Title

Data assimilation of hyper-local water level sensors for real-time monitoring of coastal inundation

Authors

Youngjun Son [a, *], Emanuele Di Lorenzo [a, b], Kyungmin Park [c], Spenser Wipperfurth [a], and Jian Luo [d]

Affiliations

[a] *Program in Ocean Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[b] *Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, RI, USA*
[c] *Coastal Sciences Division, Pacific Northwest National Laboratory, Seattle, WA, USA*
[d] *School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[*] *Corresponding Author*

- E-mail address: youngjun.son@gatech.edu (Y. Son)
- Address: *311 Ferst Drive, Atlanta, GA, USA, 30332*

Submitted to

Coastal Engineering

Highlights

- Objective Analysis is implemented to reconstruct water levels from sensor observations.
- Spatial covariance statistics are derived from coastal-ocean hydrodynamic model simulations.
- Variability in water levels is better captured by combining spatial covariance information.
- Real-time assimilation can provide situational awareness of floods in coastal communities.

Abstract

As flood events become increasingly prevalent in coastal regions with sea level rise, multiple communities have deployed water level monitoring networks across estuaries in addition to existing tide gauges located primarily at immediate coasts. Due to the spatially-distributed nature of sensor deployments, however, water level data are only available at specific sensor locations during the time of monitoring. As a result, an information gap on water levels exists along estuarine channels outside of active monitoring locations. To fill such a gap, this study presents a physics-based empirical modeling approach to assimilate coastal water levels using observations from hyper-local water level sensors. We implement an Objective Analysis (OA) procedure for sensor observation datasets obtained from the Smart Sea Level Sensors project along the U.S. Georgia coasts, based on spatial covariance structures of water levels that are extracted from high-resolution coastal-ocean hydrodynamic simulations. The approach is validated using simulated water levels, which provide basis functions for spatial covariance information. Additionally, the implementation is validated using actual water levels from active monitoring stations. The assimilated results of water levels are compared to those obtained by the same OA procedure but with a commonly-used Gaussian covariance function, which lacks prior knowledge of spatial covariance structures. To demonstrate the capability of the assimilation approach, we extend its application to a hurricane event, during which other dynamic processes may be relevant to variability in coastal water levels. Overall, the presented approach provides an accurate and efficient estimation of estuarine water levels along channels, which can support community officials to promptly identify localized flood threats to critical infrastructure systems in coastal regions.

Keywords

coastal flood; water level; data assimilation; objective analysis; monitoring network; hyper-local sensor

1. Introduction

Flooding is a growing threat to both populations and infrastructure in coastal regions (Hallegatte et al., 2013; National Academies, 2014; Allen et al., 2018). Primarily produced by a tropical storm or hurricane, a storm surge can cause catastrophic flooding and subsequent damages along coastlines, including human causalities and property losses (Grinsted et al., 2019). In addition, sea level rise due to climate change further exposes a greater number of coastal communities to an increased risk of flooding during high tides (Muis et al., 2016; Sweet et al., 2022). Over the last several decades, for example, the average number of high tide flood days per year has been steadily increasing in major cities and towns along the U.S. coasts, more than doubling particularly along the Southeast Atlantic coasts (Sweet et al., 2018; Moore and Obradovich, 2020). Recurrent flooding leads to public inconvenience and mental distress by repeatedly disrupting local transportation and business in low-lying urban areas (Moftakhari et al., 2015). Moreover, high tide flooding poses a significant risk to infrastructure systems in coastal regions, such as bridges, marinas, and stormwater drainages (Tahvildari and Castrucci, 2021; Gold et al., 2022). These structures could temporarily fail to function and even permanently lose structural integrity due to inundation by saline floodwater (Allen et al., 2018). Unfortunately, coastal infrastructure has limited adaptive capacity to cope with sea level rise and consequent flood impact due to restricted relocation options. Therefore, obtaining reliable information on water levels along major channels becomes crucial, not only for monitoring imminent flood threats but also for facilitating long-term planning to enhance flood resilience.

In many coastal regions, public and emergency management officials rely on real-time tide observations at a sparse network of NOAA stations, even though monitoring points of interest may be several kilometers away from the nearest station. In estuaries, however, water levels result from rising and falling tides and interactions with a range of channels from rivers to tributaries, landscapes, and local meteorological disturbance, which generate a complex pattern along coastlines (Gallien et al., 2011; Marsooli et al., 2016; Bilskie et al., 2021). Although advances in coastal-ocean modeling techniques have facilitated extreme water level predictions associated with hurricane-induced storm surges, operational forecasts using high-resolution models can incur high computational costs (Kerr et al., 2013; Bilskie et al., 2020). Moreover, multiple sources of uncertainties originating from input topobathymetry and boundary-forcing conditions cause gaps between model predictions and actual observations (Muñoz et al., 2022), which persist and propagate over simulation periods. Without leveraging gauge-based observations through model calibration or data assimilation, physics-based model predictions remain limited to assisting community officials in identifying immediate flood threats in coastal regions.

Recently, several coastal communities have been deploying affordable water level sensors in an urgent effort to obtain real-time water level information and better understand the hyper-local impacts of rising sea levels. As summarized in Table 1, multiple monitoring networks have been established along the U.S. coasts, using low-cost water level sensors by local researchers and communities. The sensor devices can measure water levels based on pressure, radar, and ultrasonic techniques and transmit measured data via long-range radio, cellular, or satellite communication. In the City of Norfolk, U.S. Virginia, for example, the StormSense project (Loftis et al., 2018) has deployed more than 50 water level sensors, either radar- or ultrasonic-type, to spatially complement the existing NOAA tide gauges and support predictive flood modeling. Another example is the Smart Sea Level Sensors (SSLS) project, which has installed more than 60 ultrasonic water level sensors along the U.S. Georgia coasts to augment hyper-local water level monitoring, with high priority to locations of critical infrastructure systems, such as bridges and boat ramps. These monitoring networks allow public and emergency management professionals to access hyper-local water level conditions and take prompt actions in response to potential flood risks. However, available monitoring locations are still confined by finite sensor installations (Tien et al., 2023). Furthermore, the observation availability is influenced by sensor operation that is subject to power supply or weather conditions for wireless communications. For example, sensors may cease to function due to power loss or faulty instruments, resulting in a lack of access to vital water level data, particularly during precarious events such as spring tides. Given these constraints, real-time assimilation of water level observations becomes crucial as it can facilitate extensive support for expanding sensor networks by estimating water levels at inactive sensor locations and in geographic areas beyond the immediate vicinity of sensor installations.

Table 1

Water level monitoring networks using low-cost, power-efficient sensors along the U.S. coasts.

Project	**Location**	**Number of Sensors**	**Sensor Type**	**Primary Network Communication**	**Observation Availability***	**Observation Cycle**	**Reference**
StormSense	Norfolk, Virginia	50 +	Ultrasonic and Radar	Cellular and RoLaWAN	2016 to Present	6-minute	Loftis et al. (2018)
Smart Sea Level Sensors (SSLS)	Georgia	60 +	Ultrasonic	RoLaWAN	2019 to Present	6-minute	https://www.sealevelsensors.org/
Hohonu	U.S. East Coast, Hawaii, and Alaska	90 +	Ultrasonic	Cellular	2017 to Present	6-minute	https://www.hohonu.io/
SenseStream (I-SENSE)	Florida and South Carolina	50 +	Ultrasonic	Cellular	2018 to Present	5-minute	https://www.sensestream.org/
Sensing Storm Surge	Maine	20 +	Pressure	Not capable	2017 to 2020	2-minute	Spicer et al. (2021)

* Observation availability varies depending on the deployment and operation of individual water level sensors.

Due to a short history of hyper-local water level monitoring networks, no prior study exists to assimilate coastal water levels directly from the newly emerging observation datasets. At global- or regional-scales in oceanography, however, many studies have been conducted to reconstruct a large-scale map of oceanographic data based on observations with limited spatial coverage. Since the introduction into oceanography by Bretherton et al. (1976), the Objective Analysis (OA) technique has been extensively used to interpolate spatial sampling data into a continuous field of ocean properties, including sea surface height (e.g., Ubelmann et al., 2015), temperature (e.g., Smith et al., 1996), and pressure (e.g., Kaplan et al., 2000). Smith et al. (1996) were the first to utilize Empirical Orthogonal Functions (EOFs) for spatial covariance statistics to fill spatial gaps in marine observations of sea surface temperature, using satellite-based datasets with extensive coverage. Similarly, Chambers et al. (2002) applied EOFs that were derived from satellite altimetry into historical tide gauge records to examine the effects of climate variability on global mean sea levels. To assess an emerging risk of floods in the context of a changing climate, the EOFs-based OA technique has been widely adopted in identifying global and regional variability in rising sea level trends (e.g., Church and White, 2006; Hamlington et al., 2012). The application has been further improved by reducing inherent uncertainties in observation datasets (e.g., Church et al., 2004; Church and White, 2006), by applying sophisticated basis functions, such as cyclostationary EOFs (e.g., Hamlington et al., 2011; Kim et al., 2015), or by combining with other climate variables (e.g., Hamlington et al., 2012; Kumar et al., 2020). Ocean circulation model simulations have also been used as an alternative to satellite altimetry datasets to derive the spatial patterns of physical ocean systems (e.g., Llovel et al., 2009; Meyssignac et al., 2011). Although the primary focus of these studies has been on estimating variability of sea level rise trends at large spatiotemporal scales, the application of the EOFs-based OA technique has enabled the reconstruction of water levels by combining spatial covariance structures with scattered observation datasets.

In studies related to coastal floods, the EOF analysis has been frequently employed for dimensionality reduction within surrogate models, which are commonly referred to as metamodels. For coastal flood predictions, surrogate models build a functional connection between input parameters (e.g., hurricane intensity and track) and corresponding responses (e.g., maximum storm surge) based on numerical model simulation databases for historical or synthetic flood events (Jia and Taflanidis, 2013; Al Kajbaf and Bensi, 2020). To address the challenges posed by high-dimensional spatiotemporal outputs, the EOF analysis not only reduces the dimensionality of output responses but also facilitates the extraction of response covariance patterns (Jia and Taflanidis, 2013; Jia et al., 2015), as similarly applied within the EOFs-based OA technique. In addition, various statistical representations, including Gaussian Process Regression and Neural Networks, have been utilized in surrogate models to establish relationships between the weights of response covariance patterns and metocean input parameters (see a review in Al Kajbaf and Bensi, 2020). Specifically, surrogate models have been applied to different coastal flood drivers, such as storm surge (e.g., Jia et al., 2015; Bass and Bedient, 2018; Rohmer et al., 2023; Kyprioti et al., 2023), storm waves (e.g., Jia and Taflanidis, 2013; Rohmer et al., 2023), and even tidal responses with elevated levees in estuaries (Li et al., 2020). On the other hand, the EOFs-based OA technique directly leverages spatially-distributed response observations (i.e., water levels for coastal flood monitoring) to determine the weights of response covariance patterns. Within the EOFs-based OA technique, optimal interpolation (Daley, 1991; Kalnay, 2003) is a key component that generates the most likely spatial state representation by minimizing error variances at observation locations. Regarded as a specific suboptimal variant of the extended Kalman filter (Ide et al., 1997), optimal interpolation serves a fundamental role in both real-time assimilation and sequential updates of forecast models due to its relative simplicity (e.g., Madsen et al., 2015; Asher et al., 2019 to adjust storm surge model forecasts). As water level observation datasets newly emerge at hyper-local scales for coastal flood monitoring, it becomes increasingly relevant to explore their potential applications in data assimilation, especially with established traditional methods such as the OA technique.

The objective of this study is to develop a physics-based empirical modeling approach to augment water level monitoring of a hyper-local sensor network. To achieve this objective, we apply the OA procedure to assimilate water levels in coastal regions, following a similar methodology used by Chambers et al. (2002) to reconstruct mean sea levels globally. Particularly, coastal-ocean hydrodynamic model simulations are performed to extract spatial covariance statistics of water levels which are then decomposed in EOFs. As a result, past and present water levels can be assimilated by combining the spatial covariance information with a time series of available sensor observations. Our study explores the feasibility of the empirical modeling approach for hyper-local sensor observations which integrates spatial covariance structures established from physics-based model simulations. In addition, we examine the applicability of the modeling approach using a range of scenarios, including both tides to storm events. As a computationally efficient and accessible option for real-time applications, the modeling approach allows community officials to access hyper-local water level conditions, even for temporarily inoperative sensor locations. Furthermore, the outcomes can provide valuable insights for optimizing the deployment of sensors that are constrained by limited resources. Ultimately, the assimilated water levels can benefit coastal-urban flood models that rely on accurate water level data along coastlines to predict flooding in urban systems (e.g., Smith et al., 2011; Karamouz et al., 2017; Son et al., 2023).

The remainder of the paper is organized as follows. Section 2 describes the OA procedure that combines physics-based model simulations, along with the background information on the hyper-local water level monitoring network and the high-resolution coastal-ocean hydrodynamic model. In Section 3, we examine the physics-based empirical modeling approach by using the numerical simulation results to quantify the different sources of errors. Then, the applications are extended to assimilate water levels directly using actual observations from the hyper-local sensor networks, not only for the basis simulation periods but also for a hurricane event that occurred during the operational monitoring periods. Furthermore, our study demonstrates the real-time assimilation of water levels in a pilot web-based portal for emergency management. We discuss the potential benefits and limitations of the model applications in Section 4. Finally, Section 5 provides a summary and conclusions of the present study.

