



Compounding factors for extreme flooding around Galveston Bay during Hurricane Harvey



Wei Huang ^a, Fei Ye ^a, Y. Joseph Zhang ^{a,*}, Kyeong Park ^b, Jiabi Du ^c, Saeed Moghimi ^d, Edward Myers ^d, Shachak Pe'er ^d, Jaime R. Calzada ^d, H.C. Yu ^a, Karinna Nunez ^a, Zhuo Liu ^e

^a Virginia Institute of Marine Science, College of William & Mary, Gloucester Point, VA 23062, USA

^b Department of Marine and Coastal Environmental Science, Texas A&M University at Galveston, Galveston, TX 77553, USA

^c Applied Ocean Physics & Engineering Department, Woods Hole Oceanographic Institute, Woods Hole, MA 02543, USA

^d Coast Survey Development Laboratory, NOAA, Silver Spring, MD 20910, USA

^e One Concern Inc., 855 Oak Grove Ave, Menlo Park, CA 94025, USA

ARTICLE INFO

Keywords:

Compound flooding
Hurricane Harvey
Galveston Bay
SCHISM

ABSTRACT

Coastal hazard is rarely driven by only one source, as exemplified by the compound flooding from Hurricane Harvey in Galveston Bay in 2017. A 3D creek-to-ocean model is developed to explicitly resolve, without grid nesting, the marine (combination of atmospheric forcing and tides), fluvial and pluvial extremes for this extreme event. We first thoroughly assess the model skills using all available observations in the Galveston Bay region, including High Water Marks (HWMs) and field estimates of maximum inundation extent in the watershed. Subject to uncertainties in the river flows, atmospheric forcing, initial condition of salinity and temperature, and digital elevation model of bathymetry-topography, the model is shown to generally exhibit good skills for predicting inundation and compound surges, with a hit rate for inundation extent of 0.92, average mean-absolute-errors of 0.65 m for HWMs, 1.7 psu for salinity, and 1.4 °C for temperature. We then apply the model to quantify the individual contributions from the three major forcings (ocean, river and precipitation). Comparison of results (in the form of ‘compound ratio’) from the simulations with three factors being applied individually with those from the baseline simulation with all factors included in a single model reveals the nonlinear compounding effects in most of the areas in Galveston Bay, and indicates that the compound flooding problems are best simulated using a single model that integrates across all factors because the interactions among processes are very complex and highly nonlinear; in other words, summing up the results from individual forcings would lead to large errors. For example, the hydrodynamic model results forced by river inflows at boundary and oceanic and atmospheric forcings, without explicitly accounting for the direct precipitation in the coastal watersheds, would severely underestimate the resultant flow and surge by up to 90%. ‘Regions of dominance’ are also identified for each forcing factor from the sensitivity results. These concepts are applicable to other compound flooding studies as well.

1. Introduction

Recent tropical cyclones have shown tendency to stall near the coast (Hall and Kossin, 2019) and an increasing trend of concurrent heavy precipitation (Pfahl et al., 2017), which together with the large-scale sea level rise may cause severe flooding nearshore. One of the most extreme examples is Hurricane Harvey, which caused catastrophic flood in the Texas and Louisiana coast in 2017. Harvey started as an atmospheric tropical wave between western Africa and the Caribbean on August 16–17, 2017, and developed into a stronger depression and tropical storm as it moved westward across the Caribbean before being disintegrated briefly for a few days (Valle-Levinson et al., 2020).

After the tropical wave moved northward into the Gulf of Mexico, it quickly intensified into a Category 4 hurricane by August 24, 2017. The hurricane made its first landfall on August 26 at San José Island, Texas, followed by another landfall at Holiday Beach at Category 3 intensity (Fig. 1a). From there on, the storm lingered for about a week near Texas and Louisiana coast and even briefly reemerged into the Gulf of Mexico on August 28. As a result, Harvey dumped record-breaking amount of rainfalls (Fig. 2), with the return period of the peak 3-day precipitation exceeding 1000 years (van Oldenborgh et al., 2018), and caused more than 80 fatalities and over \$150 billion economic losses, mostly due to the extraordinary flooding (Emanuel, 2017; Balaguru et al., 2018). Over the 5-day period from August 26 to 30, Harvey

* Corresponding author.

E-mail address: yjzhang@vims.edu (Y.J. Zhang).

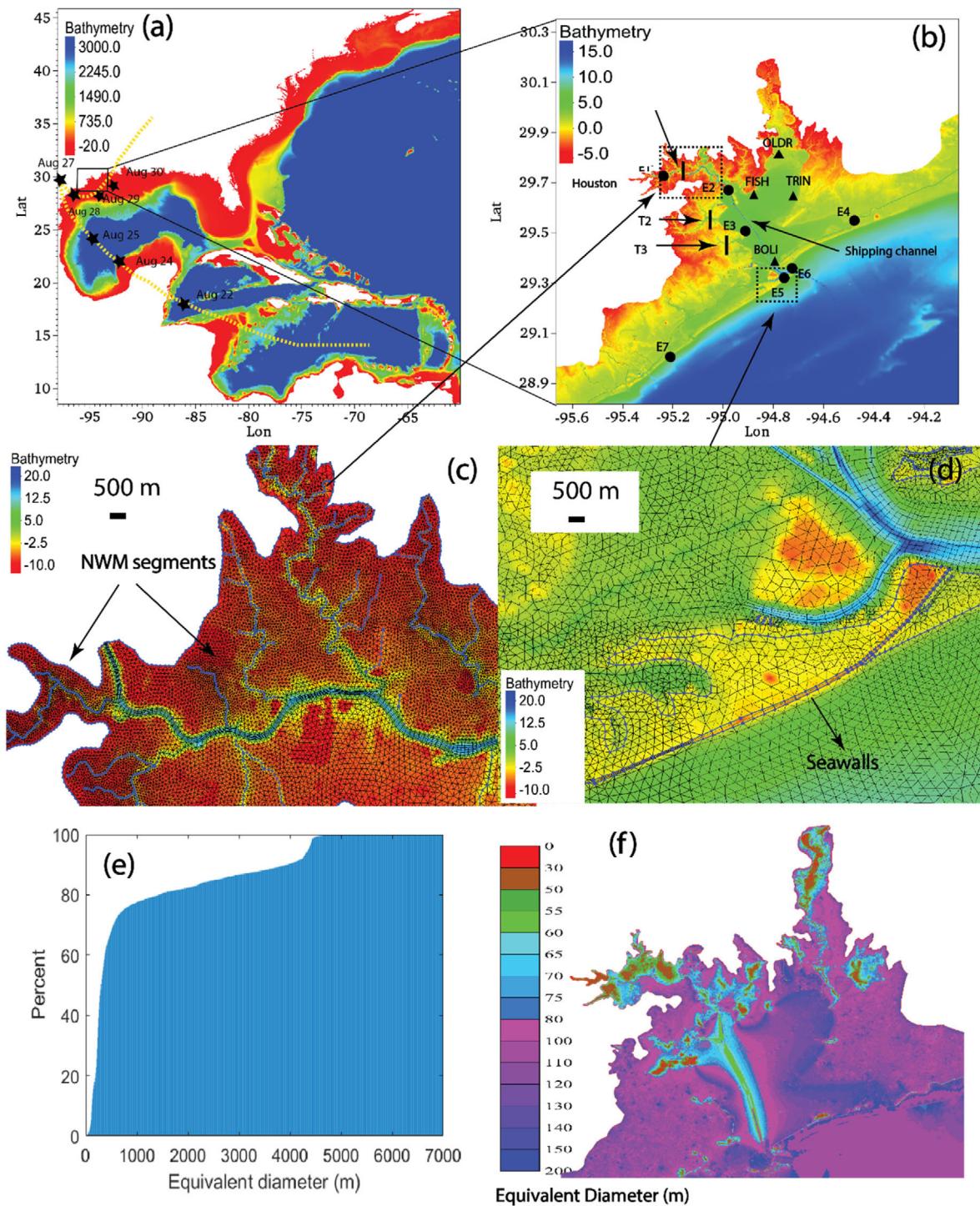


Fig. 1. Study domain with multiple zooms into Galveston Bay and its coastal watershed, including channel network, levees and seawall. (a) Grid extent and hurricane track (yellow line with stars). The hurricane lingered along the Texas-Louisiana coast for about a week. (b) Observation stations in Galveston Bay for water level (E1–E7, circles), salinity/temperature (BOLI, FISH, TRIN and OLDL), and velocity (g08010 and g06010 which are co-located with E2 and E5, respectively). T1–T3 are transects used for evaluation of flows. (c) Zoom-in of the grid near Buffalo Bayou with National Water Model (NWM) segments overlaid (the dark lines). (d) Zoom-in near bay entrance showing the channel network and seawall (represented by quad elements). (e) Cumulative histogram of grid resolution for the entire horizontal grid. (f) Grid resolution in the form of equivalent diameter for the Galveston-Houston region.

dumped $92.7 \times 10^9 \text{ m}^3$ of water across Texas and Louisiana (Fritz and Samenow, 2017), making it the wettest tropical cyclone in the U.S. history. The extraordinary amount of water load even caused up to 21 mm subsidence of Earth's crust (Milliner et al., 2018). A freshwater load of $14 \times 10^9 \text{ m}^3$ was estimated to discharge into Galveston Bay (about 3.7 times the bay volume), making the entire bay virtually

fresh for several days (Du et al., 2019a,b). The large freshwater load led to a prolonged restoration time for bay salinity (with the bay-wide average of 62 days), which had profound impact on pollutant transport and estuarine ecosystem (Du and Park, 2019; Du et al., 2019a, 2020). As many climate models predict a warmer and wetter future climate in parts of the US and the rest of the world, and as evident

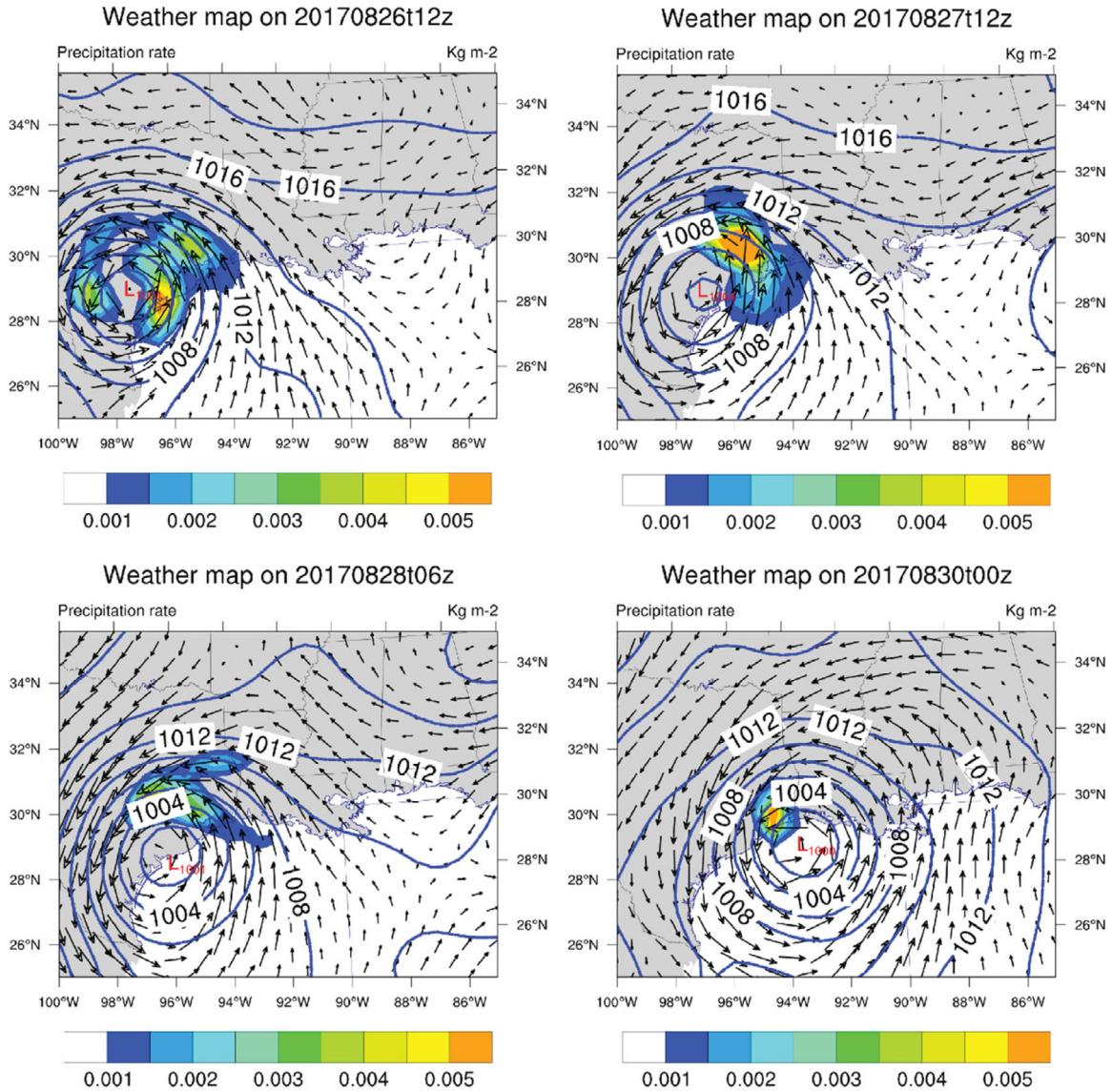


Fig. 2. Weather maps during Hurricane Harvey, showing wind vectors at 2 m above ground (arrows), air pressures in hPa (blue lines), and precipitation rate in kg/m^2 (colors). The storm stalled near the Texas-Louisiana coast for about a week. Note that the model uses 10 m above MSL wind as forcing.

from the record-breaking Atlantic hurricane season in 2020 (including a number of very wet storms), it is urgent to gain a deeper understanding of this type of extreme precipitation events as exemplified by Harvey.

As explained by Valle-Levinson et al. (2020), the compounding effects were very significant during Harvey near Houston, which had already been affected by flooding due to a smaller rainfall event before the storm. During Harvey, the Houston metropolitan area received multiple pulses ('punches') of water from various sources: the river flood from shipping channel in Buffalo Bayou (cf. Fig. 3a) that cuts through the city, the subsequent flooding in the larger tributary, San Jacinto to the north (cf. Fig. 3a), which was hypothesized to have blocked the outflowing water from Buffalo Bayou, and storm surges from the ocean that entered through the bay entrance. They also suspected that various constrictions in the bay system might have caused significant backflow, which ultimately caused a 3.5 m surge at Manchester gauge in Buffalo Bayou (E1 in Fig. 1b).

