-up time and compared with a baseline tides only simulation. Note that the change in water levels as a result of certain flood driver at the recording stations is presented as “above normal daily tides”, which is referred to a full tidal cycle including one high and one low tide.

When analyzing the influence of downstream boundary conditions on the change in total water levels at WASD, a set of storm surge simulations were performed, where downstream peak storm surge signal based on various return periods at LWTV defined by NOAA was applied. The peak surges were estimated at 0.956 m, 1.106 m, 1.236 m, and 1.346 m above NAVD88 datum at LWTV for 10, 25, 50, and 100 year, respectively. To minimize the influence from other flood drivers, no stream flow or local wind forcing was included. Whereas, to quantify the change in water levels resulting from upstream major river flows, we performed a set of simulations with flows ranging from 25 to 100 year return period. These flows are based on USGS StreamStats analysis (Ries III et al., 2017) and provided in **Table 3**. Apart from upstream river discharges, we also introduced urban runoff in the ADCIRC model at various streams along the Potomac River (**Figure 1**) and took a similar approach based on design stream flows (25 to 500 year return periods) for quantifying change in water levels. These stream flows differ from the major river flows, as they flow laterally and represent the water added to the Potomac River in case of heavy rainfall in the surrounding watersheds. Additionally, we performed a set of simulations with combined flows from major rivers (upstream) and urban runoff to quantify the change in water levels when both are contributing to Potomac River flow. Lastly, for investigating the effect of local winds on the water level variations at WASD, we performed a set of tests with local wind forcing. We reproduce eight wind directions (N, NE, E, SE, S, SW, W, and NW) with magnitudes ranging from 5-35 m/s for 12 hours, based on the observed data at WASD (2008-2020). The predominant winds observed at WASD are from NW and S. Additionally, historical NOAA records show winds from SE greater than 40 m/s and winds from NW nearly 35 m/s at Dulles Airport station (IAD) have been recorded. **Figure A2** in appendix shows a wind rose of observed wind speeds and direction at WASD.

Table 3. Flow characteristic for major rivers and urban runoffs

	Full name	Code	Drainage Area (km^2)	Max measured flow (m^3/s)	Min measured flow (m^3/s)	Average daily flow (m^3/s)	25yr return period	50yr return period	100yr return period
Major Rivers	Little Falls at Potomac River	LFMD	29940	13705	17	334	8680	10752	12908
	Bladensburg at Anacostia River	BDMD	239	349	0	4	233	317	580
Urban Runoffs	Rock Creek	RoCrk	521	77	0	2	330	417	518
	Oxon Run	OxCrk	98	-	-	-	121	160	209
	Four Mile Run	FMRun	122	36	0	0	308	390	483
	Broad Creek	BrCrk	173	35	0	1	166	218	283
	Piscataway Creek	PiCrk	420	127	0	1	231	300	384
	Little hunting Creek	LiHuntCrk	65	-	-	-	23	31	39
	Occoquan Creek	OxCrk	1538	784	0	14	623	759	921
	Cameron Run	CMRun	228	116	0	1	349	426	512

The forcing used in these scenarios are hypothetical and reflect the simplified versions of the real cases from normal daily weather to extreme weather conditions. These scenarios will allow us to better characterize contribution of each flood driver on total water levels and highlight the importance for inclusion or exclusion in a real-time integrated modeling framework.

2.5. Investigation of Operational Boundary Conditions for Real-Time Total Water Level Forecasting

Here, we discuss the various operational guidance systems (hydrometeorological and hydrodynamic) currently available that provide forecasted boundary forcing (downstream, major river flows, urban runoff and local winds) for total water level forecasting in the region. First, the original forecasts from the guidance systems were compared against observations to assess individual forecast bias. A set of events (*River*, *Coastal* and *Compound*) from the year 2020 were used as a test-bed to evaluate the real-time total water level forecasts using these predicted outputs. Since these events are reforecasted, we used the average of all the real-time forecasts issued daily (six hourly cycles). The time series of observed data for these reforecast events are shown in **Figure A3**, while **Table 4** provides the set of available guidance systems for various boundary types.

2.5.1. Downstream Boundary Conditions

The downstream boundary condition for forecasting the water levels in the Potomac River can be provided from a number of sources, including but not limited to global tidal prediction models (TPXO (Egbert & Erofeeva, 2002)), NOAA predicted astronomical tides, official water level forecasts from NWS or water level guidance provided by continental-scale storm surge guidance systems (Extra Tropical Storm Surge (ETSS) (Kim et al., 1996), Extratropical Surge and Tide Operational Forecast System (ESTOFS) (Funakoshi, Feyen, Aikman, Tolman, van der Westhuysen, et al., 2012), Chesapeake Bay Operational Forecast System (CBOFS) (Gross et al., 2000), integrated FLOOD Forecast System (iFLOOD) (Khalid & Ferreira, 2020), bias corrected iFLOOD water level (iFLOODv2) (Khalid & Ferreira, 2020) and Ensemble of all guidance systems) at LWTV station. The NOAA-predicted astronomical tides at LWTV provide an accurate estimate of astronomical tidal variation at LWTV, while the NWS or other guidance systems provide water levels that include a combination of storm surge and astronomical tides. We compared the forecasted water levels by various guidance systems and NWS against the observed water levels at LWTV (Jan 2020 to Aug 2020) to evaluate the best performing downstream boundary conditions for simulating total water levels in WASD during the reforecast *Coastal* event of 2020.

2.5.2. Upstream Boundary Conditions of Major River Flows

Currently, the NWS and the National Water Model (NWM) are the only sources that can provide the upstream river discharge boundaries (**Figure A4**) in real time for the Potomac River. The NWS stream flow forecasts are produced by the Middle Atlantic River Forecast Center (RFC), but are only provided at major discharge locations such as Little Falls (LFMD). NWS forecasts are only available for up to 72 hours in the future for these locations. No flow predictions are available from the RFC for the Anacostia River confluence (BDMD). The NWM, which is a continental scale hydrologic prediction system (Cosgrove et al., 2018), also provides flow forecasts for a number of timescales (Short, Medium and Long range) over a national coverage of nearly 2 million reaches. The NWM stream flow from the Medium range forecasts are available at LFMD and BDMD, and are used as upstream boundaries for major river flows. Since these flows influence the *River* and *Compound* events primarily, we used NWS and NWM forecasts as upstream boundaries to determine the best performing upstream boundary condition model. Note that we only analyzed the upstream boundary conditions for *River* events.

2.5.3. Upstream boundary conditions of Urban Runoff

As noted, the RFC does not provide flow forecast for all the streams in the region, and therefore the NWM is the only forecast source for almost all the urban streams (shown in **Figure A4**) in the region. The NWM stream flow forecasts were compared against the available USGS stream flow gages to understand the forecast bias. For the test period in 2020, no major urban flood event significantly affected total water levels at the WASD station; therefore, we could not validate the additional value of including the urban runoff boundary from the streams during the period. However, we used the NWM forecasted stream flows during the *Compound* reforecast event (April 2020) to analyze change in total water levels.

2.5.4. Atmospheric Forcing for Local Winds

Although a number of weather models are available for our study area, we limited our scope of evaluation for two forecasted model outputs only, North American Mesoscale (NAM) 12 km resolution and National Blend of Models (NBM) 2.5 km. In contrast to NAM, NBM is not a numerical model, instead it is a blended dataset based on NWS and non-NWS numerical weather prediction models. At least two studies (J. L. Garzon et al., 2018; Khalid & Ferreira, 2020) have shown that NAM is a highly skillful model for wind forecasts in the Chesapeake Bay, however, a NBM evaluation is not established for this region. Here we tested the accuracy in simulating the local wind impacts to total water levels based on these two wind models during a *Compound* reforecast event (April 2020). The forecasted wind speeds and direction at 10 m height were specified on the numerical modeling grid to simulate local wind impacts.

Table 4. Available boundary conditions for real-time total water level forecasting in Potomac River

Boundary Types	Forecast/Guidance System
Downstream Water Levels	NWS/ETSS/CBOFS/ESTOFS/iFLOODv2/iFLOOD
Upstream Major River Flows	NWS/NWM
Upstream Urban Runoff	NWM
Atmospheric Forcing	NAM(12km) /NBM(2.5km)

2.6. Ensemble Based Forecasting

Based on the available set of guidance systems mentioned in Table 4, we devised a full set of ensembles (> 30) containing various combinations of model input boundary forcing (downstream, upstream, urban and local winds) to simulate ensemble-based total water level forecasts. However, we then lowered the total number of ensembles to a maximum of 10 members based on the best performing guidance system for model input forcing. This set consisted of 5 downstream (CBOFS, ETSS, NWS, iFLOODv2, Ensemble), 1 upstream (NWS), 2 local wind forcing (NAM, NBM) and 1 urban runoff (NWM) boundary. In order to gain computational efficiency, we used a low-resolution mesh with defined dredged channels (*low_ch,* detailed in supplementary materials, section 1.1), therefore, the quality of the modeled total water levels was limited by the finite resolution of the model. However, it provided computational speed, while maintaining an integrated framework. These ensemble simulations were initialized from the existing deterministic forecast system for Chesapeake Bay. Our analysis of the ensemble forecasting in the Potomac River is also based on the *Compound* reforecast event of 2020. The different types of ensemble forecasts are summarized in Table 5.

Table 5. Summary of ensemble configurations

Abbreviation	Upstream	Downstream	Urban Drainage	Atmospheric Forcing
Ens1	NWS	ETSS	NWM	NAM (12km)
Ens2	NWS	CBOFS	NWM	NAM (12km)
Ens3	NWS	NWS	NWM	NAM (12km)
Ens4	NWS	iFLOODv2	NWM	NAM (12km)
Ens5	NWS	Ensemble	NWM	NAM (12km)
Ens6	NWS	ETSS	NWM	NBM (2.5km)
Ens7	NWS	CBOFS	NWM	NBM (2.5km)
Ens8	NWS	NWS	NWM	NBM (2.5km)

Ens9	NWS	iFLOODv2	NWM	NBM (2.5km)
Ens10	NWS	Ensemble	NWM	NBM (2.5km)

3. Results and Discussion

3.1. Model Validation

The ADCIRC model setup used in this study, including a combination of three boundary forcings (downstream water levels, upstream discharge and local winds) is validated on the high-resolution numerical mesh with overland areas (*high_OL*) using nine historical events (supplementary materials, section 3.1.2).

3.1.1. Historic Riverine Events

The historical validation of the extreme *River* events was predominantly dependent on the accuracy of the upstream flow observed boundary. The time series of simulated total water levels (TWL) against the observations at the WASD station are shown for all the three riverine events in **Figure 2**. The model simulated the increase in TWL as a result of large river discharge from LFMD, however, the modeled peak was 0.5 m larger than the observed peak. The simulated TWL during the 1936 Flood show an increase of 3 m above normal daily tide as a result of an approximate 100-year return period flow recorded at LFMD station. Based on upstream river flow analysis as shown in **Figure 6**, 100-year return period flow can increase the water level at WASD by almost 3.6 m; however, that increase is estimated when both upstream boundaries are flooding simultaneously. Historical simulation of Hurricane Agnes in 1972 and River Flood of 1996 also showed the same over prediction of 0.5 m at the WASD station. Note that these simulations do not include the local wind effects, as no observed wind data at WASD before 2008, and according to **Figure 8**, local wind can impact the water levels by ± 0.5 m above normal daily tide when wind speeds are nearly 15 m/s. Additionally, the lower panels of **Figure 2** shows the inland extent of flooding during these three river events.