2. Methods

2.1. Hyper-local water level sensor network: Smart Sea Level Sensors (SSLS)

The SSLS project (https://www.sealevelsensors.org/) has installed more than 40 low-cost, power-efficient, Internet of Things (IoT)-enabled water level sensors (red and pink placemarks), across Chatham County, U.S. Georgia, as shown in Fig. 1. The hyper-local sensor network spans a wide range of estuarine channels, prioritizing monitoring of coastal infrastructure that is prone to flooding, such as bridges and marinas. Prior to the SSLS deployments, the NOAA station at Fort Pulaski (blue placemark), the only tide gauge on the entire Georgia coast, provided information on coastal water levels in Chatham County. While USGS also operates multiple monitoring locations, the observation availability remains limited with a small number of stream gauges (green placemarks) located upstream of the estuaries and temporary storm tide loggers (yellow placemarks) to collect water levels during extreme weather events, such as hurricanes. Consequently, the SSLS water level monitoring network provides instantaneous access for public and emergency officials (e.g., Chatham Emergency Management Agency) to identify localized flood threats to coastal communities. The sensor devices are designed to measure the distance to water surface elevation and transmit measurement data into a monitoring server through internet-connected gateways, using ultrasonic sensors and Long Range Wide Area Network (LoRaWAN) technology. As part of the sensor deployments, two validation sensors are placed next to the NOAA tide gauge at Fort Pulaski (blue marker). A comparison of the validation sensors with the NOAA tide gauge indicated less than 0.3 *cm* differences for two years. Currently, the sensors operate with a power-saving protocol, which involves repeating one measurement cycle of collecting 18 samples for 3 seconds and transmitting the averaged signal every 6 minutes.

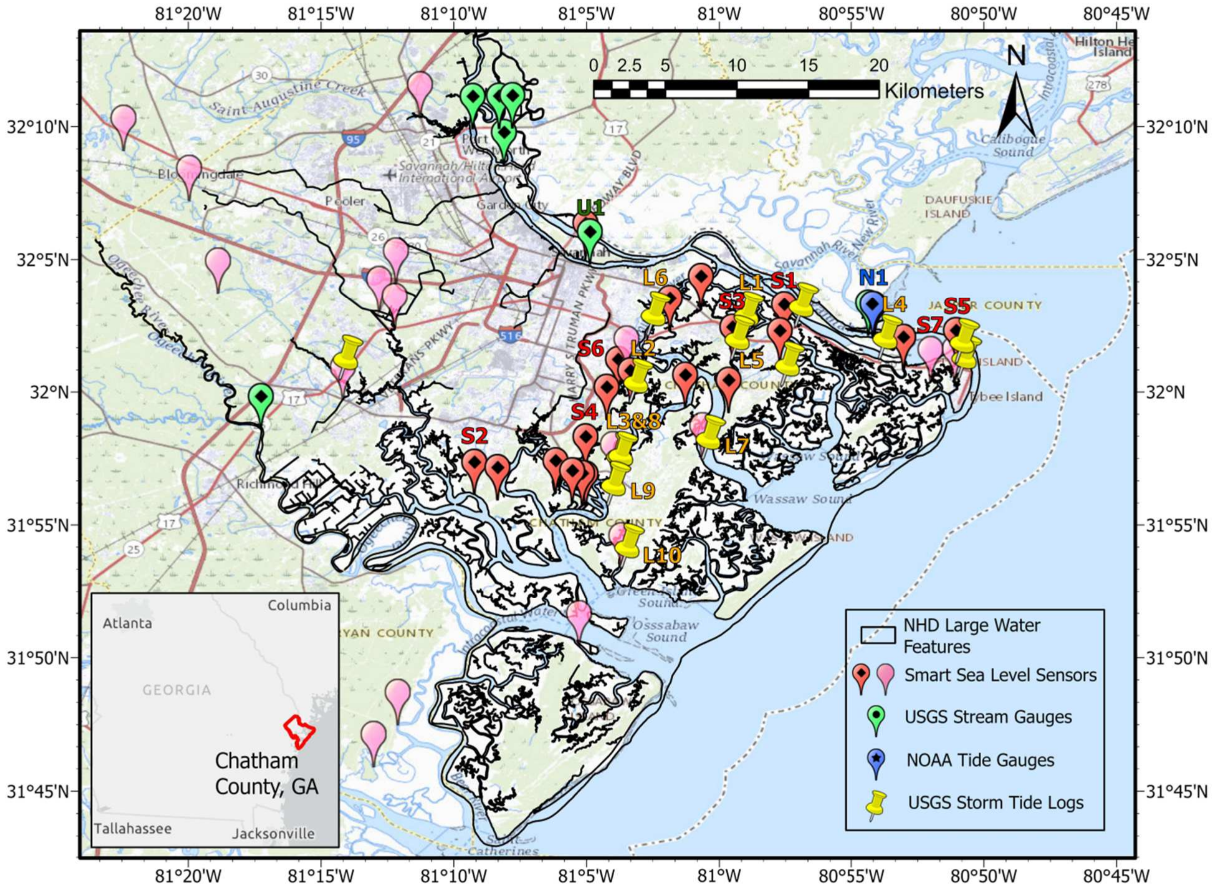

Fig. 1 Water level monitoring stations of NOAA, USGS, and SSLS with NHD Large Water Features of Chatham County in GA, USA. Among SSLS locations, the red placemarks indicate those with elevation survey information within NHD Large Water Features. The labeled locations (NOAA: N1, USGS: U1, SSLS: S1-7, and USGS Log: L1-10) will serve as points of reference in our study.

2.2. Water level simulations using a high-resolution coastal-ocean hydrodynamic model

Physics-based model simulations of coastal water levels are performed to establish spatial covariance statistics in the OA procedure. In this study, the Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM) (Zhang and Baptista, 2008; Zhang et al., 2016) is implemented to simulate water levels and inundations from the coasts to inland areas along the Georgia coasts. The SCHISM model has been widely used to simulate coastal-ocean circulation processes with an emphasis on three-dimensional baroclinic modes, including storm surge simulations (Zhang et al., 2020; Ye et al., 2020). As shown in Fig. 2a, the model domain covers the entire U.S. Georgia coast with horizontal resolutions ranging from 6 *km* (around the open boundary) to 3 *m* (along small rivers and creeks). The topobathymetry for unstructured grid systems is interpolated from the NOAA Continuously Updated Digital Elevation Model (CUDEM; 3 to 100 *m* resolutions) with local refinements using other sources of NOAA DEM (e.g., Sea Level Rise Viewer DEM; 3 to 5 *m* resolutions). The model setup integrates inland hydrologic input from the National Water Model (NOAA NWS, 2016). In the model setup, surface waves are not considered due to their marginal impacts in the study areas, which are characterized by extensive wetlands along geographically complex channels in estuarine settings (Muñoz et al., 2021). The initial and boundary conditions are provided by oceanographic datasets from Archiving, Validation and Interpretation of Satellite Oceanographic (AVISO) and Copernicus Marine Environment Monitoring Service (CMEMS). A tide model of Finite Element Solution (FES) 2014 is used for the open boundary condition and tidal potential in momentum equations, including eight major tidal constituents such as K1, K2, M2, N2, O1, P1, Q1, and S2. The ECMWF Reanalysis 5 (ERA 5) datasets are applied for atmospheric forcing.

The SCHISM model simulations are performed for October 21 to December 30, 2021, including the model spin-up phase. The simulated water levels for 30 days from November 25 to December 25 (Fig. A.1) are utilized to calculate spatial covariance statistics in the OA procedure. Fig. 2b compares the simulated water levels with the observations at the NOAA tide gauge (N1), USGS stream gauge (U1), and SSLS sensors (S1-6). The comparisons indicate that the SCHISM model setup is capable of simulating water levels accurately across estuaries, as shown for the multiple locations (N1, U1, S1, S2, and S4) in Fig. 2b. In addition, our study includes a few locations to assess the capability of the developed OA approach later in the current study: (S3), where the SSLS sensor was inactive during the simulation periods, (S5), where discrepancies may exist during low tides due to the model representation of topobathymetry, and (S6), where a slight phase lag may arise with the model uncertainty.

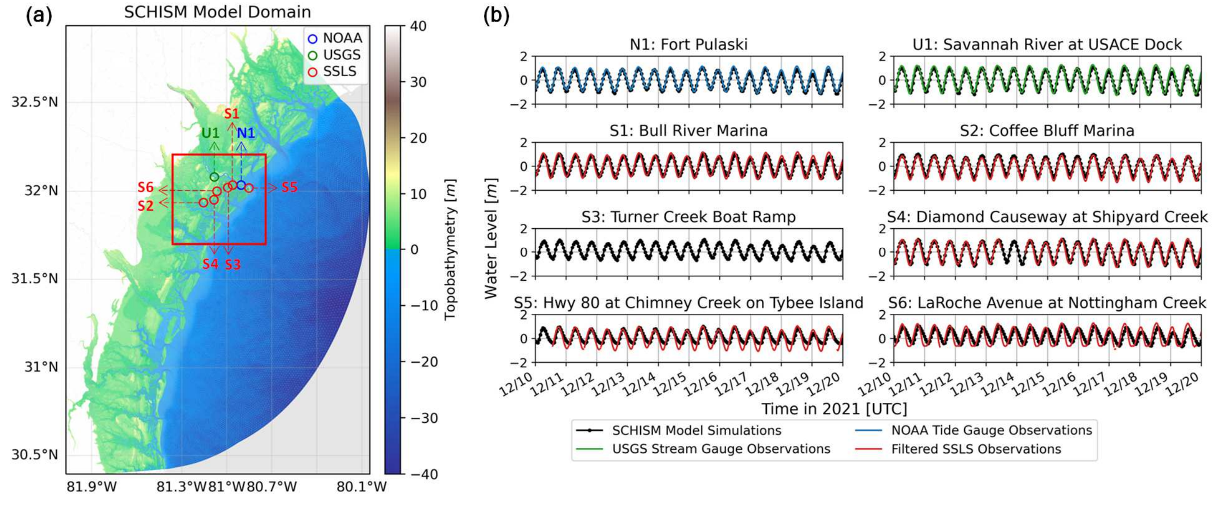

Fig. 2 SCHISM model simulations: (a) model domain and topobathymetry. The red box corresponds to the map extent in Fig. 1; (b) comparisons of simulated water levels with observations. The headings (N1, U1, and S1-6) on the titles correspond to the labels of the station locations in Fig. 1. The model evaluations, including error statistics, are listed in Appendix C.

2.3. Physics-based empirical modeling of coastal water levels

In our study, as outlined in the flowchart in Fig. 3, the reconstruction of water levels is accomplished through the OA procedure (Chambers et al., 2002) that combines the SSLS observation datasets (Section 2.1) with the spatial covariance statistics of the simulated water levels using the SCHISM model (Section 2.2). Prior to using the observation datasets, a preprocessing procedure is implemented due to different measurement timings and abnormal signals within water level records. In the simulation results, water levels in upland areas may experience wet and dry conditions, requiring a masking process for the numerical model grids to prevent incomplete statistical analysis. The EOF analysis establishes the statistical covariance structures of the simulated water levels for the masked grids. Using the OA procedure, the Principal Components time series (PCs) are calculated from the observation datasets available over a specific period of interest, which are subsequently applied to assimilate a spatial field of water levels by combining with the established EOFs. The detailed implementations will be described in the following subsections.

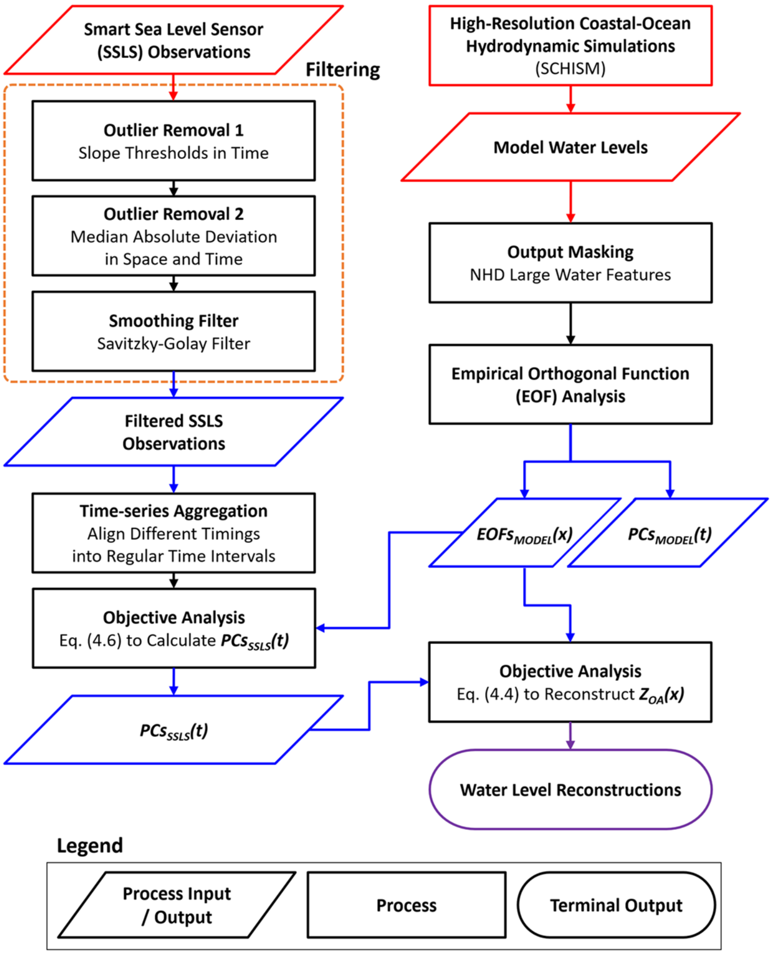

Fig. 3 Flowchart of physics-based empirical modeling of coastal water levels, using SSLS observation datasets.