Numerical models can serve as a powerful tool for the study of compound flooding and potential mitigation measures (e.g., Lewis et al., 2019; Ye et al., 2020; Zhang et al., 2020). However, because the numerical models used in previous studies of Harvey did not

directly account for precipitation and the watershed–ocean interactions, significant knowledge gaps still exist on how different processes (ocean, rivers, and rainfall) interact to produce the observed compound surges during this extreme event. For example, Wing et al. (2019) used a 2D hydrologic model coupled to National Water Model (NWM; <https://water.noaa.gov/about/nwm>, last accessed in August 2020) for Hurricane Harvey, forced by the boundary conditions from a simple storm surge model at the ocean side; as a result, the feedback from river to ocean surges was missing, which we will show to be significant in this paper. Therefore, in the present study we apply a recently developed 3D cross-scale model (Ye et al., 2020; Zhang et al., 2020) to examine the relative importance of multiple factors for the observed compound flooding in different areas of the Galveston Bay system. A rich myriad of complex nonlinear interactions among processes has so far been identified; however, in this paper we will first focus on a very thorough model validation and some results from salient sensitivity tests.

We will focus on 3D model results only in this paper and omit the discussions on the comparison of results from the 2D and 3D models; the latter has been carefully assessed in Ye et al. (2020). In general, the 3D model has the advantage of being able to capture the

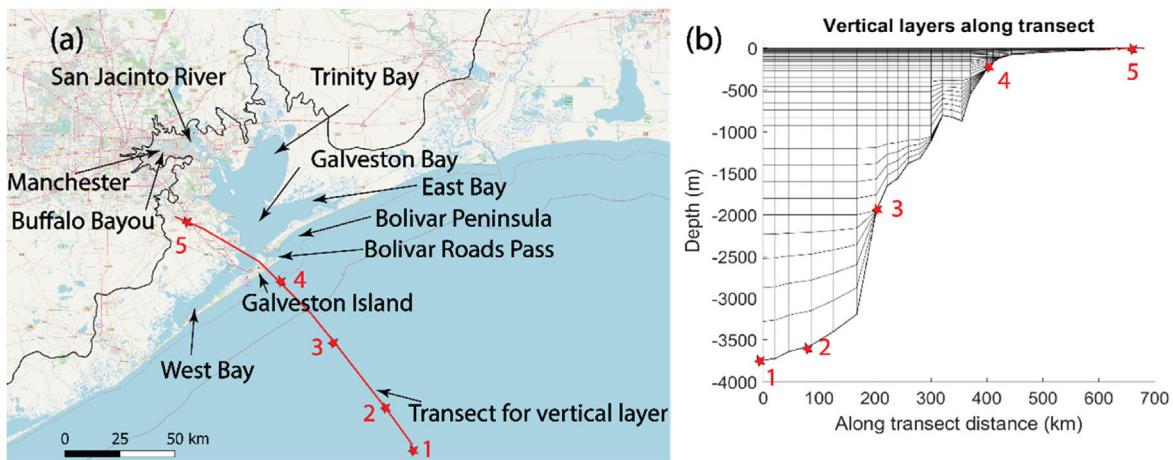


Fig. 3. (a) Geographic regions in Galveston Bay, (b) the vertical grid along the red line in (a), and (c) locations of 36 Argo profiles in Gulf of Mexico used to evaluate the model. The black lines in (a) and (c) are model land boundary.

baroclinic response of the ocean especially in the post-storm adjustment, thus giving a better prediction of the sub-tidal elevations due to the interaction between large- and small-scale processes. On the other hand, the two model configurations yield similar results in the watershed (which is not surprising since the 3D model morphs into the 2DH mode there; cf. Section 2.3). In addition, the 3D model is able to predict other relevant variables besides the surface elevation and depth-averaged velocity during the storm, such as the 3D velocity shear and tracer (salinity and temperature etc.) stratification (in the vertical and horizontal directions). In the case of Harvey, large shear and stratification in the deep shipping channel during the event are revealed by our model.

The paper is organized as follows. Section 2 describes the study area, observation assets we utilize in this paper and model setup for the baseline and sensitivity simulations. Section 3 presents a comprehensive model validation for 2D and 3D variables inside and outside the Bay as well as in the watershed. Section 4 focuses on the individual contributions from the three major forcings (ocean, rivers and precipitation), and the nonlinear compound effects as revealed from the comparison between the baseline simulation with all three forcings applied simultaneously and the sensitivity simulations with only one forcing applied at a time. A brief conclusion is summarized in Section 5.

2. Methods

2.1. Study site

Galveston Bay, located at the upper coast of Texas along the northwestern Gulf of Mexico, is the seventh largest estuary in the continental US. It is also adjacent to the heavily urbanized and industrialized metropolis of Houston, which is connected to the Gulf of Mexico through an intricate system of bayous, rivers, and bays (Dorado et al., 2015). Galveston Bay is a wide and shallow estuary with average depth of ~2 m and total area of 1600 km². Its major freshwater sources are the Trinity River (~50%), the San Jacinto River (~30%), Buffalo Bayou (~12%) and other local watersheds feeding Galveston Bay (~8%) (Guthrie et al., 2012; Lucena and Lee, 2017; D'Sa et al., 2018). During flood seasons, rain-related flooding in Houston occurs mostly due to two major tributaries: San Jacinto River (SJR) and Buffalo Bayou (B.B.) (Valle-Levinson et al., 2020); SJR delivers discharge from the Lake Houston dam to the upper Galveston Bay. Galveston Bay is connected to the open ocean by a deep (~14 m) and narrow (~200 m) channel (Houston Ship Channel) through Bolivar Roads Pass (Fig. 3a) at the entrance of the bay (Lakshmanan et al., 2010). The northern Gulf of Mexico is a microtidal region, and the major diurnal tidal range for the Bay is from 0.15 to 0.5 m during non-hurricane seasons (Liu et al.,

2019). There are several sub-bays, including Trinity Bay in the northeast, East/West Bays to the northeast/southwest that are surrounded by Galveston island and Bolivar Peninsula (Fig. 3a), and a few smaller embayments that connect B.B. and SJR to the main bay.

2.2. Observational data

Water level data were obtained from seven NOAA gauges (Fig. 1b): 8770777 (E1, Manchester), 8770613 (E2, Morgan Point), 8771013 (E3, Eagle Point), 8770971 (E4, Rollover Pass), 8771450 (E5, Galveston Pier), 8771341 (E6, Galveston Bay Entrance, North Jetty), and 8772447 (E7, Freeport Harbor). Vertical datums for stations E3–E6 are in NAVD 88. The data at other stations are in mean sea level (MSL) and therefore have been converted to NAVD88 by adding 20 cm at E1 and E2, and 11 cm at E7, as estimated by the differences between NAVD88 and MSL datums at the nearby stations. Velocity data were obtained from Texas Automated Buoy System (TABS) stations g08010 and g06010 (<http://pong.tamu.edu/tabswebsite>, last accessed in August 2020) which are co-located with water elevation stations E2 and E5, respectively. Hourly salinity and temperature data at OLDR, FISH, TRIN, and BOLI are available from Water Data for Texas (<https://waterdatafortexas.org/coastal>, last accessed in August 2020), maintained by Texas Water Development Board (TWDB). Profiles of temperature and salinity from 36 Argo floats (http://www.argo.ucsd.edu/Data_viewers.html, last accessed in August 2020) in the Gulf of Mexico near the bay were used to help assess the model skill for large-scale processes that might influence surges (Ye et al., 2020).

High Water Marks (HWMs) are available from the USGS post disaster survey in the region (<https://webapps.usgs.gov/harvey/>, last accessed in August 2020). In addition, USGS and Federal Emergency Management Agency (FEMA) cooperatively initiated a study to evaluate the magnitude of the flood and to determine the map of flood extent (Watson et al., 2018). The USGS field crews surveyed 2123 HWMs referenced to NAVD88 datum. Finally, 1258 HWMs (about half of which are outside our domain) were obtained with averaging over some locations. Inundation maps were then created to estimate the areal extent of maximum flooding based on these HWMs, and a best estimate of the maximum flood extent was compiled by FEMA based on HWMs and USGS DEMs (Digital Elevation Models) (<http://www.hydroshare.org/resource/e8768f4cb4d5478a96d2b1cbd00d9e85>, last accessed in August 2020). However, as stated in their document, areas of more widespread flooding or areas where flooding seems unusually small suggest a source data problem in the flood extents as compared with ground elevation data (FEMA, 2020), which should be treated as an uncertainty.

Bathymetry/topography information used was compiled from a variety of DEM sources with horizontal resolutions ranging from 3 arc-second (~ 90 m) to 1 m: NOAA's U.S. Coastal Relief Model (CRM; <https://www.ngdc.noaa.gov/mgg/coastal/model.html>, last accessed in August 2020), NOAA's Coastal Digital Elevation Model (CDEM; <https://catalog.data.gov/dataset/digital-elevation-models-from-noaa-ngdc>, last accessed in August 2020), Continuously Updated Digital Elevation Model (CUDEM; <https://www.ngdc.noaa.gov/mgg/coastal/model.html>, last accessed in August 2020), and the USGS Coastal National Elevation Database (CoNED; <https://www.usgs.gov/land-resources/eros/coned>, last accessed in August 2020). A 10-m resolution DEM from Lidar image (Du et al., 2019a) was used in Galveston Bay and its coastal watershed. The horizontal datums for all sources have been homogenized to WGS 84 in latitude and longitude. The vertical datums in these sources are usually given in NAVD88 except for CRM; because the vertical accuracy in CRM is often no better than the uncertainty of the datums, the vertical datum of CRM was not corrected to NAVD88.

2.3. Model setup

Building on previous successes, we have implemented a seamless creek-to-ocean SCHISM 3D model (Zhang et al., 2016) for the entire US east coast and Gulf of Mexico (Fig. 1a). SCHISM is a general-purpose 3D hydrodynamic model grounded on accurate, robust and efficient semi-implicit time stepping (with no mode splitting) and hybrid finite-element/finite-volume methods on hybrid triangular-quadrangular unstructured grids in the horizontal dimension and a highly flexible vertical gridding system in the vertical dimension (LSC²; Zhang et al., 2015). A list of worldwide applications of SCHISM can be found at schism.wiki.

Unlike in the Hurricane Irene study (Ye et al., 2020; Zhang et al., 2020), the model domain is now extended to include all major bays and estuaries in high resolution in US east coast and Gulf of Mexico. This way we can address all potential land-falling hurricanes in the US east coast and Gulf of Mexico regions. The landward boundary is set at 10 m above NAVD88 to include most of coastal watersheds (Fig. 3c) with the grid resolution of ~ 300 m. Furthermore, higher resolutions are deployed in the focus area of this study, Galveston Bay, with average resolution of ~ 150 m (Fig. 1f). Overall, the horizontal resolution varies from 1 m to 7 km with a statistic mode of 250 m and 50% of the elements having a finer resolution than 200 m (Fig. 1e). In addition, all major geometric and bathymetric features (manmade or otherwise) are resolved in the grid in order to accurately capture localized processes that may impact compound surges from all potential flood events. For example, the finest resolution of ~ 1 m is used to resolve some small-scale structures (e.g. Galveston seawall and levees in the lower Mississippi River; Fig. 1d). Strictly speaking, however, a hydrostatic model such as ours cannot accurately capture the fine details of flow over small-scale structures (levees etc.), as the latter is non-hydrostatic in nature due to large vertical acceleration. However, as explained by Zhang et al. (2020), if our focus is on the impact of such processes on the larger scales, we can still use hydrostatic models to efficiently simulate the interactions. All 284673 NWM river segments located inside our domain have been explicitly incorporated into our grid during the grid generation process (cf. Ye et al., 2020) in order to facilitate the routing of water in the watershed (cf. Fig. 1c). Altogether, the horizontal unstructured grid has 2.2 million nodes and 4.4 million triangular-quadrangular elements.

The 3D grid takes full advantage of SCHISM's 'polymorphism' (Zhang et al., 2016), with variable number of vertical layers ranging from 44 in the deep ocean (to capture the large-scale processes) to 1 in the watershed (Fig. 3b); in other words, the model is 2DH in the watersheds. The average number of vertical layers is 8, which results in greater efficiency. Customary of all SCHISM applications, no manipulation or smoothing of bathymetry was done in the computational

grid after interpolating depths from DEMs (including steep slopes in the Caribbean and all shipping channels). The vertical datum used in all simulations was NAVD88 (~ 12 cm above the local MSL in the Bay), and the 3D model was able to largely capture the sub-tidal variations locally (Ye et al., 2020).

The 3D model is initialized on August 4, 2017 and run for 40 days to cover the entire storm event; the very large freshwater outflow in this event makes the 'system memory' of initial condition short. Other aspects of the setup largely follow those in Ye et al. (2020). The initial and ocean boundary conditions are from HYCOM (<https://www.hycom.org/data/glb0pt08>) and FES2014 tidal database (Carreira et al., 2016). Since HYCOM results are not accurate nearshore, the initial conditions there are interpolated from limited observation found there. The atmospheric forcing is derived from ERA5 (~ 31 -km resolution) and a high-resolution (~ 9 km) ECMWF product (see cf. Acknowledgments); precipitation data is from NOAA's High-Resolution Rapid Refresh (HRRR) that uses a 3-km resolution, an hourly updated, cloud-resolving, convection-allowing atmospheric model, initialized by 3-km grids with 3-km radar assimilation (<https://rapidrefresh.noaa.gov/hrrr/>, last accessed in August 2020).

