

Assessing the Impacts of Coastal Flooding on Treaty of Olympia Infrastructure

*A Report to the Quinault Indian Nation, Hoh Tribe, and Quileute Tribe
Prepared by the Oregon Climate Change Research Institute*

May 2019



Assessing the Impacts of Coastal Flooding on Treaty of Olympia Infrastructure

A Report to the Quinault Indian Nation, Hoh Tribe, and Quileute Tribe

May 2019

Prepared by:

Peter Ruggiero, Interim Director
Oregon Climate Change Research Institute
College of Earth, Ocean, and Atmospheric Sciences
104 CEOAS Admin Building
Oregon State University
Corvallis, OR 97331

Contributors:

Katherine Serafin, Stanford University
Kai Parker, Universidad Técnica Federico Santa Maria, Chile
David Hill, Oregon State University

Suggested Citation:

Ruggiero, P., K. Serafin, K. Parker, and D. Hill. 2019. *Assessing the Impacts of Coastal Flooding on Treaty of Olympia Infrastructure*. A report to the Quinault Indian Nation, Hoh Tribe, and Quileute Tribe. Oregon Climate Change Research Institute, Corvallis, OR.

Acknowledgements:

This work was funded by contracted grant with the Treaty of Olympia area tribal governments (Quinault Indian Nation, Hoh Indian Tribe, and Quileute Tribe). Support was also provided by the National Oceanic and Atmospheric Administration (NOAA) Regional Integrated Sciences and Assessments Program (NA15OAR4310145) through the Pacific Northwest Climate Impacts Research Consortium. Tide gauge records are available through the NOAA National Ocean Service (NOS) website and river discharge is available through the U.S. Geological Survey National Water Information System (<https://waterdata.usgs.gov/wa/nwis/rt>). Bathymetric and topographic data were obtained from NOAA's Elevation Data viewer (DEMs). Thank you to Michael Rossotto and Garrett Rasmussen for providing updated shapefiles of the Quileute Reservation boundaries.

Cover photo: Coastal Flooding in Ocean Shores, WA due to extreme water levels at the Mouth of Grays Harbor, December 1, 2017; credit – Nick Bird, City of Ocean Shores

Table of Contents

Introduction	1
Objective 1: Assess the Relative Contributions of the Various Processes that Drive Extreme Coastal Total Water Levels	2
1.1 Emulating Extreme Water Levels in Grays Harbor.....	2
1.2 Hybrid Modeling of Compound Flooding along the Quillayute River	7
Objective 2: Quantify the Impact of a Range of Climate Change Scenarios on Each of the Various Components and on the Resulting Combined Total Water Levels.....	12
2.1 Climate Change Impact Scenarios.....	12
2.2 Impact of Climate Change on Emulated Grays Harbor Water Levels	15
2.3 Impact of Climate Change on Simulated Quillayute River Total Water Levels	21
Objective 3: Assess the Impact of Present Day and Forecasted Future Coastal Flooding Events on Infrastructure in Several Communities within the Treaty of Olympia.....	23
3.1 Impact of Flooding on Infrastructure.....	23
3.2 Climate Change Impacts, Recommendations for Adaptation, and Suggestions for Future Research.....	28
References	31
Appendix A.....	33
Appendix B.....	49
Appendix C.....	84

Introduction

Extreme coastal total water levels (TWLs) that result in flooding are the result of the complex interactions between multiple oceanographic, hydrological, geological, and meteorological forcings that act over a wide range of scales (i.e., astronomical tide, wave set-up, wind set-up, large-scale storm surge, precipitation, fluvial discharges, monthly mean sea level, vertical land motions, etc.). Coastal flooding that occurs during extreme TWLs can significantly impact communities and infrastructure resulting in substantial economic losses, even threatening human lives. Climate change may cause an increase in extreme coastal water level events driven by rising sea levels and changing patterns of storminess. An improved understanding of the physical processes during extreme coastal water level events will ultimately lead to an improved ability to predict the present day and future hazards faced by coastal communities. This information, in turn, provides the foundation for building more resilient coastal communities.

The primary objectives of this project were to 1) assess the relative contributions of the various processes that drive extreme coastal TWLs; 2) quantify the impact of a range of climate change scenarios on each of the drivers and on the resulting combined TWLs; and 3) assess the impact of present-day and forecasted future coastal flooding events on infrastructure in several communities within the Treaty of Olympia area.

This report is organized as follows: Sections 1, 2, and 3 highlight the primary results of each project objective listed above. Section 3 ends with a specific discussion of climate change impacts and some possible recommendations for adaptation. The details of the original modeling approaches specifically developed for this study have been published in the peer-reviewed literature and are given in Appendices A and B. Finally, the overarching results of a completed outer coast vulnerability assessment (Chapter 5 of Dalton et al., 2016) are reproduced in Appendix C.

Objective 1: Assess the Relative Contributions of the Various Processes that Drive Extreme Coastal Total Water Levels

Coastal flooding in the Pacific Northwest is often controlled by compound events, in which individual processes, which may or may not be extreme, combine to create extreme events (Zscheischler et al., 2018). For example, storms often generate large waves, heavy precipitation driving increased streamflow, and high storm surges, making the relative contribution of the actual drivers of extreme water levels difficult to interpret. Until recently, the compound nature of coastal flooding was not sufficiently recognized and there remains a paucity of scientific approaches for accurately characterizing risk under a compound flooding regime. Therefore, this project focused on developing new approaches for modeling and evaluating the relative contributions of the various processes that combine to generate high water levels in coastal rivers and estuaries. Due to the difference in environmental settings across the region, two new approaches have been developed, the methodologies of which have been published in the peer-reviewed literature as Parker et al. (2019) and Serafin et al. (2019) (Appendix A and Appendix B, respectively).

1.1 Emulating Extreme Water Levels in Grays Harbor

The first new approach involved developing a technique for modeling spatially variable water levels in the Grays Harbor, Washington area, per personal communication with Carolyn Kelly who encouraged a focus on this area due to the Quinault Marina and RV park, the Quinault Beach Resort and Casino, and other tribal assets in the area. To do this, a coupled hydrodynamic and spectral wave model, “Advanced Circulation Model-Simulating Waves Nearshore” (ADCIRC-SWAN), which has the ability to isolate the influence of processes such as wind set-up, pressure (through the inverse barometer effect), and wave set-up, was used (modeling grid shown in Figure 1). ADCIRC-SWAN has been extensively validated in coastal areas across the world as well as specifically in the Pacific Northwest. It is one of the leading tools for modeling flooding and coastal circulation and thus was an ideal approach to use for this project.

During model development, ADCIRC-SWAN results were first validated by confirming that the model accurately calculated observed water levels within the bay during past storm events (see Parker et al. 2019; Appendix A). In order to evaluate the contribution of processes to many different types of storm events, thousands of iterations of storm conditions would have to be simulated, which would be time consuming due to the computational expense of the model. Thus, this approach developed a ‘surrogate’ model (or emulator—see Parker et al., 2019 [Appendix A] for more details) in which any possible combination of forcing conditions (e.g., medium waves but high winds and stream flow) can be examined much more rapidly to determine whether or not they may cause flooding. To this end, once the ADCIRC-SWAN model was validated, ~500 ADCIRC-SWAN simulations representing a range of forcing conditions (taking months of computer server time) were generated to replace model runs with a spatially variable statistical representation of water levels in the bay from specified forcing conditions.

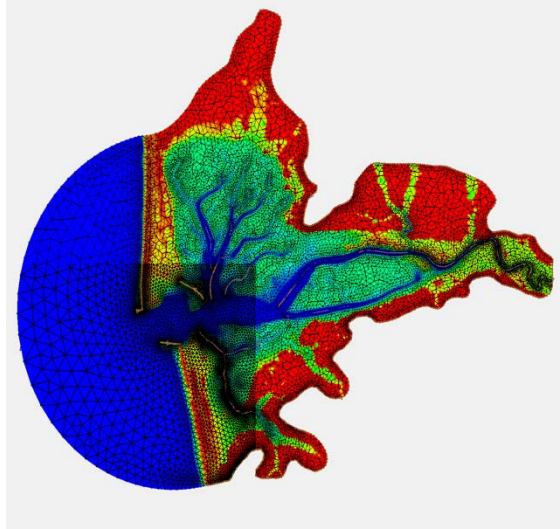


Figure 1 ADCIRC-SWAN model grid for the Grays Harbor, WA area.

This modeling approach provides a unique ability to examine the spatial variability of forcing conditions that contribute to extreme water levels in Grays Harbor. Figure 2 displays East-West and North-South transects of the average relative contribution of each forcing component during annual maxima events for the length of the record (1980s to approximately the present day). The diverse mix of processes that contribute to extreme water levels (WL) at each location, confirm that extreme WLs in Grays Harbor are compound in nature. The mean contribution of each forcing dimension to the extreme WL is significant, suggesting that including all forcing processes is necessary for the proper estimation of the magnitude of extreme WLs. The only exception is streamflow from the Chehalis River, which is found to be nominally important except near the streamflow boundary. This result is likely specific to the Grays Harbor estuary, which has a large estuary volume in comparison to the streamflow input and would be different for a more hydrologically dominated estuary system, which is further explored in Section 1.2.

The relative contribution of each variable to extreme WLs spatially varies across the estuary, leading to both an East-West and North-South gradient in the magnitude and drivers of extreme WLs. For example, waves significantly contribute to annual maximum WLs but only at stations in the northern and eastern reaches of the bay. This is likely due to little to no wave breaking induced setup occurring at the bay's entrance channel. The influence of wind on extreme WLs increases to the north due to the mean wind direction emanating from the south during storm events. The contribution of streamflow to extreme WLs decreases towards the west, moving farther away from the estuary's streamflow inlet. Finally, the influence of pressure anomalies on extreme WLs is found to be uniform. This result is likely from the spatial simplification of sea level pressure fields during the numerical procedures.

Not shown in Figure 2 is the contribution of tidal forcing to extreme WLs. This is primarily for scale reasons as the tidal component is an order of magnitude larger than any other forcing (average of 140 cm). Similar to previous results, tides also vary across the estuary (Figure 2). The contribution of tides to extreme WLs decreases by about 30 cm moving from the center of the estuary toward the North or East.

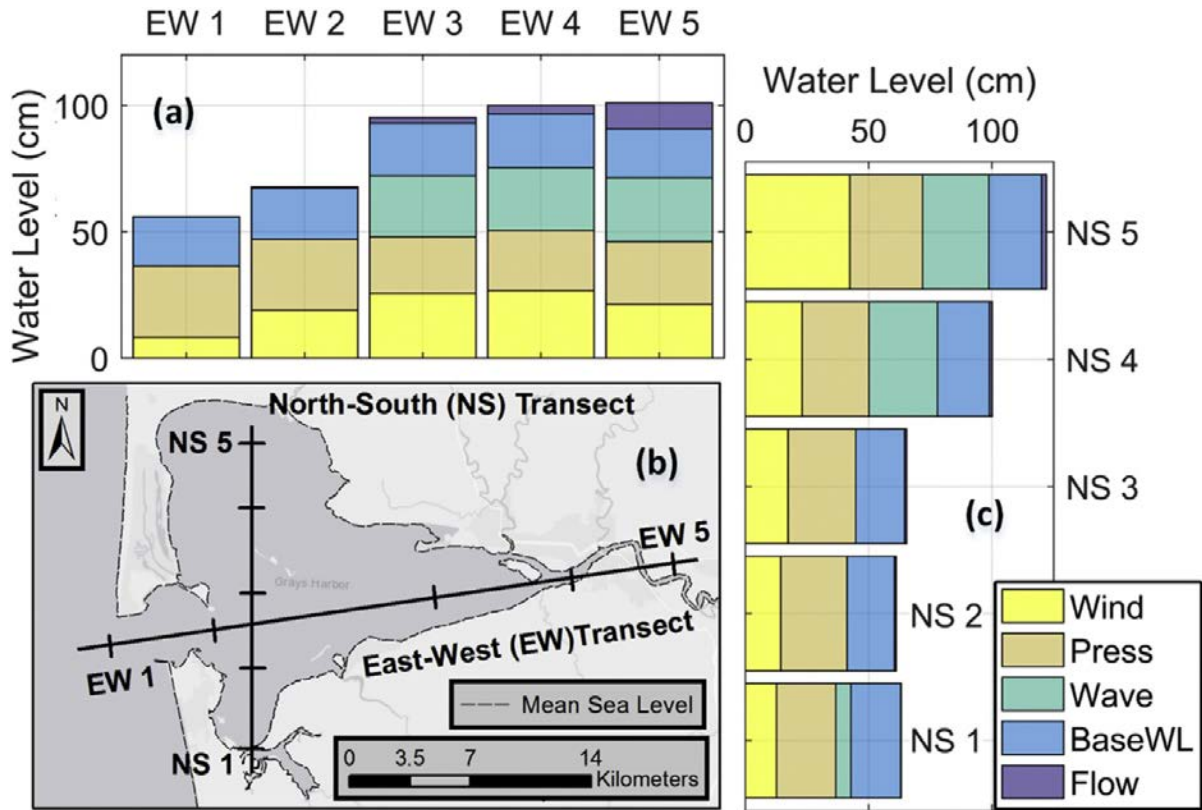


Figure 2 Average water level (WL) contribution from forcing components during extreme events (maximum annual WLs). Two transects are plotted with subplot (b) showing plotted transects, (East-West, EW) and (North-South, NS), with station locations marked as ticks. Tick locations are approximate (within 1 km) to scattered station locations. Subplot (a) is the East-West transect and subplot (c) is the North-South transect.

While this analysis shows important results for Grays Harbor as a whole, three specific locations were chosen for additional analysis due to their specific relevance to the Quinault Indian Nation. Figure 3 shows the locations of the 111 emulators that were constructed from the ADCIRC-SWAN simulations. The three locations chosen for additional analysis are plotted as a green star (Westport tide gage), pink star (near the Quinault Marina and RV park), and a cyan star (immediately offshore of the Quinault Beach Resort and Casino).

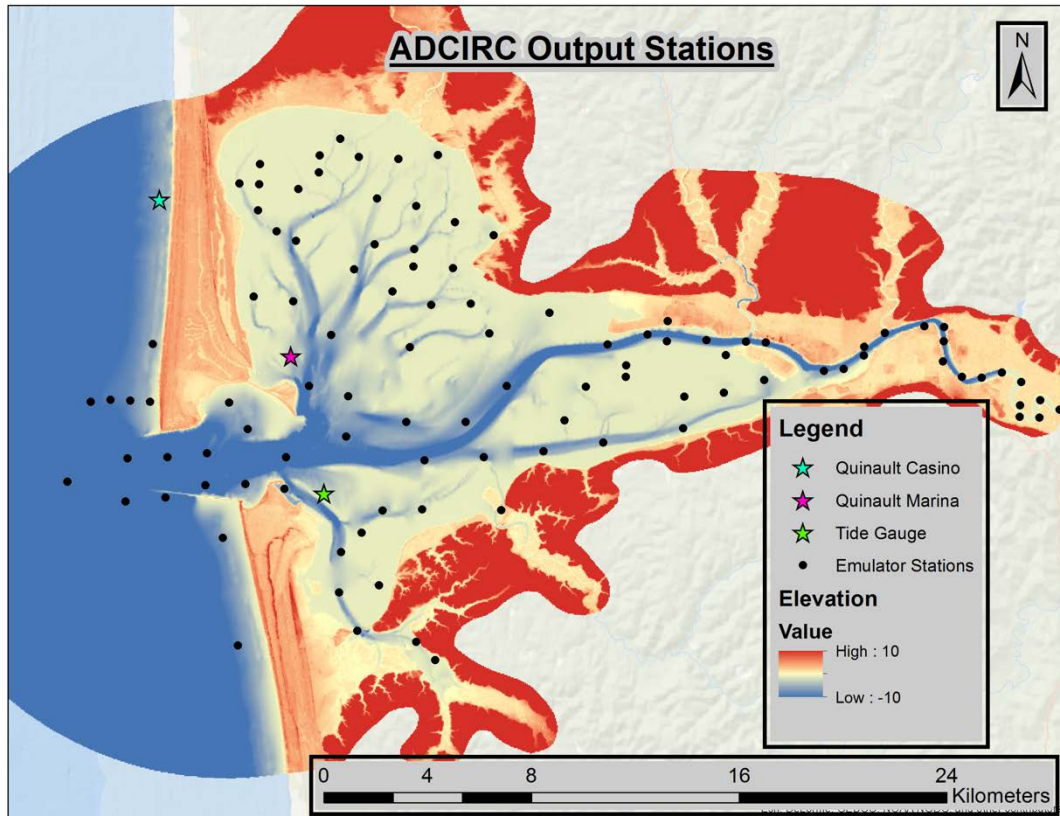


Figure 3 Location of emulators developed from the Grays Harbor ADCIRC model.

Due to differences in open coast and sheltered coast settings, it is expected that locations within and outside of the estuary would have a different composition of forcing variables contributing to coastal flooding. Figure 4 shows pie charts of contributions to extreme events at the Casino location (located along the open coast) and the Marina (within the estuary). Forcing is defined in these plots as “Tide” (deterministic and astronomically controlled), “Stream” (streamflow from the Chehalis River), “Base WL” (variations to local sea level as controlled by seasonality, monthly mean sea level anomalies, etc.), “Wave”, “Wind”, “Press” (pressure), and “LeftOvers” (nonlinear interactions between processes). These plots only show the percentage contribution from each forcing variable, not the magnitude of those contributions

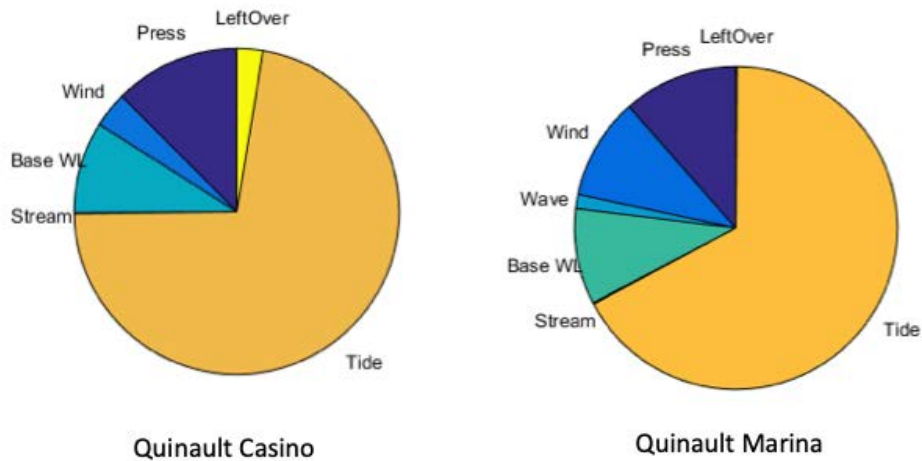


Figure 4 Contributions to extreme water levels at the Quinault Casino (left) and Marina (right)

From Figure 4 we can see that drivers of extreme WLs vary between these two locations. In this analysis, waves are an important component of extreme WLs at the Marina location while not at the Casino location. This is due to the fact that the Casino location is just offshore of the open coast beach and not under the influence of wave runup. This result is sensitive to the specific location of the station offshore, and a location closer to shore would likely show the effect of wave setup. For this reason, a total water level approach is more applicable for understanding hazard risk at this location (see section 2.2). Overall, storm surge forcing is more important at the Marina location while tides are a larger proportional contributor at the Casino. Additionally, we see that streamflow is found to be a negligible contributor in both locations.

The variability of the contribution of each process to extreme WLs for all events is shown in Figure 5 for the Marina location. Here, boxplots allow for a visual representation of how data in a sample (in this case, 30 years of extreme events), varies. The central line within the box represents the median while the bottom and top of the box represents the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme points not considered outliers.

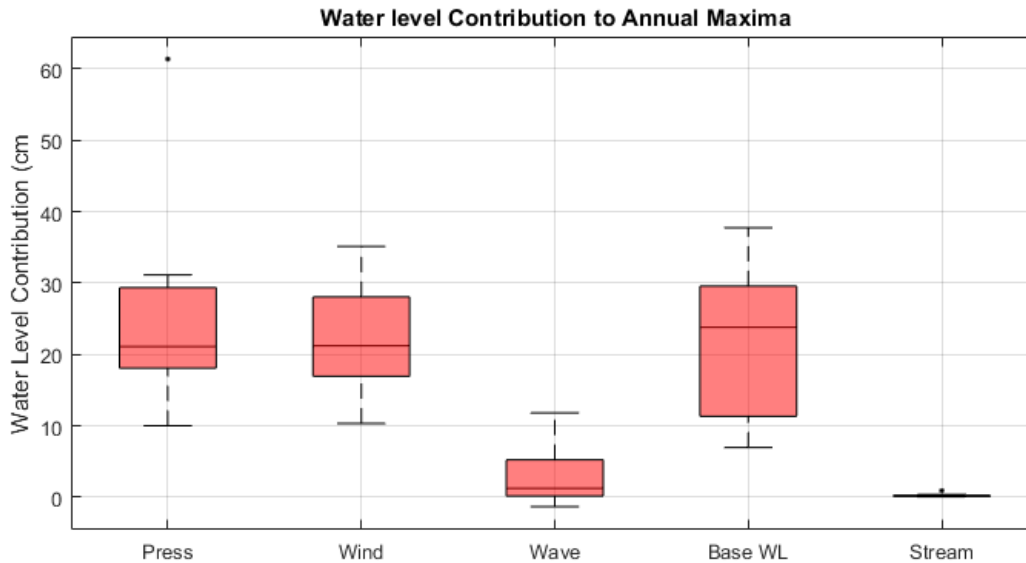


Figure 5 Boxplots showing the variability of contribution to extreme water level (WL) at the Marina location.

Figure 5 reveals that there is significant variability in what contributes to an extreme event at the Marina (shown by the height of the boxes). While one event may be wind-driven, another may be pressure or base WL-driven. This further reinforces the concept that extreme events at the study site are compound in nature. An event-based approach, like that commonly taken in hurricane influenced regions, is not likely applicable in the Pacific Northwest as it is very difficult to know a priori which events will cause coastal flooding. Furthermore, this interconnected nature of flooding means that significant care must be taken looking at climate change, which may simultaneously affect all drivers of coastal flooding as well as the relationships between them.

1.2 Hybrid Modeling of Compound Flooding along the Quillayute River

The proximity of communities like La Push and Taholah to both the ocean and river make them prone to flooding from high tides, coastal storms driving storm surges and large waves, and high streamflow events. While this is a similar situation to that explored in Grays Harbor in terms of compound forcing, these locations have a smaller estuary size (in terms of volume) and larger relative streamflow input. Furthermore, the storm surge signal measured by the tide gauge at La Push tends to coincide with high discharge events (Figure 6). This is likely to be similar at Taholah, but the limited observational data sets at this location makes it difficult to confirm.

It is hypothesized that streamflow is an important contributor to flooding in La Push. To test the hypothesis of strong fluvial control during extreme water level events, an ADCIRC model, similar to that developed in Grays Harbor (section 1.1), was constructed to model the largest streamflow event on record. Results from this modeled storm event show that the simulation including only river streamflow and tides is nearly able to recreate the measured peak storm surge signal measured at the tide gauge. This confirms that streamflow can contribute to high water levels at the study site. This simulation also confirms that the use of ADCIRC for this type of problem is limited since model simulations were found to be very unstable.

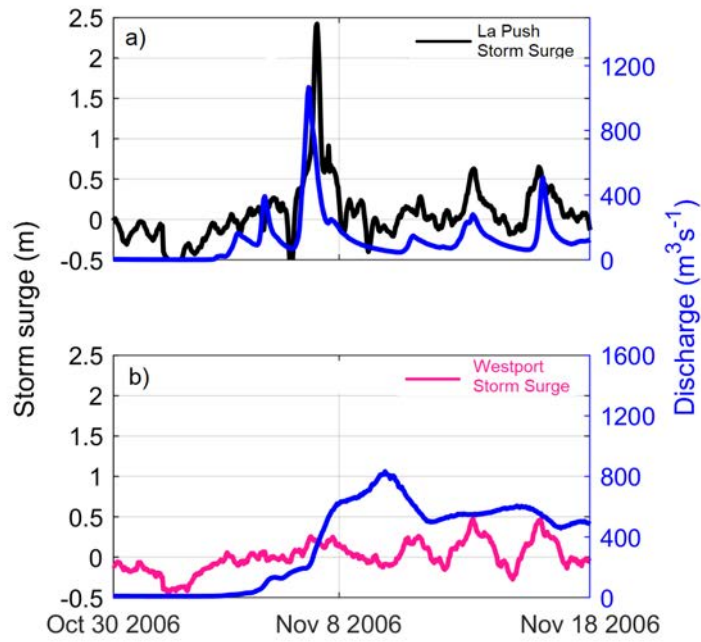


Figure 6 Example storm surge and river discharge relationship at a) La Push, Washington and b) Westport, Washington.