2.3.1. Preprocessing of SSLS observation datasets

In water level monitoring networks, the operations of sensor devices are often influenced by environmental factors (e.g., animal or human activities) and various instrument conditions (e.g., power supply or network connections) (Spicer et al., 2021). As a result, the SSLS observation datasets may contain abnormal signals, such as sporadic fluctuations, spikes, and occasional outages. To address such issues, a data-processing procedure is applied to filter the water level records, including outlier removals and smoothing, with the following steps:

1. *Slope Thresholds in Time*:
   To identify local outliers in time series data for each sensor, the time rates of change in water levels are calculated from consecutive data points within a specified timeframe to count a portion of excessive or unreasonable changes in water levels. For example, if more than 10 % of neighboring data points within a two-hour timeframe indicate unreasonable changes in water levels (e.g., > 1 *m/hour*), the corresponding measurement is deemed unstable and discarded from the analysis. The assessment of the rate of change is highly recommended by Integrated Ocean Observing System (2021).
2. *Median Absolute Deviation (MAD) in Space and Time*:

$$\mathrm{MAD} = median|z_t - \langle z_t \rangle| \tag{1}$$

   To detect outliers in both space and time, the MAD of the time rate of change in water levels, z_t, is calculated by using more than five sensor measurements for cross-comparisons in space. Then, if a measurement deviates more than three standard deviations from the calculated MAD, which is equivalent to the Hampel filter (Pearson et al., 2016), it is identified as an outlier and subsequently removed from the analysis. This step is particularly useful when nearby observations are available, as suggested by Integrated Ocean Observing System (2021).
3. *Savitzky-Golay Filter*:
   The Savitzky-Golay filter (Savitzky and Golay, 1964; Gorry, 1990) is applied to smooth time series signals, which involves fitting a piecewise polynomial function to a set of measurements. Specifically, our study uses a third-order polynomial function with a four-hour time window for each measurement.

At the beginning of each step, our study excludes measurement points with fewer than four counts within two hours due to the limited representation of water level variations. As an example, Fig. 4 shows the step-by-step preprocessing results for SSLS observation datasets.

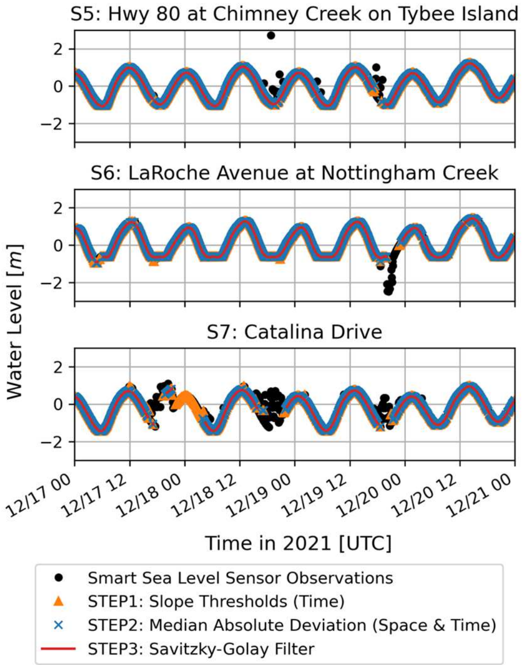

Fig. 4 Example results of step-by-step preprocessing for SSLS observation datasets: Steps 1 and 2 for outlier removals and Step 3 for smoothing.

2.3.2. OA based on spatial covariance statistics

Our study uses the same OA procedure of Chambers et al. (2002) that combines spatial covariance statistics to interpolate water levels from spatially-distributed observation datasets. The OA procedure consists of two main components: the first part applies the EOF analysis to physics-based model simulations of water levels to obtain spatial covariance structures as EOFs, and the second part determines the corresponding PCs for these EOFs based on available observation datasets to reconstruct water levels.

The EOF analysis decomposes spatiotemporal variations of water levels into a linear combination of spatial modes (EOFs), multiplied by associated time-varying amplitudes (PCs). Through a singular value decomposition, water levels, $\mathbf{Z} \in \mathbb{R}^{m \times n}$, which is a function of space, $\boldsymbol{x} \in \mathbb{R}^{m \times 1}$ at m locations, and time, $\boldsymbol{t} \in \mathbb{R}^{n \times 1}$ for n time steps, can be expressed as:

$$\mathbf{Z}(\boldsymbol{x},\, \boldsymbol{t}) = \mathbf{U\Sigma V^T} \tag{2}$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ is the spatial modes (EOFs), $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ is the diagonal matrix with the eigenvalues of $\mathbf{Z}$, and $\mathbf{V} \in \mathbb{R}^{n \times n}$ is the temporal variations of the spatial modes. Hence, Eq. (2) can be rewritten by simply replacing the latter terms with the time-varying amplitudes (PCs), $\boldsymbol{\alpha}(\boldsymbol{t}) \in \mathbb{R}^{m \times n}$:

$$\boldsymbol{\alpha}(\boldsymbol{t}) = \mathbf{\Sigma V^T} \tag{3}$$

$$\mathbf{Z}(\boldsymbol{x},\, \boldsymbol{t}) = \mathbf{U}(\boldsymbol{x})\boldsymbol{\alpha}(\boldsymbol{t}) \approx \mathbf{U}_k(\boldsymbol{x})\boldsymbol{\alpha}_k(\boldsymbol{t}) \tag{4}$$

The columns of $\mathbf{U}$ and rows of $\boldsymbol{\alpha}$ correspond to the individual modes of EOFs and PCs, respectively, and the subscript, k, denotes the number of the truncated modes. As a result, the EOF analysis enables reducing the number of modes by selecting a few low-order EOFs and PCs that explain a significant portion of variances in water levels. In addition, the leading-order EOFs and PCs may characterize distinct spatial covariance patterns and related temporal variations of water levels.

Based on Kaplan et al. (2000), the PCs can be found through a least-squares estimation process that minimizes the following objective function, S:

$$\min \mathrm{S}(\boldsymbol{\alpha}_k) = (\mathbf{H}\mathbf{U}_k\boldsymbol{\alpha}_k - \mathbf{Z}^{\mathbf{o}})^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{U}_k\boldsymbol{\alpha}_k - \mathbf{Z}^{\mathbf{o}}) + \boldsymbol{\alpha}_k^T\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\alpha}_k \tag{5}$$

$$\boldsymbol{\alpha}_k = \left(\mathbf{U}_k^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{U}_k + \boldsymbol{\Lambda}_k^{-1}\right)^{-1}\mathbf{U}_k^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{Z}^{\mathbf{o}} \tag{6}$$

where $\mathbf{Z}^{\mathbf{o}} \in \mathbb{R}^{p\times n}$ is water level observations at p locations, $\mathbf{H} \in \mathbb{R}^{p\times m}$ is a sampling operator matrix to consider observation availability, $\mathbf{R} \in \mathbb{R}^{p\times p}$ is the matrix for data error covariances due to instrumental, sampling, and truncation errors, and $\boldsymbol{\Lambda}_k \in \mathbb{R}^{k\times k}$ is the truncated diagonal matrix with the eigenvalues of the covariance matrix. As an initial study to apply the OA procedure to assimilate water levels directly from hyper-local sensor observations, our study implements the simplest approach by Chambers et al. (2002) where the identity matrix is used for $\mathbf{R}$ and the latter term is omitted as follows:

$$\min \mathrm{S}(\boldsymbol{\alpha}_k) = (\mathbf{H}\mathbf{U}_k\boldsymbol{\alpha}_k - \mathbf{Z}^{\mathbf{o}})^T(\mathbf{H}\mathbf{U}_k\boldsymbol{\alpha}_k - \mathbf{Z}^{\mathbf{o}}) \tag{7}$$

$$\boldsymbol{\alpha}_k = \left(\mathbf{U}_k^T\mathbf{H}^T\mathbf{H}\mathbf{U}_k\right)^{-1}\mathbf{U}_k^T\mathbf{H}^T\mathbf{Z}^{\mathbf{o}} \tag{8}$$

Once the PCs are obtained from Eq. (8) for available observation datasets, a spatial field of water levels can be reconstructed by combining PCs with the EOFs, as shown in Eq. (4).

In our study, the water levels simulated by the SCHISM model (over 53,000 grid points) serve as a basis dataset for the EOF analysis. Prior to performing the EOF analysis, we remove the temporal means of water levels over the simulation periods and apply the channel boundary based on the USGS National Hydrography Dataset (NHD) Large Water Features for Chatham County to mask the water level outputs along the coastal channels, as delineated in Fig. 1. This masking procedure enables us to reduce the number of grid points involved in matrix computations, with an emphasis on estuarine channels. Estuarine channels are of primary interest because the majority of SSLS have been deployed to monitor bridges and marinas. Additionally, we exclude grid points where water levels exhibit little variations and dry conditions to prevent using incomplete data for the EOF analysis. To align the different timings of the SSLS measurements in calculating the PCs, the observation datasets are aggregated by averaging them over a 15-minute interval.

2.3.3. OA using a Gaussian covariance function

The OA procedure commonly employs an analytical covariance function to interpolate a limited set of observations into a spatial field of interest. Based on the Gauss-Markov theorem, the least-square linear estimator can be calculated by

$$\mathbf{Z}^{x} = \widehat{\mathbf{Z}^{x}} + \mathbf{C}_{xo} \boldsymbol{C}_{oo}^{-1} \left(\mathbf{Z}^{o} - \widehat{\mathbf{Z}^{o}} \right) \tag{9}$$

where $\mathbf{Z}^{x} \in \mathbb{R}^{m \times n}$ is water levels at locations of interest, $\mathbf{C}_{xo} \in \mathbb{R}^{m \times p}$ is the spatial covariances between observations and locations of interest, and $\mathbf{C}_{oo} \in \mathbb{R}^{p \times p}$ is the spatial covariances between all pairs of observations. The $\widehat{\ }$ denotes the estimated mean water levels.

In the absence of spatial covariance information, it is often assumed that the covariances among the observations themselves, as well as between observations and a spatial field of interest, depend solely on the spatial distance between each pair of locations, regardless of the locations themselves. Among various analytical forms, we use the Barnes Scheme (Barnes, 1964) that adopts a Gaussian function to describe spatial covariances based on distance, which is expressed by:

$$\mathrm{C}(\boldsymbol{x}_1, \boldsymbol{x}_2) = f(d) = e^{(d/L)^2} \tag{10}$$

where d and L denote the distance between two locations and the decorrelation scaling parameter, respectively. Consistently with Section 2.3.2, we use the same SSLS measurements that are aggregated over a 15-minute interval. The least-square fitting technique is applied to estimate the linear mean trends of water levels. The decorrelation length scale of 5 *km* is used based on the preliminary implementation by Tien et al. (2023).

3. Results and discussion

3.1. EOF analysis and sensitivity tests

Through a singular value decomposition, the EOFs and PCs are obtained from the SCHISM model simulations of water levels from November 25 to December 25, 2021. Fig. 5 shows the first three spatial modes of the EOFs, along with the associated variance explained by each mode. To facilitate comparison across the different modes, the EOFs and PCs are normalized by multiplying the EOFs with the standard deviations of PCs and dividing the PCs by the standard deviations, respectively. As a result, the normalized EOFs represent the physical magnitudes of the spatial modes while retaining unit standard deviations for the normalized PCs. The lowest EOF mode in Fig. 5a has the strongest signal characterized by uniform water level patterns that can either rise or fall, primarily forced by astronomical tides. Fig. 5b shows the second EOF mode, which exhibits a linear slope pattern in water levels with strong correlations with the distance to open coasts. This linear pattern is likely driven by either flood or ebb tide flowing along the channel landscapes. The third spatial pattern in Fig. 5c is similar to the previous mode but has concentrated signals upstream in the complex channels of the estuaries. As plotted in Fig. 5d, the explained variances significantly decrease as the EOF modes increase, accounting for more than 95 % of the total variances with these three lowest modes. Therefore, our study uses the three dominant EOF modes for the reconstruction of water levels. The selection of these EOF modes implies that spatiotemporal variations in water levels result from a combination of rising or falling water levels, their flooding or receding phase, and adjustments related to upstream dynamics.