The watershed inflow from rivers is mostly derived from NWM version 2.0 and the flow is simply injected at 6752 intersecting locations between NWM segments and our land boundary (Fig. 4b; Ye et al., 2020). However, the current version of NWM tends to over-estimate the flow peaks at managed rivers (Brian Cosgrove, private comm), as was the case for Hurricane Irene (Ye et al., 2020). The errors in the river flows will greatly impact the predicted surges (cf. Figs. 4 and 5). Therefore, in the *baseline* setup we have replaced NWM flows at the two major rivers (B.B. and SJR) with either USGS observation (at B.B.) or a best estimate (at SJR) based on a flow-stage relationship developed by TWDB (Fig. 4c,d). During Harvey, significant flooding occurred at all rivers and the SJR flow, which peaked above 10 000 m³/s, is comparable to the total flow from the other 84 rivers discharging into Galveston Bay (Fig. 4d). Therefore, the uncertainty in the other river flows is also expected to impact the predicted surges. Note that the river flows injected at our model land boundary have indirectly incorporated the precipitation that occurred *outside* (but not *inside*) the model domain, and therefore, the addition of direct precipitation onto our model domain is justified and will be shown to exert a significant impact on the compound processes. Since we have no information on the scalar concentrations for river inflows and rainfall, we applied 0 psu for salinity and ambient water temperature (i.e., the temperature at the local receiving cell calculated without accounting for the rivers or raindrops) for the injected water and also for the rainfall. Obviously, the latter will impact the calculated temperature, which is another source of uncertainty for the model results (cf. Section 3.6).

Multiple simulation runs were done during the calibration process. The *baseline* setup generally follows Zhang et al. (2020) and is briefly described here. The time step was set at 100 s (although 150 s led to very similar results). The wetting and drying threshold was set at 10^{-6} m in order to capture very thin layer of water from rainfall initially (Zhang et al., 2020). The air-sea exchange scheme used was the bulk aerodynamic algorithm of Zeng et al. (1998) for heat and salt exchanges (including precipitation and evaporation). The air-sea heat exchange was found to misbehave on very thin layer of water and was therefore turned off based on the local water depth; the baseline run used 0.1 mm as the minimum threshold for the local total water depth (Section 3.7 will further discuss the sensitivity to this choice). As stated in Ye et al. (2020), the bottom friction coefficients should be larger in the watershed; for a national grid like this and in view of lack of bottom characteristic and bed form data in most bays (not to mention the vegetation and building effects in the watershed), we start from a simple friction parametrization function of bathymetry only: the friction coefficients increase linearly from a constant 0.0025 (in the ocean) to 0.025 (in the watershed) as the bathymetry changes from -1 m to -3 m. In the future, fine tuning of friction in light of

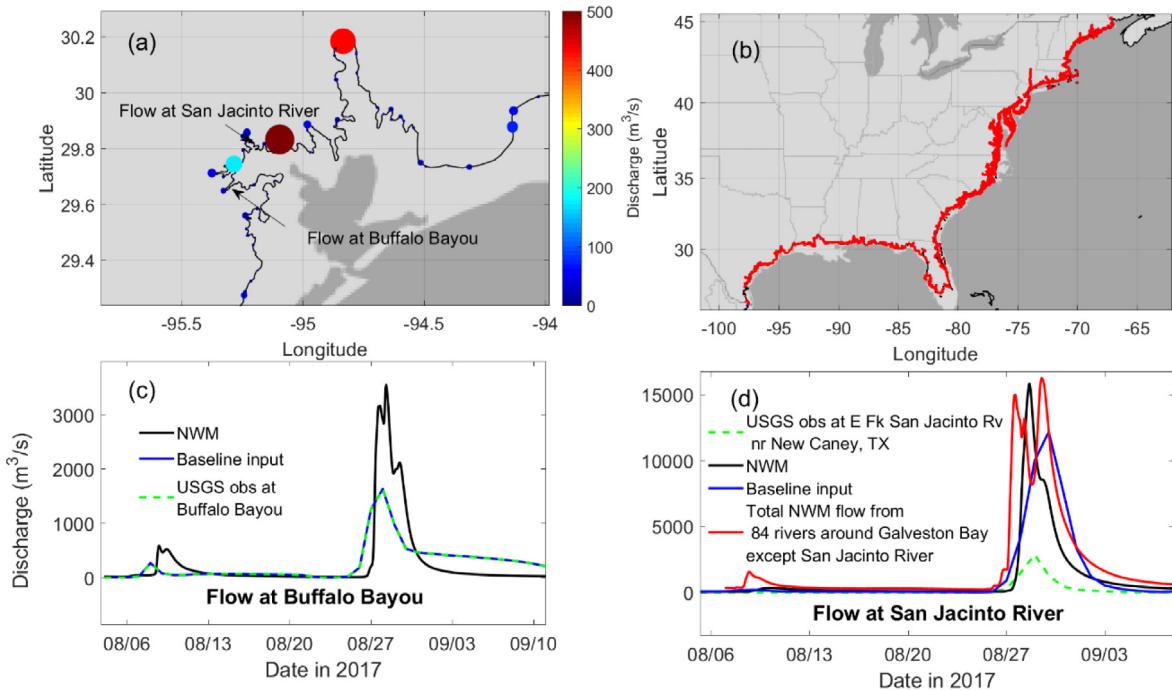


Fig. 4. (a) Averaged flows from NWM at the SCHISM land boundary in the Galveston Bay region, with the size and color of each circle representing the magnitude of flow. (b) Intersections between SCHISM boundary and NWM segments where river flows are injected, along the entire land boundary. (c) Comparison of flow at B.B. from NWM and USGS gauges (sum of White Oak 08074500 and B.B. 08074000 gauges, as used in baseline model setup, so the green and blue lines are identical). (d) Comparison of total flow from all rivers discharging into Galveston Bay except SJR (red line, from NWM) and the SJR flow, as predicted by NWM (black line). In addition, the SJR flow used in the baseline setup, which is derived from a flow-stage relationship from TWDB (Dale Crockett, private communication) is shown in blue, and the measured flow at the upstream USGS gauge (gauge 08070200 in E Fork San Jacinto River near New Caney, TX) is shown in green. The above-dam flow at 08070200 is significantly lower than the discharge from the dam (i.e., SJR flow); the latter is unknown and is approximated by the blue or black line.

bottom characteristics information may be done for specific systems. The turbulence closure scheme used Generic Length Scale scheme of $k\cdot k l$ (Umlauf and Burchard, 2003) but several other schemes yielded similar results. Observation at NDBC buoy 42019, located ~120 km southwest of the Bay in the deep ocean, indicated a relatively modest sea state with maximum wave height of ~6 m (vs. ~10 m during Hurricane Ike in 2008). The wave height decreased quickly toward the shoreline and at the buoy nearest to the Bay (42035), the maximum wave height was only ~3.2 m. Results from a coupled wave-current simulation (not shown) suggested a modest wave-induced setup of up to 30 cm in some parts of the barrier islands only. The dominant wave direction was from the south, which was not conducive for the wave energy to penetrate into the Bay through the narrow NW-SE oriented entrance. More importantly, the large freshwater outflow from the watershed, which dominates the maximum elevations, occurred a few days after the storm has made the landfall and so we found there is little difference in the predicted maximum elevation inside the Bay (i.e. the wave induced setups do not enhance the positive bias as shown in Section 3). Furthermore, no wave-induced overtopping flow across the barrier islands in the Galveston Bay was observed during the event (Anarde et al., 2020). Therefore, we will not discuss further the wave effects in this paper.

The baseline and important sensitivity runs presented in this paper are described in Table 1. The baseline is a 3D baroclinic run with river flows, oceanic and atmospheric forcing and direct precipitation onto the model domain. For river flows, the B.B. flow was from USGS gauges and SJR flow was derived from a flow-stage relationship from TWDB as discussed before. The NWM results were used for all other river flows (which will likely introduce biases also). Three sensitivity runs were designed to separately assess three major forcing factors: ocean, river flows, and direct precipitation, and a detailed description for these runs is given in Section 4. Another experiment was done with all river flows (including B.B. and SJR) using NWM values in order to demonstrate the

Table 1

The baseline and major sensitivity runs. All runs use same atmospheric forcing as explained in Section 2.3.

ID	Run name	Setup
1	Baseline	3D baroclinic run with river flows, oceanic forcing and direct precipitation; flow at Buffalo Bayou from USGS; flow at San Jacinto River derived from TWDB relationship. The cut-off depth for air-sea heat exchange is 0.1 mm
2	Ocean only	Same as baseline but only with oceanic and atmospheric forcing (but without precipitation effects)
3	Rivers only	Same as baseline but only with river flows
4	Precipitation only	Same as baseline but only with direct precipitation
5	Base_NWM flows	Same as baseline but with the NWM results for all river flows at land boundary
6	Base_hmin_airsea_ex = 1 mm	Same as baseline but with cut-off depth of 1 mm for air-sea heat exchange

sensitivity of surges to different estimates of river flows. The last sensitivity run in Table 1 was used to assess the impact of the threshold used in the air-sea heat exchange on temperature results (cf. Section 3.7). All simulation runs achieved performance of 49 times faster than real time on 800 cores of Sciclone cluster (<https://www.wm.edu/offices/it/services/researchcomputing/accreq/index.php>, last accessed in August 2020), or 80 times faster than real time on 2000 cores of NASA's Pleiades.

Another important caveat is that no groundwater runoff or infiltration is explicitly included in SCHISM (although NWM has incorporated some of these effects). A future task is to assess the 'local runoffs' provided by NWM, although these are expected to have a minor impact

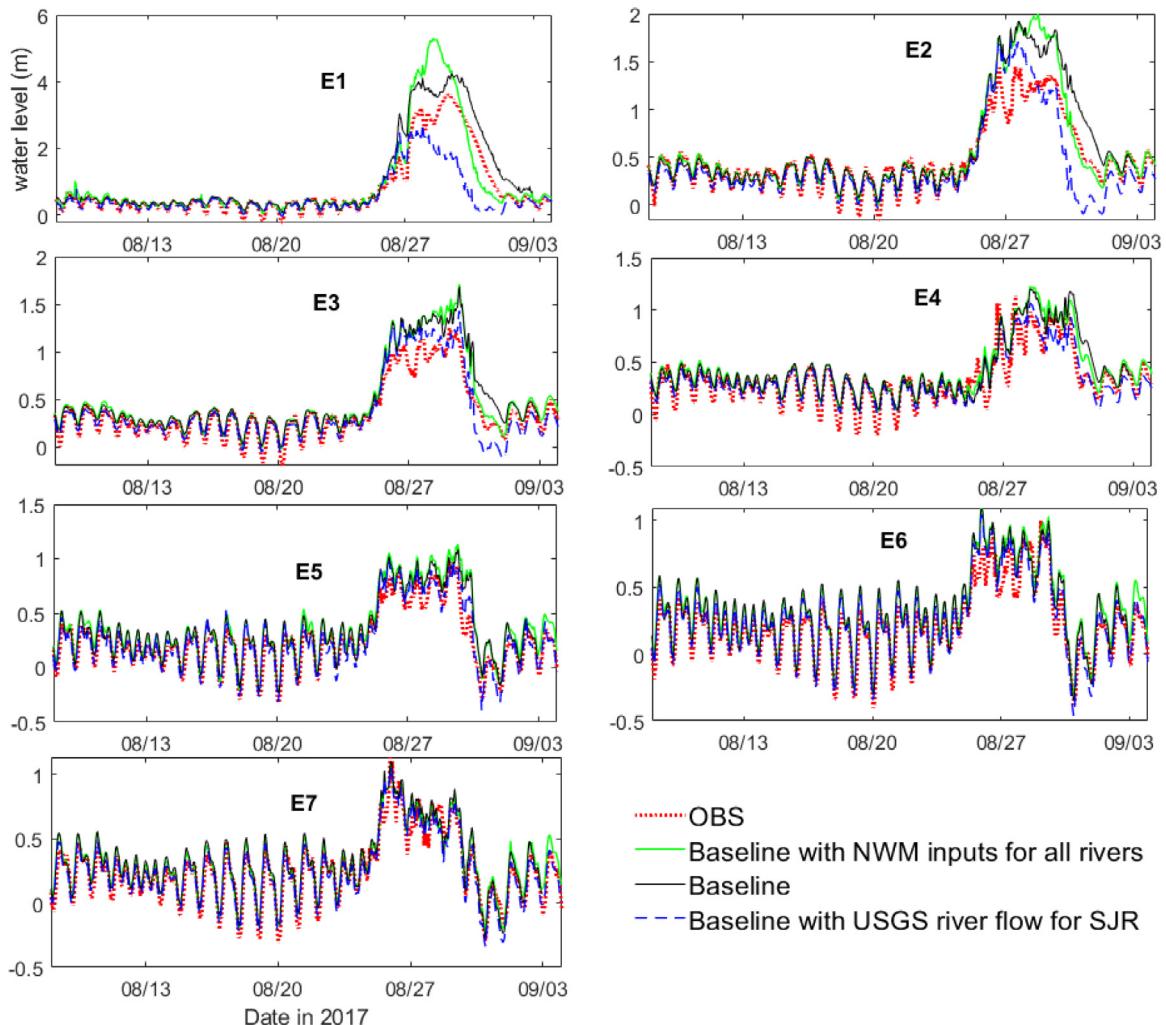


Fig. 5. Comparison of elevations (m, NAVD88) at 7 NOAA stations (E1–E7, generally from upstream to downstream; see Fig. 1b for their locations). The three SCHISM configurations are identical except for the river flows used: black line (baseline run), green line (sensitivity run #5 in Table 1) and blue-dashed line (same as the baseline run except using the USGS flow at the above-dam gauge 08070200). ‘OBS’ is data from NOAA.

for the maximum surges during Harvey. Note that the NWM domain partially overlaps our model domain, thus allowing us to inter-compare some flow results between the two models (cf. Section 3.5).

3. Model validation and discussion

Before applying the current model to specific events like Harvey, we had first validated the model for tides and sub-tides for the entire east coast and Gulf of Mexico; some results in the Galveston Bay area can be seen below (before the storm). A comprehensive presentation of the tidal results for other bays and estuaries is out of scope of this paper and will be presented elsewhere. However, to set the stage for the discussion of total elevation in the next subsection, we remark first that the averaged complex error for the most dominant frequency (K1), calculated as

$$C_e = \sqrt{(A_o \cos \phi_o - A_m \cos \phi_m)^2 + (A_o \sin \phi_o - A_m \sin \phi_m)^2} \quad (1)$$

for a non-storm period (Aug 9 to Sep 9, 2011) for the same 7 stations is only 3.7 cm, which serves as the lower bound for the model errors for all periods. In Eq. (1), A_o and A_m are amplitudes from observation and model respectively, and ϕ_o and ϕ_m are phases respectively.