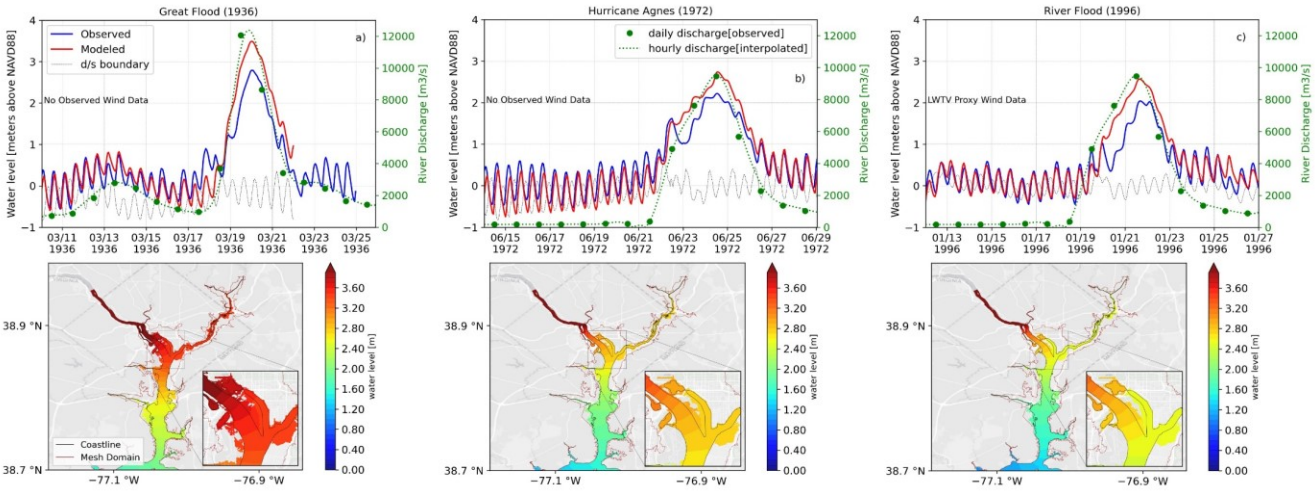


Figure 2. Time Series of total water level (TWL) and spatial map of maximum TWL during the peak of storm at WASD a) Great Flood (1936) b) Hurricane Agnes (1972) c) Blizzard (1996)

3.1.2. Historic Coastal Events

The historical validation of the *Coastal* events, on the other hand, is majorly dependent on the accuracy of the downstream water level boundary and atmospheric forcing. The time series of simulated TWL against the observations at the WASD station are shown for all the three coastal events in **Figure 3**. The model simulated the increase in TWL at the WASD station as a result of strong storm surge signal propagating upstream from LWTV. The time series of the modeled peak at WASD was over estimated by 0.5 m for Tropical Storm Ernesto (2006), while in case of Hurricane Sandy (2012) and Florence (2018), the model captured the peak more accurately. During TS Ernesto, no observed wind data was available at WASD, and instead proxy data from LWTV was used. Since local winds have potential to impact water levels by ± 0.5 m above normal daily tide when wind speeds are nearly 15 m/s (**Figure 8**), inaccurate wind data used during simulation could have resulted in noted overestimations. The section 3.2.4 further elaborates the local changes in water levels at a recording station due to local wind effects. Similar to **Figure 2**, **Figure 3** also shows the extent of flooding during these three *Coastal* events, which is not propagated overland as far as *River* based events.

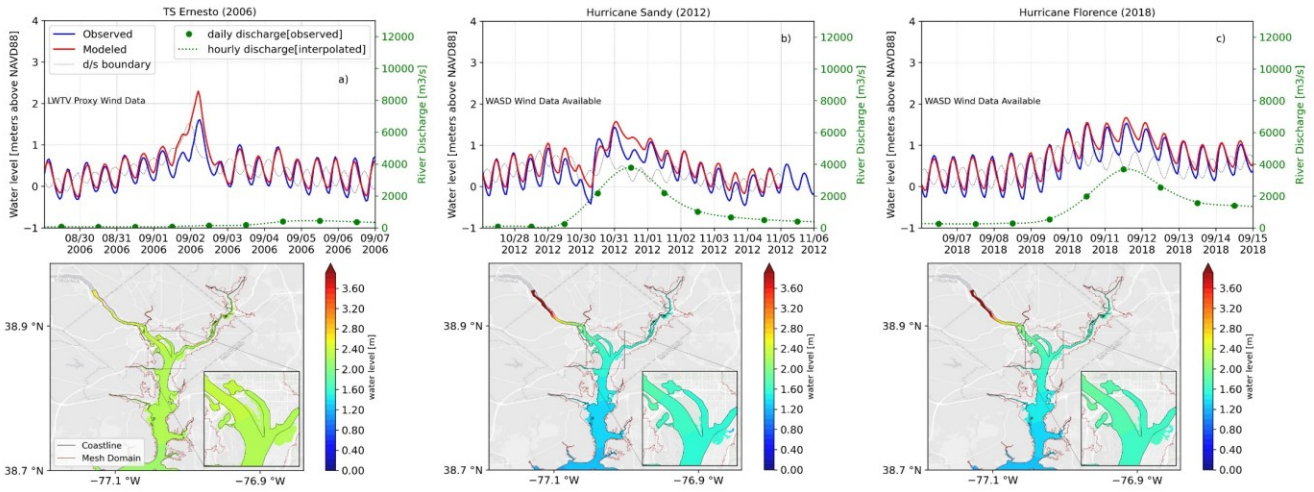


Figure 3. Time Series of TWL and spatial map of maximum TWL during the peak of storm at WASD a) TS Ernesto (2006) b) Hurricane Sandy (2012) c) Hurricane Florence (2018)

3.1.3. Historic Compound Events

The time series of TWL results from simulations of three major *Compound* events that happened in the National Capital Region are shown in **Figure 4**. The simulated results show that the model was able to reproduce the first peak of the compound events (exception of River Flood 1937) more accurately when compared to the second peak that was influenced by a large upstream river discharge at LFMD. Among all the three case studies, the first peak at the WASD was correctly modeled during Hurricane Isabel (2003) and Hurricane Fran (1996). Interestingly, the second peak followed by the large upstream flow showed an overestimation of 0.5 m during all the three events, similar to *River* and *Coastal* events validations. Similarly, since no observed wind data was available for WASD during all these three *Compound* events, one can argue that accurate local atmospheric forcing could help with accurate estimations of TWL.

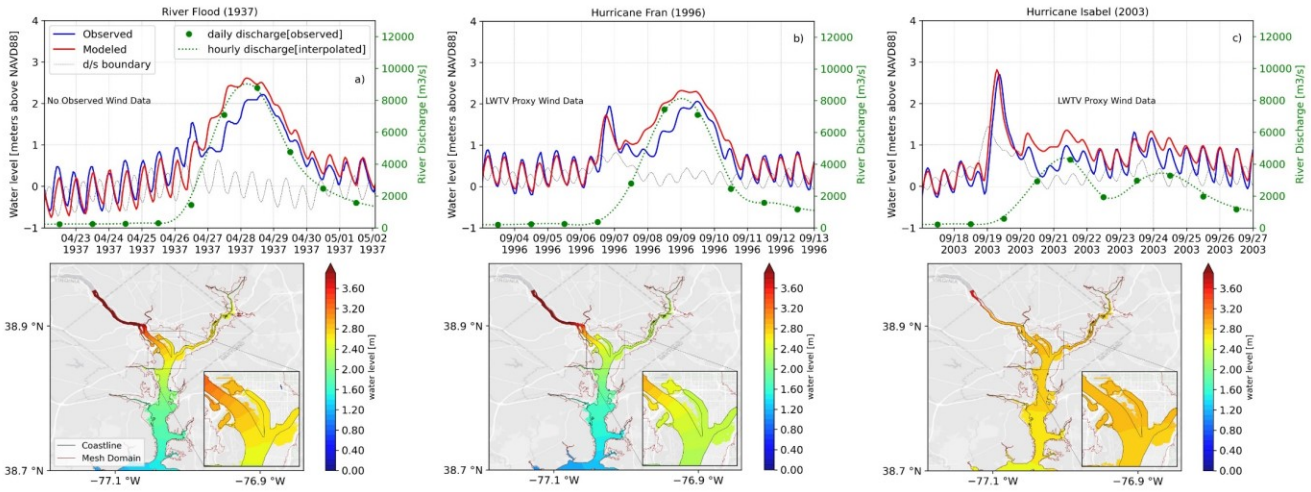


Figure 4. Time Series of TWL and spatial map of maximum TWL during the peak of storm at WASD a) River Flood (1937) b) Hurricane Fran (1996) c) Hurricane Isabel (2003)

Furthermore, a consistent over prediction of 0.5 m was noted during high upstream river discharges. Our validation results for Great Flood of 1936 and Hurricane Isabel (2003) at WASD showed slight discrepancy from the results published in an earlier study (Wang et al., 2015), where the authors used a Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM) with upstream river boundary at LFMD and downstream boundary at Colonial Beach. The simulated peak of our modeling setup showed 0.5 m overestimation during Great Flood of 1936, but only a 0.1 m over prediction during Hurricane Isabel peak. Simulated water level peak error during Isabel also compared favorably with ADCIRC model results at WASD published earlier (Mashriqui et al., 2014). It is worthwhile to note that the model tends to overestimate TWL by almost 0.5 m when stream flows higher than 3000 m^3/s are introduced at the LFMD boundary. Additionally, for some historical events, some of the observed data (LWTV water levels and WASD winds) were not available, which may have led to uncertainty in the simulated water levels. Lastly, the Potomac River channel has undergone significant changes since the 1930s that may have also influenced the simulated results at WASD. This becomes clear upon examination of the model validation results that the integrated modeling framework can simulate historic extreme water levels at WASD with slight over predictions (~0.5 m). Since the over prediction was consistent in all the case studies, proposed a systematic bias correction at WASD when Potomac River flows are above 3000 m^3/s to increase the accuracy of model results.

3.2. Influence of Input Boundary Conditions and Atmospheric Forcing

3.2.1. Downstream Boundary Conditions

Figure 5 shows the increase in water levels above normal daily tides at a number of recording stations (LWTV to WASD), when 10 to 100 year return period storm surge enters the Potomac River. The growth in magnitude becomes consistent at Alexandria (ALVA) and continues upstream. Results show that maximum increase in water levels above normal daily tides at WASD during a 100-year return period does not exceed 1.5 m. TS Ernesto in 2006 resulted in a similar magnitude of surge propagating upstream from LWTV (1.468 m above NAVD88) and elevated the water levels at WASD by almost 1.11 m above NOAA predicted astronomical tides. The discrepancy between the estimated increase and observed increase due to downstream boundary conditions could have resulted from other flood drivers during TS Ernesto. Results also show that there is almost a 26% increase in storm surge magnitude (right panel of **Figure 5**) as the storm surge signal travels from LWTV to WASD. Historical data analysis (Sumi & Ferreira, 2019) demonstrated that LWTV contributes 80% of the water level in WASD, whereas we found almost 76% in this analysis. Since such a large contribution of water level exists due to downstream boundary condition, it is very important for accurately estimating total water level.