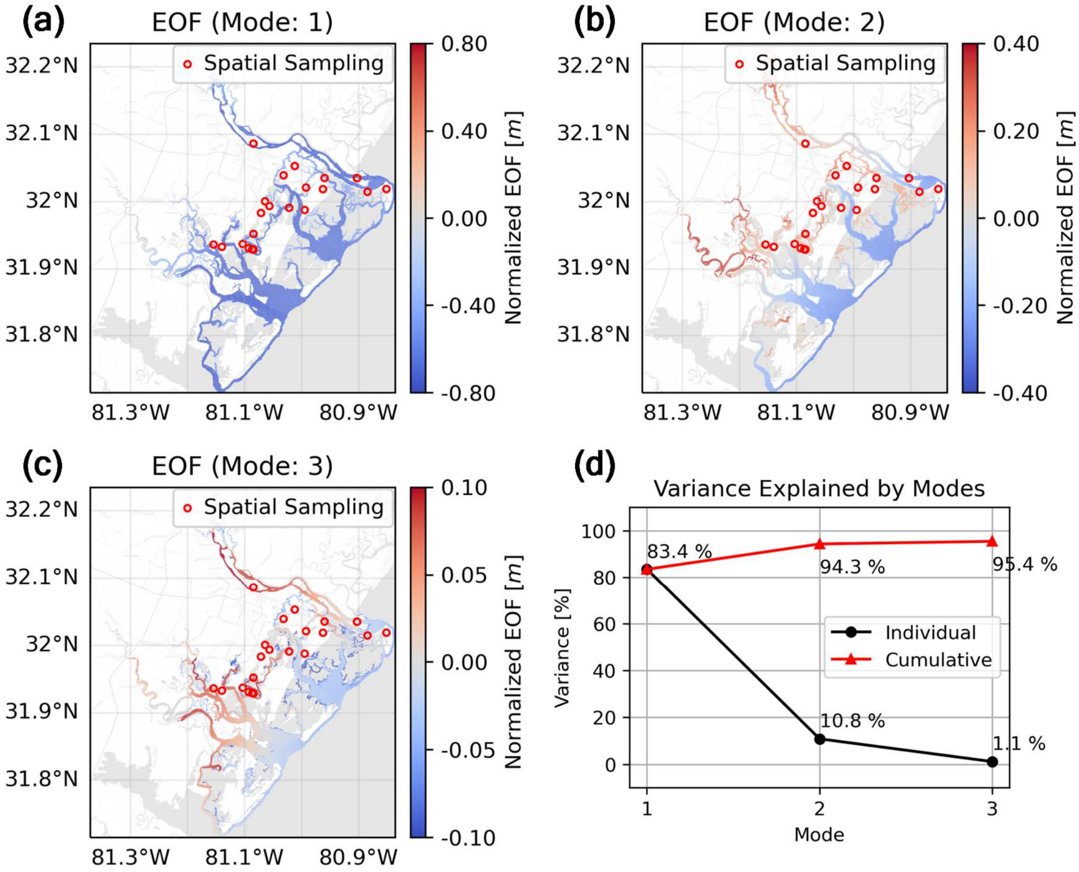

Fig. 5 Normalized EOFs and associated variances explained: (a) – (c) three lowest spatial modes of water levels. The EOFs are normalized with the standard deviations of PCs to indicate physical magnitudes. The red circles highlight SSLS that are located within the study boundaries and have sensor elevation survey information, corresponding to the red placemarks in Fig. 1; (d) variances explained by different spatial modes.

To assess the stationarity of the EOFs with respect to simulation periods, we implement a 3-fold test that involves extracting the EOFs from three separate chunks, each spanning 10 days of the simulation periods. Then, the obtained EOFs are applied to solve Eq. (8) to derive the corresponding PCs across the simulation periods. However, accurate water level observations will only be available at locations where SSLS are installed and their elevations are surveyed, which is referred to as spatial sampling in this study. The effects of spatial sampling due to the limited availability of observations are taken into consideration in the 3-fold test. Thus, we use the simulated water levels only at 22 SSLS locations within the study boundaries, which are highlighted with red circles in Fig. 5. As shown in Fig. 6, the normalized PCs are obtained from the different EOFs based on the three separate chunks of the simulated water level time series. The comparisons of the PCs within each mode show that the leading-order PCs are not significantly influenced by the selection of water level time series for the EOFs. However, minor variations are identified in the third PC, particularly around the peaks and troughs of the PC. Consequently, the EOFs and PCs marginally depend on simulation periods, even when applying the simulated water levels through spatial sampling for the valid SSLS locations.

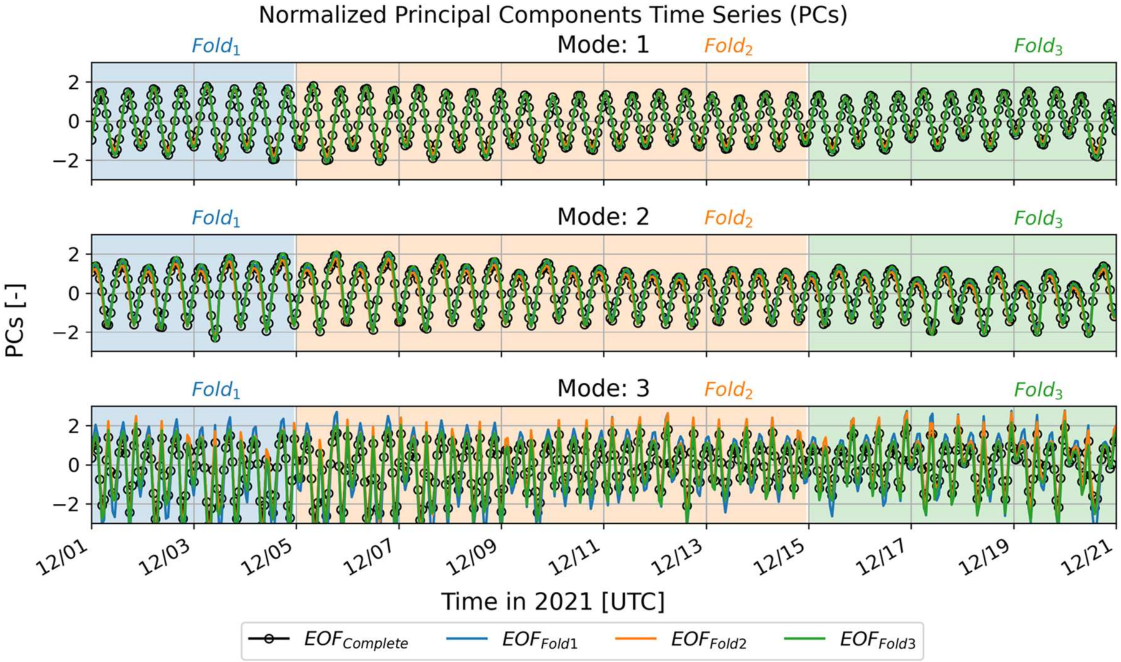

Fig. 6 Normalized PCs derived from the $\mathrm{EOF}_{\mathrm{Complete}}$ (11/25 – 12/25), $\mathrm{EOF}_{\mathrm{Fold1}}$ (11/25 – 12/05), $\mathrm{EOF}_{\mathrm{Fold2}}$ (12/05 – 12/15), and $\mathrm{EOF}_{\mathrm{Fold3}}$ (12/15 – 12/25). The PCs are calculated from spatial sampling of the simulated water levels at 22 SSLS locations within the study boundaries (red circles in Fig. 5) and normalized by their standard deviations.

Regarding the characteristics of the PCs, the first two PCs exhibit similar periodic cycles, which are highly correlated with astronomical tidal cycles. In addition, the peaks and troughs of the first PC closely align with the zero-crossings of the second PC, implying out-of-phase differences between the spatial patterns of the first and second EOF modes. The phase lag corresponds to flood or ebb tidal currents that reach their maximum at the mean water level between low and high tides. The third PC shows fluctuations with a higher frequency, suggesting potential adjustments in water levels as water enters and exits the upstream estuarine channels. These PCs represent the magnitudes, timing, and interplay between the three dominant EOF modes, assimilating the specific state of estuarine water levels.

3.2. Model verification with SCHISM model simulations

Prior to using the SSLS observation datasets, the simulated water levels from the SCHISM model are first utilized to reconstruct water levels and identify modeling gaps in the OA procedure. Fig. 7 shows example plots of water levels at a specific time during flood tides, comparing the SCHISM model simulations (Fig. 7a) with two types of reconstruction: one using all water levels within the study boundaries (Fig. 7b) and another using water levels only at 22 SSLS locations (Fig. 7c). Despite being confined to the study boundaries due to the masking procedure, the assimilated water levels are consistent with the numerical simulation results. However, the spatial sampling leads to discrepancies in the assimilation of water levels, particularly in the immediate coasts due to the limited sampling of water level data in these areas.

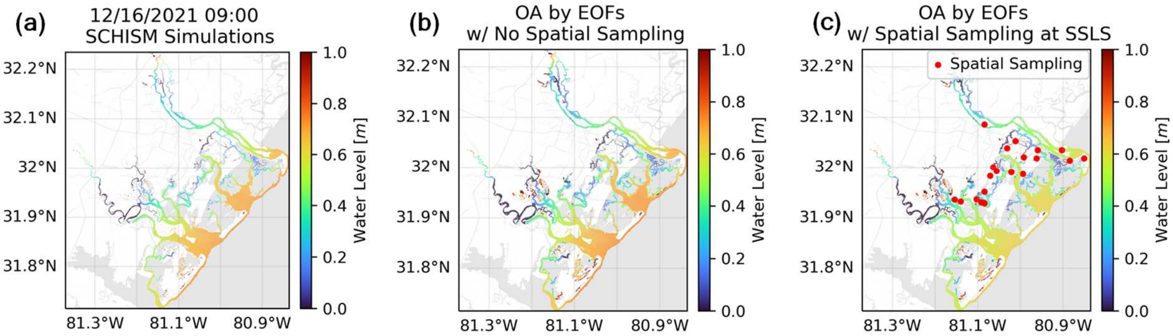

Fig. 7 Example plots of water levels at 12/16/2021 09:00 UTC during flood tide: (a) SCHISM model simulations; (b) EOFs-based reconstruction with all the simulated water levels shown in (a); (c) EOFs-based reconstruction with spatial sampling at 22 SSLS locations within the study boundaries (red circles).

Error statistics are calculated for the 30-day simulation periods, employing multiple evaluation metrics that are described in Appendix B. The SCHISM model simulations are regarded as a ground truth as the simulated water levels are currently used to compute the PCs using Eq. (8). Consequently, the evaluation metrics provide a means to assess the reconstruction capabilities of water levels, in comparison with an ideal EOFs-based OA procedure without errors. Fig. 8 shows the normalized root-mean-square errors (NRMSE) as defined in Eq. (B.2), using all information within the study boundaries (Fig. 8a) and using spatial sampling limited to 22 SSLS locations (Fig. 8b). The first NRMSE in Fig. 8a represents truncation errors resulting from using the three leading spatial modes of water levels, which remain below 0.05. However, it should be noted that some upstream channels have relatively higher NRMSE values, indicating limitations in capturing the dynamics only with a few EOFs due to complex interactions between water and estuarine landscapes. The second NRMSE in Fig. 8b corresponds to spatial sampling errors arising from the limited availability of observations, in addition to the truncation errors. The NRMSE values increase up to 0.1 for distant areas from the valid SSLS locations, particularly in proximity to the immediate coasts. The difference between the two NRMSEs, as shown in Fig. 8c, indicates the portion of the reconstruction errors originating solely from spatial sampling. The NRMSE differences are generally less than 0.04 but become more pronounced towards the outer study boundaries. This is because the majority of the SSLS deployments are centered around urban areas where coastal infrastructure is located. A similar discrepancy is also captured during flood tide, as observed in Fig. 7b and Fig. 7c.

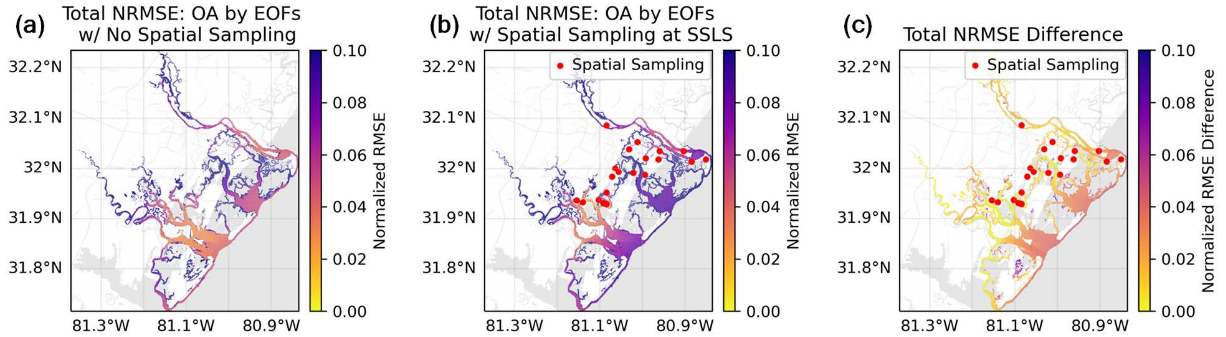

Fig. 8 NRMSE calculations based on the SCHISM model simulations as a ground truth: EOFs-based reconstructions (a) with the simulated water levels and (b) with spatial sampling at 22 SSLS locations within the study boundaries (red circles); (c) NRMSE difference between (a) and (b).

Table 2 summarizes the median and quantile values of other evaluation metrics that are defined in Appendix B. These metrics are applied to more than 53,000 SCHISM grid points within the study boundaries. Since the grid points are distributed unevenly depending on channels, the median and quantile values of the error statistics are selected to represent the relative trends in accuracy. The overall RMSE slightly increases due to the spatial sampling for 22 SSLS locations, as identified particularly for the lower quantile value from 0.039 *m* to 0.052 *m*. On the other hand, the MBE, R, and NSE show minor changes within acceptable ranges.

Table 2
Error statistics of water level reconstructions using the simulated water levels from the SCHISM model as a ground truth. For comparison, the median and quantile values are calculated from more than 53,000 SCHISM grid points.

Water level availability	**Spatial covariance**	**RMSE [*m*]**			**MBE [*m*]**			**R [-]**			**NSE [-]**		
		25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%
All locations	EOFs	0.037	0.055	0.088	0.000	0.000	0.000	0.982	0.994	0.999	0.963	0.988	0.997
22 SSLS locations	EOFs	0.050	0.064	0.083	0.000	0.000	0.000	0.978	0.993	0.997	0.956	0.987	0.994
	Gaussian	0.209	0.383	0.723	-0.223	-0.046	0.038	0.688	0.882	0.965	-1.426	0.551	0.899

[*] RMSE: root-mean-square errors, MBE: mean bias errors, R: correlation coefficients, and NSE: Nash-Sutcliffe efficiency.