3.1. Compound surges

The model-calculated water levels are compared with the NOAA observations (Fig. 5). Overall, the baseline model results show a good

agreement with the observations with an averaged MAE of 0.13 m and cross correlation of 0.94. Comparison of MAE and the complex error shown above suggests that the total error is dominated by the storm signals, not tidal signals. Note that the 3D model is able to largely capture both the tidal and sub-tidal variations (with reference to NAVD88). More importantly, the model captured the high surges that were sustained for ~5 days at all stations. The maximum surge is ~1 m in most part of the main Bay, but the compound effects have led to a 3.5 m surge at the B.B. station E1 (Valle-Levinson et al., 2020). A later and higher second surge occurred at E1 only, which is mostly due to the compound river effect (and to a lesser extent, precipitation effect; cf. Section 4).

Relatively large errors are found at upstream stations (E1–E3), and are attributed to the uncertainty in the river flows, as demonstrated by the comparison among model results driven by different flows at B.B. and SJR. In particular, the larger flows from NWM (Fig. 4) have led to higher and overestimated surges at E1–E3 (green lines in Fig. 5) but the effect of the timing differences in the flow peaks between NWM and baseline (Fig. 4d) is also clearly seen at these stations: the earlier decline in the NWM flow has produced a better phase match with the observation for the retreating wave around August 31, while the baseline results showed a delayed response in the retreat (Fig. 5). On the other hand, using the upstream flow for SJR (gauge 08070200), which almost certainly underestimates the true flow, has led to dramatically underestimated and short-lived surges especially at

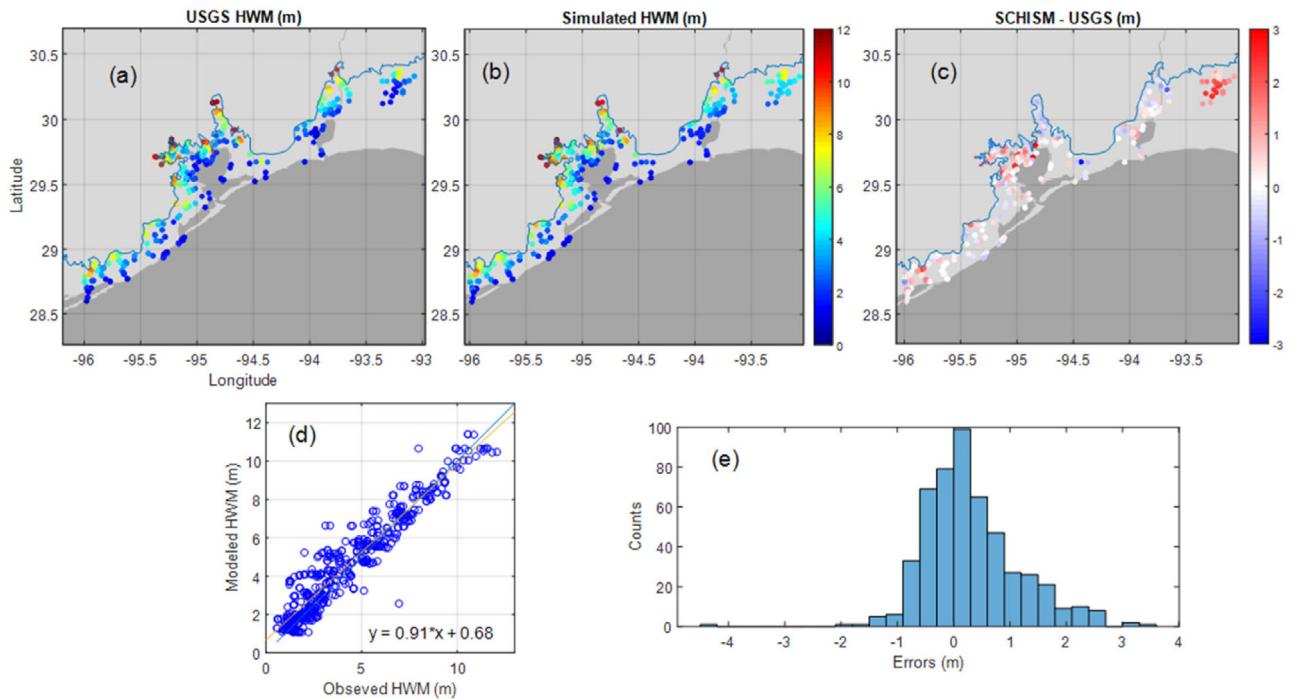


Fig. 6. Comparisons of (a) observed and (b) simulated HWMs with (c) the model errors (model – data), (d) the linear regression and (e) the histogram of errors (note the positive bias). Note in (a) that the observation locations are mostly in the watershed, where there are larger uncertainties in the vertical datums. The errors in the predicted river flow also contribute to the HWM errors. The border between the gray and dark gray regions corresponds to the shoreline.

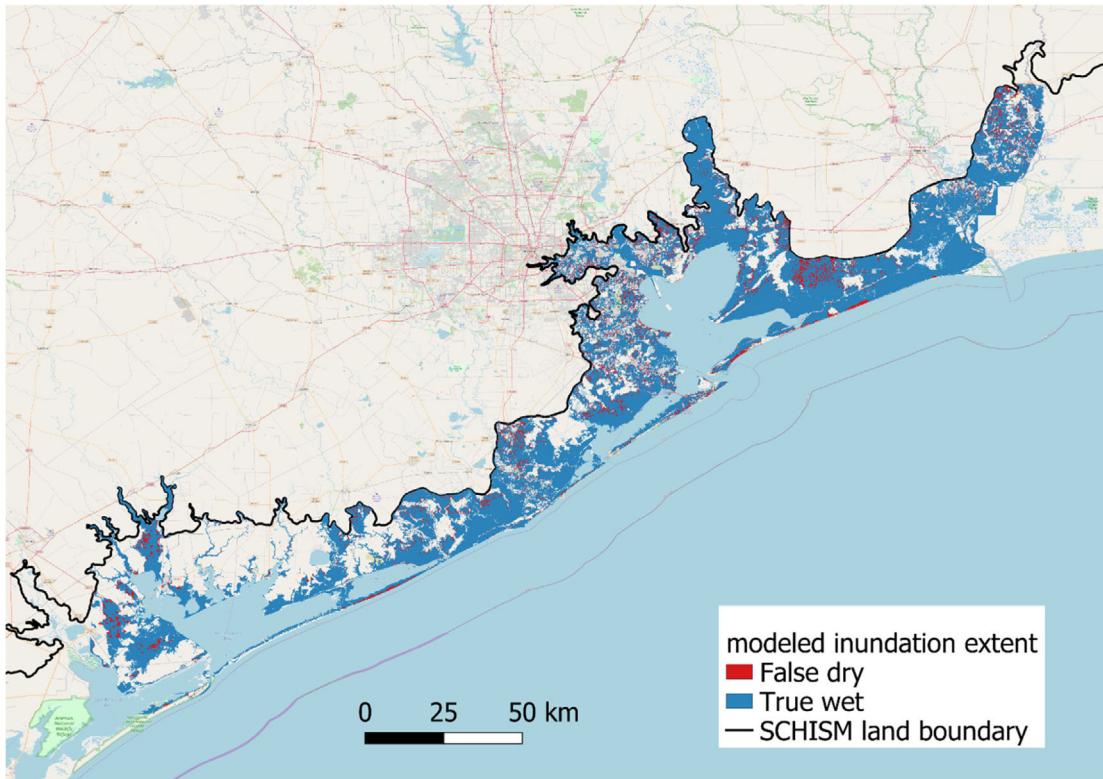


Fig. 7. Comparison of inundation extent between the baseline run and FEMA estimates, with 10 cm used as the inundation threshold in the post-processing of the model results, as typically used for nuisance flooding.

E1 (blue dashed lines in Fig. 5). Note that E2 and E3 are also influenced by other nearby rivers whose flows are also uncertain. The comparison serves as a direct evidence that the peak outflow from SJR that occurred about a day later than the B.B. peak flow is the main reason for the

later and higher surge at the B.B. station E1 on August 30, as SJR flow was blocking the outflow from B.B. (Valle-Levinson et al., 2020). The results suggest that the river flows account for the largest uncertainties for the simulated surges upstream and better estimates of SJR and other

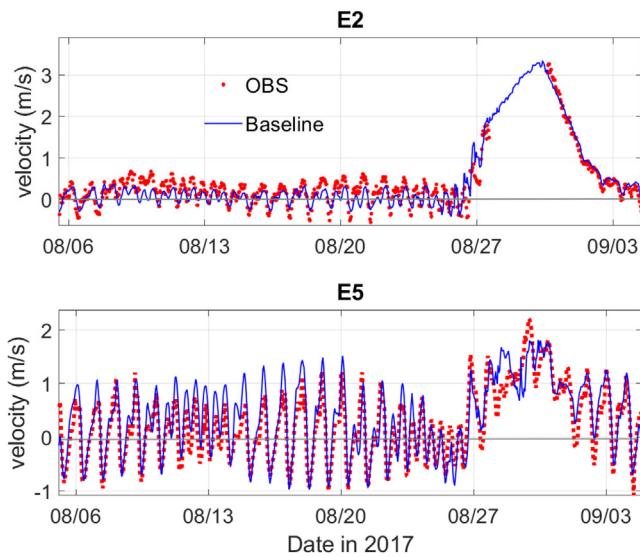


Fig. 8. Comparison of along-channel velocity at g08010 (at 3.8 m below surface; E2) and g06010 (at 6 m below surface; E5). Positive values denote seaward flow, and the thick horizontal lines correspond to zero velocity. At these two stations, the flow is highly channelized so the cross-channel velocity is much smaller.

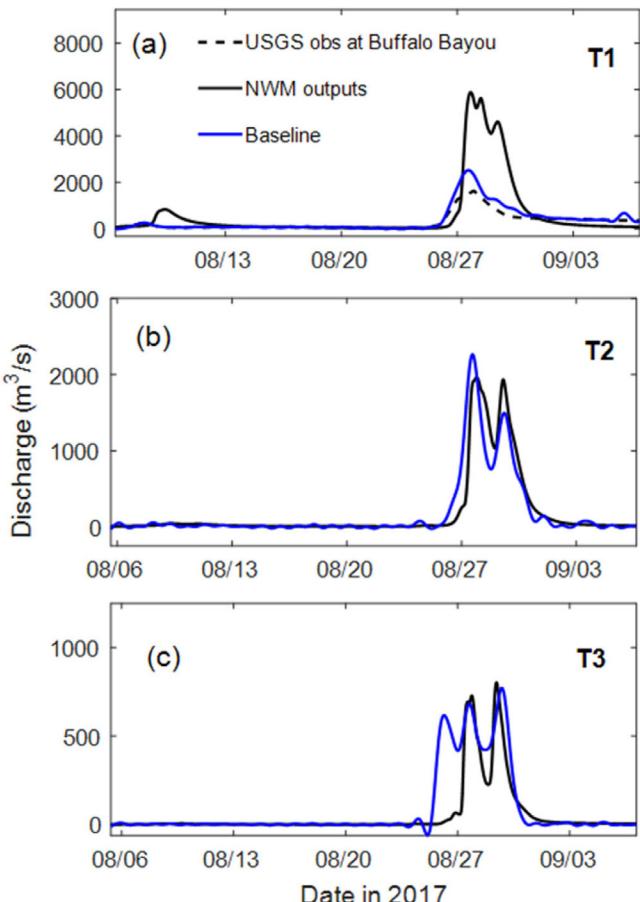


Fig. 9. Comparison of flow rates across 3 transects (see Fig. 1b for their locations) between the baseline results and the NWM outputs inside our domain (which are different from the NWM results imposed at our land boundary). The USGS data are available for the transect T1 only.

river flows should further improve the model results. Unfortunately, there is no direct measurement on the SJR outflow from the Lake Houston dam. Empirical functions have been used by TWDB to estimate the outflow based on monitored water level at Lake Houston (Dale Crockett, private communication). The function presumably only works under normal condition when water level does not exceed the dam's maximum capacity. It is doubtful the empirical function is applicable during Harvey as water was spilled over the dam during the peak flow. The overall freshwater input to Galveston Bay during Harvey has been estimated by Du et al. (2019a,b) and Thyng et al. (2020), with the values ranging widely from 12 to 22 billion cubic meters. Considering this large range of uncertainty and the fact that the total freshwater inflow into the Bay is ~3.7 times the Bay volume, the model error of ~0.5 m in the predicted maximum surges at the upstream stations E1–E3 is quite reasonable. While we acknowledge these uncertainties, we deem that the analysis below about the compounding factors for the flooding will not be greatly affected by these uncertainties, as we focus more on the difference between sensitivity runs and the baseline. On the other hand, the compound surges at other stations (including the sheltered stations E4, E7 behind the barrier islands) are all well simulated.

3.2. High water marks

HWMs represent the highest water level reached during the storm event at a specific location. Simulated HWMs are compared with ~510 USGS field estimates collected during post disaster survey in the region, including Galveston Bay, Sabin Lake, and Calcasieu Lake (Fig. 6). The vertical datum used in the HWMs is NAVD88 but the quality of the local DEMs varies, which partly contributes to the model errors. The overall MAE is 0.65 m with a positive bias of 0.32 m, indicating a general overestimation by the model. The hot spots of relatively large errors occur in the Lake Charles region, north of Calcasieu Lake (Fig. 6c), which is likely caused by the larger uncertainty in the DEMs used there. The MAE would be reduced to 0.58 m if this region were excluded. The combination of the smaller errors in most of the Galveston Bay region (Fig. 6c), expected larger model errors on high ground (24% of the bottom elevations at those points are above 5 m), and the high correlation between the observation and model (Fig. 6d; with the correlation coefficient being 0.96) indicates an overall satisfactory performance of our model. The MAE in this study appears to be lower than that in Wing et al. (2019) although they included many observation points outside our domain.