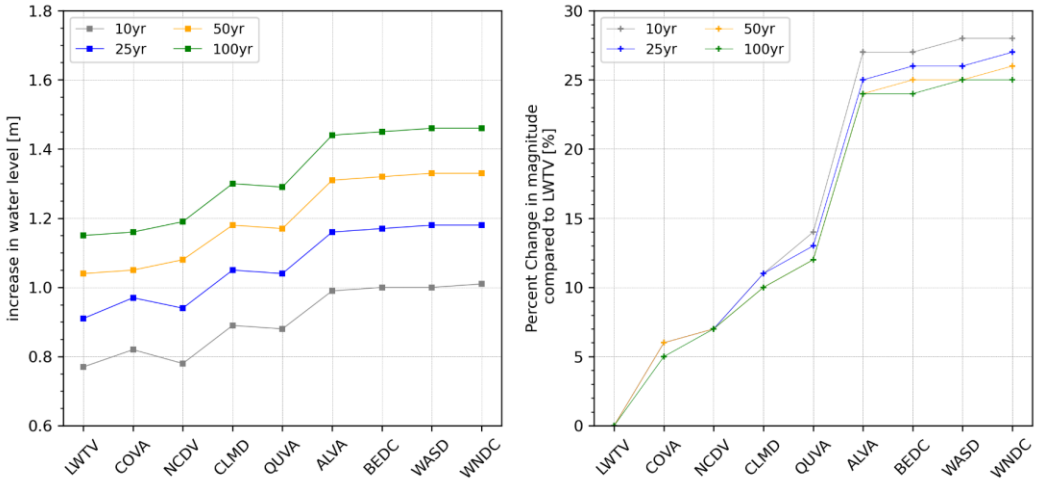


Figure 5. Change in water levels magnitude (above normal daily tides) in the Potomac River due to downstream boundary conditions.

3.2.2. Upstream Boundary Conditions of Major River Flows

Figure 6 shows the increase in water levels when a range of return period (25 to 100 year) discharges are introduced on a normal tidal simulation in January 2020. The increase resulting from individual upstream rivers discharges is also show as dashed and dotted lines in **Figure 6.** Since the drainage area for the Anacostia River is significantly smaller, the increase in water level based on Anacostia discharges were almost negligible for 25 to 100 year return period discharges. Whereas, Potomac River discharges raised water levels nearly 4 m above normal river elevation during a 25-year return period at WADC, which surpassed 5.5 m mark as the return flow increased

to 100-year return period. It is interesting to note that the increase in water levels at the WASD station as a result of 25 and 100-year return period is almost 2.5 to 4 m above normal daily tides. Our model validation for the 1936 Flood showed slightly smaller increase in water level (~3 m above normal tides) at WASD when a 100-year observed return period flow at LFMD flooded the region. Note that during validation for the 1936 Flood, only LFMD exceeded the 100-year return period flows, while BDMD had no record of observed flow until 1938. **Figure 6** also shows the increase in water level as a function of distance when various return period flows were introduced. The plot suggests that upstream discharge can have an influence as far downstream as Colonial Beach (COVA), which aligns well with the observation in Mashriqui's study (Mashriqui et al., 2014).. This analysis showed the importance of including the upstream major discharge boundaries, since in the absence of upstream boundaries, the model will not be able to capture TWL, especially when riverine flows are above 10 times average daily flow (equivalent to 2 year return period).

3.2.3. Upstream Boundary Conditions from Urban Runoff

The second panel of **Figure 6** shows the increase in water levels when a range of return period stream flows from tributaries (yellow arrows in **Figure 1**) are introduced in the Potomac River. It can be seen that the increase in water levels is not as prominent as the results caused by major river flows (25-year return period flow by major river is added as reference in second panel of **Figure 6**); however, stations close to the stream boundaries experienced an increase of water level between 0.22 m to 0.59 m above normal daily tides during a 100-year return period flow. Rock Creek, which is upstream of the WASD station, in the absence of major river discharge, only increased the water levels by 0.22 m above normal daily tides during a 100-year return period flow.

3.2.4. Upstream Boundary Conditions from Combined River flows and Urban Runoff

Additionally, we performed a set of simulations with combined major river flow and urban runoffs for 25-100 year return period flows. The lowest panel of the **Figure 6** shows the increase in water levels above normal daily tides as function of distance while arrows on the figure also indicate the location of various streams and Major River inputs. Interestingly, the increase in water level as a result of combined major river flow and urban runoff for 100 year return period raise by almost 0.4 m and becomes equal to a 200 year return period (not shown here) during a major river flow only. Similarly, the increase in water levels at WASD as a result of combined discharges for 100 year exceeds the 4.35 m height above normal daily tides. This analysis shows the importance of including the urban drainage in addition to major river flows when modeling historical events or developing an integrated total water level forecast system. Although urban runoff boundary alone may not influence the water levels significantly at WASD, when combined with large river discharge, it will significantly impact simulated TWL.

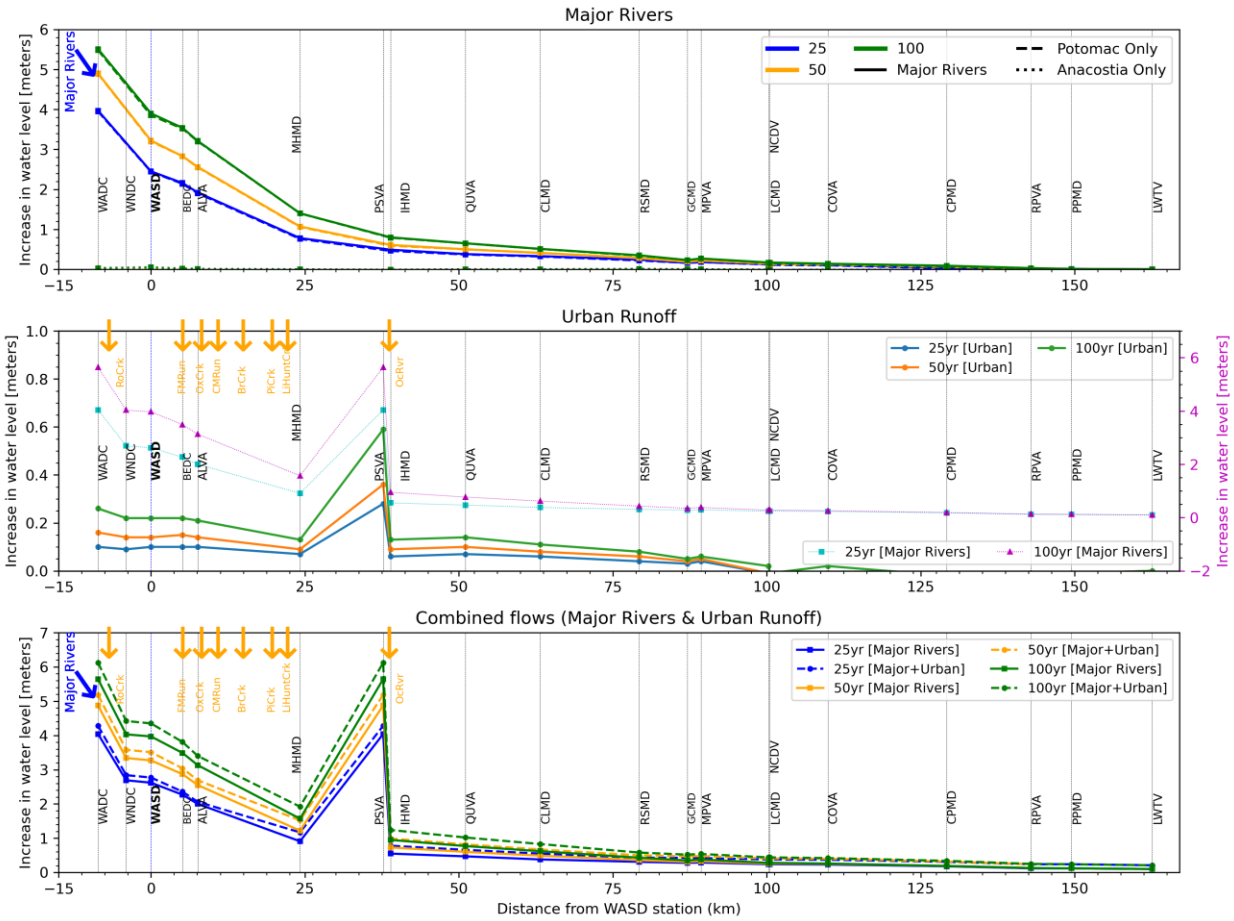


Figure 6. Change in water levels above normal daily tides in the Potomac River due to major river, urban runoff and combined discharges boundary conditions

3.2.5. Effect of Local Winds

The results of considering local effects of winds for eight primary directions at six stations surrounding WASD are shown in **Figure 7**. The positive values of change show an increase in water levels, while negative values show a decrease in water levels. Based on the location of the WASD recording station, we would expect an increase in water levels locally when winds are blowing from South and a decrease in water levels when wind blowing from NW (pushing water away from the station). **Figure 7** clearly shows that when winds are blowing from N, NE, W and NW direction they tend to decrease the water levels at the WASD station and as the magnitude of the winds increase, a large decrease is shown. An opposite trend is noted when winds are blowing from S, SW, E and SE. These local changes to total water levels as a result of local winds are shown in an earlier study using a Delft3D model, where winds >5.5 m/s from NW direction drained water out of Potomac River, therefore lowering the water levels at WASD (Mashriqui et al., 2014). Our wind forcing analysis showed the importance of including local wind forcing in excess of 5.5

m/s to accurately capture the changes to total water level locally. On the other hand, winds smaller than 5.5 m/s will not affect the water levels locally, therefore forecasts produced by ocean scale coastal guidance systems by ignoring local winds will perform similar to our dedicated Potomac forecast system. Analysis of observed winds at WASD showed that magnitudes between 5 m/s and 8 m/s are observed at least once in most months (**Figure A5** in appendix). Although events with local wind speeds greater than 10 m/s (**Figure A5** in appendix) are less common, the absence of local wind forcing in such events, will completely misrepresent the forecasted TWL.