Both Fig. 8 and Table 2 suggest that the current EOFs-based OA procedure has more prominent errors due to the leading-order truncations compared to those attributed to the limited availability of observations. Although adding higher-order EOF modes may lead to an improved representation of local dynamics, the corresponding observations in those local regions should be followed to accurately estimate the contributions through the PCs. With the error statistics, the accuracy and confidence maps can be generated to guide the practical applications of the EOFs-based OA procedure. In addition, the evaluation metrics can serve as a useful indicator to identify potential locations for additional sensor deployments to enhance the reconstruction capabilities of the OA procedure (Tien et al., 2023).

Based on a Gaussian covariance function, we perform similar reconstructions using the simulated water levels at 22 SSLS locations. As shown in Fig. 9a, the OA approach reproduces the simulated water levels for the utilized SSLS locations as expected. However, when excluding the sensor's own location for cross-validation (light blue dashed lines), the reconstructed water levels fluctuate with offsets and phase lags (S4 and S6). Similarly, as demonstrated in Fig. 9b, the NRMSE rapidly exceeds 0.1 for areas away from the valid SSLS locations. Table 2 also includes the error statistics of water level reconstructions using a Gaussian covariance function. Compared to the EOFs-based OA procedure, however, all evaluation metrics indicate inadequate performance, particularly with noticeable variations in RMSE (upper quantile value of 0.723 *m*) and NSE (lower quantile value less than zero).

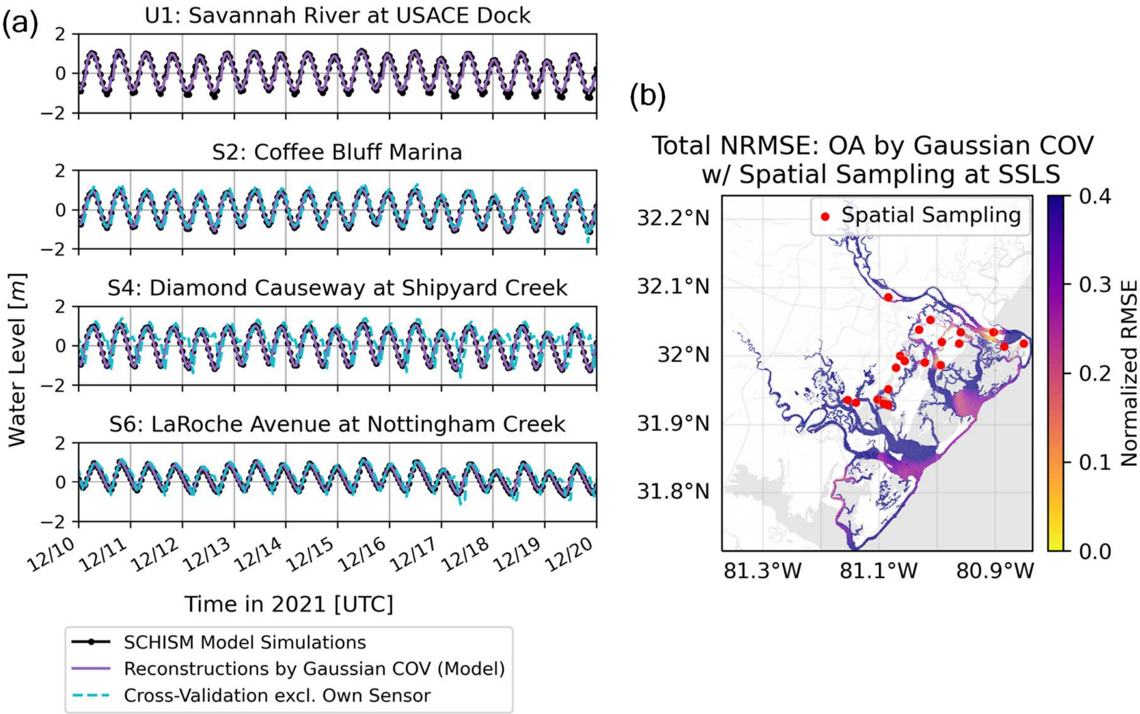

Fig. 9 OA results based on a Gaussian covariance function: (a) reconstructions of water levels at SSLS locations. The light blue dashed lines represent reconstructions with the sensor's own location excluded; (b) a spatial map of NRMSE.

The analytical OA procedure could be improved by adjusting the decorrelation scaling parameter or by incorporating a more refined analytical function to characterize spatial covariances (Verri et al., 2017). However, the large reconstruction errors can be attributed to the simplified representation of a Gaussian covariance function without physical information (Rong and San Liang, 2022), such as hydrologic connections. Additionally, using a single value for the decorrelation length scale may not be appropriate for representing the variable density of SSLS deployments across estuaries. Since the analytical OA procedure fails to accurately represent complex patterns of water levels in estuaries, we exclude its further applications in this study.

3.3. Model applications with SSLS observations

3.3.1. Validation for basis simulation periods

Using the SSLS observation datasets, we apply the OA procedure integrating the established structures of spatial covariance to assimilate water levels. It is important to note that the availability of SSLS observations depends on the operational status of individual sensors at the time of monitoring. Throughout the same simulation periods, a total of 12 SSLS devices were active although some sensors were sporadically inoperative. Fig. 10 shows comparisons of the reconstructed water levels (orange lines) with the SCHISM model simulations (black lines) and the available observation datasets (blue line for a NOAA tide gauge, green line for a USGS stream gauge, and red lines for SSLS observations). Additionally, cross-validation cases are included to evaluate the reconstruction sensitivity and robustness when excluding the corresponding SSLS measurements (light blue dashed line). The EOFs-based OA procedure enables providing access to water levels along estuarine channels even in the absence of direct observations, as demonstrated throughout the S3 periods and later portions of S1 and S6 (discontinued SSLS measurements after 12:00 on December 21). This capability to fill in missing measurements is further supported by the effective reconstructions of water levels without the SSLS's own observations. In cases where discrepancies exist between the SSLS observations and physics-based model simulations (S5 and S6), the reconstructed water levels closely track the measurements of the SSLS monitoring network. By minimizing error variances directly using actual SSLS observations, the EOFs-based OA procedure realistically determines the weights of spatial covariance structures, identifying the most probable spatial state of water levels. Therefore, the EOFs-based OA procedure has the advantage of leveraging empirical observations that closely approximate the true conditions of water levels.

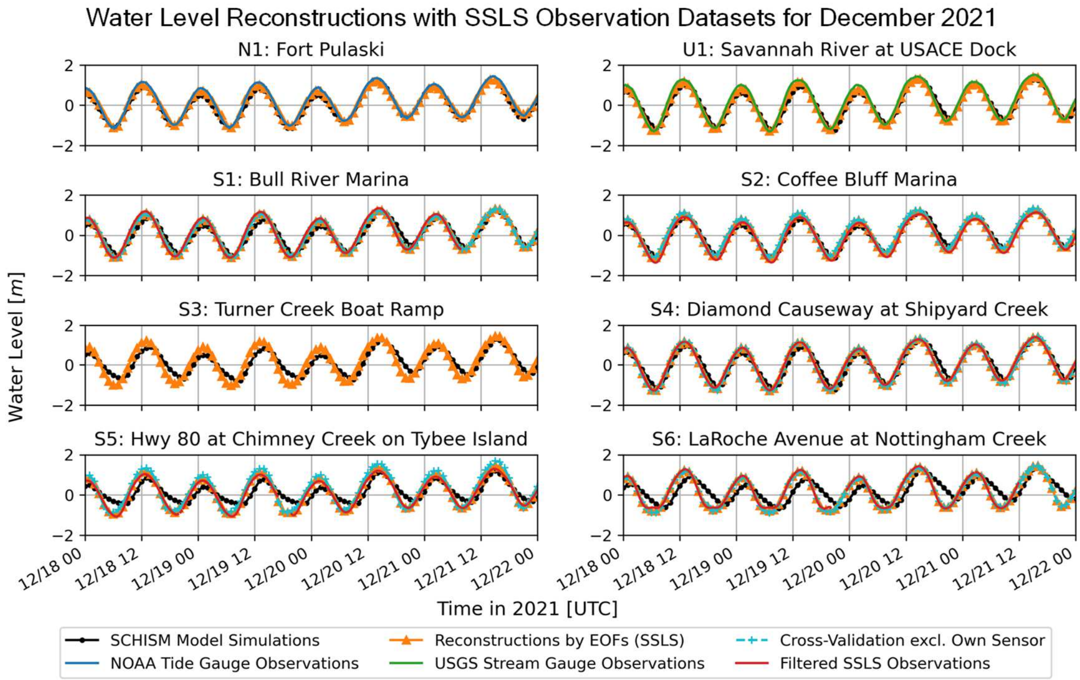

Fig. 10 Water level reconstructions using SSLS observation datasets for the basis simulation periods. The cross-validation cases (light blue dashed lines) correspond to excluding the sensor's own measurements for reconstructions. The model evaluations, including error statistics, are listed in Appendix C.

3.3.2. Application for Hurricane Dorian in 2019

Based on the spatial covariance information derived in Section 3.1, the assimilation of water levels is performed for Hurricane Dorian in 2019 to evaluate the capabilities of the OA procedure. Since the SSLS deployments in 2019, Hurricane Dorian has been identified as the most impactful hurricane that closely approached the Georgia coasts. As shown in Fig. 11, although the occurrence of the storm surge on September 05 did not coincide with the astronomical tidal peaks, the hurricane generated wind gusts of approximately 95 *km/hour* and induced a storm surge exceeding 1 *m* along the Georgia coasts (Avila et al., 2020). Consequently, the combined effects of the storm surge and prevailing winds may cause variability in water levels across the estuaries. The SSLS observation datasets are applied for the assimilation of water levels during Hurricane Dorian. 17 SSLS devices measured hyper-local water levels when the hurricane moved along the U.S. East Coast (September 02 to 06). In addition, numerical simulations are independently conducted using the SCHISM model for 30 days from August 20 to September 19, 2019. The model setup is the same as described in Section 2.2, except for using more refined atmospheric forcing obtained from the NOAA High-Resolution Rapid Refresh (HRRR) datasets.

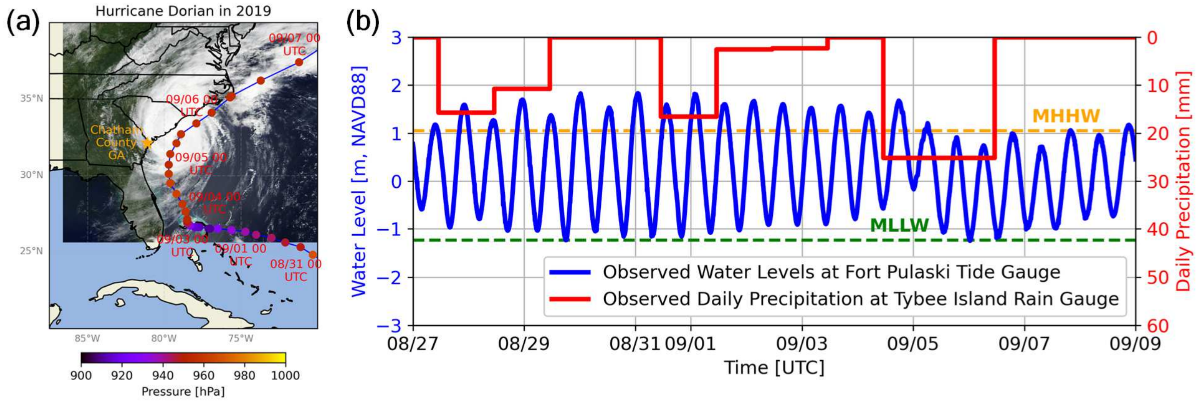

Fig. 11 Hurricane Dorian in September 2019: (a) hurricane track with intensity and (b) water levels (blue) at the NOAA Fort Pulaski tide gauge and daily precipitation (red) at the local rain gauge. Satellite image credit: MODIS Land Rapid Response Team and NASA GSFC (2019).

For comparison, we employ the USGS storm tide records (yellow placemarks in Fig. 1) that were collected during Hurricane Dorian, which include additional locations beyond the active SSLS positions. In Fig. 12, the assimilated water levels (orange lines) are compared with the USGS storm tide records (green lines), as well as the SCHISM model simulations (black lines). The USGS storm tide sensors consist of two different types: radar-based rapid deployment gauges that continuously measure water levels (L1-3) and pressure-based storm tide gauges that read water levels exceeding the sensor elevation thresholds (L4-10). The assimilated water levels for the L1 to L3 locations show a good agreement with the storm tidal cycles recorded by radar-based USGS sensors. Moreover, at the L4 to L10 locations, the peaks of assimilated water levels closely capture the instances where water levels temporarily exceed the sensor elevations, except for a slight overestimation at the L6 location. At the L8 and L10 locations where no water hit the sensors, both the assimilated and simulated water levels consistently remain below the sensor elevation thresholds.