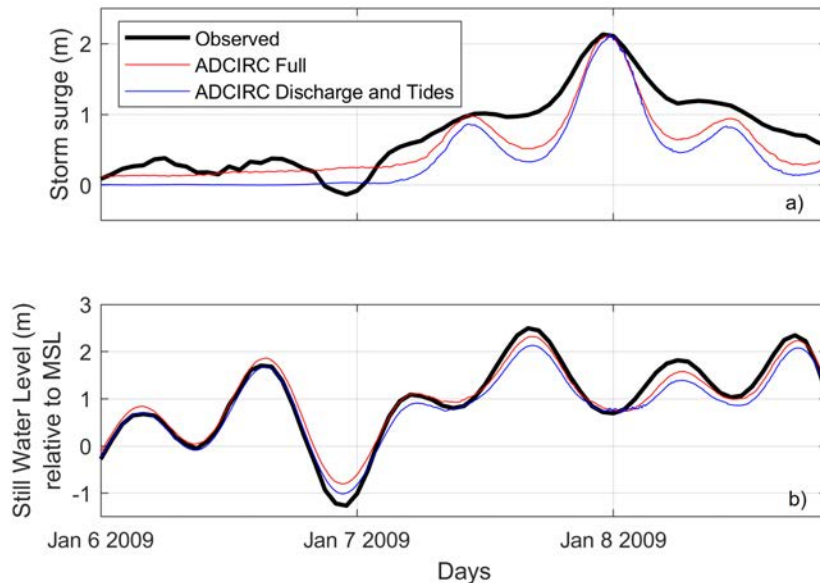


Figure 7 Simulated storm surge (a) and still water level (SWL) (b) at the La Push tide gauge. The full forcing simulation is plotted as a red line, a simulation using only streamflow and tides is plotted a blue line, and the observed storm surge or SWL is plotted as a black line.

While several Treaty of Olympia tribal communities experience these issues, our second modeling approach focused on the Quillayute River and the community of La Push (Figure 8) due to the availability of data for sufficient model calibration and validation. Early in the project we proposed to collect relevant data in the Quinault River to develop the data sets needed to build models at that location, but eventually it was decided that this was not feasible (personal communication with Carolyn Kelly). In general, some of the overall lessons learned from the Quillayute River analysis described below could be applied to other characteristically similar estuary systems, like the Quinault, Queets, and Hoh Rivers.

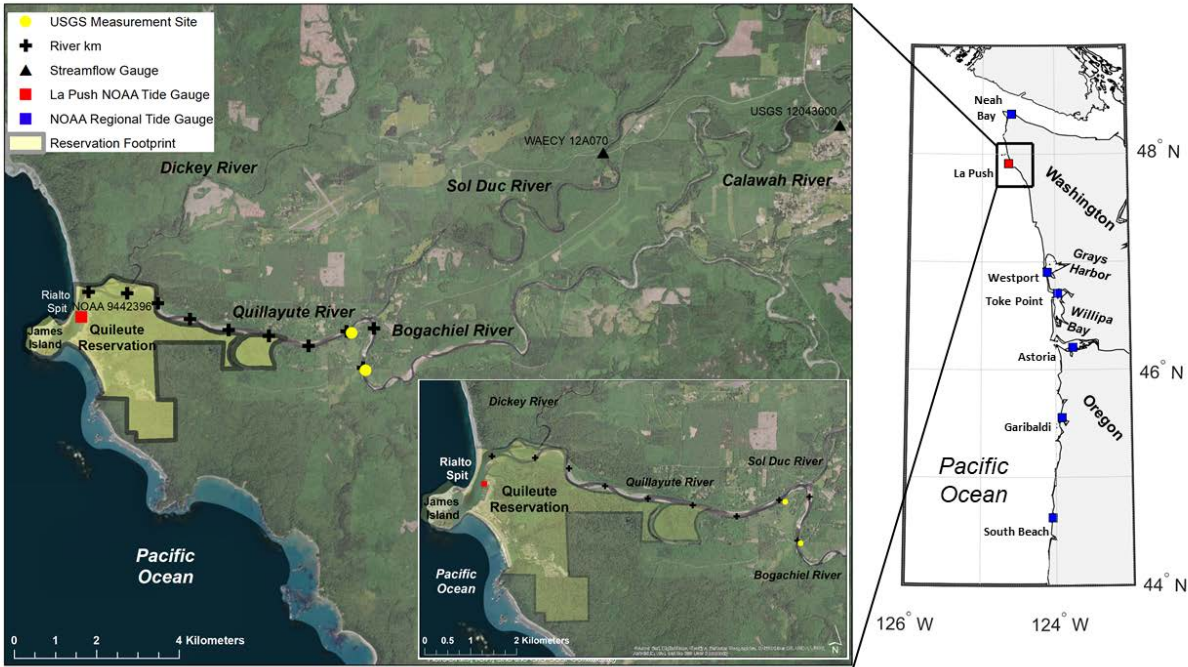


Figure 8 Map of study area (left) for our second modeling approach. The La Push tide gauge is represented as a red square while other regional tide gauges are represented as blue squares (right). The Calawah and Sol Duc river gauges are represented as black triangles and USGS measurement sites from the May 2010 survey are depicted as yellow circles. Approximate river kilometers are denoted as black crosses on the study area map.

To explore the influence of river and ocean forcing on extreme water levels (WLs), our second approach developed a hybrid modeling framework by merging the hydraulic model, Hydrologic Engineering Center's River Analysis System (HEC-RAS, model domain is displayed in Figure 9) for simulating river flow, with probabilistic simulations of co-occurring river and ocean events (Serafin et al. 2019; Appendix B). The HEC-RAS model was initially calibrated to successfully model the water surface elevation along the river during a USGS data collection effort (Czuba et al. 2010). Similar to emulating extreme WLs in Grays Harbor (section 1.1), this technique allows for insights into the extent and the relative importance of each WL component to flooding along a gradient from pure oceanographic forcing to pure riverine forcing.

The probabilistic, full simulation model of Serafin and Ruggiero (2014) was modified to incorporate the dependency between river discharge and the remainder of relevant WL components. This probabilistic model allows for the generation of multiple, synthetic WL records to produce numerous estimates of low-probability events not captured in the observational record. Modeling all of the statistically simulated boundary conditions in a hydraulic model to output along-river water levels would be prohibitively expensive. As an alternative to time consuming simulations, surrogate models are developed to approximate the response of a HEC-RAS simulation at each along-river location. This technique allows for the analysis of along-river water levels driven by a variety of boundary conditions. Long synthetic records allows for the direct empirical extraction of water level return levels rather than an extrapolation from historic observational forcing conditions.

In total, seventy 500-year simulations were run, which allowed for the extraction of the water surface elevation at every along-river location, as well as the ocean and river forcing driving that elevation (see Serafin et al. 2019; Appendix B). Understanding the relative forcing of extreme WLs moving from open ocean boundary conditions to where river processes dominate will help Treaty of Olympia communities better understand the risk of compounding impacts of various environmental forcing, which is important for increasing resilience to future events.

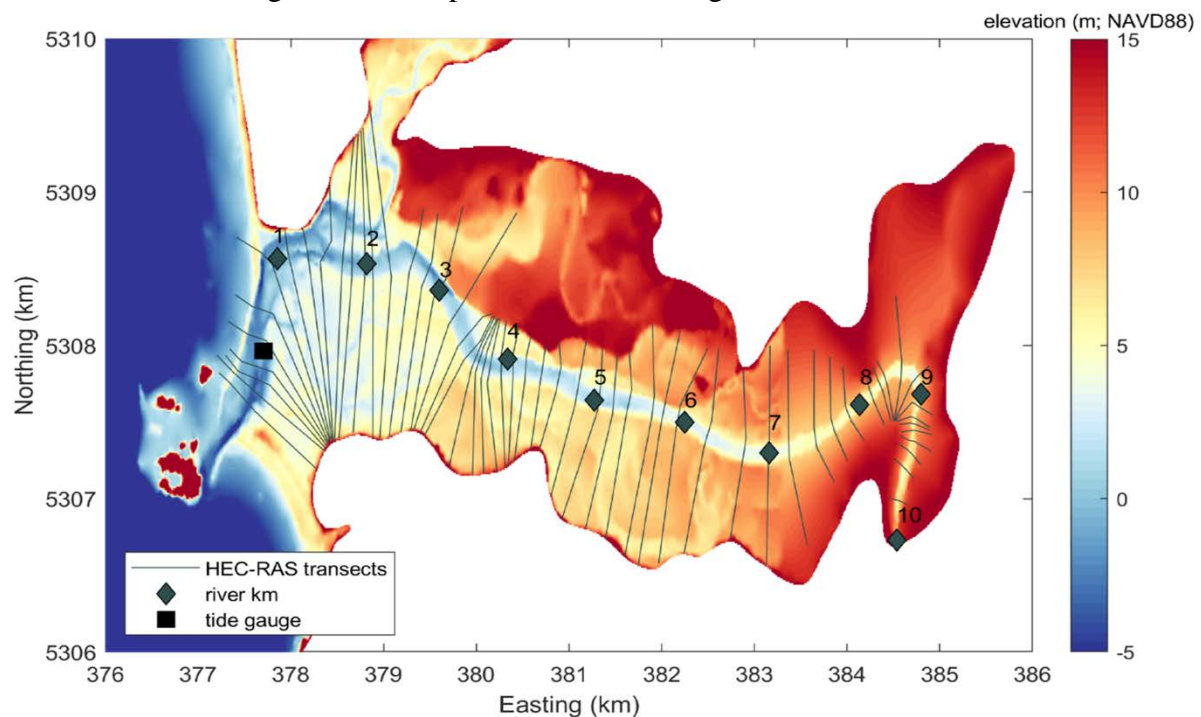


Figure 9 Digital Elevation Model used for the HEC-RAS simulations of the Quillayute River. HEC-RAS cross sections are depicted as grey lines. Approximate river kilometer and the location of the tide gauge are depicted as diamonds and a square, respectively.

The thousands of simulated upstream and downstream boundary conditions show that variability in the ocean boundary conditions impact river water levels, and therefore the potential for flooding and erosion, as far east as river kilometer 5 (Figure 10). Note that this is east of Thunder Field, an area of significant concern due to the potential of river avulsion due to channel migration and streambank erosion.

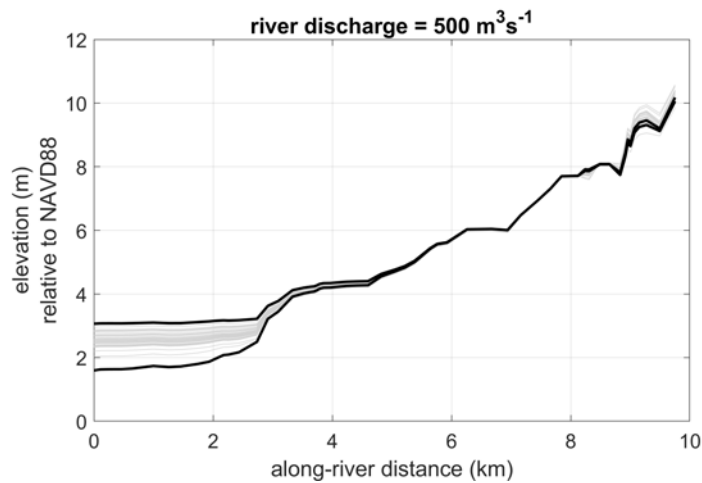
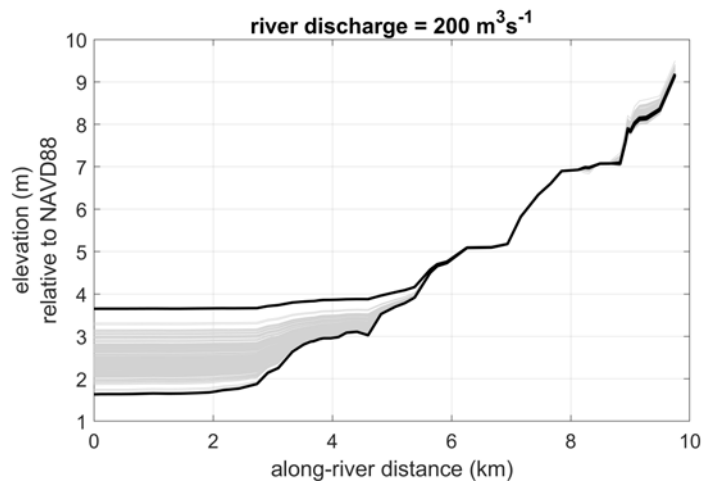
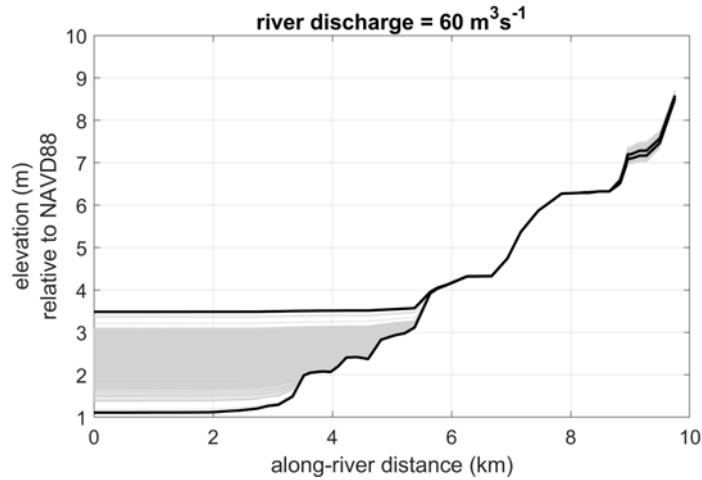


Figure 10 Results from HEC-RAS modeling on the Quillayute River showing water surface elevations for various scenarios. Three discharges are modeled for a wide range of different still water levels showing the varying influence of oceanic processes moving upstream. Each grey line represents one example simulation for the flow in the title $\pm 0.5 \text{ m}^3 \text{ s}^{-1}$ and the water surface profile extracted. Black lines represent maximum and minimum still water level boundary conditions.

Objective 2: Quantify the Impact of a Range of Climate Change Scenarios on Each of the Various Components and on the Resulting Combined Total Water Levels.

A vital consideration in any attempt to quantify climate change impacts is the associated uncertainty in the resulting estimates. The ability to predict the future climate is limited for a variety of reasons ranging from the difficulty of modeling the full earth system to the inability to know how human impact on the climate will evolve. A typical way to characterize this uncertainty is to consider “scenarios” which follow various plausible trajectories of the future climate. This project approaches Objective 2 using state-of-the-art estimates of future sea level rise as the basis for four climate change impact scenarios. These climate change impact scenarios were applied to both modeling approaches described above to explore the impacts on hazards.

An additional source of uncertainty is from the chaotic nature of the climate system. Consideration of a single climate time series is generally considered inadequate as this time series may or may not be indicative of the overall behavior of the system. For example, a single time series may be uncharacteristically extreme due to the stochastic nature of the climate system. Analysis based solely on this single time series could therefore produce biased results. To reduce this source of error, as well as to better characterize uncertainty in the analysis, this report takes the approach of considering multiple iterations (an ensemble) of climate time series.

2.1 Climate Change Impact Scenarios

Climate change impact scenarios were developed based on sea level rise (SLR) projections for the Washington coastline developed by Miller et al. (2018). At the time of the analysis for Grays Harbor, absolute SLR projections (i.e., sea level relative to a fixed, unmoving point) had been completed, and researchers were still in the process of creating relative SLR projections (i.e., sea level relative to land, which incorporates vertical land movement such as uplift and subsidence) for the Washington coast. The SLR scenarios in the Grays Harbor work use absolute sea level projections for RCP 8.5 from Appendix A of Miller et al. (2018) and incorporate the best estimates of vertical land movement specific to Grays Harbor at the time (1.5 mm/yr on average), as well as uncertainty around those estimates (Figure 11). SLR scenarios for the La Push study site (latitude of 47.9N and a longitude of 124.6W) were developed based on the newly completed relative SLR projections, which include vertical land movement developed by Miller et al. (2018) (Figure 11).

In both approaches, the low impact SLR scenario uses a low-end projection that is 95% likely to be exceeded. The medium impact SLR scenario used a mid-range projection and has a 50% exceedance probability. The high impact SLR scenario uses a high-end projection that is 5% likely to be exceeded. Finally, a “worst case” scenario was explored which has only a 0.1% chance of being exceeded.

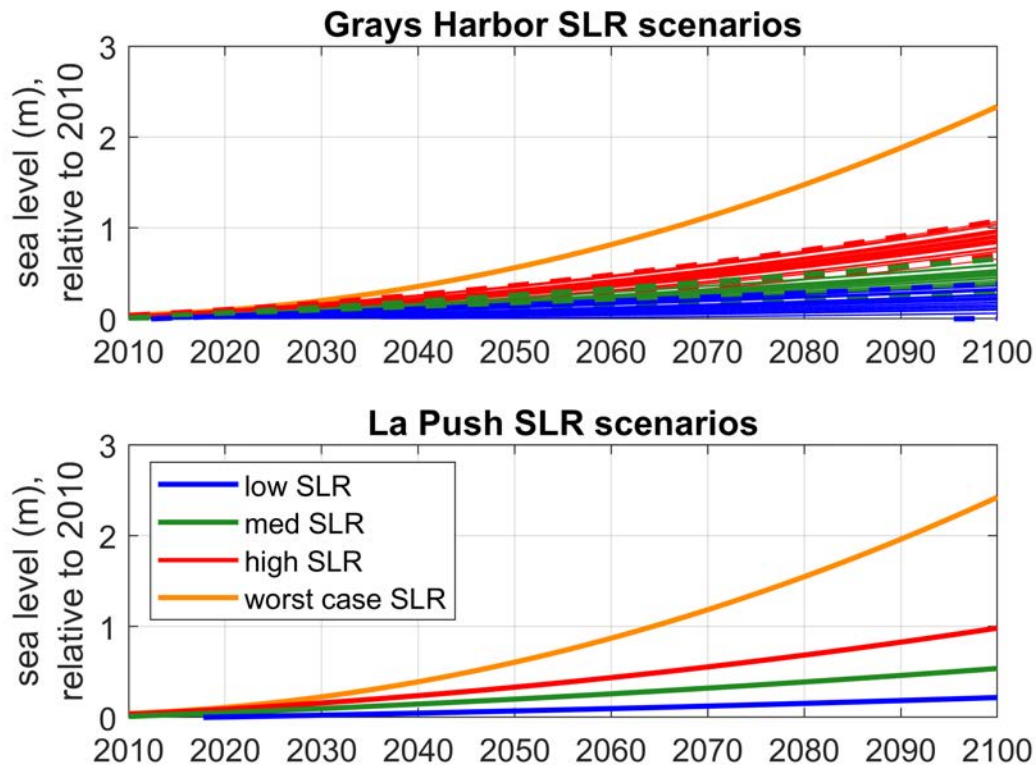


Figure 11 Envelope of four primary sea level rise (SLR) scenarios plotted as colors (blue for low, green for medium, red for high, and orange for worst case), relative to 2010 sea levels. Each SLR iteration in the top panel for scenarios low, medium, and high incorporates a range of vertical land movement, while the bottom panel SLR scenarios were developed from relative SLR scenarios.

While SLR is expected to be the largest driver of long-term increasing coastal flood hazards in the Pacific Northwest (PNW), other controls on flooding are also expected to change. Research has shown that changing wave climate has recently been a major driver of intensifying coastal hazards in the PNW (Ruggiero 2013). This is because for PNW open coastlines, waves are one of the largest contributors to extreme total water levels (TWLs) (Serafin et al. 2017). For this reason, we also explored the possibility of changes to the PNW wave climate. Projected changes in wave height were estimated by shifting wave height distributions (Figure 12) based on future estimates of wave height change in the Northeast Pacific from global climate model projections (Hemer et al. 2013; Wang et al. 2014). Water levels and wave heights are also affected by major El Niño events, which have been associated with severe flooding and erosion in the PNW (Barnard et al. 2017). The frequency of major El Niño events was allowed to double and halve the present-day frequency, which is approximately once every 7–10 years. In our Grays Harbor work, thirty-three probabilistic TWL simulations for each high, medium, and low climate change impact scenario, resulted in 99 different 100-year projections of daily maximum TWL (Figure 13). One TWL simulation was also completed for the “worst case” climate change impact scenario.

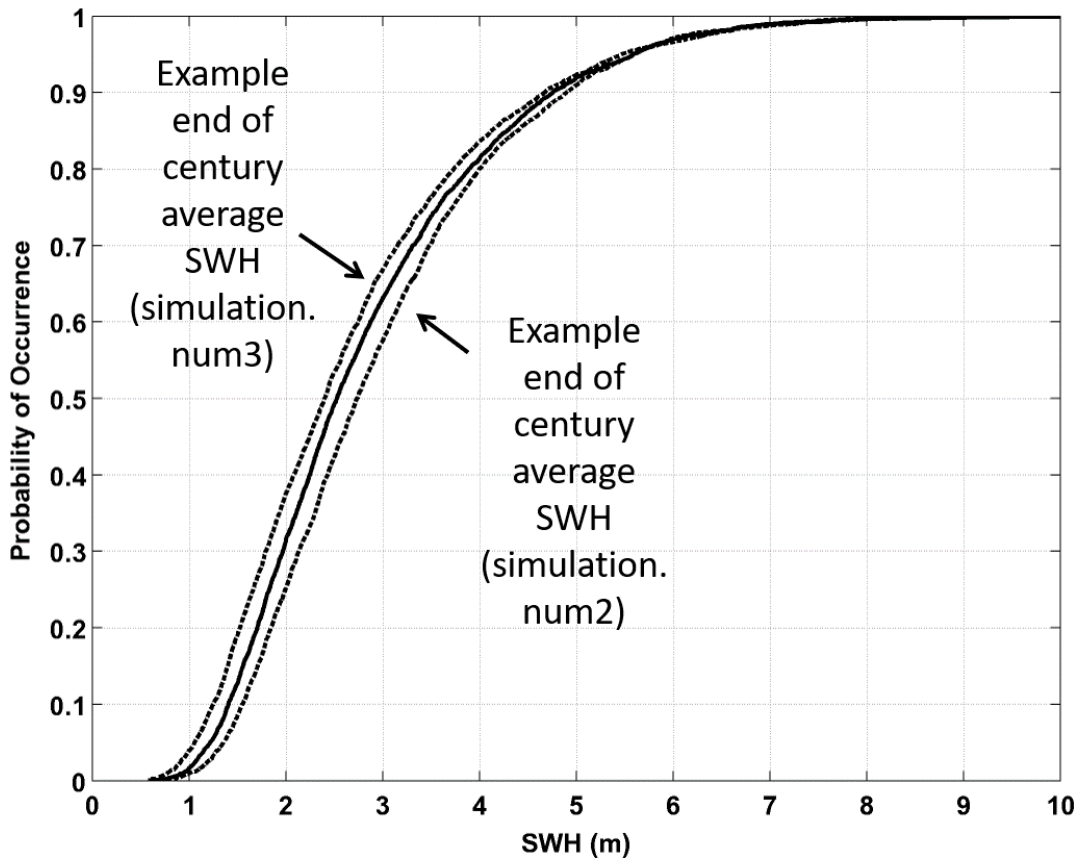


Figure 12 Envelope of wave climate variability in terms of shifting wave height (SWH)

For the Quillayute River site, climate change-controlled modifications to streamflow were also incorporated within the uncertainty. The western Olympic Peninsula is projected to experience increased winter precipitation (Mote et al. 2013), which could subsequently increase the frequency or intensity of high streamflow events along the Quillayute River. However, there are currently no specific estimates of changes to future streamflow specific to the Quillayute watershed. We therefore completed a sensitivity analysis where we allow for the average winter streamflow to increase by 2, 5, 10, and 20%.

A key strength common to both of the new approaches developed by this project is the ability to consider future uncertainty in projections. This allows a more robust estimation of the most likely future hazard as well as an understanding of how drastically the hazard could vary from this most likely outcome. Figure 13 gives an example of how this project has approached uncertainty through the perspective of TWLs. The colors show variability as a function of SLR scenario while the shaded regions show variability as a function of climate iteration. From this plot it can be seen how important consideration of uncertainty is, with future predicted TWLs in the year 2100 differing at the scale of meters and predicted average annual TWLs for each scenario varying by about half a meter.