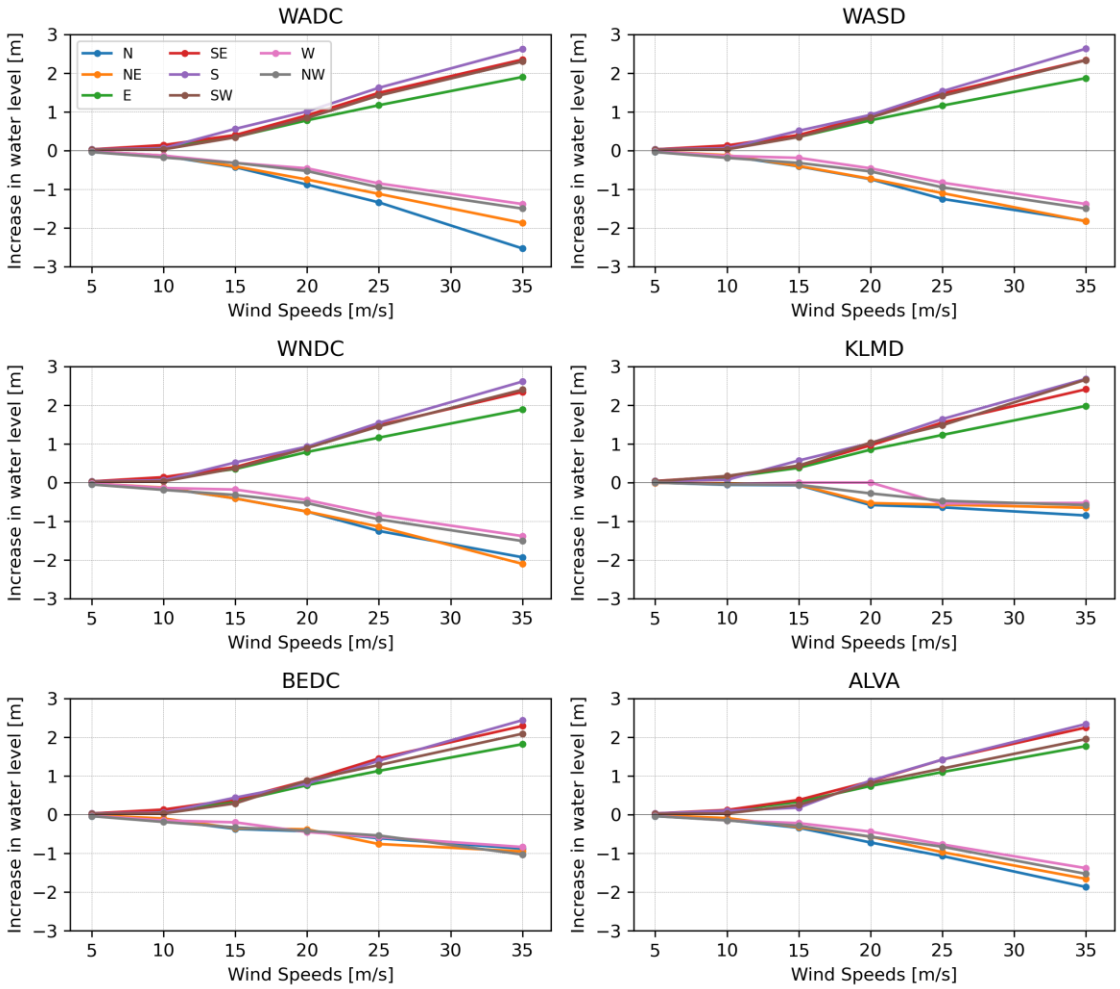


Figure 7. Change in water levels above normal tides at WASD and nearby stations due to local winds forcing in eight directions.

3.2.6. Summary of flood driver's influence

Figure 8 provides a summary for the relevance of individual flood drivers at WASD. Note that the systematic error (+0.5 m) at WASD noted during the validation of historic river events (section 3.1.1) is corrected when estimating the maximum water levels shown in **Figure 8**. The

urban runoff boundary does not increase the water levels above "Action" stage as a result of 25 to 100 year return period. Storm surges, on the other hand, can increase the maximum water levels at WASD above "Moderate" flooding level, i.e. as noted during TS Ernesto 2006. Local wind forcing greater than 20 m/s from South direction is also shown here to increase the maximum water levels from "Minor" to "Major" flooding level. Among all the individual flood drivers, major rivers discharges of 100-year return period are shown to cause the maximum water levels at WASD, i.e. as noted during Great Flood of 1936. Lastly, maximum water levels caused by combined river flows provides an example of combination of flood drivers that leads to significant influence on the water levels at a given recording station.

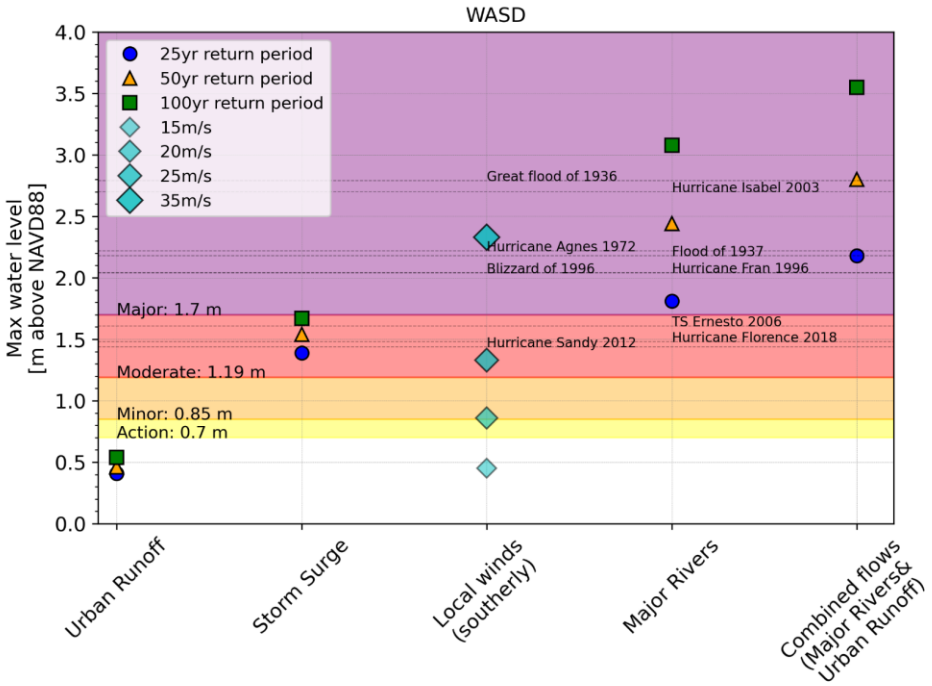


Figure 8. Maximum water levels at WASD resulting from various flood drivers. Horizontal lines represent maximum water levels recorded at WASD during historical events.

3.3. Investigation of Operational Boundary Conditions for Real-Time Total Water Level Forecasting

Sample forecast graphics of water levels at LWTV from all the existing coastal guidance systems is provided in the upper left panel of **Figure 9**, while the upper right panel shows the average forecast error (bias) at LWTV over 7 months period (Jan to Aug, 2020). Clearly, average biases for NWS, iFLOODv2, ETSS and CBOFS are the smallest compared to all the other guidance systems. Similarly, the forecasted stream flows of NWS and NWM for upstream major

rivers at LFMD is given in left lower panel of **Figure 9**, where forecast bias of NWM over 7 months period (Jan to Aug, 2020) is shown higher than NWS during observed peak stream flows. The forecasted bias of urban runoff at the major streams over 7 months period (Jan to Aug, 2020) in the Potomac River is also shown in Figure 9 (lower right panel), where, on average, the bias is less than 2.5 m3/s. Forecasted wind speeds from NAM and NBM were not included in this analysis due to lack of real-time outputs retrieved for NBM during this 8 months analysis period. This initial assessment of forecast bias for individual guidance systems helped identify consistently high performing guidance systems, i.e. downstream boundary (NWS, iFLOODv2, ETSS and CBOFS), upstream major river boundary (NWS), urban runoff (NWM).

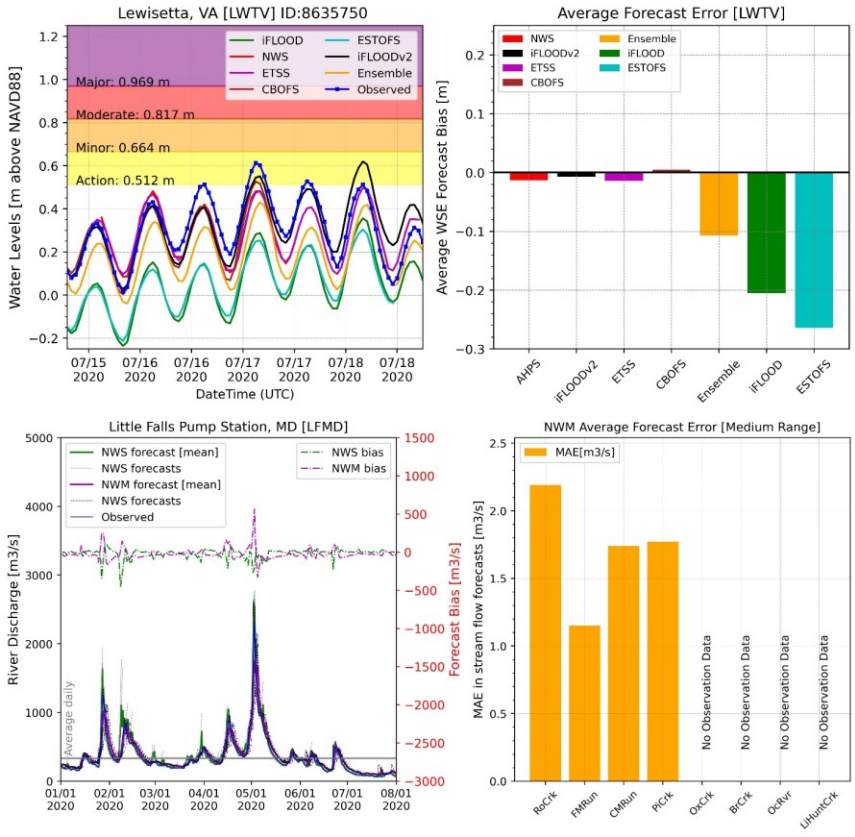


Figure 9. Sample forecast advisories and error plot (average bias and MAE) for various forecast guidance systems a period of 8 months (Jan 2020 to Aug 2020).

Secondly, we performed a set of validations based on reforecast events of 2020 to demonstrate that the above-mentioned guidance systems reduce the forecast error when modeling TWL predictions at WASD. For the reforecast *Coastal* event, the time series of simulated total water levels at WASD against observations (**Figure 10,** panel a) shows that the downstream boundary provided by the NWS predicted water levels more accurately when compared to other guidance systems (iFLOODv2, ETSS and CBOFS). The MAE resulting from the simulation using NWS as

boundary condition was smaller compared to other guidance systems (**Figure 10,** panel e). For the reforecast *River* event, simulated total water levels using NWS and NWM as upstream boundary conditions for major river flows showed almost the same pattern at WASD (underestimation) (**Figure 10,** panel b), however, MAE resulting from the simulation using NWS as upstream boundary condition was slightly lower than with the NWM (**Figure 10,** panel e). From the above *Coastal* and *River* reforecast analysis, upstream and downstream boundaries forecasted by the NWS are considered the best for the selected events, and further utilized to simulate a *Compound* event.

Panel c of **Figure 10**, shows the time series of simulated *Compound* event with and without urban drainage. MAE using the urban drainage was slightly smaller than "No Urban" flow boundary. For this specific event, the observed urban runoff was much smaller than the 25-year return period; therefore, we would not have expected any significant influence on the water levels. However, our analysis (section 3.1.3) showed that the addition of urban runoff can certainly help capture the increase in water levels around National Capital Region in the case of large urban events. Lastly, to find the best performing weather forecast model, we again simulated the *Compound* event (April 2020) without the local winds and in the presence of forecasted winds from NAM and NBM atmospheric models (**Figure 10,** panel d). Based on our hypothetical analysis (3.1.4) and the observed winds (8 m/s from SE) during this particular event, we would expect a small increase of 0.08 m to water levels at WASD. The simulated TWL using NAM weather forcing showed the expected increase (~0.1 m) based on local wind forcing, while NBM weather forcing based simulation did not show an increase. Further analysis of the wind magnitude and direction forecasted by NAM and NBM (not shown here) revealed the under prediction at WASD by NBM, while NAM accurately forecasted the high winds during *Compound* event. The MAE in the absence of local wind forcing was nearly 0.25 m, which decreased to 0.22 m using NAM wind forcing.