3.3. Inundation spatial extent

The HWMs serve as a more rigorous validation of modeled water levels and (given accurate DEMs) flood depths, but shed little light on the spatial limit of inundation during the storm event. The latter was estimated by FEMA using HWMs from Harris County Flood Control District, USGS surveys, and other inspection data. Elevation data comprised a mosaic of 3-m resampled elevations from 1 and 3-m LiDAR and IFSAR data with one erroneous IFSAR data replaced with a blended 10-m section; note that this estimation inherently contains errors and uncertainties as described by Watson et al. (2018). Therefore, in assessing the modeled inundation, we will only assess the extent to which the model captured the overall spatial pattern of the flooding by comparing the Hit Rate (H_R), defined as:

$$H_R = \frac{M_1}{B_1} \quad (2)$$

where M_1 indicates the total area of 'observed' wet cells correctly captured by the model ('true wet') and B_1 represents the total area of wet cells from the field estimates ('observation'). The H_R , therefore, can be interpreted as the proportion of 'observed' flooded areas that is successfully reproduced by the model (Wing et al., 2019). It ranges from 0 to 1 with 0 indicating no match between model and 'observation' and

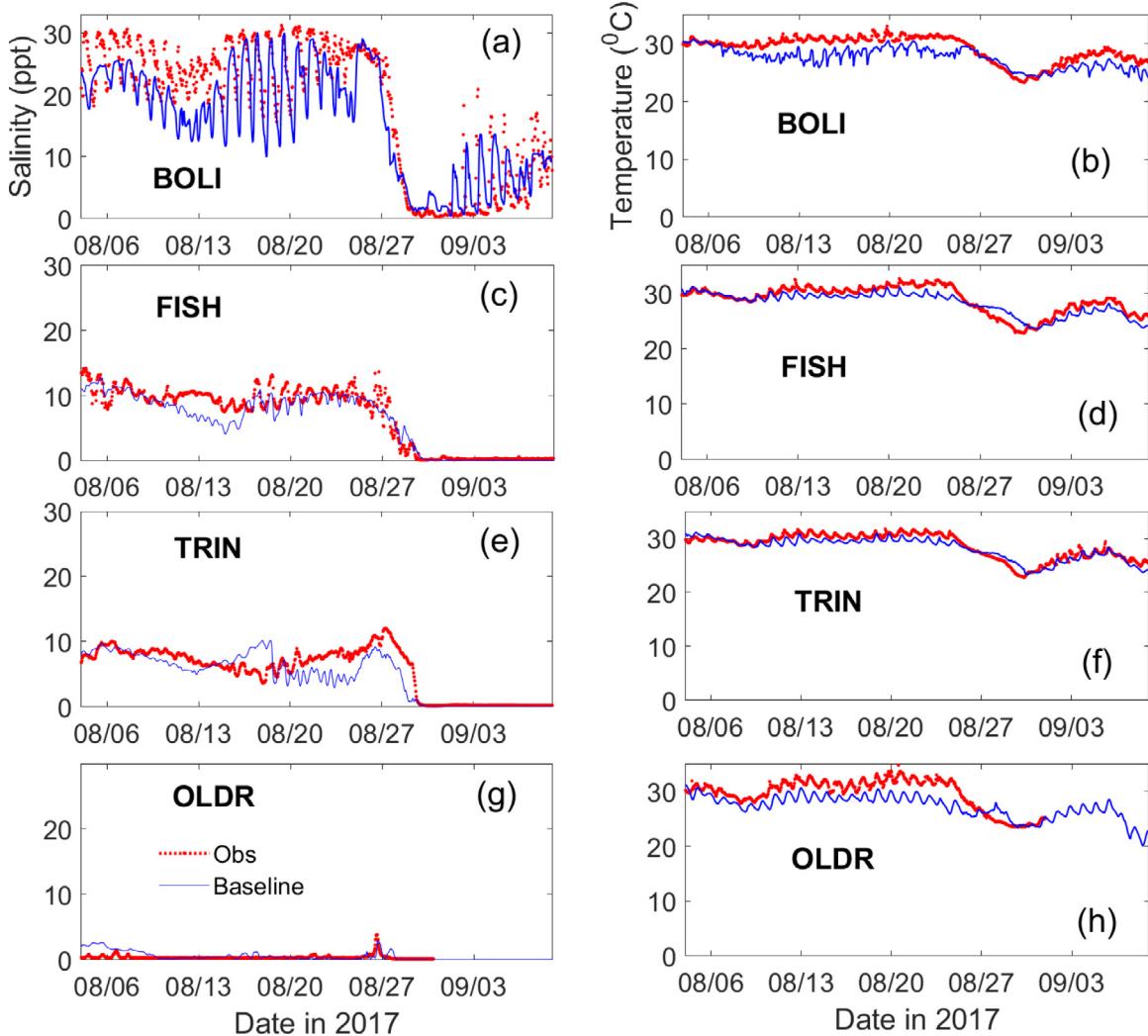


Fig. 10. Comparison of mid-depth (1.5 m below surface for BOLI and FISH, 0.5 m below surface for TRIN, and 3 m below surface for OLDR) salinity (left panels) and temperature (right panels) at 4 stations (generally from the bay entrance to upstream).

1 indicating a perfect match. It penalizes ‘false dry’ errors and is thus a measure of the model’s tendency to underpredict the flood extent (Wing et al., 2019). Note that the ‘observation’ here contains no dry spots so we cannot assess the model error for ‘false wet’.

Comparison of the inundation extent from model and ‘observation’ indicates that the model mostly reproduces the ‘observed’ flooded areas, with ‘false dry’ areas found in the watershed of Matagorda Bay, Brazoria River (west of Galveston Bay) and north of Anahuac National Wildlife Refuge in the Galveston Bay (Fig. 7). The overall hit rate, however, is very high (0.92), indicating that the model is capable of reproducing the observed flooding pattern.

3.4. Velocity

Having thoroughly validated the model for water level, we now turn our attention to 3D variables in the next subsections. As stated in Introduction, comparing to 2D models, 3D models like SCHISM have the advantage of being able to provide better sub-tidal elevations and more complete 3D information about the event. As shown in Du and Park (2019), some of the 3D information is critical for our understanding of how the ecosystem is altered by the storm beyond inundation. Observation at two TABS stations in Galveston Bay offers rare insight into the sheer magnitude of the ‘punches’ from the river flood (Valle-Levinson et al., 2020): station g08010 (Fred Hartman Bridge) at the upper shipping channel downstream of the confluence of SJR and B.B.

and station g06010 at the bay entrance (E2 and E5, respectively, in Fig. 1b). The observed peak outflow exceeded 3 m/s at g08010 despite data gap and 2 m/s at the bay entrance station g06010 (Fig. 8). The upper bay velocity data showed no inflows over ~8 days between August 26 and September 3 and no tidal fluctuations between August 28 and September 2. At the bay entrance station, there were tidal fluctuations but no inflows for ~5 days. All of these observed features are successfully captured by the model (Fig. 8). In addition, there is also occasionally large shear (and stratification) at the deep station E5 particularly during the large outflow (not shown), which obviously cannot be captured by a 2D model.

3.5. Watershed flows

NWM provides flow outputs for part of the coastal watershed within our model domain, which can be qualitatively compared to SCHISM flow results across the river channels (note that NWM is essentially a 1D model while SCHISM is 3D). As NWM and SCHISM use very different flow routing algorithms (Muskingum-Cunge channel routing for NWM and shallow-water equations for SCHISM), a qualitative agreement between the two modeled flows would be reassuring. Since NWM results do not yet include tidal signal, a 30-hour low-pass filter was applied to the SCHISM results before the comparison.

Comparison of the flow rates at T1 (B.B. near Manchester), at which the USGS observed flow data are also available, indicates that the

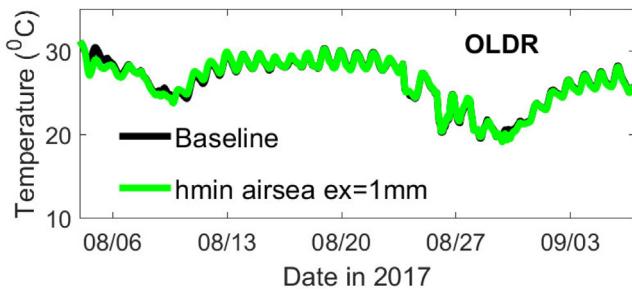


Fig. 11. Comparison of temperature at 3 m below surface from a sensitivity test #6 (Table 1) and baseline run.

SCHISM computed flow reproduces the USGS observation well with some overestimation (Fig. 9a), likely due to precipitation added inside our domain. On the other hand, NWM output at this transect is more than twice of the observed.

Closer to the main Bay into the tidal environment, we found a general agreement between NWM and SCHISM flows at T2 (Clear Lake) and T3 (Dickinson Bayou) (Fig. 9b,c); the differences are attributed to the different flow routing algorithms used. As shown in Section 4.2,

pluvial effect at T3 may also play a role in making the difference. It is reassuring that the SCHISM and NWM flows generally agree with each other at these two interior locations.

3.6. Salinity and temperature

Salinity and temperature are important environmental variables for both large- and small-scale processes. Ye et al. (2020) examined the relationship between the large-scale Gulf Stream and compound flooding inside Delaware Bay during Hurricane Irene and found that the prevailing surface slope induced by the Gulf Stream played an important role in ‘blocking’ the outflow during the post-storm water level re-adjustment. Du and Park (2019) used SCHISM to study the recovery time for bay salinity after Harvey and found that it took an average of two months for the salinity over the entire bay to recover.

The model validation for the large scale is shown in Appendix, and here we focus on the response in the Bay. During Harvey, the large freshwater load (~3.7 times the Bay volume) has led to freshening of the bay (Du and Park, 2019) and created large salinity gradients afterward (Fig. 10). The enormous pulse of freshwater drastically transformed the bay salinity condition, dropping the salinity from 30 psu to 0 within ~3 days even at the bay entrance and leaving the whole Bay fresh for ~2 days. Starting from around September 1, the ocean water

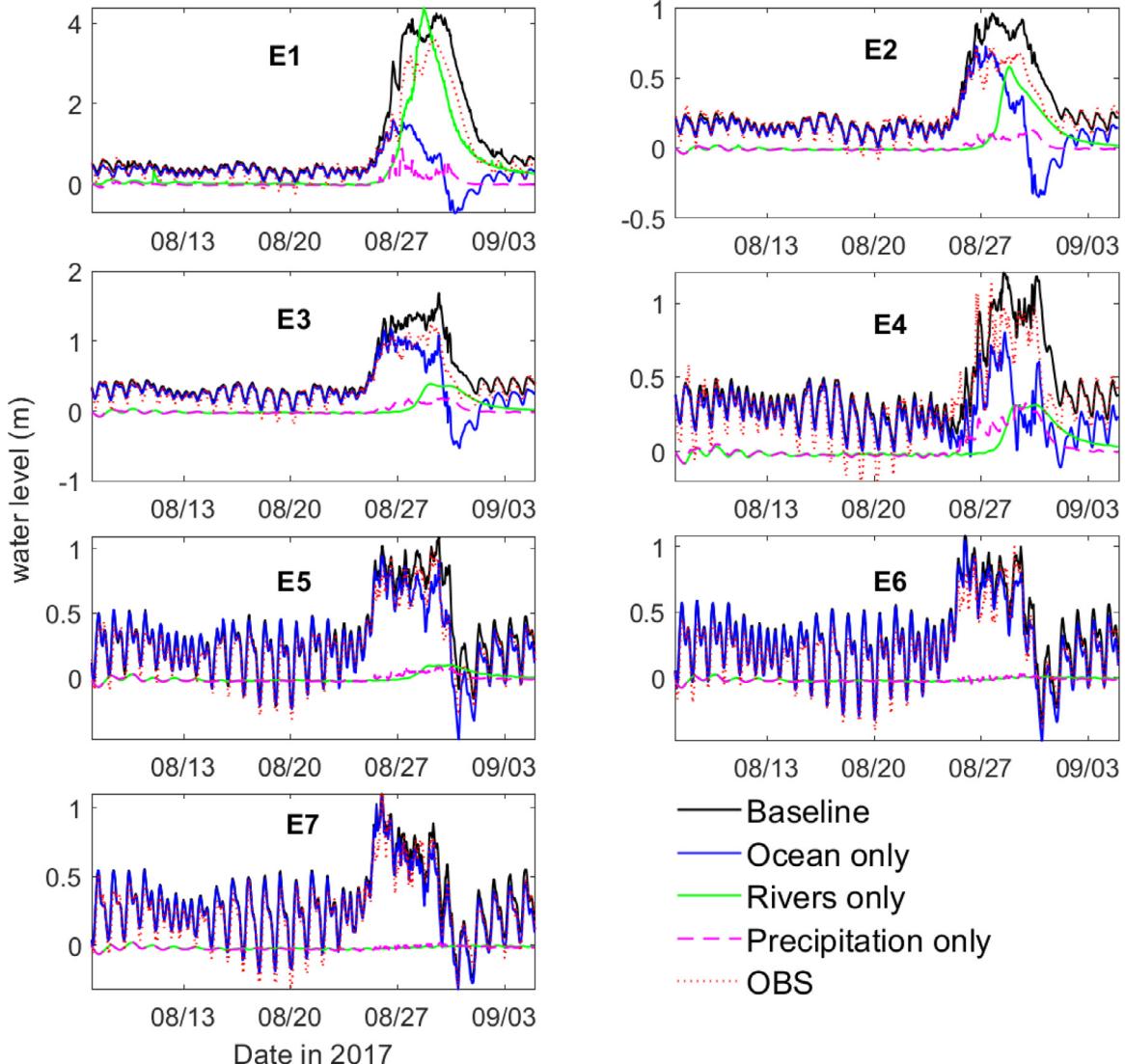


Fig. 12. Individual effects from three factors (ocean, river flows and precipitation) on water level as compared to the baseline results at the 7 NOAA stations.