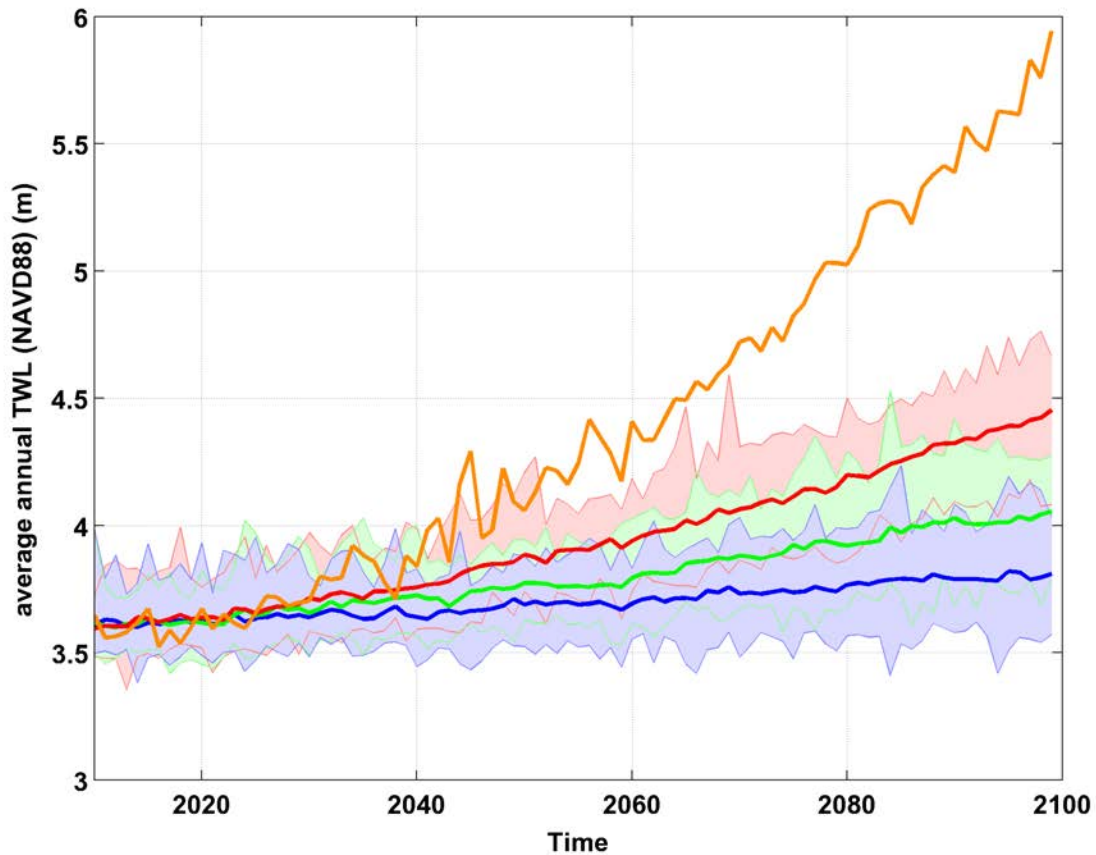


Figure 13 Influence of climate change on average annual total water levels (TWLs). The solid line depicts the average for each climate impact scenario, while the shaded regions depict the variability due to changes in the wave climate, frequency of major El Niño events and vertical land movement.

2.2 Impact of Climate Change on Emulated Grays Harbor Water Levels

Once emulators of the Grays Harbor ADCIRC model were built, time series of forcing variables were passed through to estimate extreme water levels (WLs) at a variety of locations. Here we refer to WLs as the integrated effect of a variety of forcings/processes on the water level of a location. For emulating WLs within Grays Harbor, this does not include surface waves or any type of runup but rather can be thought of as the water level that would be measured if there were a tide gauge at this location. Within bays this is the water level metric that best corresponds to inundation. To explore the influence of climate change on WLs relevant to tribal infrastructure, we focused our efforts on the three emulators specified earlier (see Figure 3): the Westport tide gage, near the Quinault Marina and RV park, and immediately offshore of the Quinault Beach Resort and Casino.

Using the emulators, we explored the time varying WLs at these stations for a range of sea level rise (SLR) scenarios. Figure 14 shows an example synthetic WL time series at the Marina station for the four climate change impact scenarios.

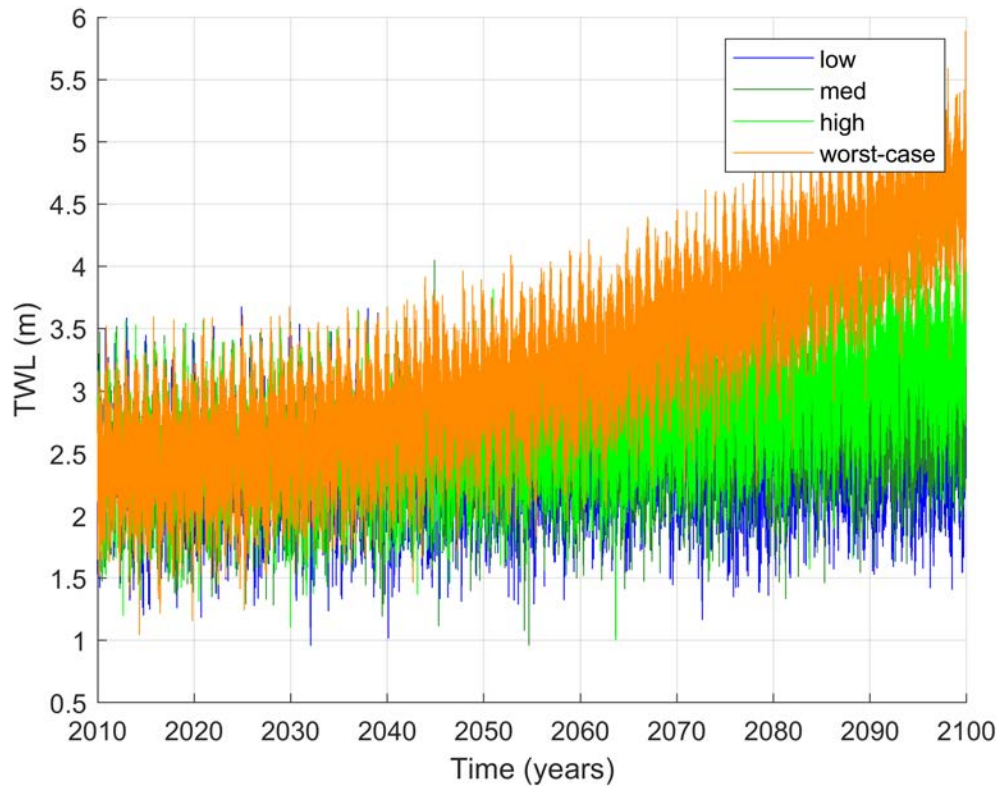


Figure 14 Example time series of total water levels (TWL) at the Marina for a range of sea level rise scenarios.

Of note, Figure 14 shows only one of the more than 30 climate iterations analyzed for each climate change impact scenario (other than Worst-Case which only was considered with one iteration). Figure 14 displays the seasonal and interannual variability in WLs at the Marina location, as well as how these may change moving forward with possible changes to the climate. However, the impact of these results on flooding hazards can be better understood by viewing TWLs from a threshold exceedance perspective.

Figure 15 illustrates the number of days per year that WLs at this location exceed an elevation of 3.2 m (NAVD88), which is approximately the elevation associated with minor nuisance flooding at this location. Figure 15 (bottom) includes the worst-case climate change impact scenario and reveals that by the end of the 21st century nuisance flooding may start occurring every day under this scenario. Figure 16 displays the same plot but for the location of the Westport tide gauge, a proxy for flooding in and around the Westport Marina.

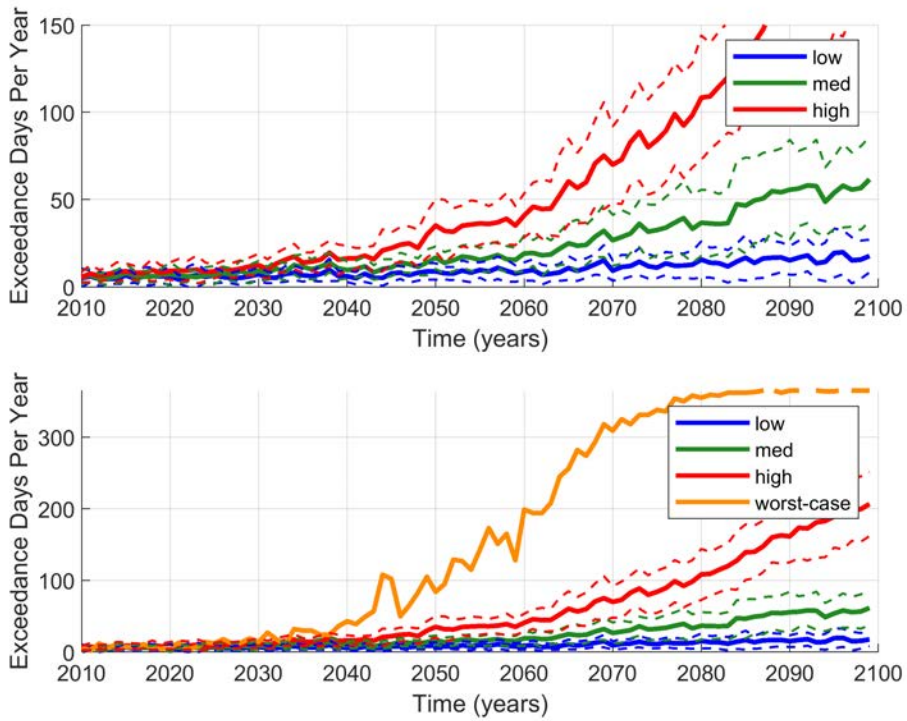


Figure 15 Exceedance days per year when total water levels at the Quinault Marina and RV Park exceed an elevation of 3.2 m NAVD88.

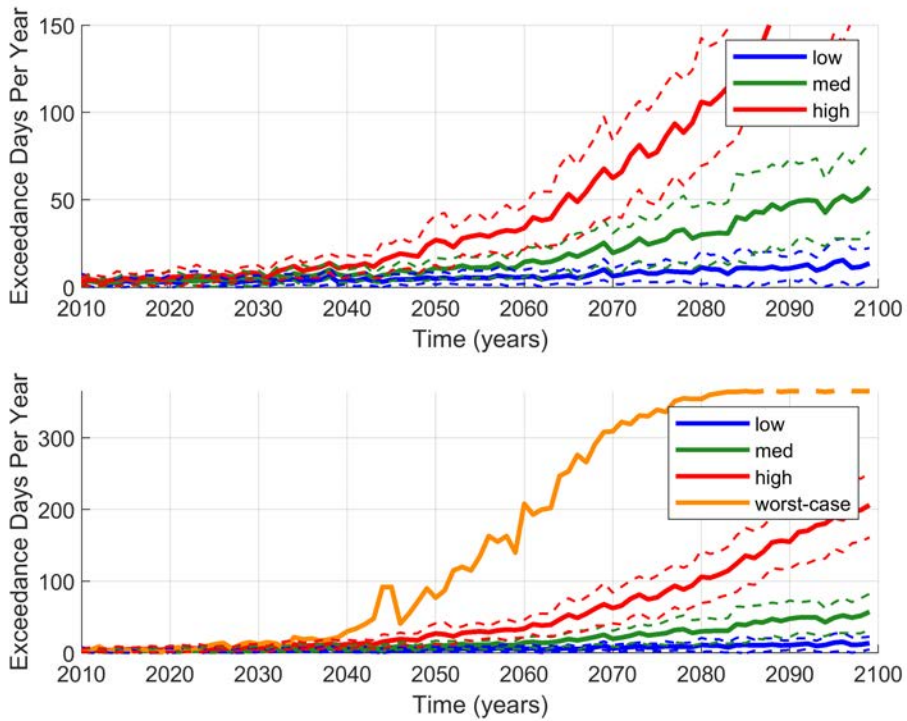


Figure 16 Exceedance days per year when total water levels at the Westport Marina (tide gauge) exceed an elevation of 3.2 m NAVD88.

The final example of total water level (TWL) exceedance elevations in and around Grays Harbor is along the outer coast fronting the Quinault Beach Resort and Casino. Similarly to locations within the bay, extreme coastal TWLs along the outer coast are the result of interactions between multiple oceanographic, hydrological, geological, and meteorological forcings that act over a wide range of scales (e.g., astronomical tide, wave set-up, large-scale storm surge, monthly mean sea level, vertical land motions, etc.). However, the additional presence of large breaking waves means that, at any given time, the elevation of the TWL, relative to a fixed datum, is comprised of two components such that

$$TWL = SWL + R \quad \text{Eq. 1}$$

where the *SWL* is the still water level, or the measured water level from tide gauges, and *R* is a wave induced component, termed the wave runoff (Figure 17). The wave runoff calculation is often dependent on the wave height, wavelength, and the local beach morphology (e.g., Ruggiero et al. 2001), making it a highly site-specific computation. Because we are interested primarily in extreme events, *R* is parameterized using *R*2%, defined as the 2% exceedance percentile of runoff maxima.

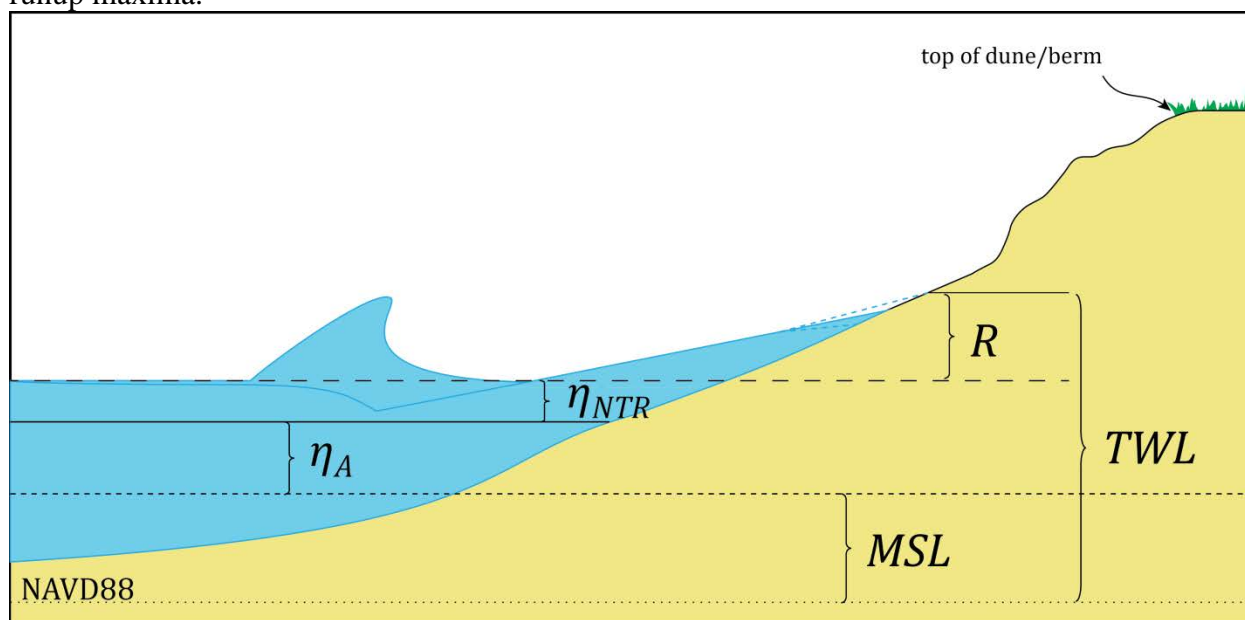


Figure 17 Definition sketch of total water levels (TWL). Dune/bluff erosion or infrastructure damage occurs when the TWL, relative to a datum such as the land based NAVD88 datum, exceeds the elevation of the dune/structure toe, and overtopping/flooding occurs when the TWL exceeds the elevation of the dune or structure crest (figure from Serafin and Ruggiero, 2014).

TWLs can then be compared to backshore morphology to estimate the percentage of time certain contours are inundated as well as the risk of coastal flooding and erosion. Figure 18 shows topographic beach profiles collected just in front of the Casino (<http://nvs.nanoos.org/BeachMapping>). From these data we extracted quantitative information such as the beach slope, the average elevation of the dune toe, and the average elevation of the dune crest relative to NAVD88. Figure 19 shows one example time series of forcing conditions, including TWLs calculated using equation 1, fronting the Casino. Similar to our analyses for stations within Grays Harbor, Figures 20 and 21 show days of exceedance in which TWLs are projected to be above the dune toe (a proxy for coastal erosion, Figure 20) and above the dune crest (a proxy for inland flooding, Figure 21).

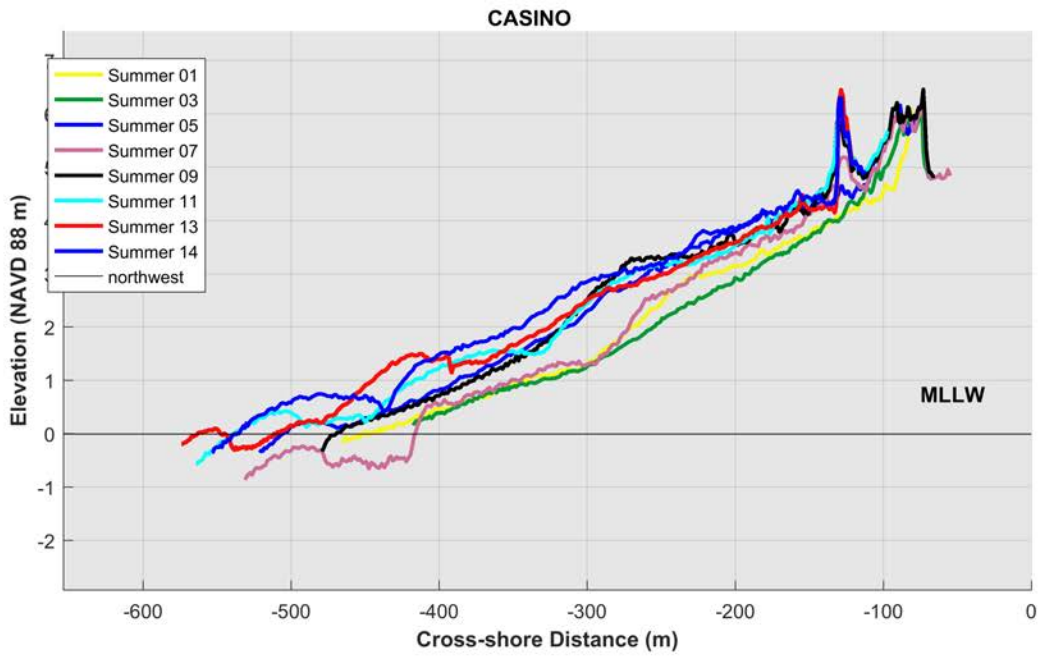


Figure 18 Beach profile fronting the Quinault Beach Resort and Casino

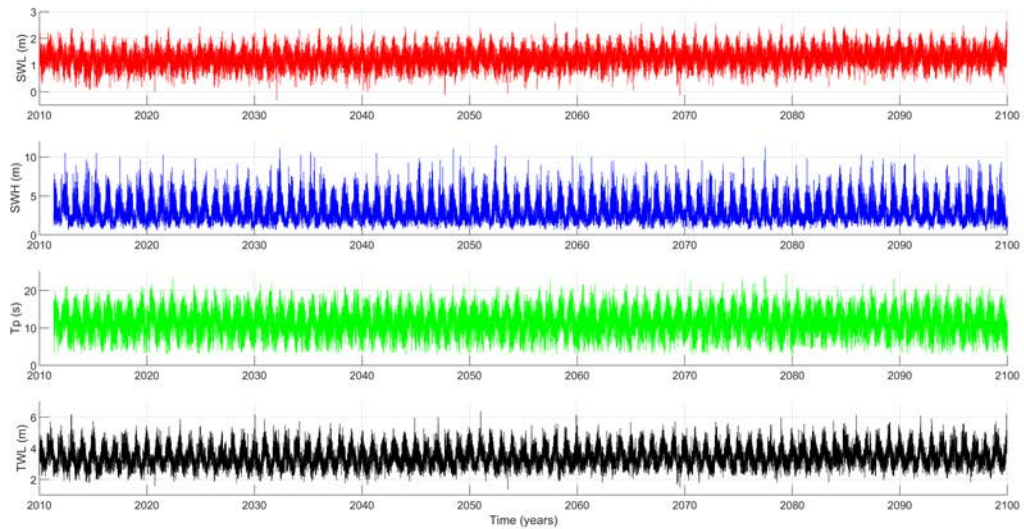


Figure 19 Forcing conditions for computing total water levels fronting the Quinault Beach Resort and Casino.

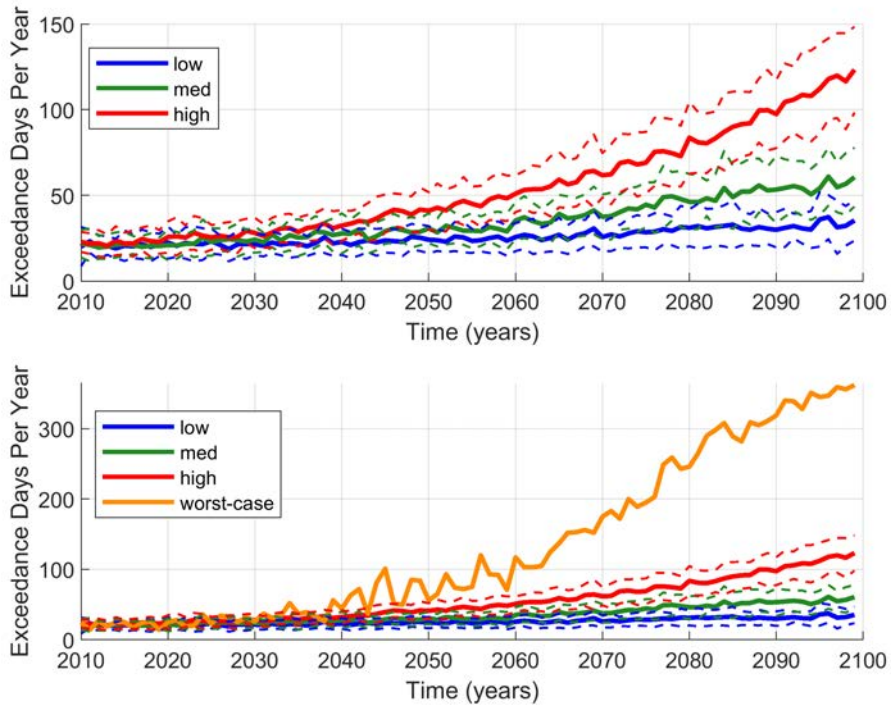


Figure 20 Exceedance days per year when total water levels at the Quinault Beach Resort and Casino exceed an elevation of 4.5 m NAVD88 – approximately the elevation of the dune toe and hence a proxy for possible erosion.

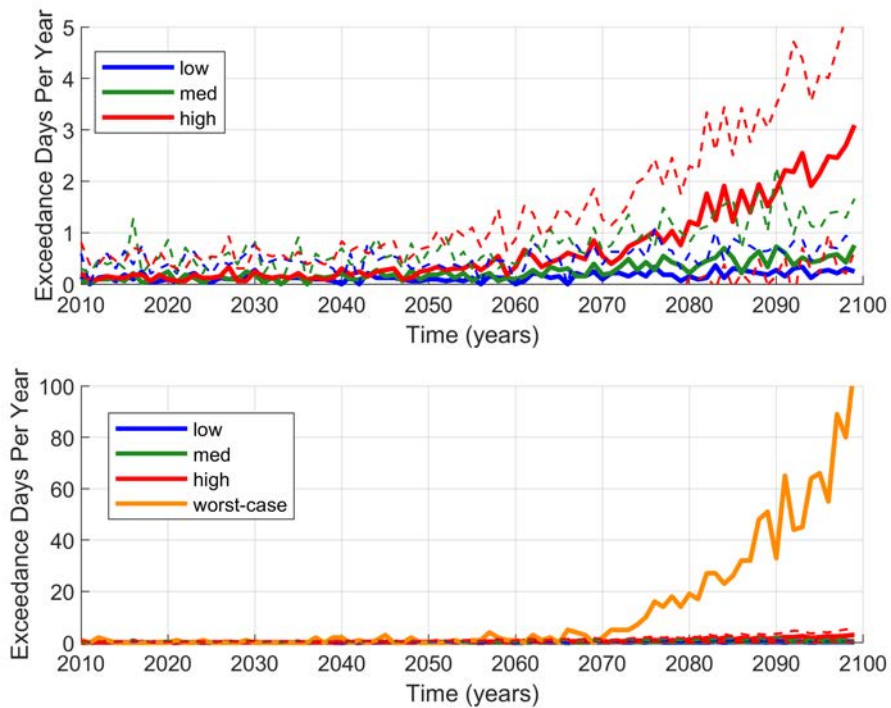


Figure 21 Exceedance days per year when total water levels at the Quinault Beach Resort and Casino exceed an elevation of 6.0 m NAVD88 – approximately the elevation of the dune crest and hence a proxy for possible overtopping and inundation.

2.3 Impact of Climate Change on Simulated Quillayute River Total Water Levels

Models for the Quillayute River were set up to explore the influence of climate change through increasing sea level and increasing winter river discharge on coastal flood probabilities. With increasing sea levels, flooding along the first 2.5 km of the Quillayute River shifts dramatically. For example, at river km 1.5, less than one day per year in present day, low, and medium impact climate scenarios changes to 3 and 95 days per year in high and worst case impacts scenarios, respectively (Figure 22). Flooding in the worst-case scenario could increase to one-third of the year in certain along-river locations (not shown). For all scenarios, river km 5 and upstream is not affected by changing ocean water levels. An increase to winter streamflow alters flooding the most between river km 3 and 4 (Figure 23), and minimally across other locations of the river.