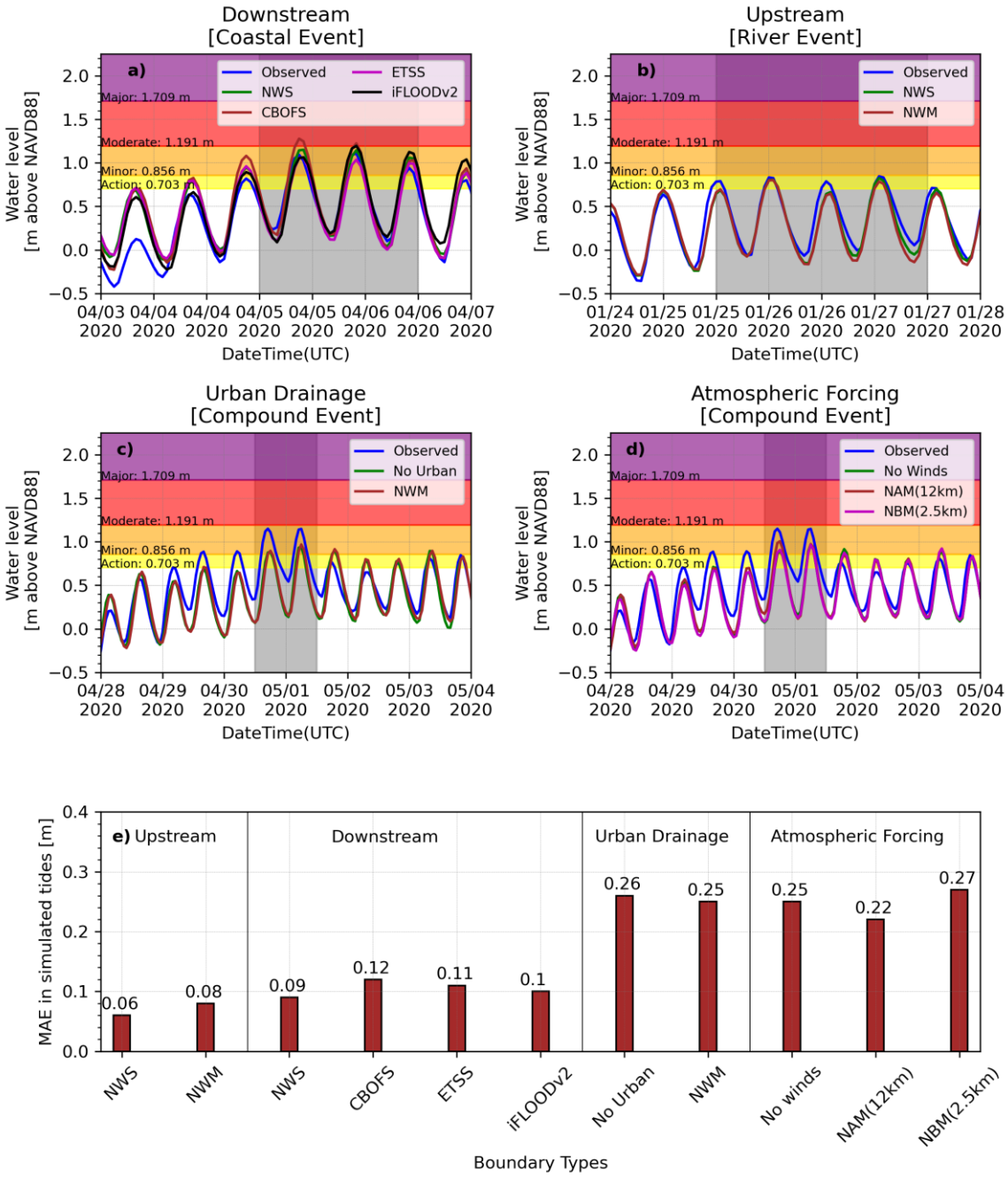


Figure 10. Reforecast analysis of *River*, *Coastal* and *Compound* events of 2020 using various boundary types at WASD station.

Although some guidance systems performed better than others for given reforecast events, a larger set of events will be required to confirm a best set of boundary guidance systems. Based on our given set of reforecast events, we proposed a recommended set of guidance systems (Table 6) for development of TWL forecast system in the Potomac River.

676

677 **Table 6.** Recommended set of boundary conditions for total water level forecasting in Potomac
678 River based on reforecast events of 2020

Boundary Types	Suggested Forecast/Guidance System
Downstream Water Levels	NWS
Upstream Major River Flows	NWS
Upstream Urban Runoff	NWM
Atmospheric Forcing	NAM

679

680

681

682 **3.4. Ensemble Based Forecasting**

683 The use of multi model ensembles has been shown in numerous applications (Hagedorn et
684 al., 2005; Kirtman et al., 2014; Krishnamurti et al., 2000; Weigel et al., 2008) to improve forecast
685 skill. Our results showed the advantage of using ensemble-based flood prediction in comparison
686 to single value flood forecasts. Ensemble forecasts of TWL predictions during *Compound*
687 reforecast event of 2020 using a set of 10 simulations is shown in **Figure 11**. Coastal downstream
688 boundaries provided by CBOFS forecasted the highest water levels at LWTV, followed by NWS
689 and then the Ensemble mean. Interestingly, significant variability in the predicted TWL, shown at
690 LWTV was reduced as it propagated upstream at WASD. **Figure 11** shows that on 30th April 1800
691 UTC, 6 out of 10 ensemble simulations predicted water levels at WASD exceeding the 0.856 m
692 flood threshold, resulting in 60% chance of "Minor" flooding. Several hours later, on 1st May 0600
693 UTC, the chance of flooding reduced to 40%, however, two ensemble simulations based on the
694 CBOFS boundary condition, forecasted the peak TWL at WASD with high accuracy (bias less
695 than 0.03 m).

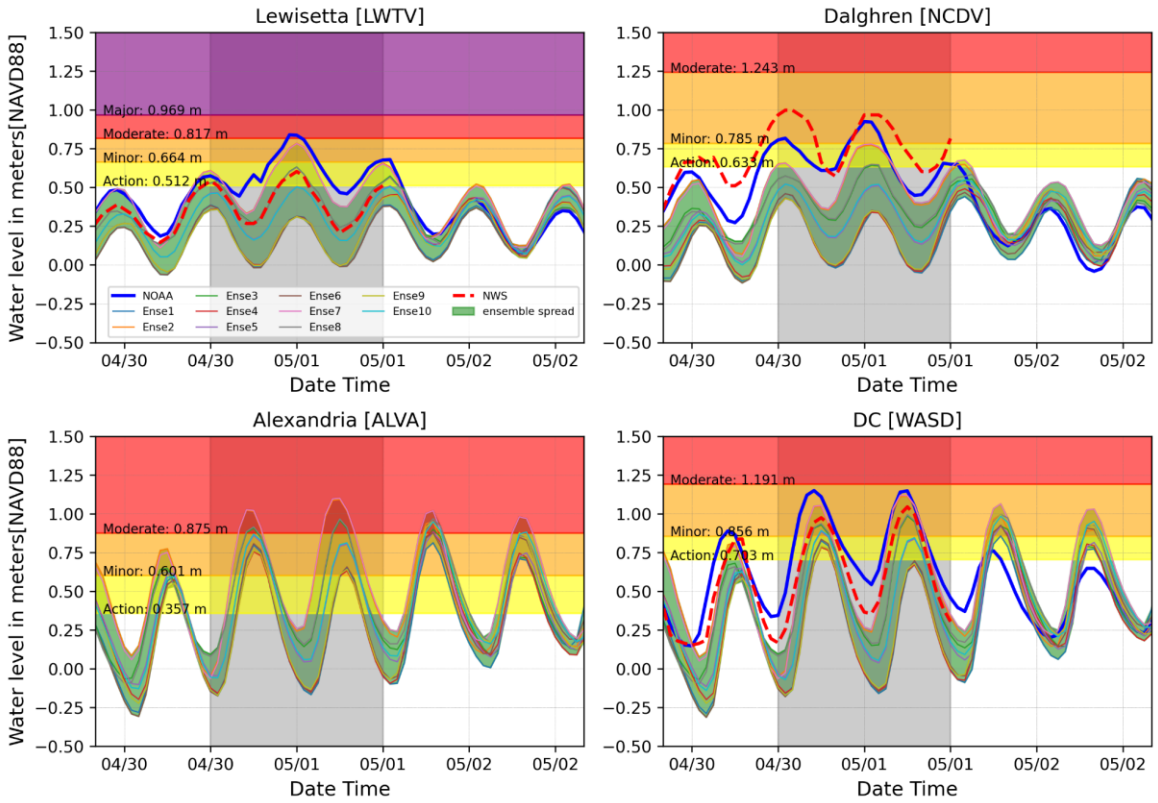


Figure 11. Ensemble-based forecasting of total water levels for the *Compound* reforecast event of 2020 in Potomac River.

The upper bound of the ensemble distribution predicted total water levels at nearly all the stations with high accuracy. **Figure 11** also show that the ensemble spread grows during high flood events leading to higher forecast uncertainty and shrinks back under normal daily conditions (i.e. after 2nd May). Ensemble predictions serves as a compromise to high-resolution numerical modeling, due to its ability to capture the forecast uncertainty. This analysis showed the ability of this integrated framework to utilize ensemble-based TWL forecasting, while providing representation of uncertainty originating from selected boundary guidance systems. Although, similar to last section, one test case may not be sufficient, this briefly highlights the value of ensemble forecasting over deterministic forecasts to develop high accuracy official flood forecasts in the Upper Tidal Potomac Region using model-generated guidance.

4. Conclusion

Real-time flood forecasting in upstream tidal areas is challenging due to the complex and dynamic interaction of several flood drivers. This study presents a detailed assessment of various flood drivers required for accurate total water level (TWL) forecasting in upstream tidal rivers.

The Tidal Potomac River is a representative example of a tidal river that has complex physical interaction of ocean tides, freshwater inflows, urban runoff, and local wind impacts in hydrodynamics. The existing operational coastal guidance systems frequently underestimate water levels predictions in such complex environments. This study area is of national importance, since the National Capital Region is located at the confluence of the Potomac River and the Anacostia River, and it is susceptible to an increasing threat from flooding. In this study, we utilized the ADCIRC-2DDI model to simulate these interactions and implemented a calibrated and validated model set up for the Potomac River, which was further assessed to quantify the contribution of each flood driver on the TWL at Washington, DC (WASD). Model validation results indicated that with riverine flows greater than 3000 m^3/s overestimate the water levels at WASD by almost 0.5 m, which was corrected before further evaluation.

Using a range of hypothetical boundary forcing, we have shown that the influence of downstream boundary, upstream river discharge, local urban runoff and wind forcing are important and must be considered while forecasting total water levels in the region. For instance, the downstream boundary at LWTV represents three-fourths of observed water levels at WASD, upstream major river flows as low as a 25-year return period flow can increase the water levels by almost 2.5 m above normal daily tides at WASD, and local urban runoff combined with major river flows can raise the flooding levels by almost 0.5 m (100-year return period increase equals 200-year return period). Similarly, the influence of local "impact winds" (> 10 m/s) is noticeable on water levels, as it can elevate water levels (± 0.15 to 2.25 m) above normal daily tides at WASD. Furthermore, results based on reforecast events of 2020 showed that upstream flow and downstream water level boundary forcing based on the NWS system can reasonably forecast water levels at WASD. Likewise, forecasted urban flows provided by NWM and wind forcing from NAM weather model improved the TWL estimates at WASD.