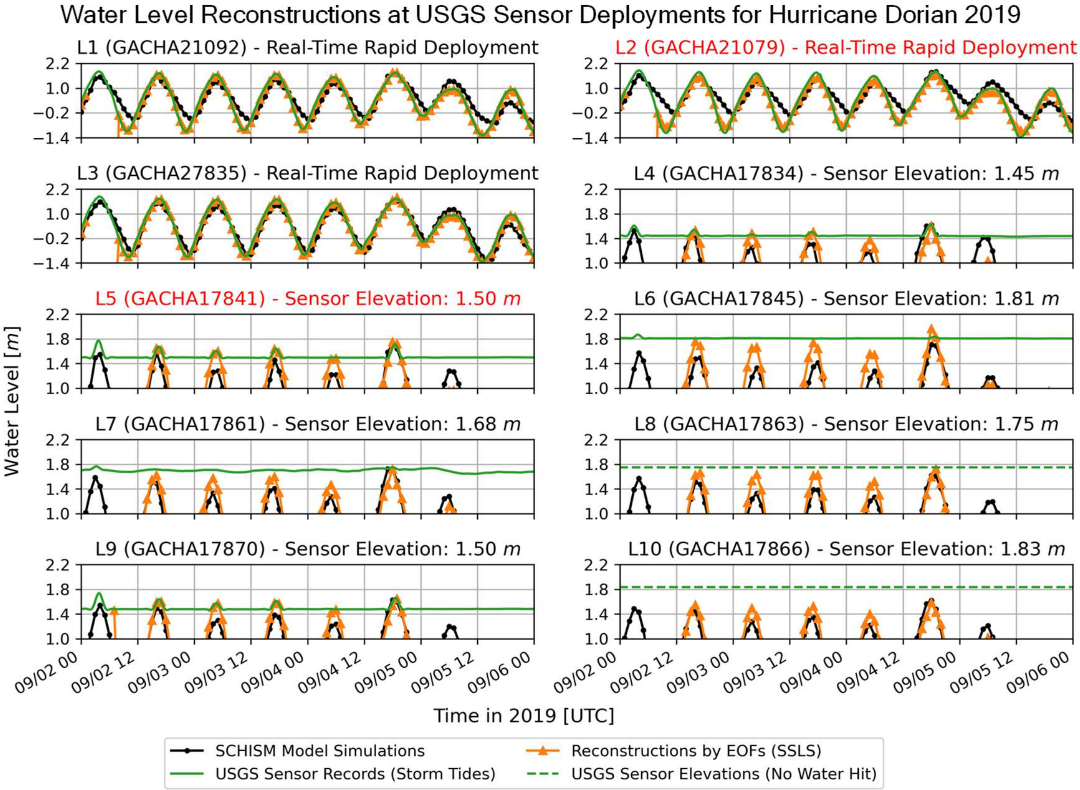

Fig. 12 Comparison of water levels at USGS storm tide monitoring locations during Hurricane Dorian in 2019. The recorded data from radar-based sensors (L1-3) have continuous measurements while those from pressure-based sensors (L4-10) provide actual readings only when water levels reach the specific elevation thresholds. The headings (L1-10) on the titles correspond to the labels in Fig. 1. The titles highlighted in red indicate the locations where the SSLS measurements are available nearby.

The application of the EOFs-based OA procedure reveals several limitations. During the initial comparison periods (09/02 02:00 to 08:00), no assimilation outputs are available due to the complete unavailability of observation datasets from the SSLS monitoring network. Although water levels can be directly assimilated based on at least one SSLS observation, there is currently a lack of information regarding the minimum number of SSLS observations required at different locations to ensure accuracy. In addition, it should be noted that the assimilated water levels are spatially confined to coastal channels and temporally limited to historical and real-time applications, in comparison with the simulation results. Nevertheless, the physics-based empirical modeling of water levels demonstrates its capability to reconstruct water levels along coastal channels in an acceptable manner, even for extreme weather events where various dynamic factors can influence water levels.

3.3.3. Real-time application in a web-based portal for emergency management

By adopting spatial covariance statistics derived from numerical simulations, the physics-based empirical modeling approach maintains computational efficiency, which enables real-time assimilation of coastal water levels and subsequent applications for monitoring coastal inundation. As demonstrated in Fig. 13, real-time visualizations of inundation depths are implemented within a pilot web-based portal for the Chatham Emergency Management Agency (CEMA), along with simultaneous access to SSLS observations. The operational platform uses the most recent 15-minute SSLS observations to calculate the real-time state of water levels, leveraging the established spatial covariance structures. Then, the corresponding inundation depths, which are of primary interest to public and emergency officials, are determined by subtracting the topobathymetry data. In addition, the NRMSE statistics (Fig. 8) are employed to extract areas with sufficient accuracy (e.g., NRMSE < 0.3). As a result, the spatial maps of water levels can provide information even for inactive sensor locations (gray circles) and areas beyond the immediate vicinity of sensor installations (in between circles). The expanded spatial coverages allow community officials to promptly identify potential flood threats and take appropriate actions for effective emergency management. Moreover, the increased access to real-time water levels contributes to enhancing situational awareness of floods in coastal communities.

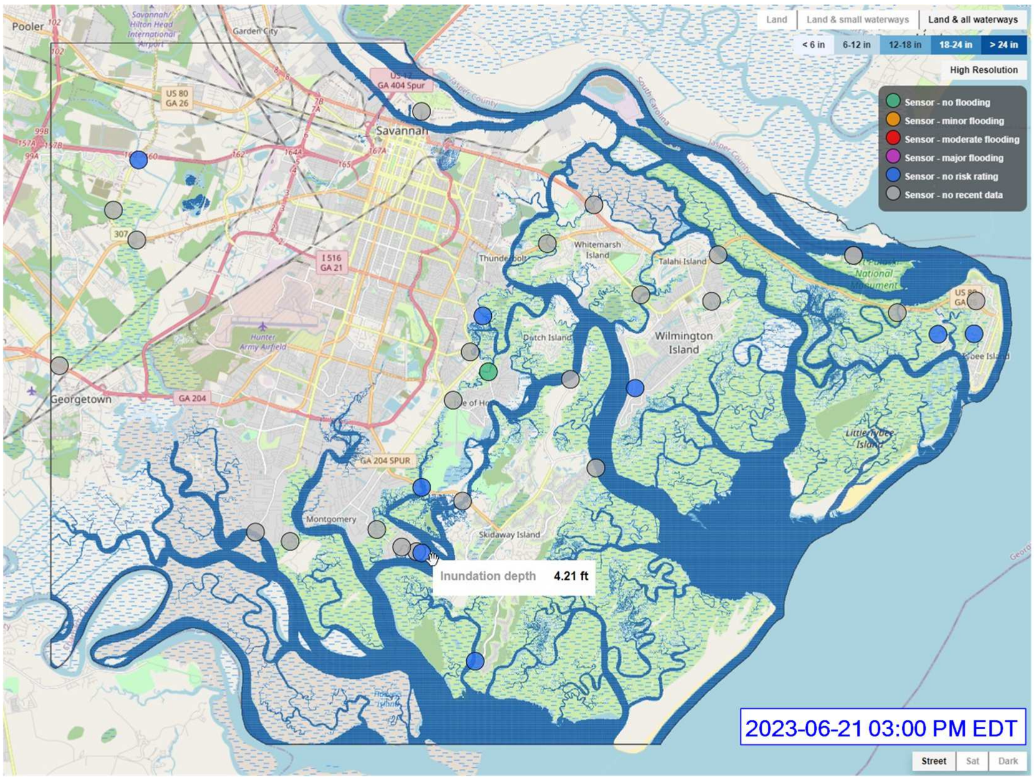

Fig. 13 Real-time application in a pilot web-based portal for the Chatham Emergency Management Agency in U.S. Georgia. The blue contour overlay depicts inundation depths, which represent water heights above topographic elevations. The circles with color indicate active SSLS locations at the time of monitoring.

4. Discussion

Our study demonstrates the implementation of the EOFs-based OA procedure to enhance hyper-local water level monitoring in complex estuarine settings, informed by geospatial statistics from physics-based numerical model simulations. The EOF analysis, which is applied to simulated water levels, enables the extraction of water level patterns that represent the complex dynamics of water flows and interactions within estuarine channel networks, as well as the substantial reduction of the output dimensionality into a few principal modes. Once spatial covariance structures are established through the EOF analysis, the integration with IoT-enabled sensor observations of water levels facilitates the generation of either a retrospective or real-time map of assimilated water levels (Fig. 13).

The enhanced access to hyper-local water level conditions can benefit community officials by providing reliable estimates of water levels beyond active sensor locations (e.g., blue or green circles in Fig. 13), extending to coastal infrastructure with neither sensor installation nor observation (e.g., S3 in Fig. 10 and gray circles in Fig. 13). By using a real-time map of assimilated water levels during uncertain flood situations, for example, public and emergency management officials can anticipate potential overtopping at different bridges and flooding at piers, docks, and boat ramps. Given the limited adaptability of coastal infrastructure in response to rising sea levels, the continuous monitoring of increased flood risks to coastal infrastructure is significant for coastal communities to develop strategies for long-term flood resilience (Habel et al., 2020; Son et al., 2023).

While coastal-ocean hydrodynamic models are capable of effectively simulating water levels (Fig. 2), the operation of forecasting systems demands substantial computational resources (Kerr et al., 2013; Bilskie et al., 2020), which may be prohibitive for coastal communities with limited access. Nonetheless, the EOFs-based OA procedure can efficiently provide timely updates on real-time water level information. This efficiency and utility are particularly evident during non-extreme flood events, when the effectiveness of applying a coastal-ocean hydrodynamic model is in question. Moreover, it is important to note that even calibrated numerical model simulations retain inherent uncertainties that persist throughout the simulation periods (e.g., S5 and S6 in Fig. 2) (Muñoz et al., 2022). The empirical modeling component of the EOFs-based OA procedure facilitates not only the emulation of available water level observations (e.g., S5 and S6 in Fig. 10) but also the generation of the most likely spatial state representation across estuaries. Furthermore, the differences between assimilated water levels and operational predictions produced by coastal-ocean hydrodynamic models can be integrated to update the initial or forcing conditions within operational forecasting system through additional assimilation steps (e.g., Madsen et al., 2015; Asher et al., 2019), which can consistently reduce uncertainties in water level predictions.

The geospatial evaluations of modeling errors (Fig. 8) play an important role in informing model accuracy and confidence across various areas and subsequently optimizing sensor deployments to constrain modeling errors. If these modeling errors are linked to a specific EOF mode, its prominent amplitudes and nodal structures can provide valuable insights into ideal locations for additional sensor deployments to ensure high signal-to-noise ratios. In the current modeling framework, which lacks weighting for different sensor locations (Eqs. (5) and (7)), adding a few sensor locations into the existing arrangement is less likely to reduce modeling errors noticeably. Therefore, rigorous optimization techniques, such as the Monte Carlo simulation, can provide a systematic approach to configuring optimal sensor deployments with different weighting. Furthermore, optimizing sensor deployments needs to take into consideration various practical factors. For instance, although the immediate coastal regions currently exhibit higher spatial sampling errors (Fig. 8c), most of these areas consist of extensive wetlands, which are little relevant to real-time flood monitoring. Additionally, other practical factors, such as social vulnerability priorities and potential flood impacts due to critical infrastructure failures, should be incorporated into the optimization process to maximize the marginal benefits of additional sensor deployments. For the SSLS monitoring network, Tien et al. (2023) applied a multi-objective optimization approach based on these practical criteria, in addition to the modeling errors of the OA technique.

There are several limitations identified in the applications of the EOFs-based OA procedure, which will be discussed in the subsections below.

4.1. Limitations

4.1.1. Hyper-local water level sensor observations

The EOFs-based OA procedure is inherently constrained by the limitations of hyper-local water level sensor observations in monitoring networks. Currently, the community-driven monitoring networks (Table 1) are designed to augment water level monitoring, primarily for high tides, storm surges, and rising sea levels. To address the sparse availability of observations from the NOAA tide gauges and USGS stream gauges in coastal communities, these monitoring networks deploy cost-effective, power-efficient sensor instruments, which are engineered to measure water levels at an operational cycle of 2- to 6-minute intervals (Observation Cycle column in Table 1), similar to the NOAA (6-minute intervals) and USGS (5-minute intervals) gauges. However, due to these observation cycles, the current sensor observations and their data-processing (e.g., Fig. 4) do not adequately capture water level fluctuations caused by certain flood drivers occurring at shorter timescales, including those associated with short and infragravity waves. To accurately resolve shorter-period water level fluctuations in observations and utilize observation datasets for reconstruction, it is necessary to implement a dedicated observation campaign and specialized operation protocols (e.g., Sweet et al., 2015), which differ from those of the community-driven monitoring networks.

4.1.2. EOFs: spatial covariance statistics

Our study uses a month-long dataset of simulated water levels for the EOF analysis to derive spatial covariance statistics (Fig. 5). The stationarity of the EOFs is evaluated with a 3-fold test, which compares the time-varying amplitudes of the EOFs computed from the different chunks of the simulation time series (Fig. 6). In our study, the EOFs-based OA procedure primarily aims to assimilate real-time water levels for extensive flood monitoring, including both nuisance flooding (e.g., high tides with rising sea levels) (Moftakhari et al., 2018; Li et al., 2021) and extreme flooding (e.g., storm surges). Consequently, the EOF and stationarity analysis focuses on identifying robust spatial covariance statistics from water level simulations over a longer period than typical flood timescales. This approach allows capturing quasi-static responses of water levels that predominantly balance in estuarine systems by forcings and facilitating real-time assimilation of water levels, rather than flood prediction under anticipated metocean conditions (e.g., Rohmer et al., 2023 using a surrogate model with multivariate statistical analysis). While our study can reproduce water level variations during a moderate storm surge event (e.g., Fig. 12 for Hurricane Dorian in 2019) using the established EOFs, the stationarity of the EOFs may not hold under different climates, seasons, and weather conditions. Therefore, future research should prioritize exploring the applicable ranges of the EOFs-based OA procedure in these diverse conditions. In particular, it is important to further understand the limitations of the modeling approach during various hurricane and hydrologic runoff events, which are recognized as extreme flood drivers for coastal communities. Then, researchers can identify specific conditions under which the modeling approach may perform less accurately and work towards improving its robustness in such scenarios. As operational forecasts of coastal-ocean hydrodynamic models become increasingly accessible across different climates, seasons, and weather conditions, future research should aim to provide a more comprehensive assessment of stationarity and improve representative spatial covariance statistics of water levels along estuarine channels.