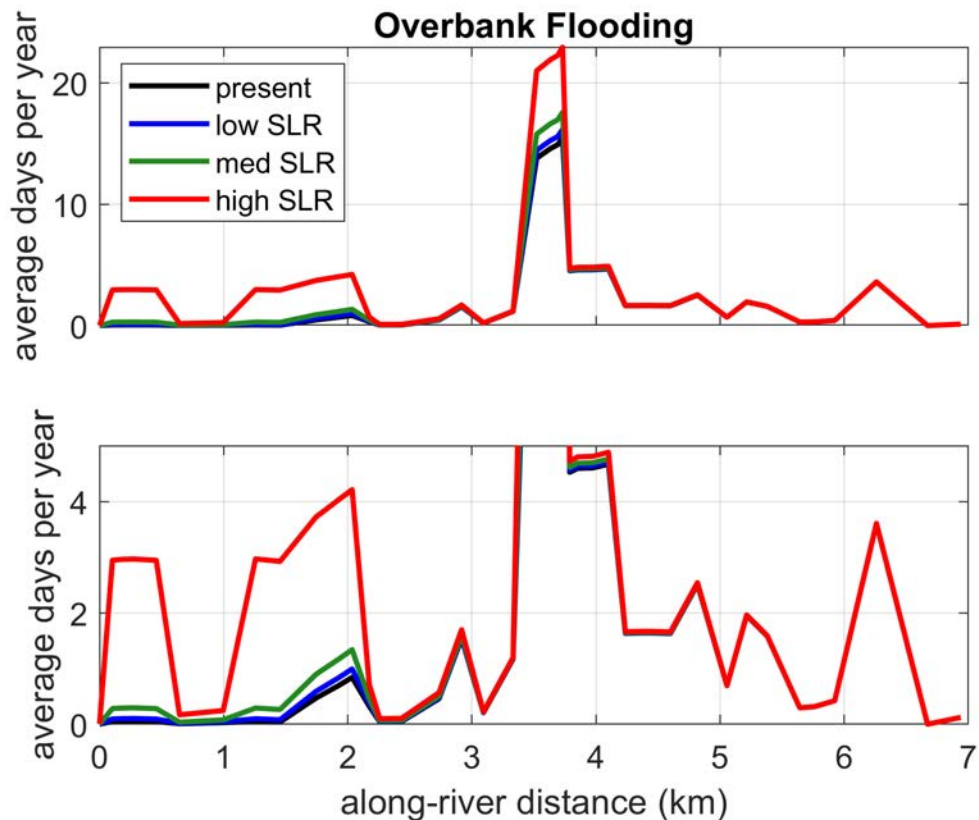


Figure 22 Average number of days per year of possible flooding as a function of river kilometer. For present-day climate, low sea level rise (SLR), medium SLR, and high SLR. Both panels show similar data but at different scale.

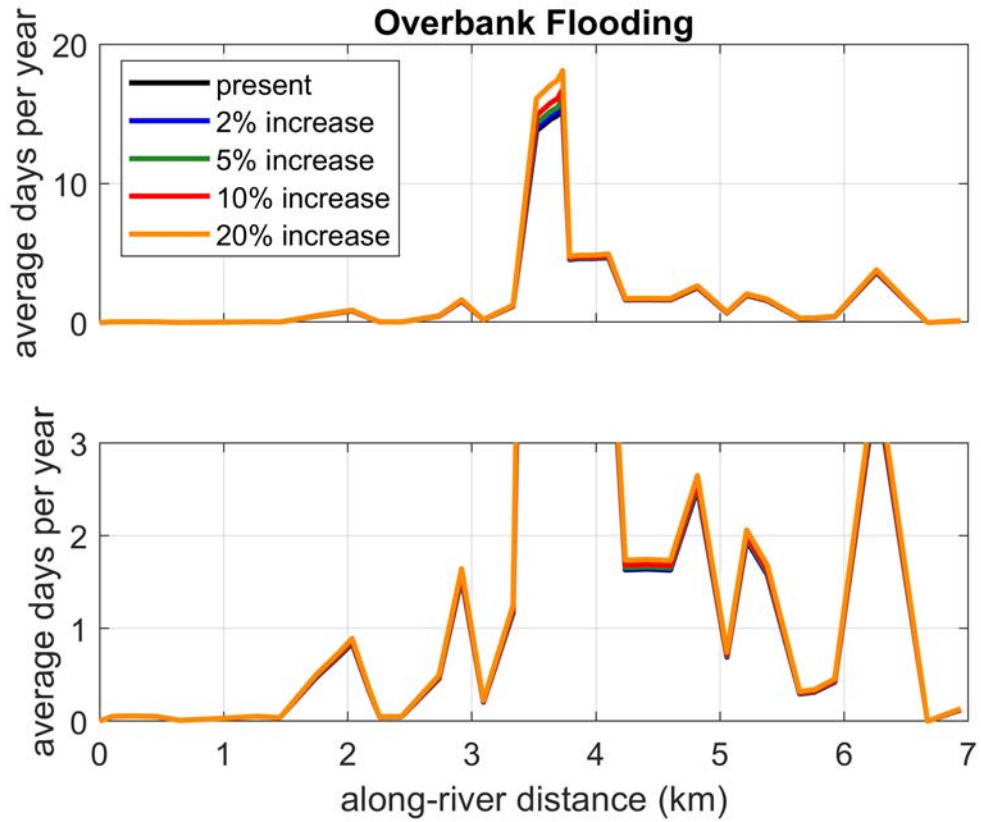


Figure 23 Average number of days per year of possible flooding as a function of river kilometer. For present-day climate, a 2%, 5%, 10%, and 20% increase in winter river discharge.

Objective 3: Assess the Impact of Present Day and Forecasted Future Coastal Flooding Events on Infrastructure in Several Communities within the Treaty of Olympia

3.1 Impact of Flooding on Infrastructure

One valuable application for the surrogate model approach developed in this project is the ability to produce more robust estimates of the 100-year flood event for Grays Harbor. The 100-year flood event is the most common engineering tool for determining flood risk in communities. This surface represents flooding that has a 1% chance of occurring in any given year. The 100-year flood event is often used for community development planning, flood insurance premiums, and many other applications. Therefore, any improvements to how this metric is computed can provide a wide range of benefits.

An example of the 100-year flood event results from this study is shown below in Figure 24. This is a unique application as many extreme flood inundation studies assume a spatially constant 100-year water level (known as a bathtub model as the assumption that water raises statically like a bathtub). In contrast, this approach is able to calculate spatial variability in this flooding surface. This is found to be a significant factor with some locations experiencing 100-year water elevations that are half a meter higher than other locations. Additionally, these results can be considered relatively more robust than traditional techniques since they are calculated using extreme events from 100-year simulated time series. Generally, 100-year flooding events are based on statistical extrapolations from short simulations or observed events. This results in significant uncertainty as a result of the extrapolation. This new approach therefore allows a significant reduction in this uncertainty, at least in terms of the present-day climate. Future projects could fully explore how this refined flooding estimate may affect specific infrastructure in and around Grays Harbor, WA.

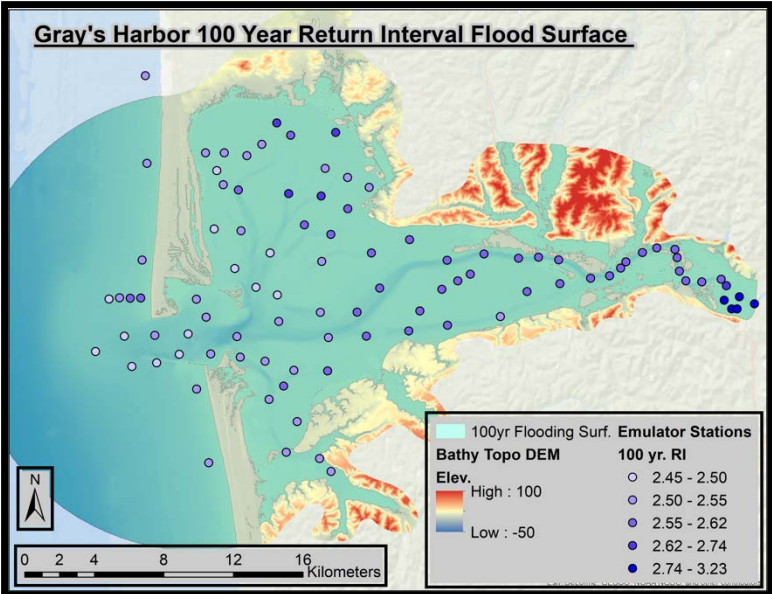


Figure 24 Example 100-year return flood surface produced from ADCIRC results. Vertical datum is relative to mean sea level.

Similarly, the 100-year (and other) return level event can be extracted along the Quillayute River and the drivers of these events can be assessed. The 100-year event is typically assumed to be driven by a specific forcing event, such as the 100-year rainfall or storm surge. However, for processes driven by multiple dimensions, different sizes and combinations of forcing conditions could potentially generate extreme flood magnitudes. The relative importance of both oceanic and riverine forcing to extreme water levels (WLs) emerges when averaging the magnitude of the drivers of the water level return levels at each transect from all seventy 500-year simulations (Figure 25). The magnitude of the average streamflow (Q) driving water level return levels gradually increases over river km 0–2 and then is consistent from river km 2 to 10. Downstream, between river km 0 and 0.25, the magnitude of the average still water level (SWL) driving water level return levels is consistent and then gradually decreases over a 1 km zone. This approach confirms the presence of an oceanographic-fluvial transition zone, where traditional methodologies for defining return level events based on a single driver are insufficient for defining water level return levels. Between river km 1 and 2, a range of SWL and Q conditions drive all return level events, and extreme water levels are driven by neither the individual SWL or Q return level event.

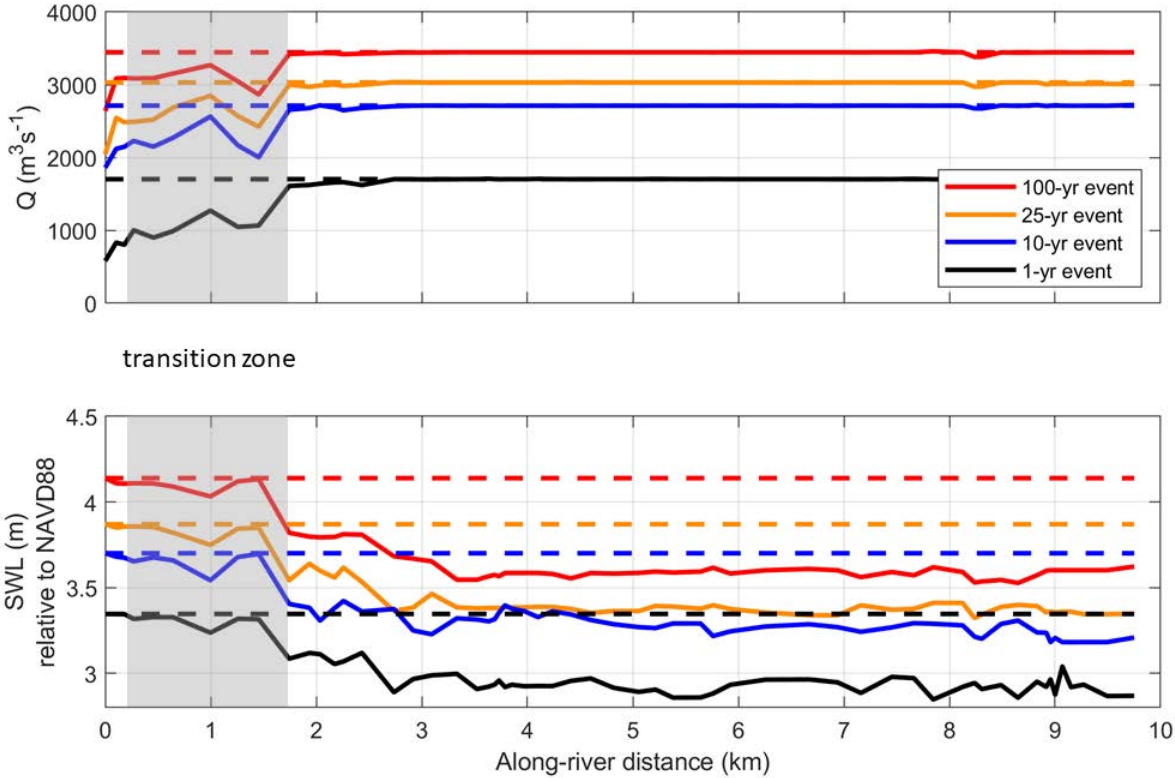


Figure 25 The average forcing condition driving along-river return levels at each transect where a) displays the Quillayute streamflow (Q) conditions and b) displays the still water level (SWL) conditions. The dashed lines depict the individual forcing conditions, where the along-river return level is assumed to be driven by either Q or SWL. Red, orange, blue, and black lines represent the 100, 25, 10, and annual return level event. The grey shaded area represents a transition zone, where the water level is driven by a combination of SWL and Q events.

The approach taken for the Quillayute River also allows consideration of flooding hazards along the river basin. As an example, the developed modeling setup can be used to assess important metrics such as the number of days per year that river stage exceeds bank elevation as a proxy for flooding potential (shown schematically for two locations in Figure 26). Figure 27 is an estimate of the average number of days per year in which the river elevation exceeds its banks for the present-day climate.

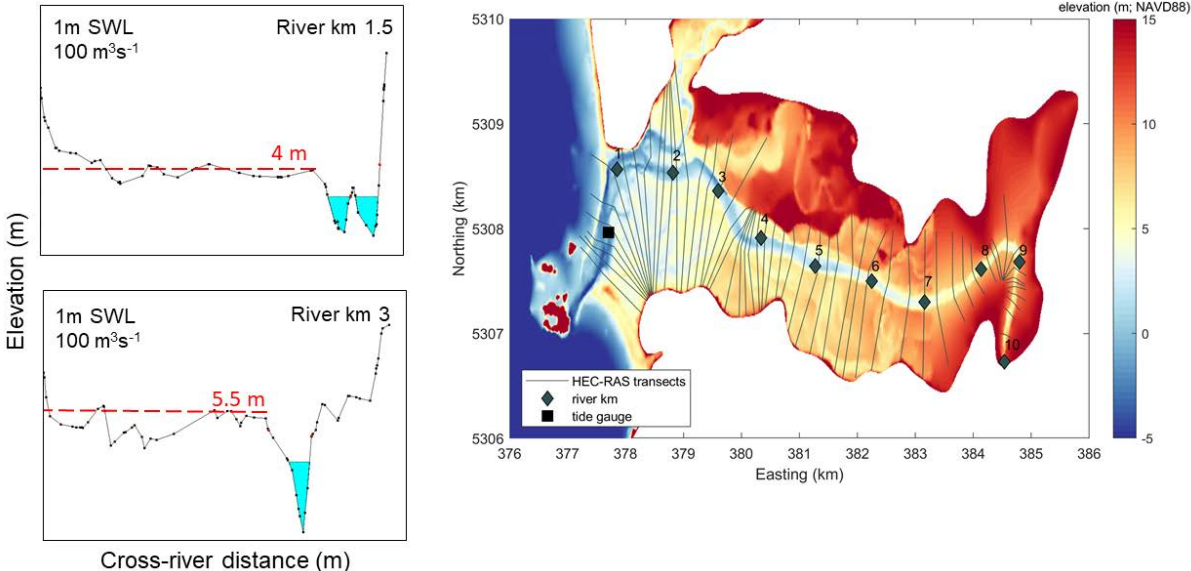


Figure 26 Example results from HEC-RAS modeling of the Quillayute River showing water surface elevations at model topography/bathymetry at River km 1.5 and River km 3 for low flow conditions.

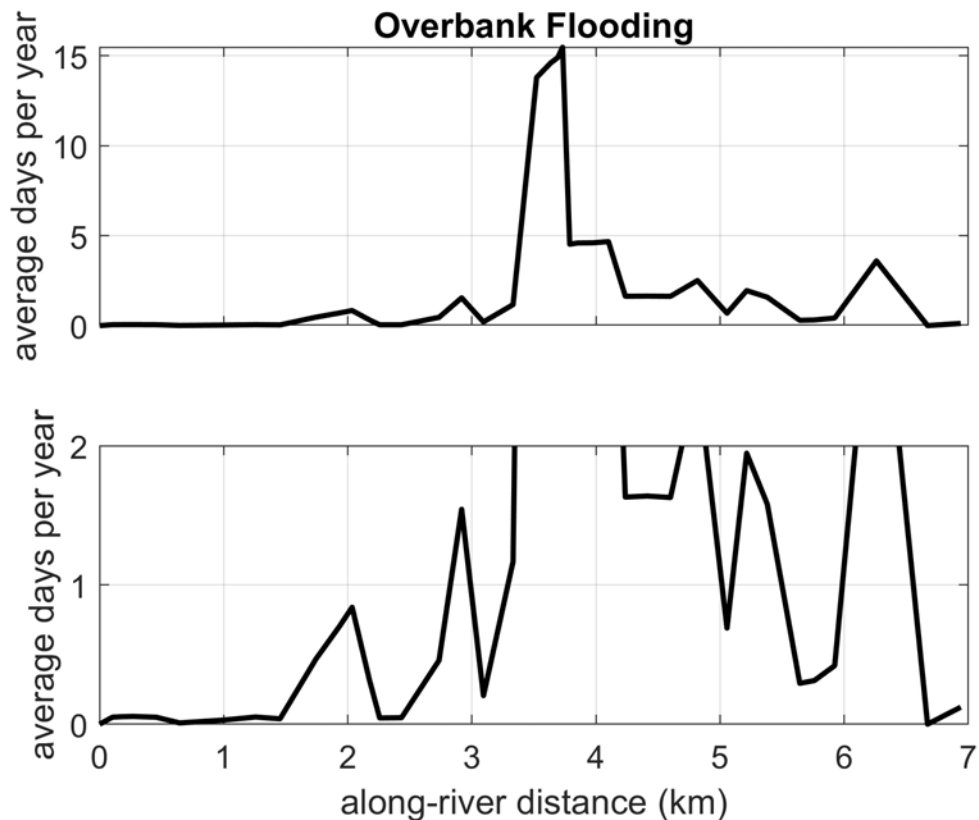


Figure 27 Average number of days per year of possible flooding as a function of river kilometer for present-day climate. The bottom panel displays the same results, but zoomed in.

We also evaluated the change in exceedance days per year at specific locations near the tide gauge as well as near river km 3.5 near the Thunder Field area (Figure 28). Here, we counted the number of days per year water levels exceeded 4m NAVD88 and 5m NAVD88, the approximate elevation of the bank on the Quileute Reservation side of each transect. Under low and medium climate change impact scenarios, the bank near the tide gauge is exceeded every few years, but not consistently. For the Thunder Field area, we see that while important, sea level rise has a lesser influence on flooding, which means increases to discharge will matter comparatively more for this location. For the worst-case scenario, sea levels flood the tide gauge area more than half of the year by mid-21st century and then every day of the year by the end of the 21st century (Figure 29). For this upper end SLR scenario, sea levels influence flooding at the Thunder Field area as well.

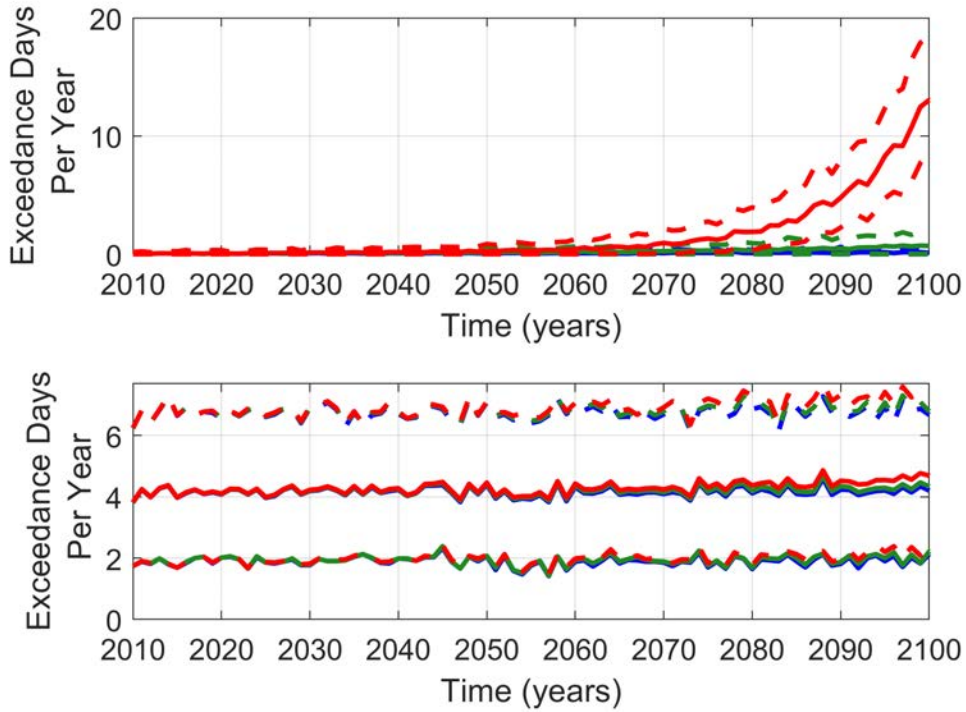


Figure 28 Average number exceedance days per year for the tide gauge (top) and Thunder Field area (bottom) during low (blue), medium (green), and high (red) sea level rise scenarios. The solid lines indicate the average while the dashed lines indicate bounds around the average.

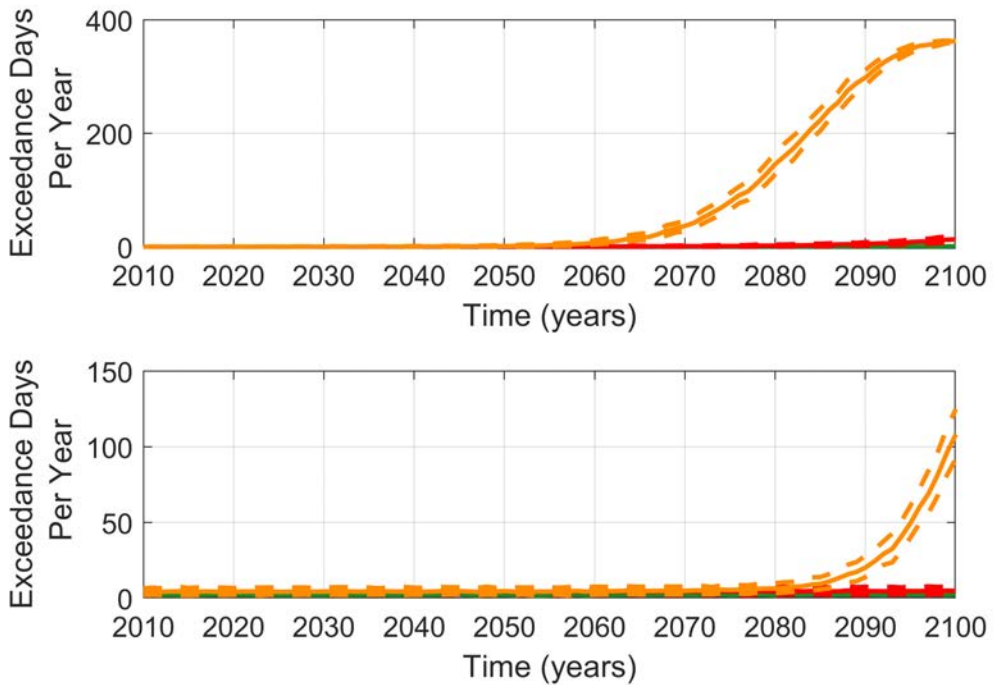


Figure 29 Average number exceedance days per year for the tide gauge (top) and Thunder Field area (bottom) during low (blue), medium (green), and high (red), and worst case (orange) sea level rise scenarios. The solid lines indicate the average while the dashed lines indicate bounds around the average.

3.2 Climate Change Impacts, Recommendations for Adaptation, and Suggestions for Future Research

As explored while emulating extreme water levels in Grays Harbor, analysis indicates that infrastructure along the bay will be exposed to significant flooding risk from climate change. That said, this risk will be sensitive to climate change impact scenario with characteristically different outcomes between the low and worst-case scenarios. Overall, the low climate change impact scenario is associated with modest impacts to infrastructure at the locations of interest. Nuisance flooding for the Quinault and Westport Marinas were projected to remain on the order of 10 days per year even by the year 2100. A similar result was found for extreme total water levels, including wave runup, at the beach fronting the Quinault Beach Resort and Casino (although with exceedance of the dune toe approaching 20 days per year). While this represents a change to the current hazard status quo, it might not be a drastic outcome.

The low climate change impact scenario is 95% likely to be exceeded (see section 2.1). Therefore, impacts to flooding associated with climate change are likely to be greater than this outcome. The high climate change impact scenario, which is 5% likely to be exceeded, shows a very different outcome to infrastructure with nuisance flooding exceeding 150 days a year. This could drastically affect the functionality of these marinas with some form of adaptation being required. The fronting beach for the Quinault Beach Resort and Casino would likely become erosional with over 100 dune toe impact events a year. This said, overtopping projections remain low even for the high scenario at around 3 days per year by 2100. This result may change though as the beach profile adjusts to an erosional regime. Even this outcome is mild in comparison to the worst-case scenario which sees both marinas becoming continuously inundated (365 days a year) and overtopping/flooding of the Quinault Beach Resort and Casino approaching one-third of the year. While this outcome is unlikely, it is within the range of plausible future climates.