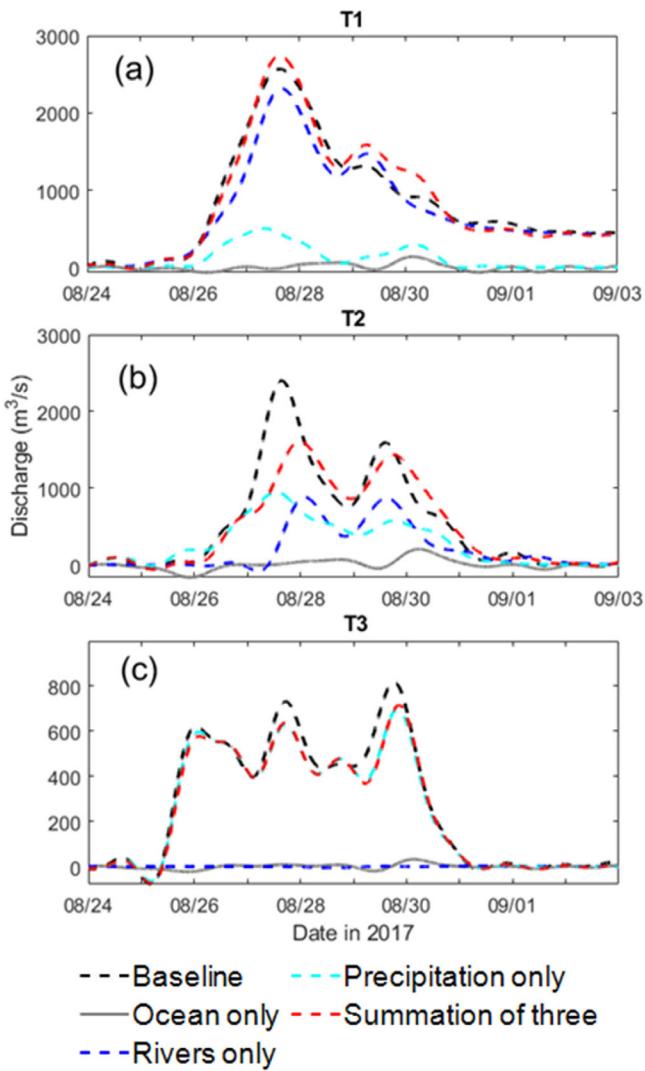


Fig. 13. Comparison of flows across 3 transects of T1–T3 (see Fig. 1b for their locations) between the baseline and three sensitivity runs.

started to re-intrude into the Bay (Fig. 10a) but the full restoration of salinity there did not happen until two months later (Du and Park, 2019). The Bay temperature, on the other hand, exhibits a synchronous drop of ~ 5 °C during the passage of the cold wake of the storm around August 30 (Fig. 10); the most upstream station (OLDR), just downstream of the Trinity River mouth, however, showed a slightly delayed response (Fig. 10h), likely due to different heat content of river water that was eventually mixed with the Bay water. The lack of information on the water temperature in the rivers makes it difficult to assess its impact.

The current model application did a good job in capturing the above-mentioned processes, with validation of temperature and salinity as summarized in Fig. 10. In particular, the dramatic drop and gradual recovery of the Bay salinity were captured in terms of both magnitude and timing. The synchronous drop in the Bay temperature was also captured albeit with small cold biases. The largest error occurred at the most upstream station (OLDR), where the model at times showed a ~ 2 °C cold bias. This bias was attributed to the uncertainties in the water temperature boundary conditions, specifically the unknown temperature in the river flows injected at the model land boundary and to a lesser extent, the unknown temperature of the rainwater. Both are assumed to be at the ‘ambient’ model temperature, which is not a good assumption. Overall, the averaged MAEs for temperature

and salinity are 1.4 °C and 1.7 psu, respectively; normalized by the mean observation, the scaled MAEs are respectively 0.06 and 0.22 . Addressing the uncertainties in the watershed temperature should be a priority in the future.

3.7. Sensitivity of temperature to threshold depth for air-sea exchange

The temperature skill was also found to be sensitive to the heat fluxes calculated from the air-sea exchange model used in SCHISM (Zeng et al., 1998). This is because large areas of watershed are covered with thin layer of water (e.g. from rainfall initially), where small errors in the calculated heat fluxes would lead to large temperature swings due to the small water mass there. Therefore, the exchange is best to be deactivated when the water depth falls below a threshold. To demonstrate that the choice of the threshold used in the baseline is robust, an additional simulation was conducted to test the sensitivity of modeled temperature to this choice (sensitivity run #6 in Table 1). Results suggest that thresholds lower than 1 mm would have a relatively minor influence on the calculated temperature (Fig. 11), and therefore the choice of 0.1 mm in the baseline run is acceptable.

4. Compound effects of oceanic, fluvial and pluvial origins

In this section, we will assess the individual contributions from three forcing factors: ocean, river and rainfall. Three sensitivity runs were conducted using only one factor in each run (sensitivity runs #2–#4 in Table 1). Specifically, ‘ocean only’ simulation includes all processes of oceanic origin (tides, Gulf Stream, atmospheric forcing, etc.), but excludes the effects from river flows injected at the land boundary and precipitation. ‘Rivers only’ simulation turns off all oceanic and atmospheric forcing (‘ocean as lake’) and precipitation, with the only forcing coming from river flows (at our land boundary). Similarly, the only forcing in ‘precipitation only’ simulation is the precipitation. Note that the ‘rivers only’ simulation indirectly includes precipitation that occurred *outside* the model domain, which contributed to the river flows injected at the land boundary, but not the direct precipitation *inside* the model domain. Strictly speaking, turning on/off precipitation or rivers affects the Bay salinity, but this effect is deemed minor for surges. Therefore, the three factors are mostly orthogonal to each other and it is a common practice in many other models (as reviewed by Santiago-Collazo et al., 2019) to use the sum to represent the compound flood by implicitly assuming linearity. Comparison of this sum to the current baseline model results thus reveals the nonlinear compound effects.

4.1. Impact of compounding factors on the total water level

A calibrated model like ours allows us to assess the contributions from each of the three forcing factors on the total water level (Fig. 12). The ocean-induced storm surge is on the order of 1 m at all stations except at E4, which is located in a small sheltered embayment on the eastern end of East Bay (Rollover Bay) behind Bolivar Peninsula, where the surge is slightly lower. Note that there is a small man-made inlet near E4 connecting the Rollover Bay to the Gulf, which sometimes gets clogged. Hence, the water level at E4 tends to be more erratic. All of the ocean-induced surges reached their peaks on Aug. 26 followed by a steady decline afterward, and all showed a negative set-down around Aug. 31 as the water withdrew from the Bay and tidal oscillations were restored, which is typical in all storm surges. However, the more sheltered station E4 showed more complex processes of wave reflection and water stacking. Comparison between the ‘ocean only’ results and the observation indicates that the ocean alone can adequately explain the surges at the coastal stations (E5–E7), and also the initial surges at the upstream stations E1–E3, but the ocean-induced surges there were not sustained as long as the observation; in addition, part of the initial surge at E1 is not of ocean origin (Fig. 12).

The fluvial effect is very significant at E1–E4; not surprisingly, the effect increases toward upstream rivers and watershed (Fig. 12). At E1,

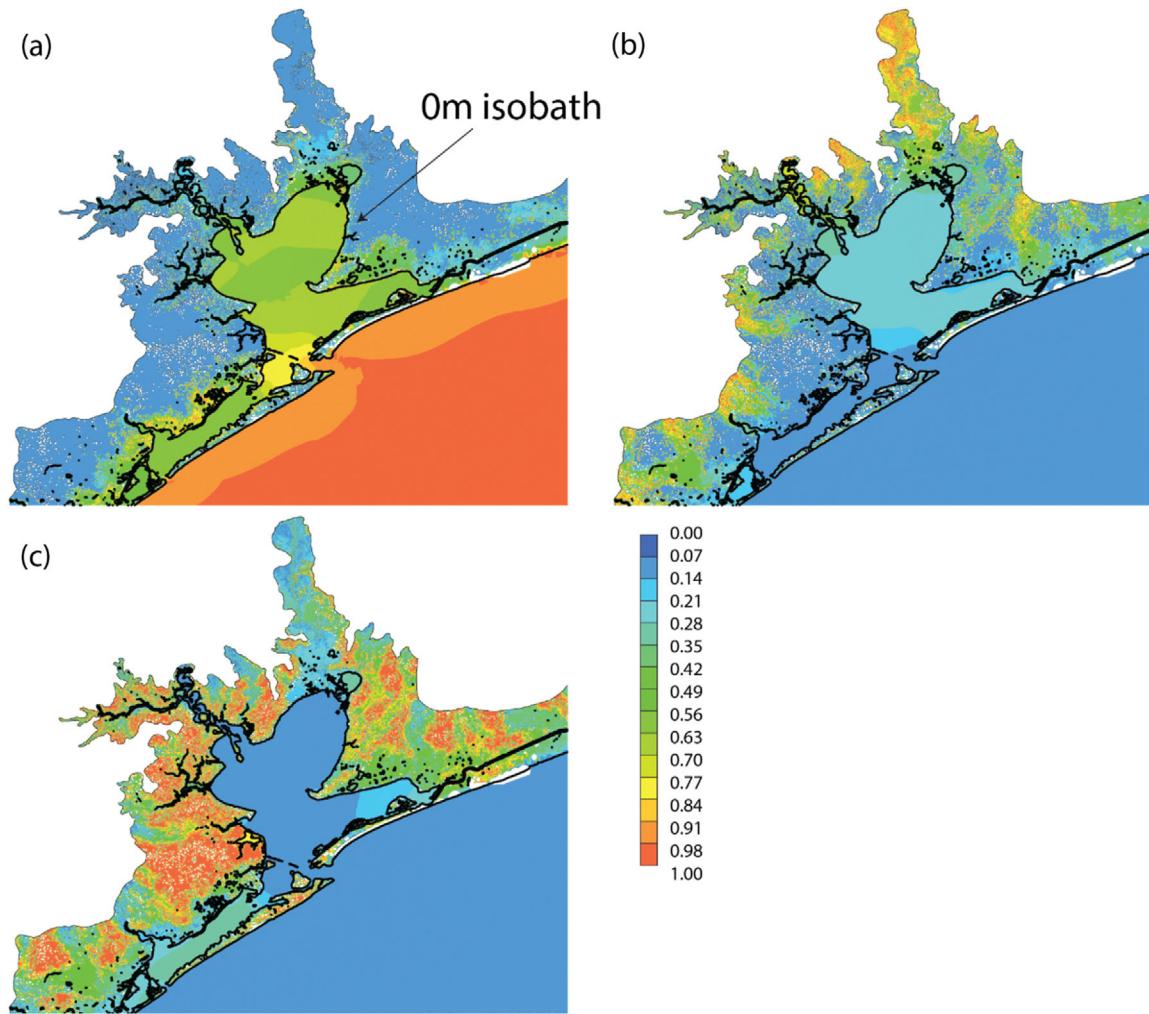


Fig. 14. Ratios of maximum disturbances to the total from each of the three factors (neglecting compound effects in the total disturbance) of (a) ocean only, (b) rivers only, and (c) precipitation only. A cut-off depth of 10 cm is used to exclude regions with low inundation (the white space). The black lines are 0 m isobath. Any values less than 1 (i.e. non-red regions) indicate locations with potentially significant compounding effects.

this effect dwarfs the other two factors. Even at the entrance station E5 the fluvial effect is still visible. The fluvial peaks lag the storm surge by ~ 2 days. The pluvial effect is smaller than the other two factors at all stations except E4, where local stacking of water seems to be severe (possibly due to the very small inlet there). Still, the pluvial effect accounts for up to 30 cm of surge at the upstream stations E1–E4, and tends to be synergistic with the fluvial forcing in raising the surge level (Fig. 12).

In summary, E1 is heavily dominated by river flows and all three factors are important at E1–E4. The downstream stations E5–E7 are heavily dominated by the ocean factor. The three factors work in tandem to sustain the surge for over 5 days due to their different arrival times: as the ocean surge is retreating the arrival of river and rainfall induced flood is crucial in re-raising the water level (Fig. 12). The compound effects are most severe at upstream stations. However, the factors do not always reinforce each other; a prominent example is seen around August 29 at E1 when the river alone would have led to a higher surge than the baseline, but the latter is partially offset by the withdrawing flow (decreasing water level) from the ocean source during this period. Therefore, the time history of different forcing factors is very important for compound flooding processes.

4.2. Impact of compounding factors on flow

Like the water levels, the flow in the Bay and watershed is expected to be dominated by different forcings at different locations: river and/or

precipitation may dominate watershed locations while oceanic forcing is expected to be important for Bay locations. This hypothesis is tested here using the 3 transects (cf. Fig. 9). The oceanic effects turn out to be relatively minor at all 3 transects when compared with the other two factors (Fig. 13). Both river and precipitation effects are significant at T1 and T2. At T3, the river factor is minor and most of the flow there came from local rainfall. At all locations, the sum of the flows calculated from the three factors is close to the baseline results albeit with some discrepancies (in the form of over-estimation at T1 and under-estimation at T2 and T3), which indicates nonlinear compound effects for flow. It is likely that the situation at T1 is more complex than the other two transects because it is located near the two largest rivers and additional blocking of the outflow from B.B. by SJR provided by the inclusion of precipitation effects in the baseline might have reduced the T1 flow as compared to the linear superposition. An important finding here is that imposing inflow boundary condition alone in a typical coastal watershed domain (while neglecting local precipitation) may considerably underestimate the resultant flow.