Prior to the EOF analysis, the masking process is applied to prevent incomplete statistical analysis, which results in water level assimilations that are spatially confined to coastal channels. To address the spatial confinement, additional mapping techniques, such as downscaling extrapolations (e.g., Rucker et al., 2021), can be integrated to extend assimilated water levels from coastal channels into inland regions. Moreover, by imposing assimilated water levels as spatiotemporal boundary conditions, coastal-urban flood models can simulate subsequent flooding within urban systems (e.g., Karamouz et al., 2017;

Son et al., 2023). These integrations can enhance spatial coverages of assimilated water levels, thus facilitating more comprehensive and efficient flood predictions beyond coastal channels.

5. Summary and conclusions

With a growing number of water level monitoring networks, coastal communities have more access to real-time water level information at sensor deployment locations. These hyper-local water level observations enable public and emergency management officials to identify and assess localized flood threats to coastal infrastructure, particularly in areas that have limited capacity to adapt to sea level rise. In line with these community-driven efforts, our study develops a physics-based empirical modeling approach that can efficiently assimilate estuarine water levels from hyper-local sensor observations to enhance water level monitoring capabilities.

Our approach uses the OA procedure that combines hyper-local sensor observations with spatial covariance statistics of water levels in coastal regions. Prior to the OA procedure, outlier removal and smoothing techniques are implemented to mitigate the impacts of anomalous signals, including occasional spikes and fluctuations, in the hyper-local sensor observations (Fig. 4). The representative patterns for spatial covariances are established through the EOF analysis on numerical simulations using the high-resolution coastal-ocean hydrodynamic model, SCHISM. The significant spatiotemporal variations in water levels are decomposed into a combination of rising or falling movements, transitional phases between them, and adjustments in the upstream channel (Fig. 5), all of which interact within the complex estuarine settings. Through the 3-fold cross-validations, we demonstrate minor sensitivity to the selection of basis simulation periods for extracting the spatial mode patterns (Fig. 6). In addition, we investigate the assimilation modeling errors associated with the leading-order truncations of the spatial mode patterns and the limited availability of the SSLS observations. Although the truncation errors tend to be larger than the spatial sampling errors, the overall modeling errors remain within acceptable ranges (Fig. 8). When compared to the common OA procedure using a Gaussian covariance function, the EOFs-based OA procedure shows enhanced performance and robustness in assimilating water levels. The improvement is attributed to the inherent limitations of simplified distance-based analytical functions that inadequately capture the complex dynamics of estuaries, including hydrologic connections. Using the actual SSLS observations, the applications of the EOFs-based OA procedure successfully reproduce the temporal patterns and peak magnitudes of water levels along coastal channels for both the basis simulation periods (Fig. 10) and Hurricane Dorian (Fig. 12), demonstrating the capability of our approach to leverage the empirical observations in combination with physics-based statistical information. The assimilation results for Hurricane Dorian show promise in capturing water level states during unusual weather events.

However, further validations are necessary to evaluate the extensibility of the modeling approach for different storms with large forcing signals and significant impacts on water level variations.

With the established spatial covariance statistics, the real-time assimilation of water levels (Fig. 13) can provide community officials with a broader understanding of hyper-local water level conditions, including temporarily inactive monitoring locations and areas that are not covered by sensors. Thus, the assimilation data can serve as a supplementary tool for consistent monitoring of immediate flood risks to coastal infrastructure systems, such as bridges and marinas. Ultimately, coastal communities can make advanced use of water level monitoring networks by integrating the real-time assimilation framework to develop adaptive strategies for flood resilience planning in response to rising sea levels.

Appendix A

Simulated water levels for 30 days at Fort Pulaski, U.S. Georgia, for the EOF analysis

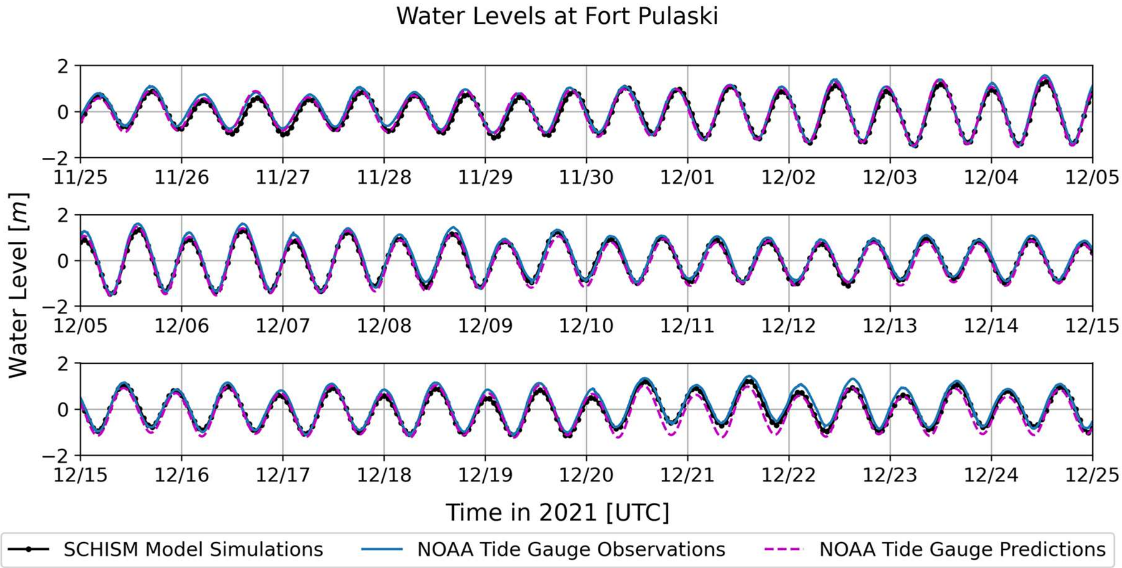

Fig. A.1 Simulated water levels for 30 days at Fort Pulaski, U.S. Georgia, for the EOF analysis and comparisons with NOAA tide gauge observations and predictions. The model evaluations, including error statistics, are listed in Appendix C.

Appendix B

Model evaluation metrics for error statistics

$$\text{Root-mean-square error, RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(s_t - o_t)^2} \tag{B.1}$$

$$\text{Normalized RMSE, NRMSE} = \frac{RMSE}{\sigma_o} \tag{B.2}$$

$$\text{Mean bias error, MBE} = \frac{1}{N}\sum_{t=1}^{N}(s_t - o_t) \tag{B.3}$$

$$\text{Pearson's correlation coefficient, R} = \frac{\sum_{t=1}^{N}(s_t-\bar{s})(o_t-\bar{o})}{\sqrt{\sum_{t=1}^{N}(s_t-\bar{s})^2\sum_{t=1}^{N}(o_t-\bar{o})^2}} \tag{B.4}$$

$$\text{Nash-Sutcliffe efficiency, NSE} = 1 - \frac{\sum_{t=1}^{N}(s_t-o_t)^2}{\sum_{t=1}^{N}(o_t-\bar{o})^2} \tag{B.5}$$

where s is the model estimations, o is the actual values, such as observations, and N is the number of samples. The $^{-}$ and σ denote the mean and standard deviation, respectively.

Appendix C

Model evaluations with observations for Fig. 2b, Fig. 10, and Fig. A.1

Table C.1

Model evaluations with observations for Fig. 2b, Fig. 10, and Fig. A.1 (11/25 – 12/25 in 2021). The table index (N1, U1, and S1-6) corresponds to the labels of the station locations in Fig. 1.

Observation Location	**SCHISM Model Simulation (Fig. 2b and Fig. A.1)**				**Physics-based Empirical Modeling Approach (Fig. 10)**							
					Reconstructions by EOFs (SSLS)				Cross-Validation excl. Own Sensor			
	RMSE [*m*]	MBE [*m*]	R [-]	NSE [-]	RMSE [*m*]	MBE [*m*]	R [-]	NSE [-]	RMSE [*m*]	MBE [*m*]	R [-]	NSE [-]
NOAA: N1	0.142	-0.129	0.989	0.947	0.084	-0.057	0.995	0.979	-	-	-	-
USGS: U1	0.222	-0.104	0.963	0.907	0.116	-0.110	0.998	0.977	-	-	-	-
SSLS: S1	0.312	-0.012	0.914	0.829	0.096	0.030	0.993	0.984	0.091	0.023	0.990	0.979
SSLS: S2	0.197	0.136	0.984	0.928	0.208	0.178	0.989	0.920	0.228	0.217	0.982	0.877
SSLS: S3	0.359	0.002	0.801	0.642	0.188	0.184	0.998	0.902	0.187	0.187	0.998	0.898
SSLS: S4	0.287	0.016	0.934	0.871	0.048	0.026	0.999	0.996	0.057	0.014	0.997	0.992
SSLS: S5	0.384	0.095	0.851	0.672	0.228	0.205	0.994	0.884	0.297	0.260	0.977	0.751
SSLS: S6	0.539	0.059	0.646	0.407	0.209	0.196	0.995	0.911	0.193	0.181	0.991	0.914

[*] RMSE: root-mean-square errors, MBE: mean bias errors, R: correlation coefficients, and NSE: Nash-Sutcliffe efficiency.

Data availability

The SSLS measurement data can be accessed publicly at https://www.sealevelsensors.org/. The DEM, water level, oceanographic, and meteorological data are available through the following links:

- DEM data: NOAA Digital Coast (https://coast.noaa.gov/dataviewer/)
- Water level data: NOAA Tide & Currents (https://tidesandcurrents.noaa.gov/), USGS National Water Dashboard (https://dashboard.waterdata.usgs.gov/), and USGS Flood Event Viewer (https://stn.wim.usgs.gov/)
- Oceanographic data: AVISO (https://www.aviso.altimetry.fr/en/data/products.html) and CEMES (https://data.marine.copernicus.eu/products)
- Meteorological data: ECMWF ERA 5 (https://www.ecmwf.int/en/forecasts/datasets) and NOAA HRRR (https://rapidrefresh.noaa.gov/hrrr/)

The SCHISM simulation and Objective Analysis data can be provided upon request to the corresponding author.

CRediT author statement

Youngjun Son: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Emanuele Di Lorenzo**: Conceptualization, Methodology, Investigation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Kyungmin Park**: Software, Validation, Writing – review & editing. **Spenser Wipperfurth**: Software, Validation, Writing – review & editing. **Jian Luo**: Methodology, Writing – review & editing, Supervision.

Acknowledgments

This research is funded by the NOAA Coastal Infrastructure and Resilience Research Initiative: The Georgia Coastal Equity and Resilience (CEAR) Hub project (NA22NOS4690219). We would like to thank Dr. Russell Clark, Kait Morano, Mangesh Atpadikar, Bhargav Singam, Vaishnavi Kulkarni at Georgia Institute of Technology, and Randall Mathews of the Chatham Emergency Management Agency (CEMA) for helping in establishing the real-time application in a pilot web-based CEMA portal. The authors are grateful to the anonymous reviewers for providing insightful and constructive feedback.

References

Al Kajbaf, A. and Bensi, M., 2020. Application of surrogate models in estimation of storm surge: A comparative assessment. Applied Soft Computing, 91. https://doi.org/10.1016/j.asoc.2020.106184

Allen, T.R. et al., 2018. Linking Water Infrastructure, Public Health, and Sea Level Rise: Integrated Assessment of Flood Resilience in Coastal Cities. Public Works Management & Policy, 24(1): 110-139. https://doi.org/10.1177/1087724x18798380

Asher, T.G., Luettich, R.A., Fleming, J.G. and Blanton, B.O., 2019. Low frequency water level correction in storm surge models using data assimilation. Ocean Model, 144. https://doi.org/10.1016/j.ocemod.2019.101483

Avila, L.A., Stewart, S.R., Berg, R. and Hagen, A.B., 2020. Hurricane Dorian 2019, National Hurricane Center. https://www.nhc.noaa.gov/data/tcr/AL052019_Dorian.pdf

Barnes, S.L., 1964. A Technique for Maximizing Details in Numerical Weather Map Analysis. Journal of Applied Meteorology, 3(4): 396-409. https://doi.org/10.1175/1520-0450(1964)003<0396:Atfmdi>2.0.Co;2

Bass, B. and Bedient, P., 2018. Surrogate modeling of joint flood risk across coastal watersheds. Journal of Hydrology, 558: 159-173. https://doi.org/10.1016/j.jhydrol.2018.01.014

Bilskie, M.V., Hagen, S.C. and Medeiros, S.C., 2020. Unstructured finite element mesh decimation for real-time Hurricane storm surge forecasting. Coast Eng, 156. https://doi.org/10.1016/j.coastaleng.2019.103622

Bilskie, M.V. et al., 2021. Enhancing Flood Hazard Assessments in Coastal Louisiana Through Coupled Hydrologic and Surge Processes. Frontiers in Water, 3. https://doi.org/10.3389/frwa.2021.609231

Bretherton, F.P., Davis, R.E. and Fandry, C.B., 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep Sea Research and Oceanographic Abstracts, 23(7): 559-582. https://doi.org/10.1016/0011-7471(76)90001-2

Chambers, D.P., Mehlhaff, C.A., Urban, T.J., Fujii, D. and Nerem, R.S., 2002. Low-frequency variations in global mean sea level: 1950-2000. J Geophys Res-Oceans, 107(C4). https://doi.org/10.1029/2001jc001089

Church, J.A. and White, N.J., 2006. A 20th century acceleration in global sea-level rise. Geophys Res Lett, 33(1). https://doi.org/10.1029/2005gl024826

Church, J.A., White, N.J., Coleman, R., Lambeck, K. and Mitrovica, J.X., 2004. Estimates of the Regional Distribution of Sea Level Rise over the 1950–2000 Period. Journal of Climate, 17(13): 2609-2625. https://doi.org/10.1175/1520-0442(2004)017<2609:Eotrdo>2.0.Co;2

Daley, R., 1991. Atmospheric Data Analysis. Cambridge Atmospheric and Space Science Serie. Cambridge University Press, 457 pp.