As shown by the wide range of uncertainty in future outcomes, appropriate adaptation measures will be very dependent on which scenario comes to pass. Traditionally, responses to climate change can be divided into groups of general strategies or grouping of adaptation measures such as “realign”, “protect”, and “restore” (e.g., Mills et al. 2018; Lipiec et al. 2018). A “realign” scenario involves changing human activities to suit the changing environment (e.g., relocation of people and infrastructure). A “protect” scenario involves resisting change through the maintaining/strengthening of current infrastructure, mainly through engineering solutions (e.g., shoreline armoring, seawalls, etc.). A “restore” scenario could involve implementing natural or nature-based solutions to accommodate environmental change (e.g., wetland restoration or dune building to reduce flooding). Each of these approaches most likely would bring out a set of tradeoffs that are highly dependent on the specific context of the application. While a detailed cost/benefit analysis of various adaptation measures was not part of the scope of this study, some general points can be made.

While our analysis suggests that overtopping of the beach fronting the Quinault Beach Resort and Casino will likely remain modest, this could change drastically if the fronting beach profile changes significantly. Projected increases in impacts to the dune toe could provide the mechanism for this to occur. Therefore, it is recommended that careful monitoring of the fronting beach be maintained (e.g., <http://nvs.nanoos.org/BeachMapping>). If the beach appears to be shifting towards an erosional regime, then further detailed research into adaptation measures may

be warranted. Based on the significant hard infrastructure investment in the Quinault Beach Resort and Casino, relocation is unlikely. Therefore, some form of hard or soft (natural) protection may be necessary. The current location of the Resort is well set back from the shoreline providing with a healthy dune system. This potentially provides the opportunity for a natural solution although this will have to be explored more fully as the beach system changes.

With projected increases to flooding at the investigated Marina locations, likely some form of adaptation will be required (although the timing will be dependent on climate scenario). Realignment will likely be difficult due to the high cost of moving infrastructure, especially hard engineering structures. Rather a protective adaptation approach will likely be required through increasing armoring height and reinforcing the waterfront. Some infrastructure may be able to be moved back but a full analysis will be necessary to determine the optimal path forward.

For the La Push study site, results indicate that climate change impacts may significantly change the frequency with which the Quillayute River floods. Changes to high water levels will impact river stage as far as 5 km inland, most likely significantly worsening the existing Thunder Field erosion/avulsion threat (Figure 30). Assuming channel morphology stays similar, higher flows would increase flooding along some sections of the river and may influence more flooding closer to the river mouth. At the same time, changes in mean sea level will likely drive the oceanographic influence upstream. While we have characterized the spatial variability in driving processes to flooding in the present day, there is a high likelihood that changes in the future climate will shift the importance of these interacting processes. It is important to acknowledge that the un-stabilized nature of the Quillayute River means that its channel bathymetry and path are very likely to change moving forward. This is both a source of uncertainty in the results from this study (which uses a combined bathymetry from multiple channel geometries) as well as for future projections. Channel geometry and estuary bathymetry are first order controls on flooding processes so any changes in the Quillayute River would likely result in corresponding change to flooding vulnerability. Any considerations of adaptation measures should consider this significant source of uncertainty in its analysis.

While several Treaty of Olympia tribal communities experience compound flooding issues associated with both riverine and oceanographic forcing, our second modeling approach focused on the Quillayute River and the community of La Push due to the availability of data for sufficient model calibration and validation. Collecting relevant data to build similar models in the Quinault, Queets, and Hoh Rivers would be useful for a wide range of analyses. Installing tide gages, at least temporarily, at some of these streams will help to elucidate when river signals are influencing observed non-tidal residuals (see Serafin et al. 2019 [Appendix B] for details on an approach for quantifying this). Finally, developing sophisticated process-based models (e.g., CoSMoS, Barnard et al. 2014) for the region will allow for detailed exploration of extreme events in the present day as well as under climate change. Coupling models like CoSMoS with stochastic approaches (so called hybrid modeling approaches such as those described in this report) is recommended.

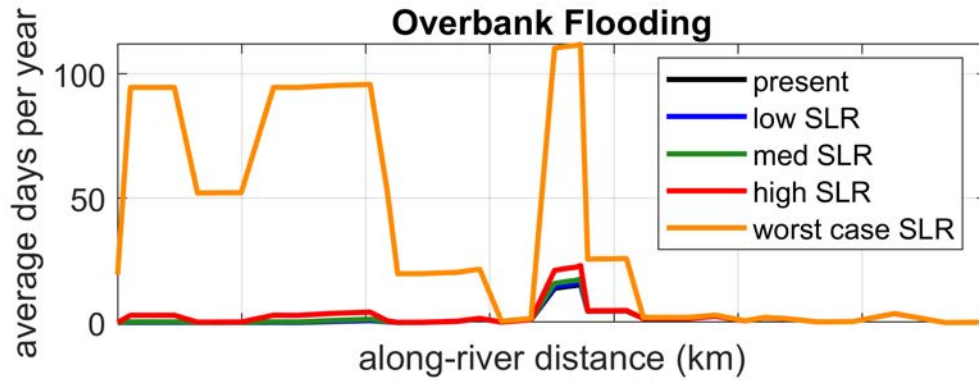


Figure 30 Average number of days per year of possible flooding as a function of river kilometer. For present-day climate, low sea level rise (SLR), medium SLR, and high SLR.

References

The majority of citations for the methods developed are in Appendices A and B.

- Barnard, P., D. Hoover, D. Hubbard, A. Snyder, B. Ludka, T. Gallien, J. Allan, G. Kaminsky, P. Ruggiero, T. Gallien, L. Gabel, D. McCandless, H.M. Weiner, N. Cohn, D. Anderson, and K. Serafin. 2017. “Extreme Oceanographic Forcing and Coastal Response due to the 2015-2016 El Niño.” *Nature Communications*, 8: 14365. <https://doi.org/10.1038/ncomms14365>.
- Barnard, P. L., M. van Ormondt, L. H. Erikson, J. Eshleman, C. Hapke, P. Ruggiero, P. N. Adams, and A. C. Foxgrover. 2014. “Development of the Coastal Storm Modeling System (CoSMoS) for Predicting the Impact of Storms on High-Energy, Active-Margin Coasts.” *Natural Hazards*, 74 (2): 1095–1125. <http://dx.doi.org/10.1007/s11069-014-1236-y>.
- Dalton, M. M., L. Benda, M. Case, S. Chisholm Hatfield, N. Cohn, M. Conlin, J. Lawler, P. Mote, D. Sharp, G. Reeves, P. Ruggiero, and K. Serafin. 2016. *Climate Change Vulnerability Assessment for the Treaty of Olympia Tribes*. A Report to the Quinault Indian Nation, Hoh Tribe, and Quileute Tribe. Oregon Climate Change Research Institute, Corvallis, OR.
- Czuba, J. A., C. R. Barnas, T. E. McKenna, G. B. Justin, and K. L. Payne. 2010. *Bathymetric and Streamflow Data for the Quillayute, Dickey, and Bogachiel Rivers, Clallam County, Washington, April–May 2010* (Vol. 537). U.S. Department of the Interior, U.S. Geological Survey.
- Hemer, M. A., Y. Fan, N. Mori, A. Semedo, and X. L. Wang. 2013. “Projected Changes in Wave Climate from a Multi-Model Ensemble.” *Nature Climate Change* 3 (5): 471–476. <https://doi.org/10.1038/NCLIMATE1791>.
- Lipiec, E., P. Ruggiero, A. Mills, K. Serafin, J. Bolte, P. Corcoran, J. Stevenson, and C. Zanolco. 2018. “Mapping out Climate Change: Assessing How Coastal Communities Adapt Using Alternative Future Scenarios.” *Journal of Coastal Research* 34 (5): 1196–1208. <https://doi.org/10.2112/JCOASTRES-D-17-00115.1>.
- Miller, I. M., H. Morgan, G. Mauger, T. Newton, R. Weldon, D. Schmidt, M. Welch, and E. Grossman. 2018. *Projected Sea Level Rise for Washington State – A 2018 Assessment*. A collaboration of Washington Sea Grant, University of Washington Climate Impacts Group, Oregon State University, University of Washington, and US Geological Survey. Prepared for the Washington Coastal Resilience Project.
- Mills, A. K., J. P. Bolte, P. Ruggiero, K. A. Serafin, E. Lipiec, P. Corcoran, J. Stevenson, C. Zanolco, and D. Lach. 2018. “Exploring the Impacts of Climate and Management on Coastal Community Resilience: Simulating alternative future scenarios.” *Environmental Modeling and Software* 109: 80–92. <https://doi.org/10.1016/j.envsoft.2018.07.022>.
- Mote, P. W., J. T. Abatzoglou, and K. E. Kunkel. 2013. “Climate: Variability and Change in the Past and the Future: Chapter 2.” In *Climate Change in the Northwest: Implications for Our*

Landscapes, Waters, and Communities, edited by M. M. Dalton, P. W. Mote, and A. K. Snover, 25–40. Washington, DC: Island Press.

- Parker, K., P. Ruggiero, K. Serafin, and D. Hill. 2019. “Emulation as an Approach for Rapid Estuarine Modeling.” *Coastal Engineering* 150: 79–93. <https://doi.org/10.1016/j.coastaleng.2019.03.004>.
- Ruggiero, P. 2013. “Is the Intensifying Wave Climate of the U.S. Pacific Northwest Increasing Flooding and Erosion Risk Faster Than Sea-Level Rise?” *Journal of Waterway, Port, Coastal, and Ocean Engineering* 139 (2): 88–97. [https://doi.org/10.1061/\(ASCE\)WW.1943-5460.0000172](https://doi.org/10.1061/(ASCE)WW.1943-5460.0000172).
- Ruggiero, P., P. D. Komar, W. G. McDougal, J. J. Marra, and R. A. Beach. 2001. “Wave Runup, Extreme Water Levels and the Erosion of Properties Backing Beaches.” *Journal of Coastal Research* 17 (2): 407–419. <http://www.jstor.org/stable/4300192>.
- Serafin, K. A. and P. Ruggiero. 2014. “Simulating Extreme Total Water Levels Using a Time-Dependent, Extreme Value Approach.” *J. Geophys. Res. Oceans* 119: 6305–29. <https://doi.org/10.1002/2014JC010093>.
- Serafin, K. A., P. Ruggiero, and H. F. Stockdon. 2017. “The Relative Contribution of Waves, Tides, and Non-Tidal Residuals to Extreme Total Water Levels on US West Coast Sandy Beaches.” *Geophys. Res. Lett.* 44: 1839–1847. <https://doi.org/10.1002/2016GL071020>.
- Serafin, K., P. Ruggiero, K. Parker, and D. Hill. 2019. “What's Streamflow Got To Do With It? A Probabilistic Simulation of the Competing Oceanographic and Fluvial Processes Driving Extreme Along-River Water Levels.” *Nat. Hazards Earth Syst. Sci. Discuss.*, in review. <https://doi.org/10.5194/nhess-2018-347>.
- Wang, X. L., Y. Feng, and V. R. Swail. 2014. “Changes in Global Ocean Wave Heights as Projected using Multimodel CMIP5 Simulations.” *Geophysical Research Letters* 41 (3): 1026–1034. <https://doi.org/10.1002/2013GL058650>.
- Zscheischler, J., S. Westra, B. J. Hurk, S. I. Seneviratne, P. J. Ward, A. Pitman, A. AghaKouchak, D. N. Bresch, M. Leonard, T. Wahl, and X. Zhang. 2018. “Future Climate Risk from Compound Events.” *Nature Climate Change* 8: 469–477. <https://doi.org/10.1038/s41558-018-0156-3>.

Appendix A

The following is a reproduction of the publication:

Parker, K., P. Ruggiero, K. Serafin, and D. Hill. 2019. “Emulation as an Approach for Rapid Estuarine Modeling.” *Coastal Engineering* 150: 79–93.
<https://doi.org/10.1016/j.coastaleng.2019.03.004>.



Emulation as an approach for rapid estuarine modeling

Kai Parker^{a,*}, Peter Ruggiero^b, Katherine A. Serafin^c, David F. Hill^a

^a School of Civil and Construction Engineering, Oregon State University, Corvallis, OR, 97330, USA

^b College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, 97330, USA

^c Department of Geophysics, Stanford University, Stanford, CA, 94305-2215, USA

ARTICLE INFO

Keywords:

Emulation
Gaussian process regression
ADCIRC + SWAN
Estuary
Probabilistic modeling
Water levels

ABSTRACT

Probabilistic flood hazard assessment is a promising methodology for estuarine risk assessment but currently remains limited by prohibitively long simulation times. This study addresses this problem through the development of an emulator, or surrogate model, which replaces the simulator (in this case the coupled ADCIRC + SWAN model) with a statistical representation that is able to rapidly predict estuarine variables relevant to flooding. Emulation of water levels (WLs), non-tidal residual, and significant wave height, is explored at Grays Harbor, Washington (WA) USA using Gaussian process regression. The effectiveness of the methodology is validated at various model simplification levels to determine where error is being sourced. Emulated WLs are found to be skillful when compared to over a decade of tide gauge observations (root mean square error, RMSE, < 15 cm). The largest loss of skill in the method originates with ADCIRC + SWAN attempting to reproduce observations, even when the majority of relevant physics are included. Subsequent simplifications to the simulator (input reduction techniques) and the emulator itself are found to introduce a trivial amount of error (average increase in RMSE of 1 cm). Emulated WLs are also compared to spatially varying observations and found to be equally skillful throughout the estuary. An example emulation application is explored by decomposing the relative forcing contributions to extreme WLs across the study site. Results show a compound nature of extreme estuarine WLs in that all forcing dimensions contribute to extremes, with streamflow having the least influence and tides the largest. Overall the approach is shown to be both skillful and efficient at reproducing critical hydrodynamic variables, suggesting that emulation may play a key role in improving our ability to probabilistically assess flood risk in complex environments as well as being promising in a range of other applications.

1. Introduction

Modeling estuarine hydrodynamics remains both a challenge and a goal for the scientific community. Estuaries and bays are often densely populated with significant economic and cultural investment (Pendleton, 2010). They are also subject to a unique flood hazard environment, with high water levels (WLs) driven by numerous contributing processes including both offshore and local waves, storm surge, and river inflows, among others. Over the past several decades, research efforts have led to improved computational models and increased physical understanding of estuarine flood dynamics (Bode and Hardy, 1997; Kantha and Clayson, 2000; Ganju et al., 2015). However, increasing hydrodynamic model predictive skill is generally coupled to increasing complexity within numerical models and a correspondingly larger computational load. This has led to computational time, rather

than a physical understanding of the problem, being a limiting control on our ability to answer questions about estuarine flooding.

Increasing computer processing power and code parallelization has pushed the boundary for what can be explored with complex computer codes. However, even with these advances, many questions still cannot be comprehensively addressed due to computational limitations. One example is the recent focus by the scientific community on uncertainty in model results (Mastrandrea et al., 2010; Green et al., 2011). In the field of flood hazards, a major thrust area has been probabilistic assessments, which brings the benefits of uncertainty quantification, utility as a stakeholder-centered decision making tool, better handling of extreme events, and more skillful flooding estimates (Cloke and Pappenberger, 2009; Di Baldassarre et al., 2010; Dale et al., 2014). However, the combination of multiple model iterations (required for probabilistic modeling) and large per-run computational costs has

* Corresponding author. Permanent Address: 1717 7th Street, Los Osos, CA, 93402, USA.

E-mail addresses: kaiparker@gmail.com, parkerk@oregonstate.edu (K. Parker), pruggier@ceoas.oregonstate.edu (P. Ruggiero), kserafin@stanford.edu (K.A. Serafin), david.hill@oregonstate.edu (D.F. Hill).

<https://doi.org/10.1016/j.coastaleng.2019.03.004>

Received 5 November 2018; Received in revised form 5 March 2019; Accepted 18 March 2019

Available online 23 March 2019

0378-3839/ © 2019 Elsevier B.V. All rights reserved.

remained a barrier for moving forward.

Often the solution to long simulation times is a compromise, such as simplifying or eliminating various forcing components (Purvis et al., 2008; Lin et al., 2010); using smaller ensemble sizes (Davis et al., 2010); or simplifying model physics (Dawson et al., 2005; Moel et al., 2012). A promising recent development has been to implement variable model complexity, with a fast model determining relevant or extreme events and a more highly-resolved, accurate model being used to simulate the extremes (Lin et al., 2010, 2012; Orton et al., 2016). This technique has been successfully demonstrated for hurricane-induced flooding but is potentially problematic for other regions. For example, environments not dominated by tropical cyclones often are defined by compound events where combinations of non-extreme forcings can combine to create extremes (Leonard et al., 2014; Wahl et al., 2015; Mofakhari et al., 2017; Zscheischler et al., 2018). In addition, event based techniques can still be considered computationally limited as the full parameter space cannot usually be explored. There remains a need for a modeling technique that can bridge the gap between time-intensive, complex models and fast simulation times.

This paper investigates emulation as a technique for the efficient prediction of estuarine hydrodynamic variables in Grays Harbor, Washington (WA) USA. The foundational idea of emulation (also referred to as surrogate modeling, response surface modeling, and meta-modeling, among others) is the replacement of a slower processes-based model (a simulator) with a fast, statistical model (an emulator) (O'Hagan, 2006; Razavi et al., 2012). In the standard modeling paradigm, the map between simulator inputs and outputs is based on the laws of physics as implemented within a process-based model (Castelletti et al., 2012). In emulation, this map is approximated using a statistical model. The benefit is that, following an upfront computational expense to create a training dataset and train the emulator, applying the emulator is nearly instantaneous. Thus, emulation represents a tradeoff between short simulation times and errors associated with the approximation. This tradeoff suggests that emulation may be ideal for probabilistic flood modeling along with many other potential applications including assessments of model uncertainty, model optimization, sensitivity analysis, real time forecasting, and extreme event analysis (Oakley, 1999; Kennedy et al., 2006; Levy and Steinberg, 2010).

The general concept of emulation originated in the 1980s through the idea of computer experiments (Sacks et al., 1989). Since then, emulation ideas have spread widely resulting in a rich literature of applications, emulator formulations, and theories from numerous fields. Razavi et al., (2012) reviews emulation in the field of water resources, with over 30 studies revealing a wide range of applications and emulation approaches. As a brief overview of coastal applications, Gouldby et al., (2014), Malde et al., (2016b) and Rueda et al., (2016) successfully implemented emulators for wave prediction problems using SWAN (Booij et al., 1997) as a simulator. The pairing of SWAN and emulation was extended to delineating offshore conditions causing wave induced coastal flooding by Rohmer and Idier, (2012) by using kriging and an adaptive sampling technique. Timmermans (2015) used emulation to explore how tuning parameters affect uncertainty in results from the Wave Watch III (Tolman, 2009) wave model. Liu and Guillas (2017) investigated the effect of uncertainty in bathymetry on tsunami height predictions using a novel merging of Gaussian process regression (GPR) emulation with dimensional reduction techniques.

In the context of flooding, emulation has been applied to river channel flooding (Apel et al., 2008) and coastal dyke systems (Moel et al., 2012), although from the relatively simplistic perspective of lookup tables. Surge response functions (SRF; Resio et al., 2009; Song et al., 2012) can be considered a specific case of emulation through regression of dimensionless cyclone scaling terms. However, SRFs are limited in application to tropical cyclones, and have been shown to perform poorly in complex environments (Taylor et al., 2015). As an alternative to SRFs, Kim et al., (2015) used an artificial neural network

to emulate coupled ADCIRC+STWAVE calculated surge from tropical cyclones. This approach was enhanced by Bass and Bedient (2018) who used a similar strategy but with the addition of a coupled hydrologic model and GPR as the emulator formulation. Jia and Taflanidis (2013) and Jia et al. (2016) used GPR emulation for predicting tropical cyclone surges.

Overall, multiple studies have demonstrated the potential of emulation in a coastal hazard setting. Surge from tropical cyclones has, in particular, seen a variety of successful emulator implementations. This study builds on these recent efforts but explores an estuary in the USA Pacific Northwest (PNW) that does not experience tropical cyclone forcing. This results in a unique challenge in terms of handling diverse forcings and a potentially larger input parameter space, since there is no dominant forcing dimension. Other studies focused on predicting WLs, such as those by Jia and Taflanidis (2013), Jia et al. (2016), Kim et al., (2015), and Bass and Bedient (2018), reduce input dimensionality through considering only cyclones and using discrete cyclone characteristics as input dimensions. This study, however, considers a general application of emulating the coupled ADCIRC+SWAN (ADCIRC+SWAN; Dietrich et al., 2011) simulator in which any combination of forcings can be used to calculate WLs. This paper is intended as a rigorous investigation into the applicability of emulation in this new context. Therefore, the focus here is primarily on describing the methodology and validation and only a single application, decomposing extreme estuarine water levels, is presented.

2. Study sites and observations

2.1. Study site

Grays Harbor, WA (Fig. 1) is an excellent candidate for testing emulation as it exhibits many of the complexities that make estuarine modeling difficult. Grays Harbor is predominantly shallow, dominated by depths averaging less than 5 m, but also contains a maintained (United States Army Corps of Engineers; USACE) deep-water navigation channel giving it significant depth variability (Fig. 1). The bay exhibits spatial variability in WLs (Gialone and Kraus, 2001) as a result of its size (approximately 235.3 km², Engle et al., 2007), shape, and gradients in forcing. Grays Harbor is located in the PNW (Fig. 1) and is therefore subject to an energetic storm and wave climate. A Global Ocean Wave 2 (GOW2) reanalysis (Perez et al., 2017) near the study site (see Fig. 1) reveals a mean offshore significant wave height (Hs) of 2.5 m with events exceeding 7.5 m annually. Extreme storm events are generally associated with extratropical cyclones that can produce strong winds, pressure differentials, and precipitation (Allan and Komar, 2002a; Mass and Dotson, 2010). These events are often associated with significant non-tidal residuals (NTR) (Allan and Komar, 2002a, 2006; Allan et al., 2011; Serafin et al., 2017), although of a smaller magnitude than locations impacted by tropical cyclones or with broader continental shelves (Zhang et al., 1999). Within this study, NTR is defined as an observed or modeled WL with tides removed (with the specifics of how NTR is calculated detailed in section 4.3). Grays Harbor has significant hydrological input from the Chehalis, Humptulips, Hoquiam, Elk, and Johns Rivers which collectively drain a watershed of over 7 000 km² for an average monthly runoff volume of 22 million m³/month (Engle et al., 2007).

2.2. Observational data

This study utilizes a variety of observational datasets ranging from instrument deployments to reanalysis products. Forcing and model development datasets are explained in the following section (2.2.1), while section 2.2.2 details observations specifically used for model validation.

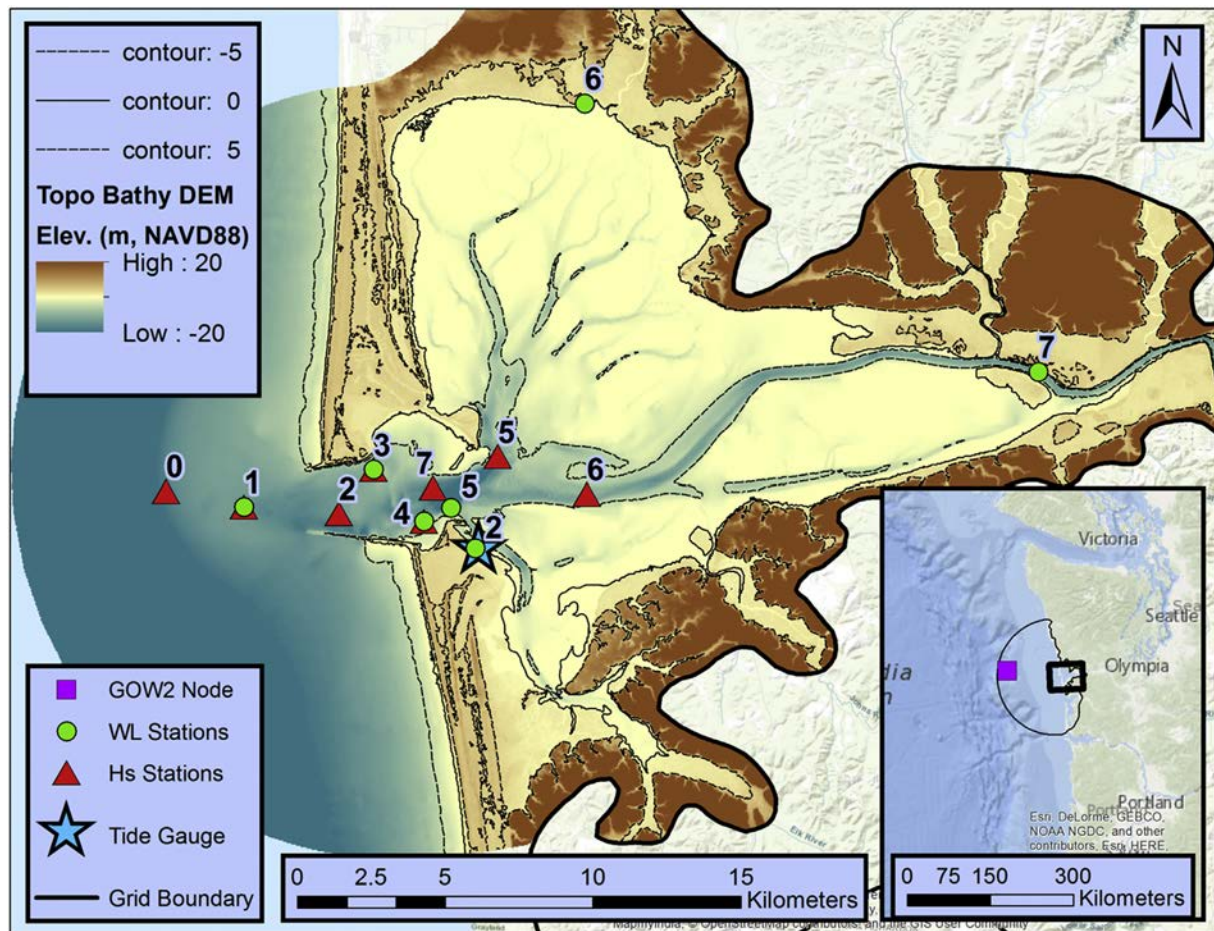


Fig. 1. Grays Harbor, WA study site and locations of observational datasets. Circles and triangles represent USACE deployments with co-located instruments labeled with a single number representing both WL and Hs stations. The main panel shows the bathymetry and topography of the estuary in the NAVD88 Datum. The inset panel shows the larger geographical context of the estuary with the thin black line delineating the domain of the hydrodynamic model. The purple square within the inset is the location of the utilized GOW2 node (located at 47° N, 125° W). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.2.1. Forcing and model development datasets

Wave forcing for the model was obtained from the GOW2 reanalysis of Perez et al., (2017) with output selected from a node located at (Lat: 47° N, Lon: 125° W; Fig. 1). Atmospheric forcing was provided by the North American Regional Reanalysis (NARR) (Mesinger et al., 2006). NARR provides a wide range of gridded atmospheric variables from which the 3-hourly 10 m wind fields and 3-hourly surface pressure fields were utilized. Streamflow was obtained from USGS river gauges with total estuary inflow constructed as the sum of three gauged rivers, the Chehalis, Satsop, and Wynoochee (USGS stations 12031000, 12035000, and 12037400 respectively). The Satsop and Wynoochee rivers are tributaries to the Chehalis river which join the Chehalis below the Chehalis gauge. Therefore, the sum of these three gauges reproduces the majority of the Chehalis flow into the Grays Harbor estuary. While Grays Harbor has other river inlets, the majority of the input flow is concentrated at the Chehalis River which captures around 80% of the watershed area. For simplicity, as well as due to temporal availability of gauge data, only the Chehalis input (as constructed from the three gauged rivers) is included in the study with all other streamflow inputs assumed to be minimal with only local influences on variables of interest.