4.3. Contributions from each forcing in different regions

To better understand the spatial variability of the impacts of the three factors on water levels, we first introduce the concept of 'disturbance'. For compound flooding processes that involve both ocean and watershed, neither the water surface elevation nor the water depth is a

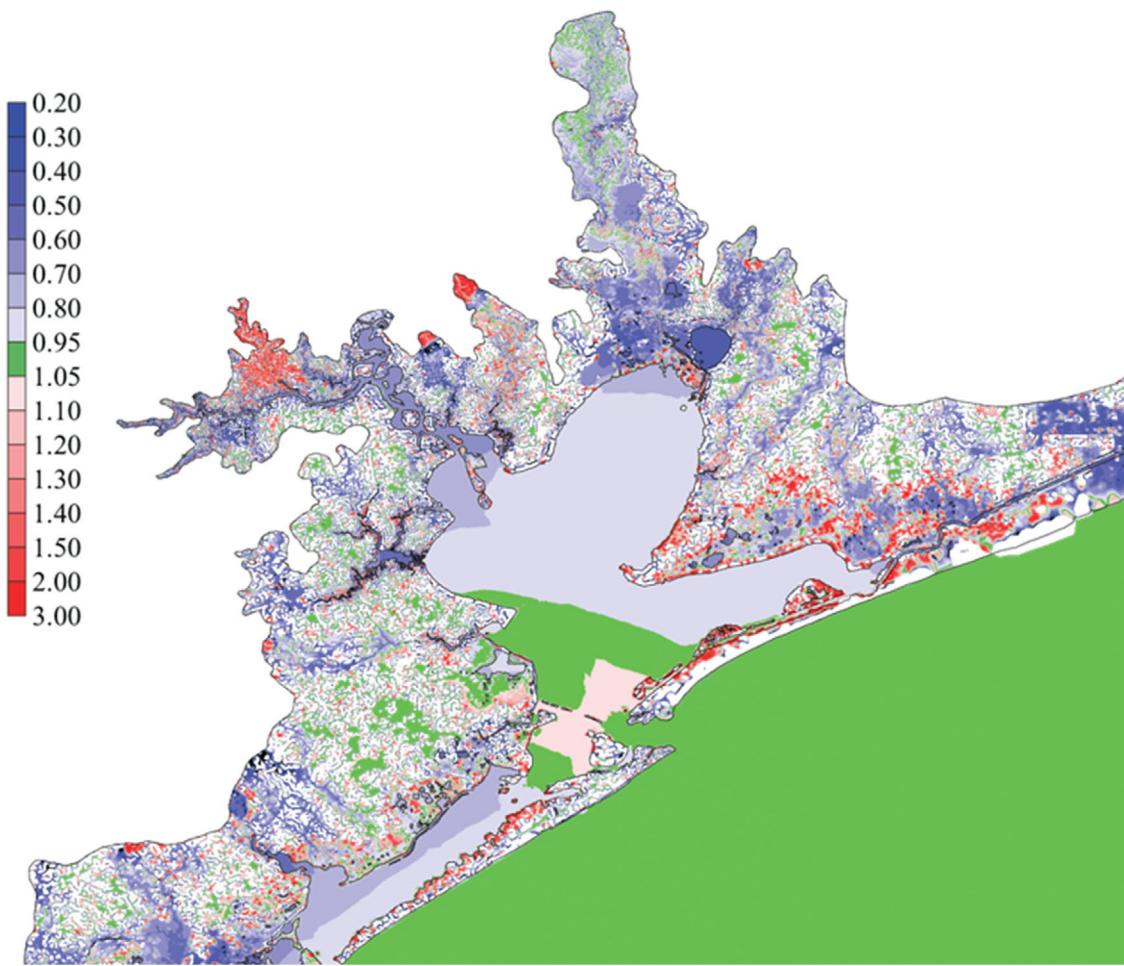


Fig. 15. Spatial distribution of the compound ratio. A cut-off depth of 10 cm is used to exclude regions with low inundation (white space). Any regions other than green indicate non-negligible compound effects.

satisfactory metric, because the nominally large water elevations on the high ground of watershed are dominated by the high bottom elevation there, and the large water depths in the bays and ocean are dominated by the local bathymetry. On the other hand, what we often need is the ‘departure’ of water level from the initial state. Therefore, here we define *disturbance* (D) as:

$$D = \begin{cases} \eta, & \text{if } h \geq 0 \\ \eta + h, & \text{if } h < 0 \end{cases} \quad (3)$$

where η is the water surface elevation and h is the bathymetry (with $h > 0$ for ocean, and $h < 0$ for watershed). Basically, D represents the departure from ‘initial condition’ (either the initial water surface or bottom, whichever is higher). Note that D is continuous across $h = 0$ and dependent on the vertical datum used. On the initially ‘dry’ ground of watershed D simply represents the local water depth (which is usually smaller than $|\eta|$), whereas at initially ‘wet’ locations D is simply the surface elevation. D is also a smoother metric to measure the compound effects as one transitions from oceanic into riverine regimes.

In the analysis here, we will first neglect the nonlinear compound effects and examine the contributions of each factor to the total disturbance, in the form of the ratio of the maximum disturbance from each factor to the total sum (so that the sum of the 3 ratios is equal to 1). Although the compound effects are ignored here, the ratios can still shed light on places where the compound effects may be potentially important (i.e. any non-red regions in Fig. 14). The oceanic factor dominates in the coastal ocean (>98%), but accounts for less than 80% of the total in the lower Bay, less than 60% in the upper Bay and low-lying rivers. The river factor dominates in the watershed

regions with existing river network, but also accounts for 30% and 10% in the upper and lower Bay, respectively. The direct precipitation is significant mostly on high ground in the watershed, but still accounts for ~10% in the Bay and more near the river network. These findings are consistent with the previous analysis for the impact of the three factors on water level (Section 4.1), which showed small influence from the ocean at the upstream stations E1–E3, where the pluvial and fluvial effects play significant roles. It is also in line with the conclusion on the impact of three factors on the flows (Fig. 13); in particular, T2 and T3 happened to be located inside a region heavily influenced by precipitation (Fig. 14c).

4.4. Compound ratio

The results shown in the previous section did not account for the nonlinear compound effects. The latter can be assessed by comparing the sum of the results from the three factors to the baseline. We define the compound ratio as the ratio between the maximum disturbance from the baseline and the sum of the three maximum disturbances at a given location:

$$C_r = \frac{\max(D)}{\sum_{j=1}^3 \max(D_j)} \quad (4)$$

The denominator is a conservative estimate of the total maximum disturbance because the maxima from each factor may occur at different times (cf. Fig. 12); therefore, the compound ratio calculated this way is generally no greater than 1. Compound ratios close to 1 indicate negligible compound effects, and any values far from 1

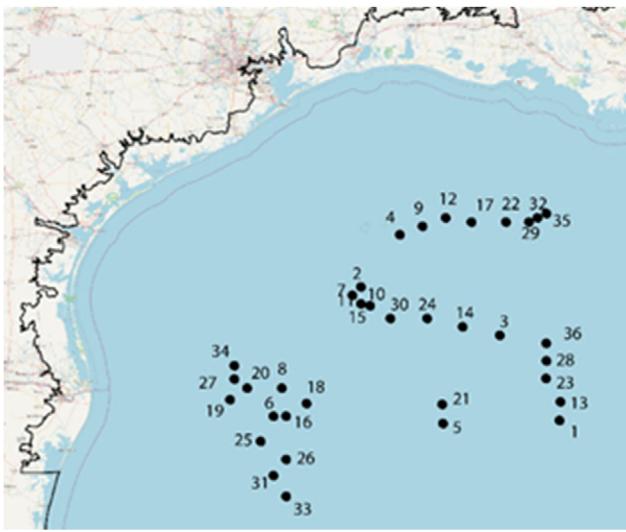


Fig. A.1. Locations of Argo profiles.

indicate significant nonlinear compound effects. As shown in Fig. 15, the compound effects are generally weak in the coastal ocean and portions of lower Bay (where the ocean factor dominates), and on the high ground in the watershed (where the precipitation is the sole source), but are significant elsewhere: in the East and West Bays, upper Bay and watershed areas with river network, etc. Near the maximum

inundation extent in the barrier islands, in the wildlife refuge north of East Bay and in parts of the northern watershed of B.B., the baseline maximum actually exceeds the sum of three factors, suggesting localized amplifications there; one possible explanation is that these regions correspond to where two factors are about equally important (Fig. 14). These results demonstrate that summing up the results from model runs forced by individual factors would overestimate in most regions by at least 20%, and thus a single model that incorporates all factors should be used to avoid the large biases.

5. Conclusions

Hurricane Harvey (2017) represents an extreme case for compound flooding due to record-breaking precipitation and ensuing river flood in the Texas and Louisiana coasts. The excess freshwater dumped into Galveston Bay also profoundly changed the physical and biological systems, with a prolonged recovery time. To fill in our knowledge gaps for this important event, we have successfully validated a 3D creek-to-ocean model and used the model to assess the contributions from all major forcing factors (ocean, rivers and precipitation). The major findings from this study are summarized as follows:

- (1) Coastal watersheds are considerably more complex than other types of watersheds with a myriad of competing factors from ocean, rivers, atmosphere (wind, heat fluxes and precipitation) and groundwater (not accounted for in this paper);
- (2) Compound effects are highly nonlinear and so simulating the impact of each source (ocean, river and rainfall) separately and summing them up would lead to large errors. Therefore, the

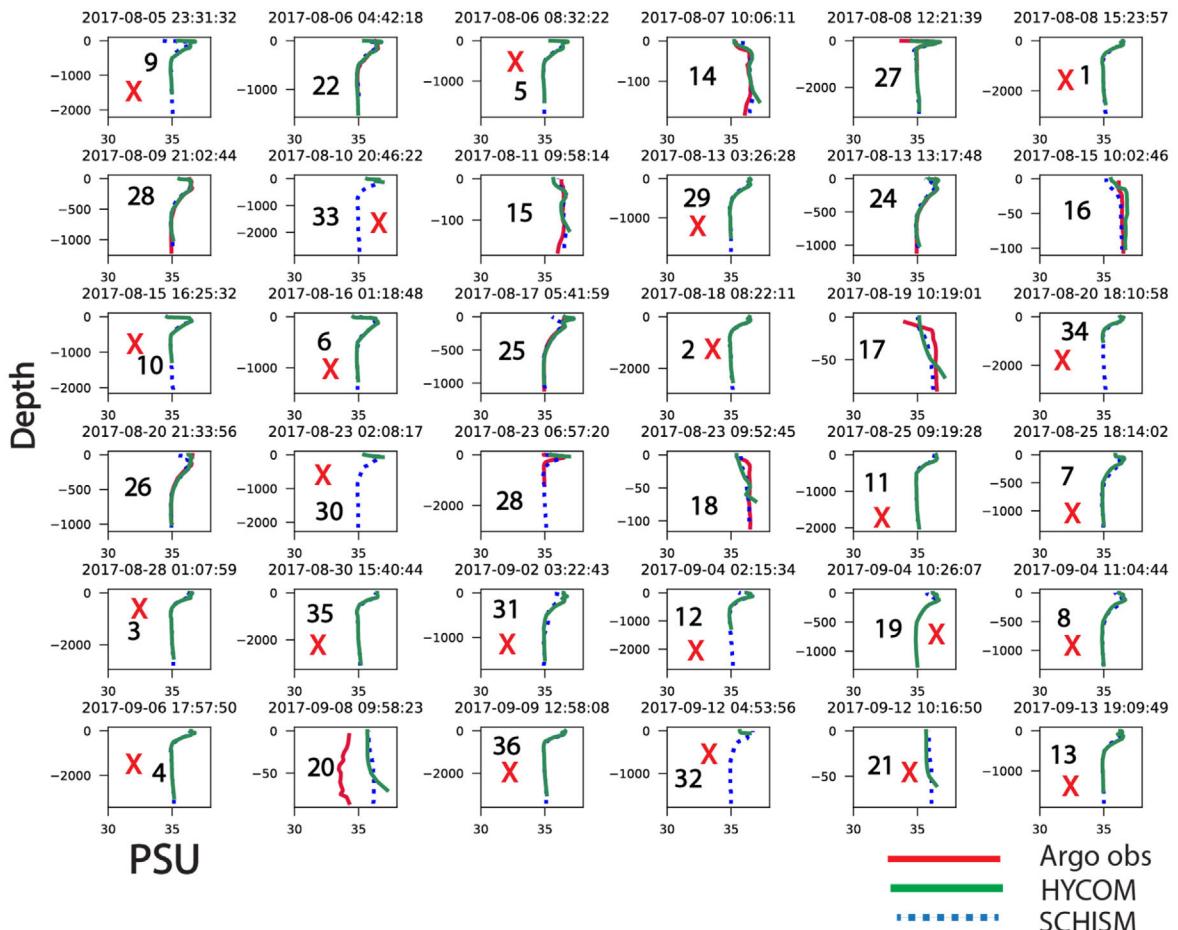


Fig. A.2. Comparison of the modeled salinity with observation from Argo floats (\times indicates no data; the numbers in the subplots indicate the profile locations as shown in Fig. 3c). Note that HYCOM uses data assimilation while SCHISM does not; also note the different depths used in the two models.

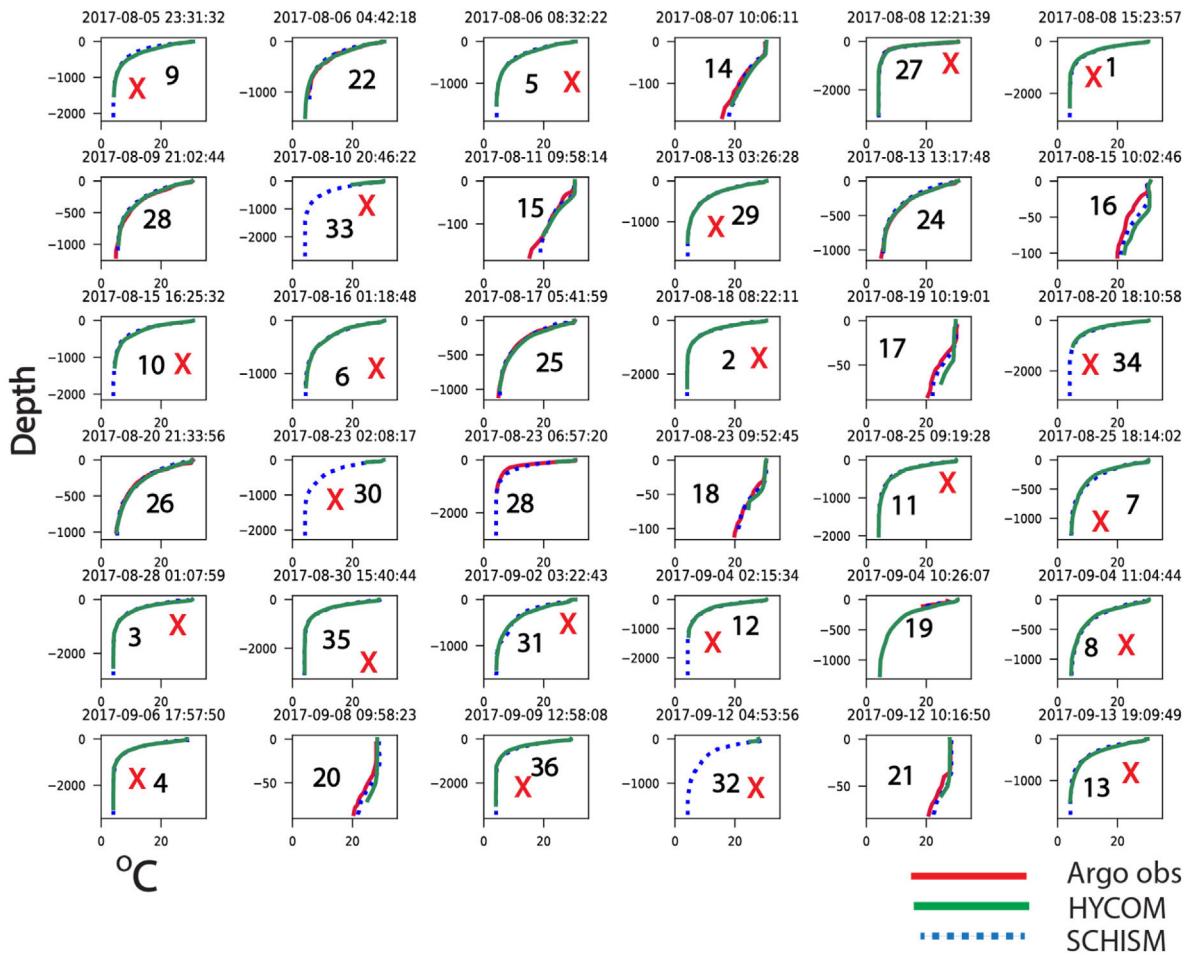


Fig. A.3. Comparison of the modeled temperature with observation from Argo floats (× indicates no data; the numbers in the subplots indicate the profile locations as shown in Fig. 3c). Note that SCHISM occasionally was able to correct HYCOM errors (e.g. profiles 17 and 20).

compound flooding studies are best carried out using a single model that integrates across all factors;

- (3) Treating watershed flows as boundary conditions and neglecting direct precipitation in a typical coastal watershed may result in large errors in downstream flow delivery and surges.