Gallien, T.W., Schubert, J.E. and Sanders, B.F., 2011. Predicting tidal flooding of urbanized embayments: A modeling framework and data requirements. Coast Eng, 58(6): 567-577. https://doi.org/10.1016/j.coastaleng.2011.01.011

Gold, A.C., Brown, C.M., Thompson, S.P. and Piehler, M.F., 2022. Inundation of Stormwater Infrastructure Is Common and Increases Risk of Flooding in Coastal Urban Areas Along the US Atlantic Coast. Earth's Future, 10(3). https://doi.org/10.1029/2021ef002139

Gorry, P.A., 1990. General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. Analytical Chemistry, 62(6): 570-573. https://doi.org/10.1021/ac00205a007

Grinsted, A., Ditlevsen, P. and Christensen, J.H., 2019. Normalized US hurricane damage estimates using area of total destruction, 1900-2018. Proc Natl Acad Sci U S A, 116(48): 23942-23946. https://doi.org/10.1073/pnas.1912277116

Habel, S., Fletcher, C.H., Anderson, T.R. and Thompson, P.R., 2020. Sea-Level Rise Induced Multi-Mechanism Flooding and Contribution to Urban Infrastructure Failure. Sci Rep, 10(1): 3796. https://doi.org/10.1038/s41598-020-60762-4

Hallegatte, S., Green, C., Nicholls, R.J. and Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. Nat Clim Change, 3(9): 802-806. https://doi.org/10.1038/nclimate1979

Hamlington, B.D., Leben, R.R., Nerem, R.S., Han, W. and Kim, K.Y., 2011. Reconstructing sea level using cyclostationary empirical orthogonal functions. Journal of Geophysical Research, 116(C12). https://doi.org/10.1029/2011jc007529

Hamlington, B.D., Leben, R.R., Wright, L.A. and Kim, K.Y., 2012. Regional Sea Level Reconstruction in the Pacific Ocean. Marine Geodesy, 35(sup1): 98-117. https://doi.org/10.1080/01490419.2012.718210

Ide, K., Courtier, P., Ghil, M. and Lorenc, A.C., 1997. Unified Notation for Data Assimilation : Operational, Sequential and Variational (gtSpecial IssueltData Assimilation in Meteology and Oceanography: Theory and Practice). Journal of the Meteorological Society of Japan. Ser. II, 75(1B): 181-189. https://doi.org/10.2151/jmsj1965.75.1B_181

Integrated Ocean Observing System, 2021. Manual for Real-Time Quality Control of Water Level Data. https://cdn.ioos.noaa.gov/media/2021/04/QARTOD_WaterLevelManual_Update_V2.1_Final.pdf

Jia, G. and Taflanidis, A.A., 2013. Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. Computer Methods in Applied Mechanics and Engineering, 261: 24-38. https://doi.org/10.1016/j.cma.2013.03.012

Jia, G. et al., 2015. Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms. Nat Hazards, 81(2): 909-938. https://doi.org/10.1007/s11069-015-2111-1

Kalnay, E., 2003. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, 341 pp.

Kaplan, A., Kushnir, Y. and Cane, M.A., 2000. Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854–1992*. Journal of Climate, 13(16): 2987-3002. https://doi.org/10.1175/1520-0442(2000)013<2987:Rsoioh>2.0.Co;2

Karamouz, M., Razmi, A., Nazif, S. and Zahmatkesh, Z., 2017. Integration of inland and coastal storms for flood hazard assessment using a distributed hydrologic model. Environmental Earth Sciences, 76(11). https://doi.org/10.1007/s12665-017-6722-6

Kerr, P.C. 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

Kim, K.Y., Hamlington, B.D. and Na, H., 2015. Theoretical foundation of cyclostationary EOF analysis for geophysical and climatic variables: Concepts and examples. Earth-Science Reviews, 150: 201-218. https://doi.org/10.1016/j.earscirev.2015.06.003

Kumar, P., Hamlington, B., Cheon, S.H., Han, W. and Thompson, P., 2020. 20th Century Multivariate Indian Ocean Regional Sea Level Reconstruction. Journal of Geophysical Research: Oceans, 125(10). https://doi.org/10.1029/2020jc016270

Kyprioti, A.P. et al., 2023. Spatio-temporal storm surge emulation using Gaussian Process techniques. Coast Eng, 180. https://doi.org/10.1016/j.coastaleng.2022.104231

Li, M., Wang, R.-Q. and Jia, G., 2020. Efficient dimension reduction and surrogate-based sensitivity analysis for expensive models with high-dimensional outputs. Reliability Engineering & System Safety, 195. https://doi.org/10.1016/j.ress.2019.106725

Li, S. et al., 2021. Evolving tides aggravate nuisance flooding along the U.S. coastline. Sci Adv, 7(10). https://doi.org/10.1126/sciadv.abe2412

Llovel, W., Cazenave, A., Rogel, P., Lombard, A. and Nguyen, M.B., 2009. Two-dimensional reconstruction of past sea level (1950–2003) from tide gauge data and an Ocean General Circulation Model. Climate of the Past, 5(2): 217-227. https://doi.org/10.5194/cp-5-217-2009

Loftis, D. et al., 2018. StormSense: A New Integrated Network of IoT Water Level Sensors in the Smart Cities of Hampton Roads, VA. Mar Technol Soc J, 52. https://doi.org/10.4031/MTSJ.52.2.7

Madsen, K.S., Hoyer, J.L., Fu, W. and Donlon, C., 2015. Blending of satellite and tide gauge sea level observations and its assimilation in a storm surge model of the North Sea and Baltic Sea. J Geophys Res-Oceans, 120(9): 6405-6418. https://doi.org/10.1002/2015jc011070

Marsooli, R., Orton, P.M., Georgas, N. and Blumberg, A.F., 2016. Three-dimensional hydrodynamic modeling of coastal flood mitigation by wetlands. Coast Eng, 111: 83-94. https://doi.org/10.1016/j.coastaleng.2016.01.012

Meyssignac, B. et al., 2011. Two-dimensional reconstruction of the Mediterranean sea level over 1970–2006 from tide gage data and regional ocean circulation model outputs. Global and Planetary Change, 77(1-2): 49-61. https://doi.org/10.1016/j.gloplacha.2011.03.002

MODIS Land Rapid Response Team and NASA GSFC, 2019. September 7, 2019 - Hurricane Dorian. https://modis.gsfc.nasa.gov/gallery/individual.php?db_date=2019-09-07 (Accessed 05/23/2023).

Moftakhari, H.R., AghaKouchak, A., Sanders, B.F., Allaire, M. and Matthew, R.A., 2018. What Is Nuisance Flooding? Defining and Monitoring an Emerging Challenge. Water Resour Res, 54(7): 4218-4227. https://doi.org/10.1029/2018wr022828

Moftakhari, H.R. et al., 2015. Increased nuisance flooding along the coasts of the United States due to sea level rise: Past and future. Geophys Res Lett, 42(22): 9846-9852. https://doi.org/10.1002/2015gl066072

Moore, F.C. and Obradovich, N., 2020. Using remarkability to define coastal flooding thresholds. Nat Commun, 11(1): 530. https://doi.org/10.1038/s41467-019-13935-3

Muis, S., Verlaan, M., Winsemius, H.C., Aerts, J.C. and Ward, P.J., 2016. A global reanalysis of storm surges and extreme sea levels. Nat Commun, 7: 11969. https://doi.org/10.1038/ncomms11969

Muñoz, D.F., Abbaszadeh, P., Moftakhari, H. and Moradkhani, H., 2022. Accounting for uncertainties in compound flood hazard assessment: The value of data assimilation. Coast Eng, 171. https://doi.org/10.1016/j.coastaleng.2021.104057

Muñoz, D.F., Muñoz, P., Moftakhari, H.R. and Moradkhani, H., 2021. From local to regional compound flood mapping with deep learning and data fusion techniques. Science of The Total Environment, 782. https://doi.org/10.1016/j.scitotenv.2021.146927

National Academies, 2014. Reducing Coastal Risk on the East and Gulf Coasts, Washington, DC, USA. https://nap.nationalacademies.org/catalog/18811/reducing-coastal-risk-on-the-east-and-gulf-coasts

NOAA NWS, 2016. National Water Model. https://water.noaa.gov/documents/wrn-national-water-model.pdf (Accessed 11/20/2022).

Pearson, R.K., Neuvo, Y., Astola, J. and Gabbouj, M., 2016. Generalized Hampel Filters. EURASIP Journal on Advances in Signal Processing, 2016(1). https://doi.org/10.1186/s13634-016-0383-6

Rohmer, J., Sire, C., Lecacheux, S., Idier, D. and Pedreros, R., 2023. Improved metamodels for predicting high-dimensional outputs by accounting for the dependence structure of the latent variables: application to marine flooding. Stochastic Environmental Research and Risk Assessment, 37(8): 2919-2941. https://doi.org/10.1007/s00477-023-02426-z

Rong, Y.N. and San Liang, X., 2022. An Information Flow-Based Sea Surface Height Reconstruction Through Machine Learning. Ieee Transactions on Geoscience and Remote Sensing, 60: 1-9. https://doi.org/10.1109/Tgrs.2022.3140398

Rucker, C.A. et al., 2021. Downscaling of real-time coastal flooding predictions for decision support. Nat Hazards, 107(2): 1341-1369. https://doi.org/10.1007/s11069-021-04634-8

Savitzky, A. and Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36(8): 1627-1639. https://doi.org/10.1021/ac60214a047

Smith, R.A.E., Bates, P.D. and Hayes, C., 2011. Evaluation of a coastal flood inundation model using hard and soft data. Environmental Modelling & Software. https://doi.org/10.1016/j.envsoft.2011.11.008

Smith, T.M., Reynolds, R.W., Livezey, R.E. and Stokes, D.C., 1996. Reconstruction of Historical Sea Surface Temperatures Using Empirical Orthogonal Functions. Journal of Climate, 9(6): 1403-1420. https://doi.org/10.1175/1520-0442(1996)009<1403:Rohsst>2.0.Co;2

Son, Y., Di Lorenzo, E. and Luo, J., 2023. WRF-Hydro-CUFA: A scalable and adaptable coastal-urban flood model based on the WRF-Hydro and SWMM models. Environmental Modelling & Software, 167. https://doi.org/10.1016/j.envsoft.2023.105770

Spicer, P., Schlichting, D., Huguenard, K., Roche, A.J. and Rickard, L.N., 2021. Sensing storm surge: A framework for establishing a citizen scientist monitored water level network. Ocean & Coastal Management, 211. https://doi.org/10.1016/j.ocecoaman.2021.105802

Sweet, W.V., Dusek, G., Obeysekera, J. and Marra, J., 2018. Patterns and Projections of High Tide Flooding along the U.S. Coastline using a Common Impact Threshold, National Oceanic and Atmospheric Administration. https://www.tidesandcurrents.noaa.gov/publications/techrpt86_PaP_of_HTFlooding.pdf

Sweet, W.V. et al., 2022. Global and Regional Sea Level Rise Scenarios for the United States: Updated Mean Projections and Extreme Water Level Probabilities Along U.S. Coastlines, National Oceanic and Atmospheric Administration. https://oceanservice.noaa.gov/hazards/sealevelrise/noaa-nostechrpt01-global-regional-SLR-scenarios-US.pdf

Sweet, W.V., Park, J., Gill, S. and Marra, J., 2015. New ways to measure waves and their effects at NOAA tide gauges: A Hawaiian-network perspective. Geophys Res Lett, 42(21): 9355-9361. https://doi.org/10.1002/2015gl066030

Tahvildari, N. and Castrucci, L., 2021. Relative Sea Level Rise Impacts on Storm Surge Flooding of Transportation Infrastructure. Natural Hazards Review, 22(1). https://doi.org/10.1061/(asce)nh.1527-6996.0000412

Tien, I., Lozano, J.-M. and Chavan, A., 2023. Locating real-time water level sensors in coastal communities to assess flood risk by optimizing across multiple objectives. Communications Earth & Environment, 4(1). https://doi.org/10.1038/s43247-023-00761-1

Ubelmann, C., Klein, P. and Fu, L.-L., 2015. Dynamic Interpolation of Sea Surface Height and Potential Applications for Future High-Resolution Altimetry Mapping. Journal of Atmospheric and Oceanic Technology, 32(1): 177-184. https://doi.org/10.1175/jtech-d-14-00152.1

Verri, G. et al., 2017. A meteo-hydrological modelling system for the reconstruction of river runoff: the case of the Ofanto river catchment. Nat Hazard Earth Sys, 17(10): 1741-1761. https://doi.org/10.5194/nhess-17-1741-2017

Ye, F. et al., 2020. Simulating storm surge and compound flooding events with a creek-to-ocean model: Importance of baroclinic effects. Ocean Model, 145. https://doi.org/10.1016/j.ocemod.2019.101526

Zhang, Y. and Baptista, A.M., 2008. SELFE: A semi-implicit Eulerian–Lagrangian finite-element model for cross-scale ocean circulation. Ocean Model, 21(3-4): 71-96. https://doi.org/10.1016/j.ocemod.2007.11.005

Zhang, Y.J., Ye, F., Stanev, E.V. and Grashorn, S., 2016. Seamless cross-scale modeling with SCHISM. Ocean Model, 102: 64-81. https://doi.org/10.1016/j.ocemod.2016.05.002

Zhang, Y.J. et al., 2020. Simulating compound flooding events in a hurricane. Ocean Dynamics, 70(5): 621-640. https://doi.org/10.1007/s10236-020-01351-x