The bathymetry data for the simulator grid were developed by blending two National Oceanic and Atmospheric Administration (NOAA) digital elevation models (DEMs): the Astoria, OR tsunami DEM (1/3 arc second) and the coastal relief model (3 arc seconds) (NOAA National Centers for Environmental Information, 2003; Love et al.,

2012). Bay topography was sourced from Oregon Department of Geology and Mineral Industries (DOGAMI) LiDAR (DOGAMI, 2010).

2.2.2. Validation datasets

In addition to forcing, a series of observational datasets were used to validate simulated and emulated variables within the study site. The first dataset is the Westport, WA tide gauge (NOAA station ID # 9441102) which provides continuous hourly WL data beginning in 2006. WL observations were decomposed into constituent components (e.g., deterministic tide, monthly mean sea level anomalies (MMSLA), storm surge etc.) using the approach described in Serafin and Ruggiero (2014). The five largest NTR events on record were extracted for testing model skill. A brief summary of these storm events is provided in Table 1.

Water level observations at the tide gauge were supplemented by a field campaign carried out by the USACE from September–December 1999 (Fig. 1, Cialone and Kraus, 2001, 2002). This dataset includes seven locations near the inlet with bottom mounted tripods measuring wave characteristics, WLs, tidal currents, and suspended sediment concentrations. Additionally, five surface stations were distributed throughout the bay measuring WLs, conductivity, and temperature. The USACE field campaign was broken up into two deployments (with a small maintenance/data collection break between the two). Instruments were replaced in approximately the same location except for Hs station 0 which was moved to location 7 for the second deployment (Cialone et al., 2002). Fig. 1 illustrates the spatial distribution of the

Table 1
Summary of forcing for the five largest NTR events at Grays Harbor, WA. Forcing values are reported at the occurrence of maximum NTR.

Date	Storm 1	Storm 2	Storm 3	Storm 4	Storm 5
	12/15/06	12/3/07	01/12/14	12/12/14	12/11/15
Non-Tidal Residual (m)	0.73	0.93	0.66	0.58	1.11
Significant Wave Height (m)	8.0	11.1	8.5	6.2	11.1
Peak Wave Period (sec)	12.7	15.4	14.3	12.2	18.2
Wave Direction (deg.)	243	195	259	229	248
Surface Pressure (hPa.)	977	989	989	986	984
Wind Speed (m/s)	17.5	19.2	14.9	9.1	12.3
Wind Direction (deg.)	217	201	238	218	198
Streamflow (m ³ /s)	1270	2020	860	670	1510

various observation stations which have been renamed in this paper for clarity.

3. Methods

3.1. Simulator configuration

This study utilizes the coupled Advanced Circulation (ADCIRC; Luettich et al., 1992) and unstructured Simulating Waves Nearshore (SWAN, Zijlema, 2010) simulator (ADCSWAN; Dietrich et al., 2011). ADCSWAN has seen extensive validation and success in predicting WLs and NTR at various estuaries around the world (Dietrich et al., 2012; Bhaskaran et al., 2013; Krien et al., 2015). Recently, ADCSWAN has been successfully implemented in the PNW with good agreement between simulator output and observations of WLs, NTR, and currents (Cialone et al., 2002; Cheng et al., 2015b). ADCSWAN is implemented in the 2D depth-integrated barotropic mode which has been shown to perform with acceptable error for computing WLs and depth integrated currents in estuaries (Resio and Westerink, 2008; Weaver and Luettich, 2010). ADCIRC is run in the fully 2-way coupled implementation with SWAN, which has been shown to be critical for resolving interactions between waves and nearshore hydrodynamics (Cialone et al., 2002; Funakoshi et al., 2008; Dietrich et al., 2010, 2011;). ADCSWAN is run on an unstructured mesh that extends beyond the continental shelf (approximately 115 km offshore; Fig. 1). Unstructured meshes provide flexibility in simulator resolution with the utilized model grid having element sizes ranging from around 7 000 m offshore to under 20 m within the inner Grays Harbor channel.

3.2. Dimensional reduction and levels of simplification

Emulator construction requires sampling the full input parameter space. This constraint dictates that the number of times the simulator must be run to create the training dataset is proportional to the number of dimensions included as inputs. In general, process-based hydrodynamic simulators are based on many inputs making some form of dimensional reduction necessary. Emulator construction thus requires finding a balance between minimizing the number of inputs and maintaining sufficient complexity to acceptably resolve output variables of interest.

Fig. 2 provides a conceptual model of the dimensional reduction approach taken in this study (through simplifications), transforming the full process-based simulator (ADCSWAN) into an emulator. Each of the simplifications, noted on the right side of Fig. 2, theoretically introduces some level of error into the output, noted on the left side of Fig. 2. These errors are discussed in this paper both as individual contributions, and in the cumulative (sum of all errors up to a given

level) sense. When discussed explicitly in this paper, simplification levels will be capitalized. For example, a comparison of model output from the level 3 simplification (Stationary Simulator) to Observations (no simplification) quantifies the cumulative level 3 error. The following sections (3.2.1–3.2.3) explain each simplification in this hierarchy while corresponding error is quantified in the Results section.

3.2.1. Simulator simplifications

The first level of simplification is simply that of using a process-based simulator. Simulators are unable to exactly reproduce observations for a variety of reasons ranging from incorrect or unresolved physics (e.g., assumptions, parameterizations, etc.) to numerical approximations (truncation errors, etc.) to incorrect or biased input forcing. The x induced by this simplification is primarily a function of the chosen model, model tuning, and the quality of forcing/bathymetric information. Research has shown that errors in model inputs such as bathymetry and mesh resolution (Bunya et al., 2010; Weaver and Slinn, 2010) and forcing fields (Madsen and Jakobsen, 2004; Lewis et al., 2013; Lakshmi et al., 2017) are significant sources of model error. Therefore, the specific configuration and choice of ADCSWAN (section 3.1) and the quality of observational data (section 2.2.1) are the primary controls on the impact of this simplification.

This study considers emulation of a specific implementation of the ADCSWAN model and therefore the model grid (bathymetry, resolution, etc.) is held constant. Additionally, ADCSWAN contains a large number of input switches, tuning parameters, forcing options, numerical configurations, and other choices (Westerink et al., 1992). This study holds all general model configuration parameters constant leaving the various forcing components of WL variability as the sole driver of input dimensionality within the emulator.

3.2.2. Forcing simplifications

Even with the simplification of holding the model configuration fixed, the input dimensionality remains high, due to the numerous physical forcing mechanisms. Below we describe simplifications that reduce the model dimensionality to 16. This reduction is desirable since it requires a smaller training dataset and therefore produces a more efficient emulator construction.

3.2.2.1. Wave simplification. It is well known that offshore wave energy can impact water levels within bays such as Grays Harbor (Olabarrieta et al., 2011; Cheng et al., 2015b). Wave forcing is implemented in the simulator using a JONSWAP spectrum fitted to peak wave period (T_p), H_s , mean wave direction (MWD), and directional spread parameters. While research has shown the importance of forcing with full directional spectra for reproducing wave observations (Rogers et al., 2007; Montoya et al., 2013), most studies accounting for wave influence on WLs use simpler bulk parameter-based formulations. Therefore, a fitted JONSWAP spectrum is used for both the Full (level 1) and Simplified Simulator (level 2) comparisons. Based on previous research in the PNW (Cheng et al., 2015a), directional spread is held constant at 20° , and wave forcing is applied uniformly along the Full Simulator open boundary (Fig. 1). With these simplifications, wave forcing is included in the emulator as three dimensions: H_s , T_p , and MWD.

3.2.2.2. Atmospheric simplification. Atmospheric forcing represents a unique challenge for emulation due to the spatial variability of wind and pressure fields. Gridded inputs represent a high degree of dimensionality, with every node potentially representing an input dimension. For this reason, a sensitivity study was undertaken to see if spatially constant atmospheric forcing could be used as an approximation of the full forcing fields. WL output from simulator runs with full gridded forcing were compared to runs with spatially constant forcing. Results indicated (not shown) that the error introduced in predicted WLs by the spatially constant assumption was

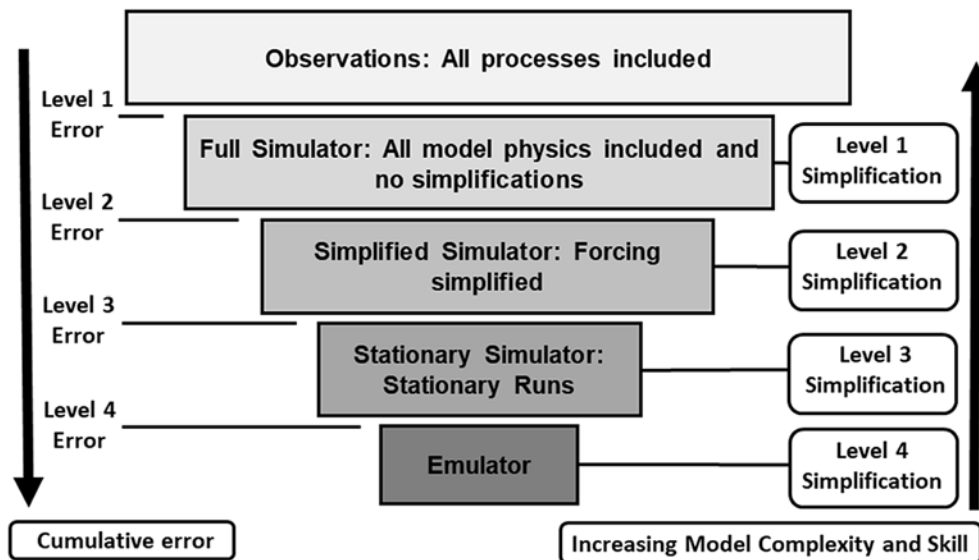


Fig. 2. Hierarchy of model simplifications between observations (top) and emulator output (bottom). Each simplification is associated with a level (right side of the figure) and some amount of error (defined on the left side of the figure).

acceptable in comparison to the corresponding reduction in dimensionality. This error is quantified in the Results section (along with other simulator simplifications) as level 2 error. Adopting the spatially constant assumption, atmospheric forcing is reduced in the emulator framework to three dimensions: wind speed, wind direction, and sea level atmospheric pressure.

3.2.2.3. Tidal simplification. Tidal forcing is generally represented in hydrodynamic models through harmonic constituents. Many studies using ADCIRC are forced with eight or fewer constituents, mainly because global databases of tidal constituents (e.g., TPXO (Dushaw et al., 1997), or LeProvost (Le Provost et al., 1994)) are typically limited to that number. Despite this, simulations using this small number of constituents are typically found to agree well with both harmonic analysis derived and observed tidal elevations (Westerink et al., 1992; Blain and Rogers, 1998; Blain et al., 2001). ADCIRC simulates tidal forcing as a boundary elevation time series (Luettich et al., 1992) determined by a spatially variable, temporally constant phase and amplitude and a temporally variable, spatially constant equilibrium argument and nodal factor. Amplitudes and phases are determined by the simulator boundary location and are therefore not an emulator input dimension when considering a fixed study site. The nodal factor represents adjustments of the amplitude/phase of each constituent that results from the nodal tide cycle. The equilibrium argument (deterministic based on date and time) controls the timing of the harmonic.

While tides are deterministic, they are included within the emulator as forcing for a variety of reasons which will be described in section 3.2.3. In approaching simplifications, a sensitivity test was performed to determine the tidal dimensionality required for accurately reproducing maximum WLs during storm events (Table 1). It was found that removing the nodal factor did not significantly change simulated WLs. After this simplification, results showed that eight harmonics (without nodal factors) were sufficient for accurately producing WLs. This allows tides to be included in the emulator as eight input dimensions: 8 harmonic equilibrium arguments, each ranging from 0 to 360°.

3.2.2.4. Streamflow simplification. Streamflow is represented in ADCIRC as a flux of water into the domain (specified as a normal flow per unit width of boundary). This allows the simulation of large rivers that have significant cross-channel velocity profiles and for calibration where data on these cross-channel profiles are available. For this study, we

instead specify a laterally constant velocity profile across each river boundary. This simplification is common (Bunya et al., 2010; McKay and Blain, 2010), especially if the boundary is far enough away from the area of interest that a natural flow profile can develop. This allows streamflow to be represented as a single input dimension (the total volumetric flow rate) for each river inlet.

3.2.2.5. Base water level simplifications. A final input dimension is considered within the emulator framework as a “Base WL” parameter. This is included to account for large scale changes to estuary sea level, as is experienced through MMSLAs, seasonal variability, and sea level rise (SLR) (Serafin and Ruggiero, 2014). These forcing dimensions are defined in the simulator simply as a static change to mean sea level and are therefore included in the emulator as a single input dimension.

3.2.3. Simulator stationarity simplification

ADCIRSWAN and other process-based hydrodynamic simulators are dynamic in that both inputs and outputs are functions of time and the simulator state is determined, in part, by previous states. Seeking simplicity, this study makes the assumption that the dynamic system can be approximated using a series of stationary simulations. Precedents for such an assumption exist for coastal systems, including spectral evolution in wave modeling (SWAN) approximated using a series of steady-state simulations (Rogers et al., 2007; Rusu and Pilar, 2008).

Simplifying tidal forcing with stationary simulations is difficult since there is no tidal equilibrium in WLs. One approach would be to consider tides as a series of horizontal water surfaces of different elevations (corresponding to tidal phases). This would reduce tidal forcing dimensionality to a single value (tidal WL), but at the cost of losing spatial variability. Testing showed that, for the Grays Harbor study site, tidal wave evolution and propagation across the estuary results in significant spatial variability in tidally forced WLs. A second approach would be to decouple NTR and tidal WLs and add the two as a linear summation. However, further testing confirmed that this simplification results in significant error. Therefore, a hybrid solution was developed in which all non-tidal forcing is stationary, but tides are computed dynamically with model output recorded only at the specific moment of interest. This approach is appropriate since tides are deterministic and, for a specific set of equilibrium arguments, the previous state of tide induced WLs will always be the same. This approach allows tidal forcing to be simplified but retains the spatial variability in tidal WLs and

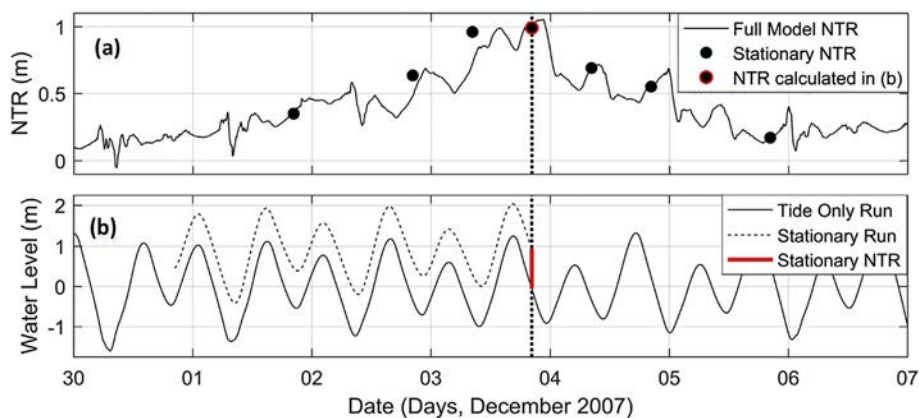


Fig. 3. Panel (a): Comparison of NTR during storm 2 from a fully dynamic simulation (black line) and simplified stationary simulations (black dots). Panel (b): Example stationary run (at the peak of the storm) showing how the stationary NTR is calculated. The horizontal bold dotted line represents the time of the stationary run. At this time, NTR is calculated by subtracting the value of a tide only run from the value of the stationary run (Bold red line). This NTR value is plotted as a red outlined dot in panel (a). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the nonlinear interactions with other processes.

Fig. 3 illustrates how stationary runs compare to the Full Simulator (dynamic). Fig. 3a compares NTR from the fully forced ADCSWAN (simplification level 1; black line) and seven stationary ADCSWAN runs (simplification level 3; black dots) during storm 2 (Table 1). NTRs are computed for both cases by subtracting a ‘tides only’ simulation from the fully forced model. Fig. 3b demonstrates how the stationary NTR is computed for the peak of storm 2. This NTR value is plotted in Fig. 3a as a red outlined dot. The agreement between the fully dynamic run and the seven stationary runs was found to be sufficient, with an RMSE error for storm 2 of 11 cm.

3.3. Experimental design

A conceptual overview of the process used for constructing an emulator, in the context of this study, is provided in Fig. 4.

The first step in building an emulator is the selection of design points (experimental design) to create the training dataset. This study implements a design from the commonly utilized Latin Hypercube sampling (LHS) family of schemes first explored by McKay et al., (1979). LHS is one of the oldest and most popular experimental designs and has been found to perform well for complex simulators (Jones and Johnson, 2009). The specific experimental design for this study was created using a “maximin” LHS design (Johnson et al., 1990; Morris and Mitchel, 1995) from the LHS package in R (Carnell, 2017).

Parameters required for a LHS design are the number of dimensions to be included, the range of each dimension, and the number of design points. As detailed in section 3.2, this study used an input parameter dimensionality of 16, including wind speed and direction, sea surface pressure, Hs, Tp, MWD, streamflow, base WL, and eight tidal equilibrium arguments. LHS considers only the maximum and minimum values of each dimension with design points spaced approximately uniformly across dimensions. Ranges were chosen for each parameter in an attempt to span all plausible forcing scenarios. This was determined by looking at 100-year return level events as calculated from the observational records. The size of the training dataset is typically controlled by the cost of running the simulator, but Loepky et al., (2009) provide the general guidance that the training dataset should be approximately 10 times the number of dimensions of the input space. Given the 16 input dimensions of this study, this suggests a theoretical

training dataset size of 160 runs. To explore the relationship between training dataset size and emulator skill and to validate the emulator’s overall effectiveness, this study conservatively developed a larger training dataset consisting of 480 ADCSWAN runs.

3.4. Emulator configuration

A variety of formulations have previously been used in an emulation context, including support vector machines, artificial neural networks, radial basis functions, and many others (Jin et al., 2001; Gano et al., 2006; Razavi et al., 2012). This study uses GPR, (also referred to as Kriging), a Bayesian statistical non-parametric regression model well suited to this particular application as it scales well to high-dimensional input and intrinsically considers model uncertainty (O’Hagan, 2006; Levy and Steinberg, 2010). Furthermore, GPR is a general and flexible framework that can be optimized for a variety of modeling problems (Rasmussen and Williams, 2006). For example, many other common emulator formulations, such as neural networks (Rasmussen and Williams, 2006) and radial basis functions (Anjyo and Lewis, 2011), can be shown to be equivalent to GPR under specific conditions.

The foundational definition of a Gaussian process is that of an infinite collection of variables for which any finite subset is described by a multivariate Gaussian distribution. Every point in the input space can be modeled as a random variable (due to uncertainty about the functional response to inputs). A Gaussian process governs how these variables are related. A common way of thinking about GPR is as a distribution over functions (Rasmussen and Williams, 2006). This is mathematically tractable as a GPR can be completely defined by a mean and covariance function (due to being modeled as a multivariate Gaussian distribution). From a Bayesian perspective, this means a GPR is specified using a prior mean and covariance function. The data then updates this prior, using Bayesian inference, with information about the true form of the function to develop the posterior. The mean posterior function is then the most probable function (considering all possible functions) given the data that has been observed.

This process is conceptualized for a one-dimensional case in Fig. 5. The effect of the Bayesian conditioning on the emulator can be seen as “anchoring” the posterior sample functions (and uncertainty) at locations of observations. This limits the possible functions to those that go through these observed points. Uncertainty is quantified by considering

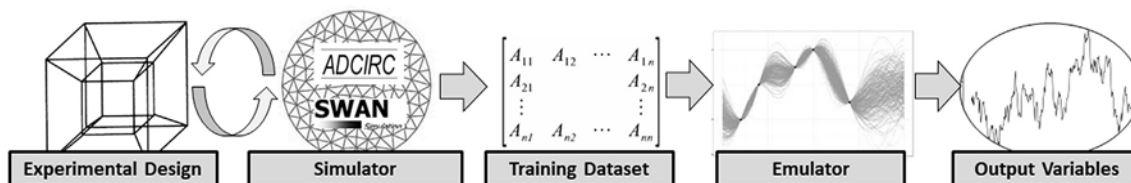


Fig. 4. Conceptual framework for developing an estuarine hydrodynamic emulator.

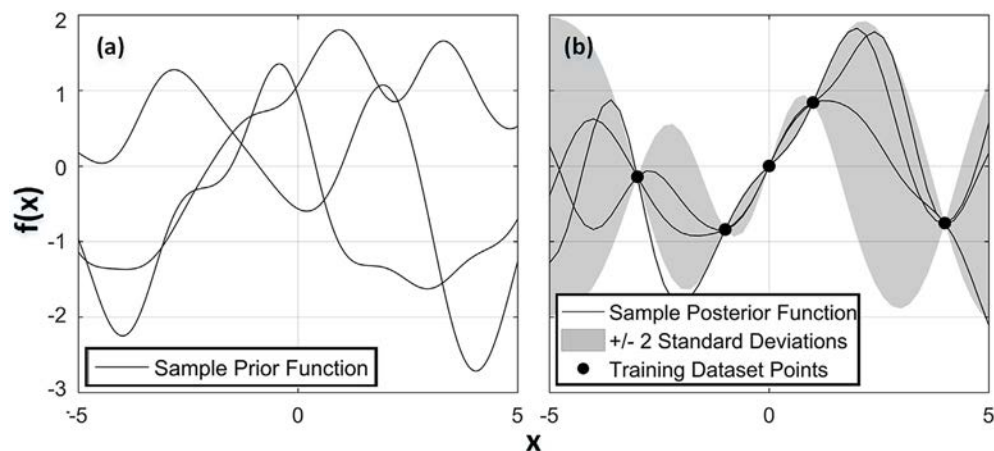


Fig. 5. Example 1-D application of GPR for determining $f(x)$ from observations. Panel (a) shows three random sample functions drawn from the prior distribution. A non-informative prior is specified so the average over functions has a zero mean. Panel (b) shows 3 random sample functions drawn from the posterior distribution after 4 training observations (dark black points). The effect of training is to constrain possible functions to only those that go through observation points. In panel (b), the shaded region represents plus and minus 2 standard deviations from the mean posterior prediction. Figure after Rasmussen and Williams (2006).

the possible functions that pass through these training points.

The first component of a Gaussian process is the mean function, which defines the mean of the infinite set of functions that are being considered. A common choice is to set the prior mean to zero, which can be thought of as a non-informative prior where the form of the function between inputs and outputs is unknown. This is demonstrated in Fig. 5, as shown by the approximate mean of the sample prior functions being zero. As an alternative, this study follows the methodology of Timmermans (2015) who used a simple linear regression to obtain information about the mean function's form. Residual analysis of our data showed a cubic relationship for the tidal equilibrium argument terms, a somewhat expected result due to the cyclic nature of tides. Based on this result, the prior mean function was defined with a cubic term for all tidal equilibrium argument inputs and a linear term for all other inputs. A k-fold cross-validation (see section 4.2) was performed to evaluate emulator skill with and without the modified mean function. Results showed a significant gain in skill (both in terms of RMSE and Determination, R^2) by using the modified mean function.