Although the contributions of each physical process in forecasting TWL are quantified and noted relevant, adding these boundaries in the large-scale coastal guidance systems increases the model complexity and operational computational time. Therefore, we demonstrate the value of a dedicated forecast system for complex tidal rivers of National importance, while including all these boundaries forcing (tides, storm surge, river discharge, urban runoff, and local winds) to accurately forecast the total water levels in the National Capital Region. Moreover, the additional benefit of using this dedicated system is the ability to run ensembles forecasts using a range of boundary forcing. Further evaluation of Potomac integrated system over a longer forecast period will provide a better assessment of its prediction capacity and the value added from the ensemble forecasts.

5. Acknowledgements

This work was made possible by Virginia Sea Grant Program [Grant # NA18OAR4170083] and the research resources from the Flood Hazards Research Lab (FHRL) at George Mason University.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NWS. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) STAMPEDE2 resources through allocation id TG-BCS130009, which is supported by National Science Foundation [grant number ACI- 1548562] (Towns et al., 2014). The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the model calibration results reported within this paper (http://www.tacc.utexas.edu).

6. References

Atkinson, L. P., Ezer, T., & Smith, E. (2012). Sea level rise and flooding risk in Virginia. *Sea Grant L. & Pol'y J.*, *5*, 3.

Austin, M. (2005). Creating a GIS from NOAA electronic navigational charts. In *Proceedings of OCEANS 2005 MTS/IEEE* (pp. 839–841).

Bacopoulos, P., Tang, Y., Wang, D., & Hagen, S. C. (2017). Integrated hydrologic-hydrodynamic modeling of estuarine-riverine flooding: 2008 Tropical Storm Fay. *Journal of Hydrologic Engineering*, *22*(8), 4017022.

Bakhtyar, R., Maitaria, K., Velissariou, P., Trimble, B., Mashriqui, H., Moghimi, S., et al. (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. https://doi.org/10.1029/2019JC015822

Blain, C. A., Linzell, R. S., Chu, P., & Massey, C. (2010). *Validation test report for the ADvanced CIRCulation Model (ADCIRC) v45. 11*. DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA518667

Bobanović, J., Thompson, K. R., Desjardins, S., & Ritchie, H. (2006). Forecasting storm surges along the east coast of Canada and the North-Eastern United States: The storm of 21 January 2000. *Atmosphere-Ocean*, *44*(2), 151–161. https://doi.org/10.3137/ao.440203

Caldwell, P. C., Merrifield, M. A., & Thompson, P. R. (2015). Sea level measured by tide gauges from global oceans--the Joint Archive for Sea Level holdings (NCEI Accession 0019568), Version 5.5, NOAA National Centers for Environmental Information, Dataset. *Centers Environ. Information, Dataset*.

Cosgrove, B. A., Gochis, D. J., Graziano, T., Clark, E., & Flowers, T. (2018). An update on the NOAA National Water Model and related activities. In *98th American Meteorological Society Annual Meeting*.

Depietri, Y., Dahal, K., & McPhearson, T. (2018). Multi-hazard risks in New York City. *Natural Hazards and Earth System Sciences*, *18*(12), 3363–3381. https://doi.org/10.5194/nhess-18-

3363-2018

Dresback, K. M., Fleming, J. G., Blanton, B. O., Kaiser, C., Gourley, J. J., Tromble, E. M., Luettich, R. A., Kolar, R. L., Hong, Y., Cooten, S. Van, et al. (2013). Skill assessment of a real-time forecast system utilizing a coupled hydrologic and coastal hydrodynamic model during Hurricane Irene (2011). https://doi.org/10.1016/j.csr.2013.10.007

Dresback, K. M., Fleming, J. G., Blanton, B. O., Kaiser, C., Gourley, J. J., Tromble, E. M., Luettich, R. A., Kolar, R. L., Hong, Y., Van Cooten, S., et al. (2013). Skill assessment of a real-time forecast system utilizing a coupled hydrologic and coastal hydrodynamic model during Hurricane Irene (2011). *Continental Shelf Research*, *71*, 78–94. https://doi.org/10.1016/j.csr.2013.10.007

Egbert, G. D., & Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic Technology*, *19*(2), 183–204. https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2

Funakoshi, Y., Feyen, J., Aikman, F., Tolman, H., Van Der Westhuysen, A., Chawla, A., et al. (2012). Development of extratropical surge and tide operational forecast system (ESTOFS). In *Proceedings of the International Conference on Estuarine and Coastal Modeling* (pp. 201–212). https://doi.org/10.1061/9780784412411.00012

Funakoshi, Y., Feyen, J., Aikman, F., Tolman, H., van der Westhuysen, A., Chawla, A., et al. (2012). Development of Extratropical Surge and Tide Operational Forecast System (ESTOFS). In *Estuarine and Coastal Modeling*. https://doi.org/10.1061/9780784412411.00012

Garzon, J., & Ferreira, C. (2016). Storm Surge Modeling in Large Estuaries: Sensitivity Analyses to Parameters and Physical Processes in the Chesapeake Bay. *Journal of Marine Science and Engineering*. https://doi.org/10.3390/jmse4030045

Garzon, J. L., Ferreira, C. M., & Padilla-Hernandez, R. (2018). Evaluation of weather forecast systems for storm surge modeling in the Chesapeake Bay. *Ocean Dynamics*, *68*(1), 91–107. https://doi.org/10.1007/s10236-017-1120-x

Garzon, Juan L., Ferreira, C. M., & Padilla-Hernandez, R. (2018). Evaluation of weather forecast systems for storm surge modeling in the Chesapeake Bay. *Ocean Dynamics*, *68*(1), 91–107. https://doi.org/10.1007/s10236-017-1120-x

Gross, T. F., Bosley, K. T., & Hess, K. W. (2000). The Chesapeake Bay operational forecast system (CBOFS): technical documentation. US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Office of Coast Survey. *Coast Center for Operational Oceanographic Products and Services*.

Hagedorn, R., Doblas-Reyes, F. J., & Palmer, T. N. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus A: Dynamic Meteorology and Oceanography*, *57*(3), 219–233.

Hanson, J., Wadman, H., Blanton, B., & Roberts, H. (2013). *ERDC/CHL TR-11-1 "Coastal*

Storm Surge Analysis: Modeling System Validation; Report 4: Intermediate Submission No. 2.0." Retrieved from www.erdc.usace.army.mil.

Herdman, L., Erikson, L., & Barnard, P. (2018). Storm surge propagation and flooding in small tidal rivers during events of mixed coastal and fluvial influence. *Journal of Marine Science and Engineering*, *6*(4), 158.

Ikeuchi, H., Hirabayashi, Y., Yamazaki, D., Muis, S., Ward, P. J., Winsemius, H. C., et al. (2017). 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*(4), 1847–1862. https://doi.org/10.1002/2017MS000943

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

Kerr, P. C., Donahue, A. S., Westerink, J. J., Luettich, R. A., Zheng, L. Y., Weisberg, R. H., et al. (2013). U.S. IOOS coastal and ocean modeling testbed: Inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico. *Journal of Geophysical Research: Oceans*, *118*(10), 5129–5172. https://doi.org/10.1002/jgrc.20376

Khalid, A., & Ferreira, C. (2020). Advancing real-time flood prediction in large estuaries: iFLOOD a fully coupled surge-wave automated web-based guidance system. *Environmental Modelling & Software*, 104748. https://doi.org/10.1016/j.envsoft.2020.104748

Kim, S.-C., Chen, J., & Shaffer, W. A. (1996). An operational forecast model for extratropical storm surges along the US east coast. In *Preprints Conference on Coastal Oceanic and Atmospheric Prediction, Atlanta, Amer. Meteor. Soc* (pp. 281–286).

Kirtman, B. P., Min, D., Infanti, J. M., Kinter III, J. L., Paolino, D. A., Zhang, Q., et al. (2014). The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society*, *95*(4), 585–601.

Kourafalou, V. H., De Mey, P., Le Hénaff, M., Charria, G., Edwards, C. A., He, R., et al. (2015). Coastal Ocean Forecasting: system integration and evaluation. *Journal of Operational Oceanography*, *8*(sup1), s127–s146. https://doi.org/10.1080/1755876X.2015.1022336

Krishnamurti, T. N., Kishtawal, C. M., Zhang, Z., LaRow, T., Bachiochi, D., Williford, E., et al. (2000). Multimodel ensemble forecasts for weather and seasonal climate. *Journal of Climate*, *13*(23), 4196–4216.

Loftis, D., & Forrest, D. (2018). Citizen-Science's Role in Flooding Resiliency: Tracing Today's King Tides to Validate a Street-Level Hydrodynamic Model and Responsibly Reduce Risk for the Floods of Tomorrow. *AGUFM*, *2018*, NH23B--06.

Luettich, R. A., Westerink, J. J., & Scheffner, N. (1992). *ADCIRC: an advanced three-dimensional circulation model for shelves coasts and estuaries, report 1: theory and*

methodology of ADCIRC-2DDI and ADCIRC-3DL. Dredging Research Program Technical Report DRP-92-6, U.S. Army Engineers Waterways Experiment Station, Vicksburg, MS,.

Lyddon, C., Brown, J. M., Leonardi, N., & Plater, A. J. (2018). Uncertainty in estuarine extreme water level predictions due to surge-tide interaction. *PLOS ONE*, *13*(10), 1–17. https://doi.org/10.1371/journal.pone.0206200

Mashriqui, H. S., Halgren, J. S., & Reed, S. M. (2014). A 1D River Hydraulic Model for Operational Flood Forecasting in the Tidal Potomac: Evaluation for Freshwater, Tidal, and Wind Driven Events 4 5 6. Retrieved from http://www.nws.noaa.gov/oh/hrl/modelcalibration/6. Hydraulic Model Calibration/potomac_modeling_JHE.pdf

McDowell, M. (2016). *Comparison of Hydrologic and Hydraulic Characteristics of the Anacostia River to Non-Urban Coastal Streams*.

Mied, R. P., Donato, T. F., & Friedrichs, C. T. (2006). Eddy generation in the tidal Potomac River. *Estuaries and Coasts*, *29*(6), 1067–1080. https://doi.org/10.1007/BF02781810

Möller, O. O., Castaing, P., Salomon, J.-C., & Lazure, P. (2001). The influence of local and non-local forcing effects on the subtidal circulation of Patos Lagoon. *Estuaries*, *24*(2), 297–311. https://doi.org/10.2307/1352953

Ray, T., Stepinski, E., Sebastian, A., & Bedient, P. B. (2011). Dynamic Modeling of Storm Surge and Inland Flooding in a Texas Coastal Floodplain. *Journal of Hydraulic Engineering*, *137*(10), 1103–1110. https://doi.org/10.1061/(asce)hy.1943-7900.0000398

Ries III, K. G., Newson, J. K., Smith, M. J., Guthrie, J. D., Steeves, P. A., Haluska, T. L., et al. (2017). *StreamStats, version 4*.