The well-validated 3D model presented in this paper has so far revealed many other interesting aspects of the compound flooding that have not been discussed in this paper: role of manmade structures (levees, dams, etc.) and constrictions on flow routing and surges; impact of the large and persistent freshwater plume on coastal regions, etc. Note that forecasts for atmospheric and hydrologic conditions already exist in NOAA and other agencies, and given the realism across a broad range of spatial scales and model efficiency as demonstrated in this paper, an operational forecast based on the current study is being tested at NOAA. This study also exposes some remaining critical knowledge gaps. The compound flooding results are shown to be particularly sensitive to several key factors: river flows, DEMs (especially bathymetry in small rivers), boundary condition for temperature in the watershed and atmospheric forcing (wind, air pressure, various heat fluxes and precipitation and evaporation). The accuracy of the predicted surges is critically dependent on the river flows, DEMs and atmospheric forcing in the ocean and bay, and therefore, a critical need in the future is to further improve the accuracy of the hydrologic and atmospheric models, as well as the nearshore seamless bathymetry-topography information that models rely on.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is funded by NOAA's Water Initiative (Grant Number NA16NWS4620043). The authors thank Dr. Linus Magnusson (ECMWF) for providing the high-resolution ERA forcing. Simulations used in this paper were conducted using the following computational facilities: (1) William & Mary Research Computing for providing computational resources and/or technical support (URL: <https://www.wm.edu/it/rc>) (2) the Extreme Science and Engineering Discovery Environment (XSEDE; Grant TG-OCE130032), which is supported by National Science Foundation grant number OCI-1053575; (3) the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at NASA Ames Research Center; (4) resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Appendix. Comparison of Argo profiles

Large-scale thermohaline structure was assessed by examining salinity and temperature profiles at 36 Argo profiles outside Galveston

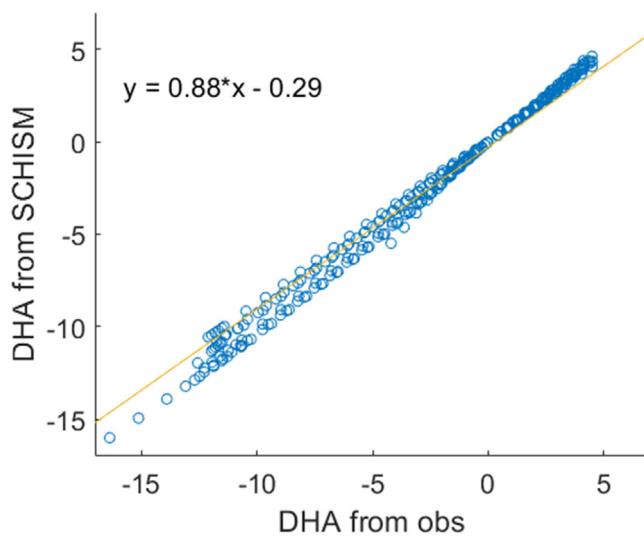


Fig. A.4. Scatter plot for dynamic height anomaly (m^2/s^2), using a reference pressure of 100 dbar. The line represents the best regression fit with equation given in the plot.

Bay (Fig. A.1). Comparison among SCHISM, (data assimilated) HYCOM and observation indicates that both models are able to capture the vertical structure for both salinity and temperature (Figs. A.2 and A.3). The averaged MAEs of SCHISM are 1.2 °C for temperature and 0.2 psu for salinity at those profiles. A synthesis plot using the Dynamic Height Anomaly (DHA) (software from <http://www.teos-10.org/>) is shown in Fig. A.4. The temperature and salinity were used to calculate the density first, and the software function then calculates the specific volume anomaly using the computationally efficient 75-term expression for specific volume (Roquet et al., 2015). Fig. A.4 suggests that the model slightly underestimates the DHA or the associated geostrophic velocity (http://www.teos-10.org/pubs/gsw/html/gsw_geo_strf_dyn_height.html) in the Gulf of Mexico.

References

- Anarde, K., Figlus, J., Sous, D., Tissier, M., 2020. Transformation of infragravity waves during hurricane overwash. *J. Marlene Sci. Eng.* 8 (545), <http://dx.doi.org/10.3390/jmse8080545>.
- Balaguru, K., Foltz, G.R., Leung, L.R., 2018. Increasing magnitude of hurricane rapid intensification in the central and eastern tropical atlantic. *Geophys. Res. Lett.* 45, 4238–4247. <http://dx.doi.org/10.1029/2018GL077597>.
- Carrere, L., Lyard, F., Cancet, M., Guillot, A., Picot, N., 2016. FES 2014, a New Tidal Model – Validation Results and Perspectives for Improvements, Presentation To ESA Living Planet Conference. Prague.
- Dorado, S., Booce, T., Steichen, J., McInnes, A.S., Windham, R., Shepard, A., Lucchese, A.E.B., Preischedl, H., Pinckney, J.L., Davis, S.E., Roelke, D.L., Quigg, A., 2015. Towards an understanding of the interactions between freshwater inflows and phytoplankton communities in a subtropical estuary in the gulf of Mexico. *PLoS One* 10, e0130931. <http://dx.doi.org/10.1371/journal.pone.0130931>.
- D'Sa, E.J., Joshi, I., Liu, B., 2018. Galveston bay and coastal ocean optical-geochemical response to hurricane harvey from VIIRS ocean color. *Geophys. Res. Lett.* 45, 10579–10589. <http://dx.doi.org/10.1029/2018GL079959>.
- Du, J., Park, K., 2019. Estuarine salinity recovery from an extreme precipitation event: Hurricane harvey in galveston bay. *Sci. Total Environ.* 670, 1049–1059.
- Du, J., Park, K., Dellapenna, T.M., Clay, J.M., 2019a. Dramatic hydrodynamic and sedimentary responses in galveston bay and adjacent inner shelf to hurricane harvey. *Sci. Total Environ.* 653, 554–564.
- Du, J., Park, K., Dellapenna, T.M., Clay, J.M., 2019b. Corrigendum to dramatic hydrodynamic and sedimentary responses in galveston bay and adjacent inner shelf to hurricane harvey [Sci. Total Environ. 653 (2019), 554–564]. *Sci. Total Environ.* 697, 134219. <http://dx.doi.org/10.1016/j.scitotenv.2019.134219>.
- Du, J., Park, K., Yu, X., Zhang, Y.J., Ye, F., 2020. Massive pollutants released to galveston bay during hurricane harvey: Understanding their retention and pathway using Lagrangian numerical simulations. *Sci. Total Environ.* 704, 135364.
- Emanuel, K., 2017. Assessing the present and future probability of hurricane harvey's rainfall. *Proc. Natl. Acad. Sci. U.S.A* 114 (48), 12681–12684. <http://dx.doi.org/10.1073/pnas.1716222114>.
- Federal Emergency Management Administration (FEMA), 2020. FEMA - Harvey flood depths grid, hydroshare. <http://www.hydroshare.org/resource/e8768f4cb4d5478a96d2b1cbd00d9e85>.
- Fritz, A., Samenow, J., 2017. Harvey Unloaded 33 Trillion Gallons of Water in the U.S. The Washington Post, https://www.washingtonpost.com/news/capital-weather-gang/wp/2017/08/30/harvey-has-unloaded-24-5-trillion-gallons-of-water-on-texas-and-louisiana/?utm_term=.0c5e249acbb7 (September 2, 2017).
- Guthrie, C.G., Matsumoto, J., Solis, R., 2012. Analysis of the Influence of Water Plan Strategies on Inflows and Salinity in Galveston Bay. Final report to the United States Army Corps of Engineers, Contract #R0100010015, Texas Water Development Board, Austin, Texas, USA, p. 71.
- Hall, T.M., Kossin, J.P., 2019. Hurricane stalling along the North American coast and implications for rainfall. *Clim. Atmos. Sci.* 2, 17. <http://dx.doi.org/10.1038/s41612-019-0074-8ence>.
- Lakshmanan, D., Howell, N.L., Rifai, H.S., Keonig, L., 2010. Spatial and temporal variation of polychlorinated biphenyls in the houston ship channel. *Chemosphere* 80, 100–112.
- Lewis, M.J., Palmer, T., Hashemi, R., Robins, P., Saulter, A., Brown, J., Lewis, H., Neill, S., 2019. Wave-tide interaction modulates nearshore wave height. *Ocean Dyn.* 69, 367–384.
- Liu, B., D'Sa, E.J., Joshi, I.D., 2019. Floodwater impact on galveston bay phytoplankton taxonomy, pigment composition and photo-physiological state following hurricane harvey from field and ocean color (sentinel-3a OLCI) observations. *Biogeosciences* 16, 1975–2001. <http://dx.doi.org/10.5194/bg-16-1975-2019>.
- Lucena, Z., Lee, M.T., 2017. Characterization of Streamflow, Suspended Sediment, and Nutrients Entering Galveston Bay from the Trinity River, Texas, 2014–2015. U.S. Geological Survey Scientific Investigations Report 2016-5177, p. 38. <http://dx.doi.org/10.3133/sir20165177>.
- Milliner, C., Materna, K., Bürgmann, R., Fu, Y., Moore, A.W., Bekaert, D., Adhikari, S., Argus, D.F., 2018. Tracking the weight of hurricane harvey's stormwater using GPS data. *Sci. Adv.* 4 (9), eaau2477. <http://dx.doi.org/10.1126/sciadv.aau2477>.
- van Oldenborgh, G.J., Wiel, K.van.der., Sebastian, A., Singh, R., Arrighi, J., Otto, F., Haustein, K., Li, S., Vecchi, G., Cullen, H., 2018. Corrigendum: Attribution of extreme rainfall from hurricane harvey, 2017 (2017 *Environ. Res. Lett.* 12 124009). *Environ. Res. Lett.* 13 (1), 019501. <http://dx.doi.org/10.1088/1748-9326/aaa343>.
- Pfahl, S., O'Gorman, P.A., Fischer, E.M., 2017. Understanding the regional pattern of projected future changes in extreme precipitation. *Nat. Clim. Change* 7, 423–427.
- Roquet, F., Madec, G., McDougall, T.J., Barker, P.M., 2015. Accurate polynomial expressions for the density and specific volume of seawater using the TEOS-10 standard. *Ocean Model.* 90, 29–43. <http://dx.doi.org/10.1016/j.ocemod.2015.04.002>.
- Santiago-Collazo, F.I., Bilskie, M.V., Hagen, S.C., 2019. A comprehensive review of compound inundation models in low-gradient coastal watersheds. *Environ. Model. Softw.* 119, 166–181. <http://dx.doi.org/10.1016/j.envsoft.2019.06.002>.
- Thyng, K.M., Hetland, R.D., Socolofsky, S.A., Fernando, N., Turner, E.L., Schoenbaechler, C., 2020. Hurricane harvey caused unprecedented freshwater inflow to galveston bay. *Estuar. Coasts* (2020), <http://dx.doi.org/10.1007/s12237-020-00800-6>.
- Umlauf, L., Burchard, H., 2003. A generic length-scale equation for geophysical turbulence models. *J. Mar. Res.* 61, 235–265.
- Valle-Levinson, A., Olabarrieta, M., Heilman, L., 2020. Compound flooding in houston-galveston bay during hurricane harvey. *Sci. Total Environ.* 747, 141272. <http://dx.doi.org/10.1016/j.scitotenv.2020.141272>.
- Watson, K.M., Harwell, G.R., Wallace, D.S., Welborn, T.L., Stengel, V.G., McDowell, J.S., 2018. Characterization of Peak Streamflows and Flood Inundation of Selected Areas in Southeastern Texas and Southwestern Louisiana from the August and 2017 Flood Resulting from Hurricane Harvey. US Geological Survey Scientific Investigations Report 2018-5070, p. 44. <http://dx.doi.org/10.3133/sir20185070>.
- Wing, O.E.J., Sampson, C.C., Bates, P.D., Quinn, N., Amith, A.M., Neal, J.C., 2019. A flood inundation forecast of hurricane harvey using a continental-scale 2D hydrodynamic model. *J. Hydrol.* X 4, 100039. <http://dx.doi.org/10.1016/j.hydra.2019.100039>.
- Ye, F., Y., Zhang, Yu, H., Sun, W., Moghimi, S., Myers, E.P., Nunez, K., Zhang, R., Wang, H.V., Roland, A., Martins, K., Bertin, X., Du, J., Liu, Z., 2020. Simulating storm surge and compound flooding events with a creek-to-ocean model: importance of baroclinic effects. *Ocean Model.* 145, 101526. <http://dx.doi.org/10.1016/j.jocemod.2019.101526>.
- Zeng, X., Zhao, M., Dickinson, R.E., 1998. Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *J. Clim.* 11, 2628–2644.
- Zhang, Y., Ateljevich, E., Yu, H.-C., Wu, C.H., Yu, J.C.S., 2015. A new vertical coordinate system for a 3D unstructured-grid model. *Ocean Model.* 85, 16–31.
- Zhang, Y., Ye, F., Staney, E.V., Grashorn, S., 2016. Seamless cross-scale modeling with SCHISM. *Ocean Model.* 102, 64–81.
- Zhang, Y., Ye, F., Yu, H., Sun, W., Moghimi, S., Myers, E., Nunez, K., Zhang, R., Wang, H., Roland, A., Du, J., Liu, Z., 2020. Simulating compound flooding events in a hurricane. *Ocean Dyn.* 70, 621–640.