The covariance function of a Gaussian process is the second necessary component for defining the emulator. The covariance function (often called the kernel) can be thought of as describing the relationship between points in the process. Practically this describes the smoothness of the resulting GPR. In general, the covariance function contains hyper-parameters describing the details of the relationship between points (e.g., parameters such as length-scale, signal variance, etc.). These parameters can be inferred from the data, which is commonly done through maximizing the marginal likelihood rather than full Bayesian inference (Schulz et al., 2018). This was the approach used for this study.

A comparison of model performance was performed using 3 commonly used covariance functions: the Gaussian, squared exponential, and Matern (Rasmussen and Williams, 2006). The Matern covariance function was tested with ν (a parameter controlling smoothing) equal to 1.5 and 2.5. The best performing model was evaluated using k-fold cross-validation and comparing model RMSE values (Kohavi, 1995; Arlot and Celisse, 2010). K-fold cross-validation breaks the total training dataset into k segments and cycles through every possible combination of withholding one segment for validation and training the emulator with the remaining segments. This results in an ensemble of skill metrics for which the mean is less biased and more robust to the training period than a standard validation methodology (Arlot and Celisse, 2010). It was found that the Matern ($\nu = 2.5$) performed the best and therefore was utilized for all results found in the following section. The training of the emulator was performed using the Managing Uncertainty in Complex Models (MUCM) package in R (Malde et al., 2016a).

4. Results

4.1. Error introduced by model simplifications

With the construction of the emulator being hierarchical (Fig. 2), it becomes important to assess the skill of the emulator at multiple simplification levels to determine where errors are being introduced. This was investigated for simulator simplifications by looking at the 5 largest storms, in terms of NTR, on record (Table 1). Storm events were chosen for this analysis as it is expected that the strong forcing gradients and rapidly changing dynamics of these events would provide the most robust test of simulator simplifications. For each storm, WLs were calculated using the Full Simulator (dynamic, non-simplified) and the Stationary Simulator (all simplifications except for emulation) to quantify the sum of level 2 and level 3 errors. This comparison was performed at six or seven (seven except for storm 2) temporally random points distributed across each storm. Fig. 6 shows the difference between calculated WLs (level 1 simplification minus level 3 simplification) at two locations: the tide gauge (Fig. 6a) and WL station 7 (Fig. 6b). This difference is denoted here as the “error” resulting from simplifying the simulator. Two locations are plotted to visually sample how error is affected by location within the estuary.

The error computed via this test was found to have a max of 25 cm

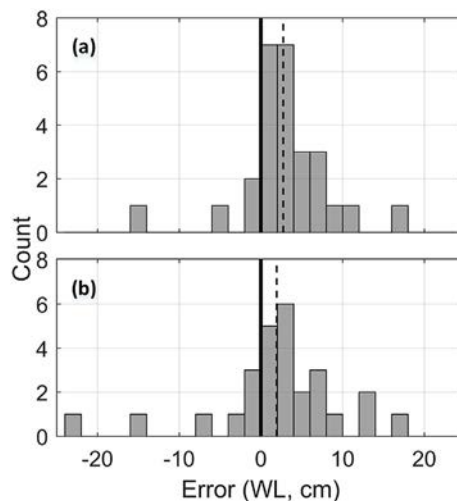


Fig. 6. Error between Full and Stationary (Level 1 and Level 3) Simulator calculated WLs. The histogram includes comparisons for storms 1–5 (Table 1). The bold line represents zero error while the dotted line is the mean error of the ensemble. Subplot (a) is computed at the tide gauge location while subplot (b) is at WL station 7 (located deeper within in the estuary; Fig. 1).

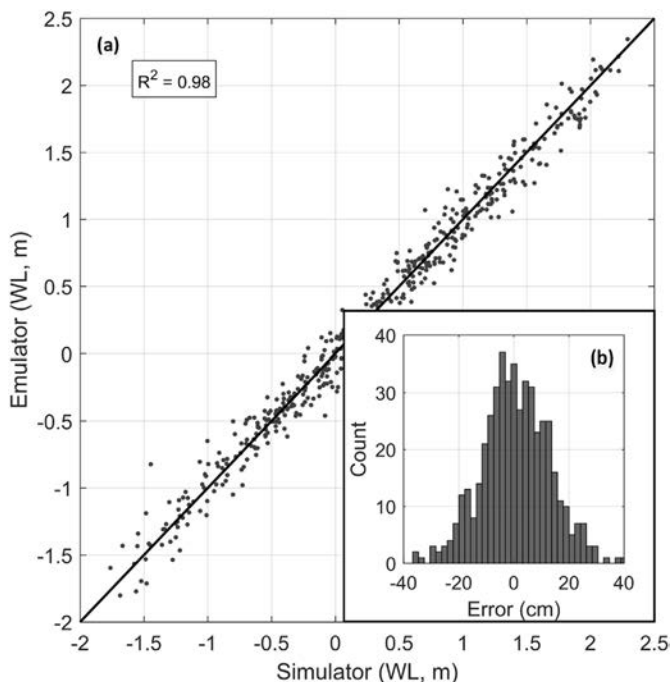


Fig. 7. Panel (a): Simplified Simulator vs. emulator WLs for the full training dataset. The comparison was performed using a 5-fold validation procedure. Panel (b): histogram of the error between Simplified Simulator and emulator WLs.

with a RMSE of 6 cm at the gauge location. The maximum RMSE for approximately 100 test stations randomly scattered across the estuary domain was found to be 9 cm. The Simplified Simulator was found to be only slightly biased with a mean approximately 2 cm lower than the Full Simulator (represented as a positive mean error in Fig. 6). Additional snapshot runs were performed to examine the model's error for non-storm conditions (not shown). Results confirmed that the Simplified Simulator, on average, performs better for non-storm conditions, suggesting that the results in Fig. 6 are likely conservative.

4.2. Emulator validation

The ability of the emulator to replicate the Stationary Simulator (level 4 error) was quantified using a k-fold cross-validation. Fig. 7 shows the results of the cross-validation with 5 segments comparing emulated WLs and simplified simulator WLs.

Overall the emulator was found to perform well at this comparison level with a high level of skill. The emulator shows little bias (mean of the residuals is less than 1 cm), and relatively even variance in residuals

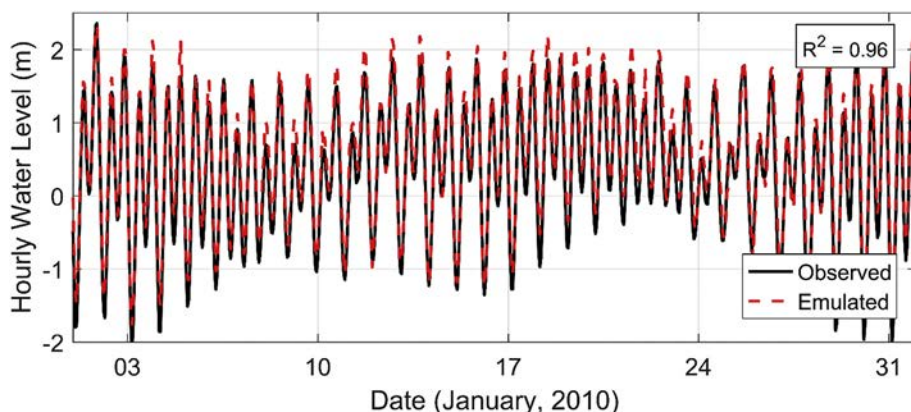


Fig. 8. Comparison of emulated (red dashed line) and observed (black line) hourly WLs at the Grays Harbor tide gauge for January 2010. Coefficient of determination is calculated using the entire tidal record (2006–2016). WLs are plotted in reference to mean sea level. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

across WL magnitude (Fig. 7a). However, the width of the histogram in Fig. 7b suggests that this step introduces more error than simulator simplifications. The RMSE was found to be around 13 cm (level 4 error) which is significantly larger than the calculated simulator simplification error (Fig. 6, sum of level 2 and level 3 error) of approximately 6 cm. However, this comparison of RMSEs is imperfect as the level 4 error assessment is based on a larger sample size and more rigorous k-fold validation and the level 2 and level 3 error assessment only examined performance during storm events.

4.3. Emulator performance: Westport, WA tide gauge

The next stage of quantifying the skill of the emulator is to compare emulated WLs to observations at the tide gauge. This test provides a measure of the cumulative level 4 error, or the total integrated error from predicting observed WLs using emulation. This analysis was performed by using the emulator to hindcast hourly WLs at the location of the tide gauge for the entire period of record (2006–2016). Comparison between tide gauge observations and hourly emulated WLs for a randomly chosen month long segment are shown in Fig. 8. Overall, hourly emulated WLs (for the over 10-year long record) compare favorably to the tide gauge with an R^2 value of greater than 0.96, RMSE of approximately 15 cm, and a bias of less than 1 cm.

As with tide gauge records, WL output from an emulator can be considered as the sum between two components: tides and NTRs. In the PNW, tides are the dominant source of WL variability (Allan and Komar, 2002b) and so the skill of the emulator in predicting WLs is primarily controlled by its ability to reproduce the deterministic tides. However, coastal hazard research often considers NTR individually as the driver of extreme WLs on top of regular, and well predicted, tidal cycles. Therefore, it is additionally important to test the emulator's skill at reproducing NTR signals. Furthermore, this provides a more robust test of emulator performance as NTR is not explicitly modeled as an output by the emulator.

NTRs are often calculated at tide gauges by subtracting the predicted tide (determined through harmonic analysis) from the measured WL. This procedure can be problematic since NTR may affect tidal phase, resulting in a false NTR signal from out of phase tidal signals (Pugh, 1996; Haigh et al., 2014; Serafin and Ruggiero, 2014). Therefore, for this study NTRs were calculated from tide gauge data using the procedure of Serafin and Ruggiero (2014) (modeled after Bromirski et al., 2003), which uses spectral filtering to remove energy from tidal bands.

The Bromirski et al., (2003) methodology is not used to determine NTRs from the ADCSWAN simulations. Storm simulations are on the scale of weeks which is too short temporally to recover energy across all tide bands of interest in the frequency domain. Instead, NTRs from the ADCSWAN simulations (at all simplification levels) and emulator simulations were calculated as a full forcing run minus a tidal only run.

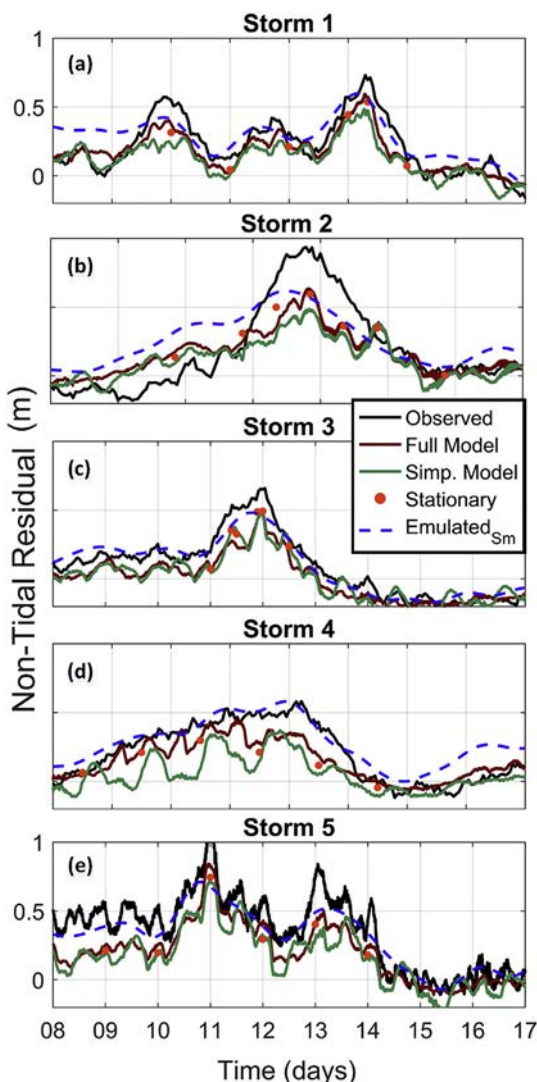


Fig. 9. Tide Gauge comparison of observed and modeled NTR at various simplification levels. Each subpanel (a through e) is one of the top 5 storms of record (see Table 1). Due to windowing for spectral filtering, storm 5's observed NTR is calculated using the subtraction method rather than the Bromirski method. All panels have the same y-axis scaling. The specific dates on the x axes vary by storm, but each tick represents a day.

Emulated NTRs were then subsequently smoothed with a loess filter to reduce noise associated with tidal phase mismatches between the tide-only and full forcing emulated time series.

Comparison of observed and hindcasted NTRs for storms 1–5 (Fig. 9) show a good overall performance of the emulator. To contextualize this comparison, the Full, Simplified, and Stationary Simulator calculated NTR time series are all plotted. A quantitative comparison of error between observed and modeled NTRs found that all simplification levels (from full ADCSWAN to emulator) have an RMSE of approximately 14 cm plus or minus 1 cm. The similar error across all simplification levels suggest that the largest source of error for NTR is in Full Simulator itself (level 1). For example, in Fig. 9b it is clear that the Full Simulator (red line) is unable to reproduce the peak NTR signal in storm 2.

4.4. Emulator performance: USACE field campaign

The tide gauge provides a rich dataset for validating the emulator due to its record length but is spatially limited to a single comparison point within the study area. One key strength of emulation, in

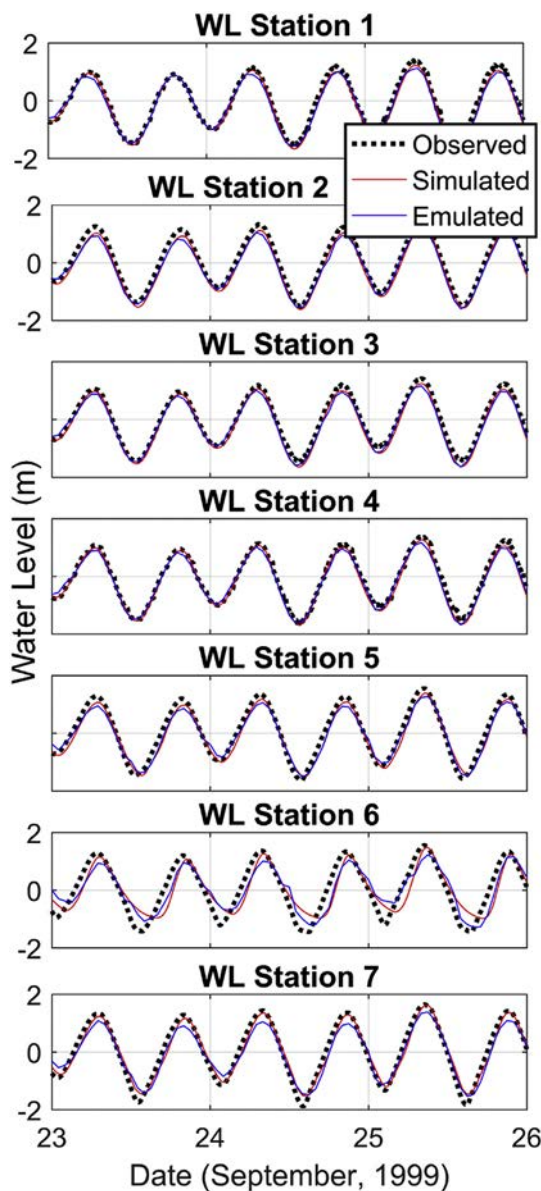


Fig. 10. Comparison of observations (USACE deployments, Fig. 1), Full Simulator and emulated WLs. All panels have the same y-axis and x-axis scaling.

comparison to a fully data driven methodology, is the ability to provide WL information across study sites where observational information may not be available. An emulator can be constructed at any location within the ADCSWAN model domain where output is provided. Fig. 10 evaluates the spatial performance of emulation through a comparison of emulated and observed WL time series for a 1999 field campaign led by the USACE in Grays Harbor (Fig. 1, Cialone and Kraus, 2001, 2002).

Fig. 10 shows good performance between observed and modeled WLs across most locations. The main exception to this is WL station 6 which displays comparatively poor agreement between the observed and emulated WLs. This lack of skill is equally shown by the Full Simulator and is therefore not a result of the emulation procedure. Table 2 gives RMSE values for a comparison between observations and modeled output at various levels of simplification (level 1 simulations and level 4 emulations are shown in Fig. 10). The levels described in the column headers are cumulative error, or a comparison of the model at that simplification level (Fig. 2) to observations. No level 3 skill estimate was developed due to the computational constraints of simulating sufficient stationary point runs to accurately quantify model skill.

Table 2

RMSE values comparing WL model output to observations at various simplification levels. Rows are locations/variables while the two column groupings represent instrument deployments. The length of the time series comparison varies depending on station deployment length.

Station	Deployment 1: RMSE (cm)			Deployment 2: RMSE (cm)		
	Level 1	Level 2	Level 4	Level 1	Level 2	Level 4
WL 1	24	24	25	21	21	22
WL 2	19	20	22	16	17	19
WL 3	26	26	27	37	37	29
WL 4	19	19	21	22	23	20
WL 5	21	22	22	19	19	20
WL 6	44	44	41	36	36	39
WL 7	21	22	28	23	21	27

Table 2 indicates that the largest drop in skill is at the level 1 simulator simplification. This corresponds to the full ADCSWAN model's inability to perfectly reproduce observations. Level 2 simplifications are found to only nominally impact modeled WLs with a small (1 cm) increase in RMSE for some stations. Level 4 simplifications additionally produce very little loss of skill.

5. Discussion

5.1. Effect of simulator simplifications

The hierarchical validation used in this study provides a unique approach to quantifying the error budget as sourced from multiple simplification levels. A comparison of model performance at various simplification levels (Table 2) found that the primary source of lost skill is from the Full Simulator rather than from simulator simplifications or from emulation. Averaging across stations and deployment periods, all simplifications and emulation only increased RMSE by 1 cm relative to the level 1 error.

This result is of particular interest when compared to the quantification of discrete error from only emulation (see the histogram in Fig. 7) which shows that emulation introduces comparatively significant error into WL estimates. A cumulative level 4 comparison at the tide gauge (Fig. 8) found a RMSE of 15 cm while Level 4 itself (Fig. 7) had a RMSE of 13 cm. This result suggests that the error from each simplification during emulator construction is not independent. In other words, the cumulative error variance is not the sum of the discrete error variances. Practically, the dominance of level 1 error is found to mask that of the other levels. This may not be true if level 1 error was able to be significantly reduced by improvements in process-based modeling at which point other simplifications may become relevant to the error budget.

In terms of quantifying model skill, the RMSE for comparisons to the 1999 USACE field data is overall larger than the RMSE in comparison to the tide gauge. A close examination of the USACE WL time series shows significant high frequency noise that is likely the cause of the overall larger RMSE values. The USACE data exhibits more high frequency variability due to a shorter averaging period (6 min for the tide gauge and 3 min for the USACE data).

This study suggests that the most effective action to improve emulated WL predictions is to reduce level 1 error. One option would be optimized tuning, a process which can be accomplished by including tuning parameters within the emulator framework (Kennedy et al., 2006; Hall et al., 2011;). ADCSWAN could also be replaced with a different simulator or physics implementation, for example ADCSWAN in 3D baroclinic mode. This would come at the cost of drastically increasing computation time and requiring additional input dimensionality in the form of density, temperature, and salinity fields.

A level 1 error reduction could also be accomplished through improving the quality of model input data, both in terms of bathymetry

and forcing. Incorrect bathymetry is likely the source of errors for WL station 6. Fig. 10 shows WLs at this station having an asymmetric tidal signal indicative of shallow water while the observations have less asymmetry. This suggests that the water depth at the time of deployment was greater than the depth in the compiled bathymetric dataset used to generate the ADCSWAN grid. Therefore, investment in more accurate or more recent bathymetry is another viable step for decreasing level 1 error. Similarly, level 1 error integrates error as a function of poor-quality forcing, making improvements to forcing another promising avenue for error reduction.

Level 2 error could be reduced by making less aggressive simplifications of forcing inputs. It is conceptually straight forward to include other input dimensionality such as spatially variable atmospheric forcing or full spectral wave forcing. A promising strategy for including field variables as input dimensions is through decomposing the field into principal components (Higdon et al., 2008; Liu and Guillas, 2017).

There are additionally a range of options for avoiding the stationarity assumption made in this study, which would eliminate or reduce level 3 error. The incorporation of temporal variability in emulators is reviewed by Reichert et al., (2011) who suggest the following strategies:

- 1) Apply a standard emulation methodology but with time as an additional degree of dimensionality (Conti et al., 2005).
- 2) Describe the time series using basis functions and then apply emulation to the basis function coefficients (Bayarri et al., 2007; Higdon et al., 2008).
- 3) Emulate the difference from one time point to the next (Bhattacharya, 2007; Conti et al., 2009).
- 4) Use a Gaussian stochastic process as a Bayesian prior (Liu and West, 2009).
- 5) Develop a hybrid dynamic/emulated model, or a “Mechanistic dynamic emulator.” (Reichert et al., 2011; Albert, 2012)

In the context of this study, strategy 1 is conceptually the simplest but it is not clear a priori how far into the past the system's memory extends and each included time step multiplies the dimensionality of the input space. Strategy 2 is complicated by identifying basis functions that adequately capture the various contributing signals. For example, a Fourier transformation is a natural solution except that storm surge is non-periodic and an important contributor to estuarine WLs. Strategies 3–5 all have potential advantages but bring additional complexity to an already complex methodology so were not explored further.

For reducing level 4 error, GPR is a flexible framework and there are likely gains to be made through a more exhaustive approach for emulator specification. In particular, handling of the periodic nature of tides within the covariance function (Roberts et al., 2013) is a promising research direction.

5.2. Computational cost considerations

Emulation is an approach to dramatically reduce simulation times and is therefore most valuable in situations where the simulator must be run for very long periods or for multiple iterations (e.g., probabilistic risk assessment). Emulation requires an upfront cost, through the running of multiple simulations to construct a training dataset, but is comparatively instantaneous after this initial investment. As the nature of the trade-off is computation time, it is useful to review the costs of building the training dataset.

The first control on computational cost for the training set is the number of input dimensions. The simplifications implemented in this study managed to reduce the input space to 16 dimensions. Each design point took approximately 66 core hours to run in parallel on a server with Intel Xeon E52450 CPUs (2.1 GHz). With this setup, a full experimental design of 160 points would require over 10.5 thousand core hours (although with parallelization the actual time was much less).

This study developed a larger experimental design (over 400 points) but this was primarily for validation rather than emulator skill (see discussion below). Full ADCSWAN required approximately 18 core hours per day of simulation time. Based off these computational costs, emulation becomes an efficient option if approximately one and a half years of simulation are required. This limit is highly situationally dependent and is controlled by processor speed, simulator, emulator, etc. and is intended only as an order of magnitude reference. Furthermore, emulation is primarily targeted at probabilistic methodologies, rather than hindcasting, for which multiple iterations of time series quickly sum to very large total simulation times.

The above analysis is based on a LHS design and the [Loeppky et al., \(2009\)](#) guideline that a training dataset should be around 10 times the number of input dimensions. However, LHS is one of many possible experimental designs ([Levy and Steinberg, 2010](#)). Significant research has focused on optimizing experimental designs beyond LHS and it is possible that a more complex design could reduce the size of the training dataset. For example, LHS does not consider the probability that a particular combination of input parameters may occur. Therefore, some design points are likely poorly utilized exploring space that is physically impossible or highly improbable (for example, high wave heights associated with low wave periods).

Finally, the above analysis did not consider the effect of training dataset size on skill. This relationship was tested by quantifying emulator performance at a variety of training dataset sizes. For this analysis, the total body of simulations (480) was partitioned into smaller dataset sizes ranging from 50 to 450 simulations for testing. For each smaller dataset, a k-fold validation with 5 segments was performed ([Fig. 11](#)) to quantify emulator skill at this smaller training dataset size. This analysis is identical to that described in section 4.2 but with an artificially decreased training dataset size.

Results from this analysis are in good agreement with the guidance of [Loeppky et al., \(2009\)](#) in that ten times the number of input dimensions is sufficient for building a skillful emulator ([Fig. 11](#)). Beyond this limit, only very small gains in skill are realized, suggesting that it is not efficient to over build the training dataset.