Roberts, K. J., Pringle, W. J., & Westerink, J. J. (2019). OceanMesh2D 1.0: MATLAB-based software for two-dimensional unstructured mesh generation in coastal ocean modeling. *Geoscientific Model Development*, *12*(5), 1847–1868. https://doi.org/10.5194/gmd-12-1847-2019

Santiago-Collazo, F. L., Bilskie, M. V., & Hagen, S. C. (2019). A comprehensive review of compound inundation models in low-gradient coastal watersheds. *Environmental Modelling and Software*, *119*(June), 166–181. https://doi.org/10.1016/j.envsoft.2019.06.002

Shen, J., Wang, H., Sisson, M., & Gong, W. (2006). Storm tide simulation in the Chesapeake Bay using an unstructured grid model. *Estuarine, Coastal and Shelf Science*, *68*(1), 1–16. https://doi.org/10.1016/j.ecss.2005.12.018

Sumi, S. J., & Ferreira, C. (2019). Compound Urban Flooding in Large Metropolitan Areas along Tidal Rivers: A Case Study for the Washington DC Region. *AGUFM*, *2019*, H13J--1821.

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

Thatcher, C. A., Brock, J. C., Danielson, J. J., Poppenga, S. K., Gesch, D. B., Palaseanu-Lovejoy, M. E., et al. (2016). Creating a Coastal National Elevation Database (CoNED) for science and conservation applications. *Journal of Coastal Research*, (76), 64–74.

Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., et al. (2014). XSEDE: Accelerating scientific discovery. *Computing in Science and Engineering*, *16*(5), 62–74. https://doi.org/10.1109/MCSE.2014.80

Tshimanga, R. M., Tshitenge, J. M., Kabuya, P., Alsdorf, D., Mahe, G., Kibukusa, G., & Lukanda, V. (2016). A Regional Perceptive of Flood Forecasting and Disaster Management Systems for the Congo River Basin. In *Flood Forecasting* (pp. 87–124). Elsevier.

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

Walsh, C. J., Fletcher, T. D., & Burns, M. J. (2012). Urban Stormwater Runoff: A New Class of Environmental Flow Problem. *PLOS ONE*, *7*(9), 1–10. https://doi.org/10.1371/journal.pone.0045814

Wang, H. V, Loftis, J. D., Forrest, D., Smith, W., & Stamey, B. (2015). Modeling Storm Surge and Inundation in Washington, DC, during Hurricane Isabel and the 1936 Potomac River Great Flood. *Journal of Marine Science and Engineering*, *3*(3), 607–629. https://doi.org/10.3390/jmse3030607

Weigel, A. P., Liniger, M. A., & Appenzeller, C. (2008). Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, *134*(630), 241–260.

Wessel, P., & Smith, W. H. F. (1996). A global, self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research: Solid Earth*, *101*(B4), 8741–8743. https://doi.org/10.1029/96jb00104

Wu, W., McInnes, K., O'Grady, J., Hoeke, R., Leonard, M., & Westra, S. (2018). Mapping Dependence Between Extreme Rainfall and Storm Surge. *Journal of Geophysical Research: Oceans*, *123*(4), 2461–2474. https://doi.org/10.1002/2017JC013472

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

Zhong, L., Li, M., & Foreman, M. G. G. (2008). Resonance and sea level variability in Chesapeake Bay. *Continental Shelf Research*, *28*(18), 2565–2573.

7. Appendix

Table A1. List of Stations in the Study area

Region	Station Full Name	Abbr-eviaton	Long	Lat	Observed Water	Observed Flow	Tide Prediction
Potomac Upper River Reach	Little Falls Pump Station, MD	LFMD	-77.13	38.95	Y	Y	-
	Wisconsin Ave, DC	WADC	-77.07	38.90	Y	-	-
Anacostia River Reach	Northeast Branch Riverdale, MD	NEMD	-76.93	38.96	-	Y	-
	Northwest Branch Hyattsville, MD	NWMD	-76.97	38.95	-	Y	-
	Bladensburg, MD	BDMD	-76.94	38.93	-	-	Y
	East Lake, MD	ELMD	-76.96	38.91	-	-	Y
	Kingman Lake, MD	KLMD	-76.97	38.89	-	-	Y
	Washington Navy Yard, DC	WNDC	-76.99	38.87	-	-	Y
Main Potomac River	Washington, DC	WASD	-77.02	38.87	-	-	Y
	Bellevue, DC	BEDC	-77.03	38.83	-	-	Y
	Alexandria, VA	ALVA	-77.04	38.80	-	-	Y
	Fourmile Run, VA	FMVA	77.05	38.84	Y	-	-
	Fourmile Run, Stream Station, VA	FSVA	77.09	38.84	-	Y	-
	Cameron St Dock at Alexandria, VA	CSVA	77.04	38.81	Y	-	-
	Cameron Run at Alexandria, VA	CRVA	77.11	38.80	-	Y	-
	Piscataway Creek, MD	PCMD	76.97	38.71	-	Y	-
	Marshall Hall, MD	MHMD	-77.10	38.69	-	-	Y
	Indian Head, MD	IHMD	-77.19	38.60	-	-	Y
	Quantico, VA	QUVA	-77.29	38.52	-	-	Y
	Liverpool Point, MD	LPMD	-77.27	38.46	-	-	Y
	Clifton Beach, Smith Point, MD	CLMD	-77.27	38.41	-	-	Y
	Riverside, MD	RSMD	-77.14	38.39	-	-	Y
	Goose Creek, MD	GCMD	-77.05	38.45	-	-	Y
	Mathias Point, VA	MPVA	-77.06	38.40	-	-	Y
	Dalghren, VA	NCDV	-77.04	38.32	Y	-	Y
	Lower Cedar Point, MD	LCMD	-76.98	38.34	-	-	Y
	Colonial Beach, Potomac River, VA	COVA	-76.96	38.25	-	-	Y
	Colton Point, MD	CPMD	-76.75	38.22	-	-	Y
	Ragged Point, VA	RPVA	-76.61	38.14	-	-	Y
	Piney Point, MD	PPMD	-76.53	38.13	-	-	Y
	Lewisetta, VA	LWTV	-76.47	37.99	Y	-	Y

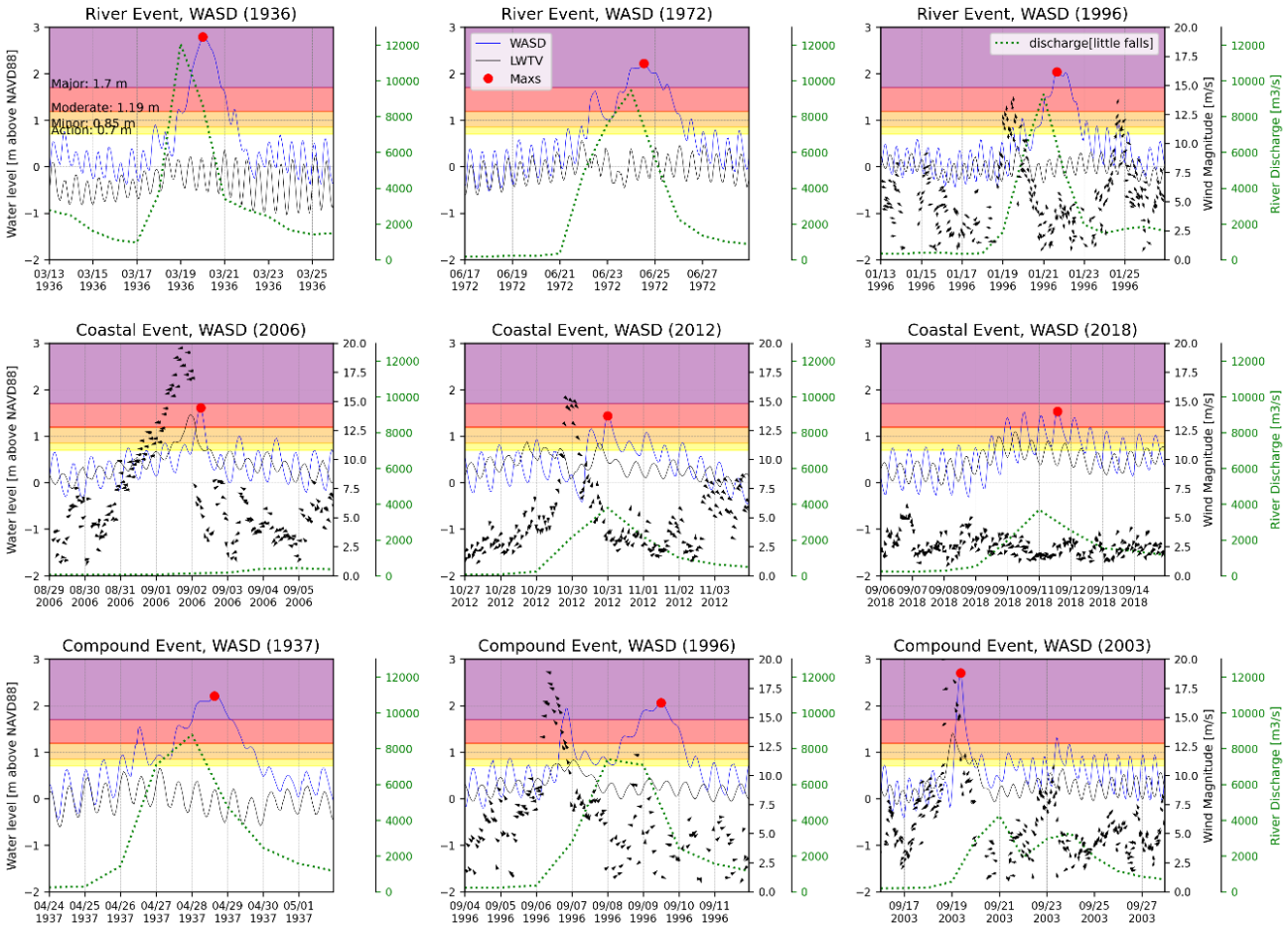
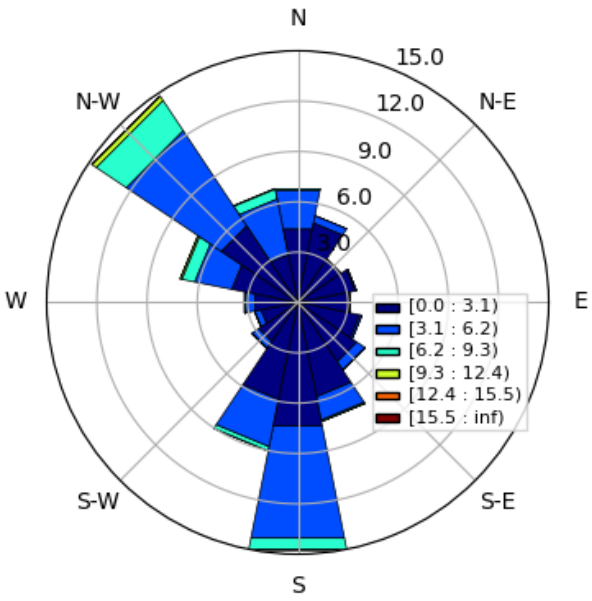


Figure A1. Time series of observed data for various case studies



955

956 **Figure A2.** Wind rose plot for the observed winds at WASD from 2008 to 2020

957

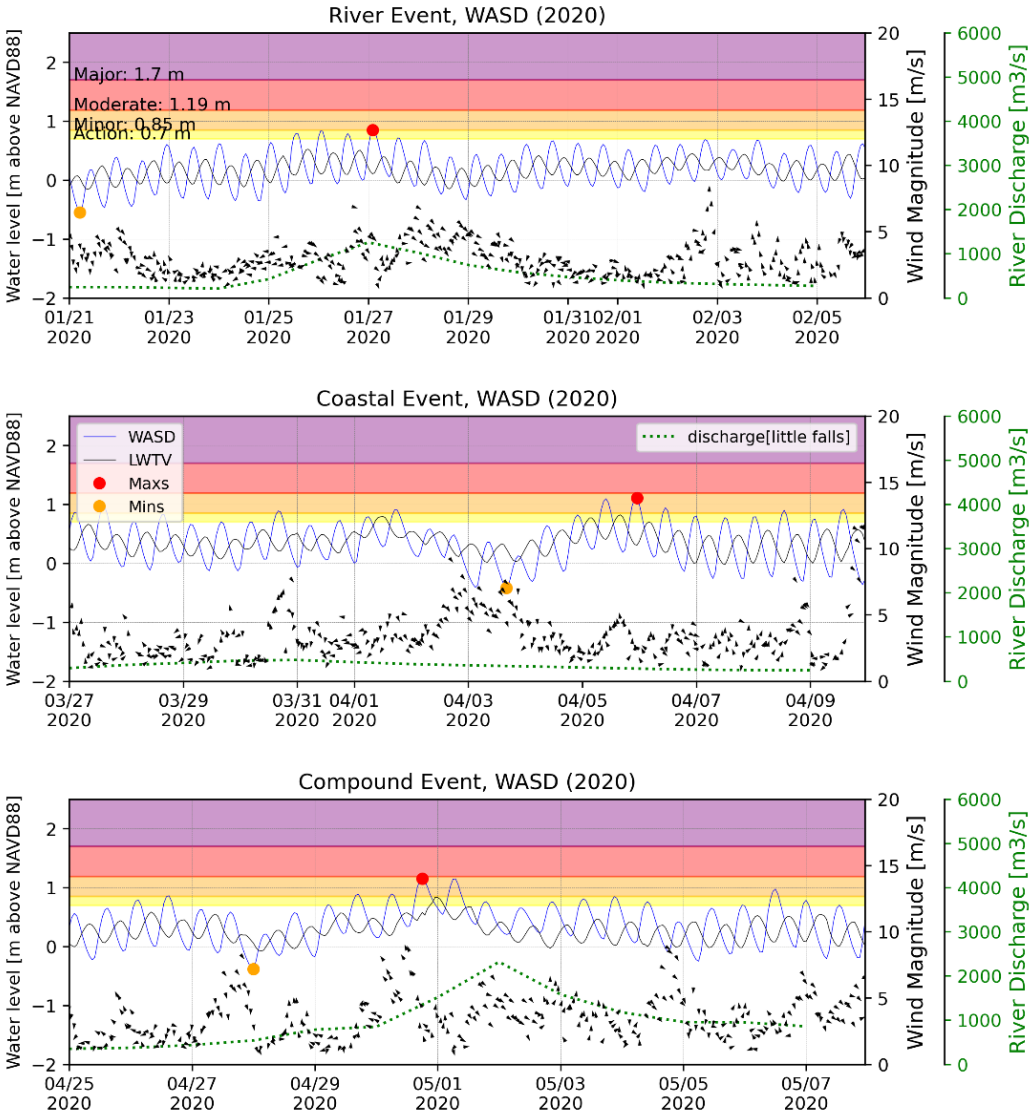
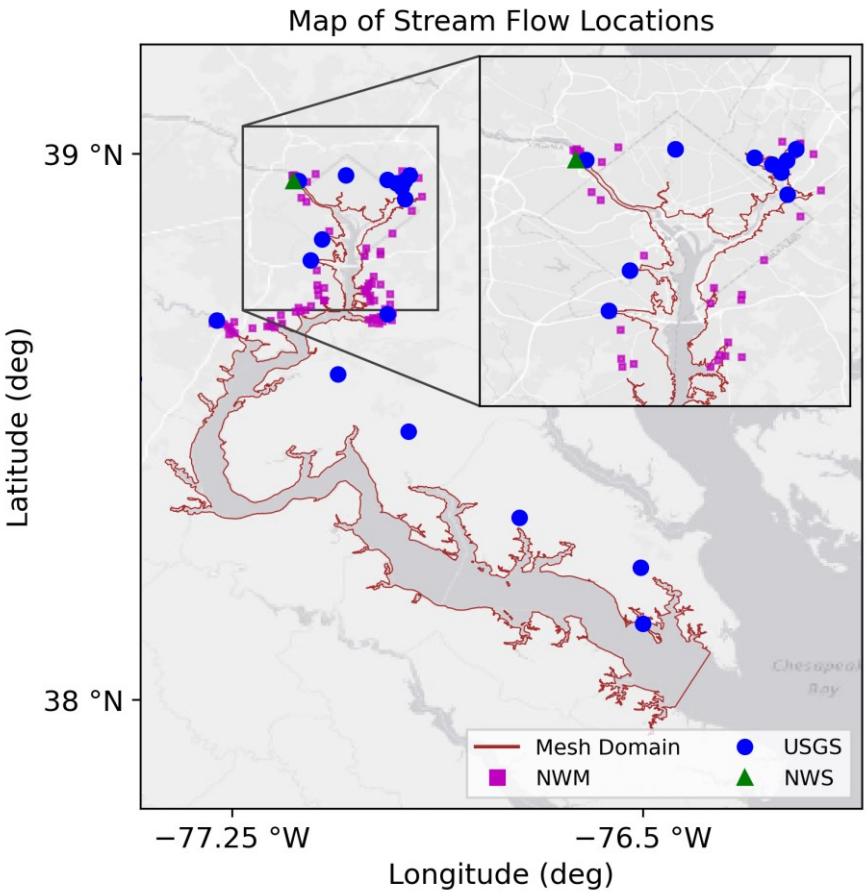


Figure A3. Time series of observed data for various case studies

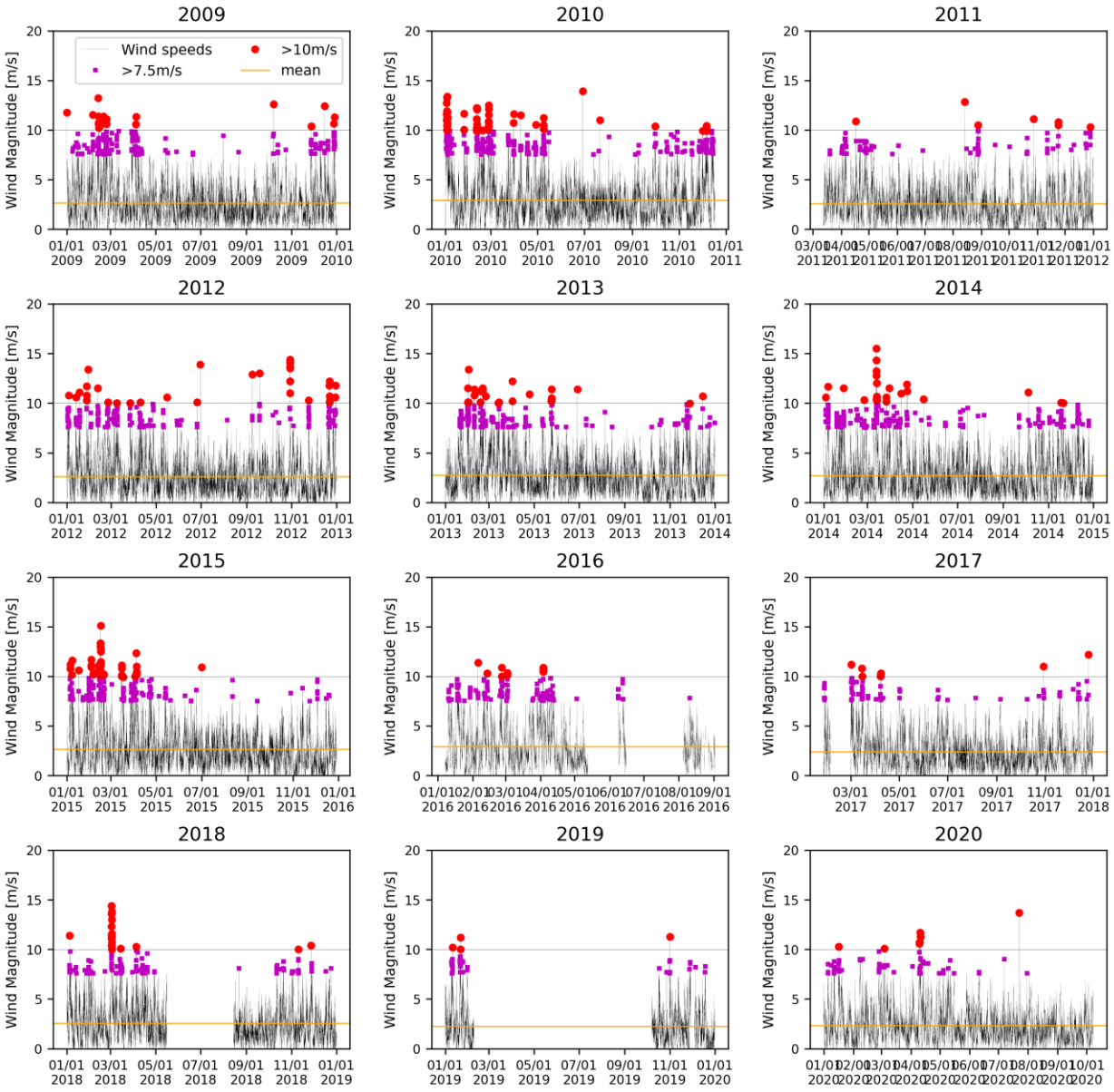


960

961 **Figure A4.** Location map of USGS, NWS and NWM observation and forecast stations.

962

963



964

965

966 **Figure A5.** Observed wind speeds (above 10m height) at WASD during the period of 12

967 consecutive years (2009 to 2020)

968

969

Data Availability Statement

All model analyses in this study were conducted on behalf of the Mason Flood Hazards Research Lab (https://fhrl.vse.gmu.edu/) and are stored on the local servers. The modeling outputs are available for non-commercial, academic research purposes, only upon reasonable request from the corresponding author. Hydrodynamic coastal storm surge model, ADCIRC, is available for non-commercial, academic research purposes, by contacting Crystal Fulcher at the University of North Carolina (cfulcher@email.unc.edu). The integrated modeling framework used in this research was based on the recently published iFLOOD paper (https://doi.org/10.1016/j.envsoft.2020.104748) and is available to view on the iFLOOD web portal (https://iflood.vse.gmu.edu/map). Historical observational data for winds and water level was retrieved from NOAA tides and currents database (https://api.tidesandcurrents.noaa.gov/api/prod/) while the streamflow data was available online at USGS water database (https://waterdata.usgs.gov/nwis). The streamflow data at a given return period was calculated using the online StreamStats server available at https://streamstats.usgs.gov/ss/. The real-time input forcing of streamflow (NWM), water level guidance (ESTOFS, ETSS) and winds (NAM, NBM) for upstream and downstream boundaries were downloaded daily from the NOAA NOMADs server available at https://nomads.ncep.noaa.gov/pub/data/nccf/com/. The water level forecasts from CBOFS model were available at http://opendap.co-ops.nos.noaa.gov/netcdf/. The iFLOODv2 forecasts were also downloaded daily from the iFLOOD online data repository available at https://data.iflood.vse.gmu.edu/?prefix=Forecast/ChesapeakeBay_ADCIRCSWAN/

The NWS forecasts of water level and stream flows were downloaded on daily basis and are available at https://water.weather.gov/ahps2/forecasts.php?wfo=lwx.