It is worth considering the cumulative computational cost of developing multiple emulators. This study takes the approach of building individual emulators at each location of interest. While emulator training and simulation is rapid for individual emulators, the sum computational cost of constructing many emulators can be significant. This is especially true considering that large estuarine hydrodynamic models can be of very high output dimensionality (the utilized Grays Harbor ADCSWAN grid has over 29,000 nodes). A common solution is to dimensionally reduce model outputs via approaches such as principal component analysis ([Chen et al., 2011](#); [Jia and Taflanidis, 2013](#); [Jia et al., 2016](#); [Bass and Bedient, 2018](#)). An alternative option uses the multivariate Gaussian process to generalize the standard GPR case to a “multi-output emulator” ([Conti and O’Hagan, 2010](#); [Fricker et al.,](#)

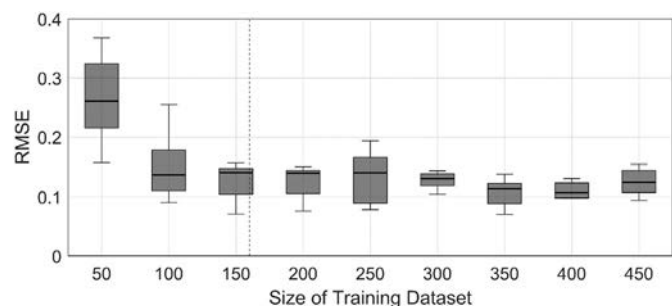


Fig. 11. Impact of training dataset size on emulator skill. RMSE is calculated with a 5-fold cross-validation and represented by a standard boxplot. The dotted vertical line represents the theoretical training dataset size from [Loeppky et al., \(2009\)](#).

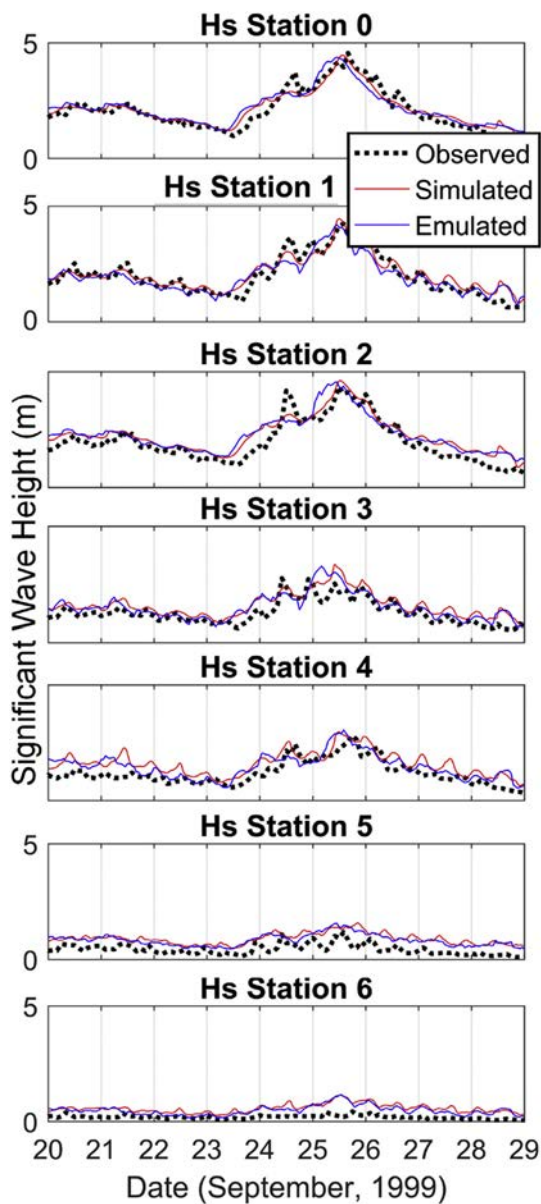


Fig. 12. Comparison of observations (USACE deployments), Full Simulator and emulated Hs time series. Symbols are used for the observations for the sake of visual clarity. All panels have the same y-axis scaling.

2010). While not considered in this study, which is primarily concerned with point assessments, these approaches could result in significant computational savings for a larger output dimensionality. Further, considering emulators individually implicitly assumes independence of output variables and ignores the inherent correlation between output variables ([Rasmussen and Williams, 2006](#)).

5.3. Emulation beyond water levels

While this study has focused primarily on emulating WLs, emulation can easily be extended to other variables in a coastal hazards framework. To explore this possibility, Hs was emulated at the observational Hs stations from the 1999 USACE field campaign ([Fig. 1](#)) ([Cialone and Kraus, 2001, 2002](#)). Hs emulators were developed using an identical approach to that of WLs except that Hs emulation was found to not need cubic terms for the prior mean function. A comparison to observations ([Fig. 12](#)) shows that GPR emulators perform well for Hs with the highest Hs (around September 26, 1999) being well reproduced by the

Table 3

RMSE values comparing Hs model output to observations at various simplification levels. Rows represent station locations while the two column groupings represent instrument deployments. Hs station 0/7 was relocated so deployment 1 values represent the Hs station 0 location and deployment 2 values representing the station 7 location.

Station	Deployment 1: RMSE (cm)			Deployment 2: RMSE (cm)		
	Level 1	Level 2	Level 4	Level 1	Level 2	Level 4
Hs 0/7	32	32	37	73	73	75
Hs 1	38	39	37	61	61	69
Hs 2	45	46	51	48	49	61
Hs 3	45	45	35	52	52	50
Hs 4	49	50	40	148	148	42
Hs 5	47	47	104	136	136	209
Hs 6	35	35	28	40	40	38

emulator at stations 0, 1, 2, 3, and 4 (Fig. 12). Performance is comparatively poor at stations 5 and 6, which are further within the estuary and less influenced by offshore waves. These results are further quantified in Table 3 which shows poor skill for interior Hs stations. It should be noted, however, that the Hs signals at these two stations have low variance and are barely above the noise floor.

Results show that, similarly to WLs, the largest loss of skill is at simplification level 1. Simplified Simulator and emulator results are found to closely track the Full Simulator. This is most evident at the bay interior stations where the Full Simulator and emulator are both found to over-predict Hs. Table 3 reveals that level 2, 3 and 4 simplifications produce little loss of skill for calculated Hs (average increase in RMSE of 2 cm). An exception to this is Hs station 5 which shows a significant increase in RMSE at the emulation level. The cause of poor emulator performance at this one location is unclear but it is likely due to a poor emulator model fit.

Overall these results suggest that emulation could be integrated into many parts of an estuary modeling system. However, a key assumption of emulation with GPR is smoothness in response characteristics, suggesting that GPR may be sub-optimal for “jumpy” variables. Not shown are similar results for Tp which can exhibit discontinuities within estuaries as Tp switches from one wave spectrum component to another. The emulator is qualitatively able to capture Tp characteristics but cannot resolve these instantaneous jumps. For this reason, it is important to carefully consider the form of the output variable being emulated and its relation to the input parameters.

5.4. Example emulator application: extreme water level decomposition

Outside of validation, it is illustrative to explore an example application of emulation. For this purpose, a decomposition of the relative forcing contributions responsible for extreme WLs within Grays Harbor was performed. Seven versions of a 31-year time series (1984–2015) were emulated under different forcing scenarios. As a baseline, a “full forcing” case time series was emulated with the observed forcing at Grays Harbor (comparable to the hindcasts in Figs. 8 and 10). Each additional forcing scenario was emulated with one forcing contribution excluded (waves, wind, pressure, base WL, streamflow, and tides) to isolate the relative contribution of individual forcings to WLs. A particular forcing contribution was calculated as the emulator output WL with full forcing minus the emulator output WL with all forcing except the component of interest. The exception to this is tides (which cannot be turned off due to how they are included in the emulator) which were calculated simply as emulator output with no other forcing but tides. WL contributions were calculated at the time of the 31 annual maximum WL events as determined by the full forcing time series. The average relative contribution of each forcing component over the 31 annual maxima are plotted along East-West and North-South transects in Fig. 13.

The diverse mix of contributions for each bar in Fig. 13 shows that extreme WL events are compound in nature. This conclusion is further supported by the variance in emulated extreme WL contributions (not shown) which reveals that the composition of each individual annual maximum event varies widely across the timeseries. The mean contribution of each forcing is found to be significant providing evidence that all included forcing processes are important for properly quantifying extreme water levels. The only exception is streamflow which is found to be nominally important except near the streamflow boundary. This result is likely specific to the Grays Harbor estuary and would be different for a more hydrologically dominated estuary system (Svensson and Jones, 2004; Lavery and Donovan, 2005; Chen et al., 2014).

The mix of contributions is found to be spatially variable across the estuary domain, leading to both an East-West and North-South gradient in contributions to WLs. For example, the streamflow contribution is found to increase moving west towards the estuary’s streamflow inlet. Wave influence is found to have a significant contribution to the annual maxima but only at stations in the northern and eastern reaches of the bay. This result is likely due to breaking induced setup not occurring at the bay’s entrance channel. The influence of wind increases to the north, due to the mean wind direction emanating from the south during storm events. The influence of pressure anomalies on extreme WLs is found to be uniform but this result is likely from the spatial simplification of sea level pressure fields.

Not shown in Fig. 13 is the contribution from tidal forcing. This is primarily for scale reasons as the tidal component is an order of magnitude larger than that from other forcing (average of 140 cm). Tides also show a gradient across the estuary although with the opposite pattern as that shown in Fig. 13. The tidal component of annual maxima WLs decreases by about 30 cm moving from the center of the estuary moving North or East. As WLs are the sum of these two components (tides and forcing driven NTR), the calculated gradient in total WLs is less than that shown in Fig. 13 (under 20 cm across the two transects).

6. Conclusions

This paper has presented an application of emulation, or surrogate modeling, to the problem of rapidly simulating hydrodynamic variables within the Grays Harbor, WA estuary. This methodology is targeted towards a variety of computationally constrained problems including probabilistic modeling, uncertainty quantification, model optimization, and non-parametric extreme event analysis. To facilitate efficiently achieving these goals, this study has focused on validating and quantifying the error induced by emulation. Additionally, a variety of simplifications to the simulator have been suggested for reducing input dimensionality, and therefore the size of the emulator training dataset.

The results from this study suggest that the Gaussian Process regression (GPR) derived emulator is skillful for calculating a variety of model output variables (WL, NTR, and Hs). A decadal-scale comparison of emulated WLs to tide gauge data showed the emulator having a RMSE of 15 cm. Emulator performance is evaluated at multiple observation points across the estuary domain providing confidence that emulation is skillful across spatial extents. Decomposing the error from different emulator construction simplification levels shows that the largest source of unexplained variance in emulator hindcasts is from ADCSWAN itself. Of particular interest, strong simulator simplifications (including that of stationarity) are a relatively low contributor to losses in emulator performance (average increase in RMSE of 1 cm). Therefore, future efforts to improve emulator performance should focus on improving the Full Simulator before reducing simplifications or optimizing the emulator.

Emulation is additionally found to be very efficient after the construction of the training dataset. Using an LHS experimental design, analysis shows that the training dataset size guidance of 10 times the number of input dimensions (Loeppky et al., 2009) is optimal in the case examined here. Overall emulation is found to have the same order

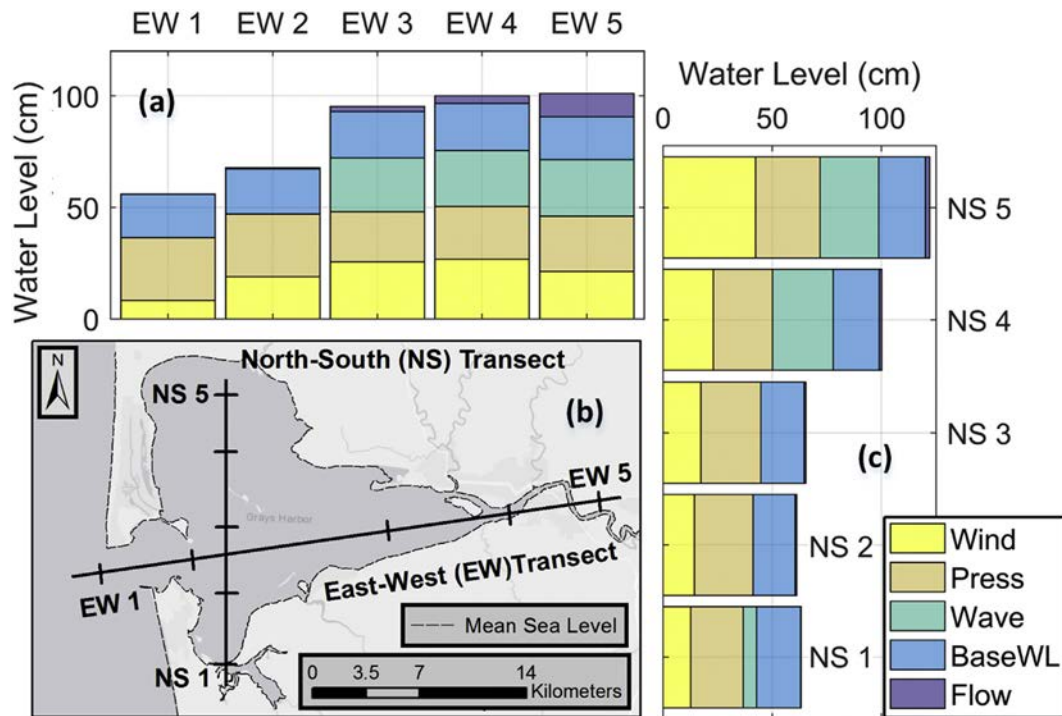


Fig. 13. Average WL contribution from forcing components during extreme events (maximum annual WLs). Two transects are plotted with subplot (b) showing plotted transects, (East-West, EW) and (North-South, NS), with station locations marked as ticks. Tick locations are approximate (within 1 km) to scattered station locations. Subplot (a) is the East-West transect and subplot (c) is the North-South transect.

of magnitude skill as process-based models as well as showing significant gains in computational efficiency. Therefore, emulation is shown to be a viable path for exploring estuarine hydrodynamic modeling problems.

Finally, the emulator was applied to investigate the relative contributions of different forcing variables to annual maxima WLs and NTR at the study site. Results show a diverse mix of forcing contributing to annual extreme WLs, indicating the importance of considering compound events for flood hazard assessments in the PNW. All forcing components, along with WL itself, were found to exhibit significant spatial variability hinting at important information for flood vulnerability assessment. Tides were found to be the largest contributor to extreme WLs with other components being of the same order of magnitude. The exception to this is streamflow which was found to be, on average, a relatively minor contributor to extremes except near the river's mouth. Additionally, waves were found to only contribute to WLs at stations near the edge of the estuary domain, a result that is likely tied to wave penetration into the estuary. While only a single example application of emulation to estuary hydrodynamics questions was explored, results signal the significant potential of emulation to a broad range of applications.

Acknowledgments

Tide gauge records are available through the National Oceanic and Atmospheric Administration (NOAA) National Ocean Service (NOS) website. River discharge is available from the USGS through the National Water Information System. The NARR climate dataset is available through NOAA's Earth System Research Laboratory website. Bathymetric and topographic data were obtained from NOAA's Bathymetric Data viewer (DEMs) and DOGAMI's LIDAR download portal. We thank Melisa Menendez and Jorge Perez at the Environmental Hydraulics Institute of the Universidad de Cantabria (IH Cantabria) for providing the Global Ocean Wave 2 (GOW2) data. We also thank Mary Cialone and Dave Michalson of the U.S. Army Corps of

Engineers for providing the Grays Harbor field observations from 1999. This work was funded by the NOAA Regional Integrated Sciences and Assessments Program (RISA) [Grant Number NA15OAR4310145] and a contracted grant with the Quinault Treaty Area (QTA) tribal governments (Quinault Indian Nation, Hoh Indian Tribe, and Quileute Tribe).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.coastaleng.2019.03.004>.

References

- Albert, C., 2012. A mechanistic dynamic emulator. *Nonlinear Anal. R. World Appl.* 13, 2747–2754. <https://doi.org/10.1016/j.nonrwa.2012.04.003>.
- Allan, J.C., Komar, P.D., 2006. Climate controls on US west coast erosion processes. *J. Coast. Res.* 223, 511–529. <https://doi.org/10.2112/03-0108.1>.
- Allan, J.C., Komar, P.D., 2002a. Wave climate change and coastal erosion in the US Pacific Northwest. In: *Ocean Wave Measurement and Analysis*. American Society of Civil Engineers, Reston, VA, pp. 680–689. [https://doi.org/10.1061/40604\(273\)70](https://doi.org/10.1061/40604(273)70).
- Allan, J.C., Komar, P.D., 2002b. Extreme storms on the Pacific Northwest coast during the 1997–98 el niño and 1998–99 La niña. *J. Coast. Res.* <https://doi.org/10.2307/4299063>.
- Allan, J.C., Komar, P.D., Ruggiero, P., 2011. Storm surge magnitudes and frequency on the central Oregon coast. In: *Solutions to Coastal Disasters 2011*. American Society of Civil Engineers, Reston, VA, pp. 53–64. [https://doi.org/10.1061/41185\(417\)6](https://doi.org/10.1061/41185(417)6).
- Anjyo, K., Lewis, J.P., 2011. RBF interpolation and Gaussian process regression through an RKHS formulation. *J. Math-for-Industry* 3, 63–71.
- Apel, H., Merz, B., Thielen, A.H., 2008. Quantification of uncertainties in flood risk assessments. *Int. J. River Basin Manag.* 6, 149–162. <https://doi.org/10.1080/15715124.2008.9635344>.
- Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. *Stat. Surv.* 4, 40–79. <https://doi.org/10.1214/09-SS054>.
- Bass, B., Bedient, P., 2018. Surrogate modeling of joint flood risk across coastal watersheds. *J. Hydrol.* 558, 159–173. <https://doi.org/10.1016/j.jhydrol.2018.01.014>.
- Bayarri, M.J., Berger, J.O., Cafeo, J., Garcia-Donato, G., Liu, F., Palomo, J., Parthasarathy, R.J., Paulo, R., Sacks, J., Walsh, D., 2007. Computer model validation with functional output. *Ann. Stat.* 35, 1874–1906. <https://doi.org/10.1214/009053607000000163>.
- Bhaskaran, P.K., Nayak, S., Bonthu, S.R., Murty, P.L.N., Sen, D., 2013. Performance and validation of a coupled parallel ADCIRC–SWAN model for THANE cyclone in the Bay of Bengal. *Environ. Fluid Mech.* 13, 601–623. <https://doi.org/10.1007/s10652-013-9284-5>.

- Bhattacharya, S., 2007. A simulation approach to Bayesian emulation of complex dynamic computer models. *Bayesian Anal* 2, 783–815. <https://doi.org/10.1214/07-BA232>.
- Blain, C.A., Preller, R.H., Rivera, A.P., 2001. Tidal prediction using the advanced circulation model (ADCIRC) and a relocatable PC-based system. *Oceanography* 15.
- Blain, C.A., Rogers, W.E., 1998. Coastal Tide Prediction Using the ADCIRC-2DDI Hydrodynamic Finite Element Model: Model Validation and Sensitivity Analyses in the Southern North Sea/English Channel.
- Bode, L., Hardy, T.A., 1997. Progress and recent developments in storm surge modeling. *J. Hydraul. Eng.* 123, 315–331. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1997\)123:4\(315\)](https://doi.org/10.1061/(ASCE)0733-9429(1997)123:4(315)).
- Booij, N., Holthuijsen, L.H., Ris, R.C., 1997. The “SWAN” wave model for shallow water. In: *Coastal Engineering 1996*. American Society of Civil Engineers, New York, NY, pp. 668–676. <https://doi.org/10.1061/9780784402429.053>.
- Bromirski, P.D., Flick, R.E., Cayan, D.R., 2003. Storminess variability along the California coast: 1858–2000. *J. Clim.* 16, 982–993. [https://doi.org/10.1175/1520-0442\(2003\)016<0982:SVATCC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0982:SVATCC>2.0.CO;2).
- Bunya, S., Dietrich, J.C., Westerink, J.J., Ebersole, B.A., Smith, J.M., Atkinson, J.H., Jensen, R., Resio, D.T., Luettich, R.A., Dawson, C., Cardone, V.J., Cox, A.T., Powell, M.D., Westerink, H.J., Roberts, H.J., 2010. A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part I: model development and validation. *Mon. Weather Rev.* 138, 345–377. <https://doi.org/10.1175/2009MWR2906.1>.
- Carnell, R., 2017. LHS: Latin Hypercube Samples.
- Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R., 2012. Data-driven dynamic emulation modelling for the optimal management of environmental systems. *Environ. Model. Softw* 34, 30–43. <https://doi.org/10.1016/j.envsoft.2011.09.003>.
- Chen, T., Hadinoto, K., Yan, W., Ma, Y., 2011. Efficient meta-modelling of complex process simulations with time-space-dependent outputs. *Comput. Chem. Eng.* 35, 502–509. <https://doi.org/10.1016/J.COMPCHENG.2010.05.013>.
- Chen, W.-B., Liu, W.-C., Chen, W.-B., Liu, W.-C., 2014. Modeling flood inundation induced by river flow and storm surges over a river basin. *Water* 6, 3182–3199. <https://doi.org/10.3390/w6103182>.
- Cheng, T.K., Hill, D.F., Beamer, J., García-Medina, G., 2015a. Climate change impacts on wave and surge processes in a Pacific Northwest (USA) estuary. *J. Geophys. Res. Ocean.* 120, 182–200. <https://doi.org/10.1002/2014JC010268>.
- Cheng, T.K., Hill, D.F., Read, W., 2015b. The contributions to storm tides in Pacific Northwest estuaries: Tillamook bay, Oregon, and the december 2007 storm. *J. Coast. Res.* 313, 723–734. <https://doi.org/10.2112/JCOASTRES-D-14-00120.1>.
- Cialone, M.A., Kraus, N.C., 2001. Engineering study of inlet entrance hydrodynamics: Grays Harbor, Washington, USA. In: *Coastal Dynamics '01*. American Society of Civil Engineers, Reston, VA, pp. 413–422. [https://doi.org/10.1061/40566\(260\)42](https://doi.org/10.1061/40566(260)42).
- Cialone, M.A., Militello, A., Brown, M.E., Kraus, N.C., 2002. Coupling of wave and circulation numerical models at Grays Harbor entrance, Washington, USA. In: *Proceedings 28th Coastal Engineering Conference*. World Scientific Publishing Company, pp. 1279–1291. https://doi.org/10.1142/9789812791306_0108.
- Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: a review. *J. Hydrol* 375, 613–626. <https://doi.org/10.1016/J.JHYDROL.2009.06.005>.
- Conti, S., Gosling, J.P., Oakley, J.E., O'Hagan, A., 2009. Gaussian process emulation of dynamic computer codes. *Biometrika* 96, 663–676. <https://doi.org/10.1093/biomet/asp028>.
- Conti, S., Anderson, C.W., Kennedy, M.C., O'Hagan, A., 2005. A Bayesian analysis of complex dynamic computer models. *Sensit. Anal. Model Output*.
- Conti, S., O'Hagan, A., 2010. Bayesian emulation of complex multi-output and dynamic computer models. *J. Stat. Plan. Inference* 140, 640–651. <https://doi.org/10.1016/j.jspi.2009.08.006>.
- Dale, M., Wicks, J., Mylne, K., Pappenberger, F., Laeger, S., Taylor, S., 2014. Probabilistic flood forecasting and decision-making: an innovative risk-based approach. *Nat. Hazards* 70, 159–172. <https://doi.org/10.1007/s11069-012-0483-z>.
- Davis, J.R., Paramygin, V. a., Forrest, D., Sheng, Y.P., 2010. Toward the probabilistic simulation of storm surge and inundation in a limited-resource environment. *Mon. Weather Rev.* 138, 2953–2974. <https://doi.org/10.1175/2010MWR3136.1>.
- Dawson, R.J., Hall, J.W., Bates, P.D., Nicholls, R.J., 2005. Quantified analysis of the probability of flooding in the Thames estuary under imaginable worst-case sea level rise scenarios quantified analysis of the probability of flooding in the Thames estuary under imaginable worst-case sea level rise scenarios. *Int. J. Water Resour. Dev.* 21, 577–591. <https://doi.org/10.1080/07900620500258380>.
- Di Baldassarre, G., Schumann, G., Bates, P.D., Freer, J.E., Beven, K.J., 2010. Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. *Hydrol. Sci. J.* 55, 364–376. <https://doi.org/10.1080/02626661003683389>.
- Dietrich, J.C., Bunya, S., Westerink, J.J., Ebersole, B.A., Smith, J.M., Atkinson, J.H., Jensen, R., Resio, D.T., Luettich, R.A., Dawson, C., Cardone, V.J., Cox, A.T., Powell, M.D., Westerink, H.J., Roberts, H.J., Dietrich, J.C., Bunya, S., Westerink, J.J., Ebersole, B.A., Smith, J.M., Atkinson, J.H., Jensen, R., Resio, D.T., Luettich, R.A., Dawson, C., Cardone, V.J., Cox, A.T., Powell, M.D., Westerink, H.J., Roberts, H.J., 2010. A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part II: synoptic description and analysis of hurricanes Katrina and Rita. *Mon. Weather Rev.* 138, 378–404. <https://doi.org/10.1175/2009MWR2907.1>.
- Dietrich, J.C., Tanaka, S., Westerink, J.J., Dawson, C.N., Luettich, R.A., Zijlema, M., Holthuijsen, L.H., Smith, J.M., Westerink, L.G., Westerink, H.J., 2012. Performance of the unstructured-mesh, SWAN+ADCIRC model in computing hurricane waves and surge. *J. Sci. Comput.* 52, 468–497. <https://doi.org/10.1007/s10915-011-9555-6>.
- Dietrich, J.C., Zijlema, M., Westerink, J.J., Holthuijsen, L.H., Dawson, C., Luettich, R.A., Jensen, R.E., Smith, J.M., Stelling, G.S., Stone, G.W., 2011. Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. *Coast. Eng.* 58, 45–65. <https://doi.org/10.1016/J.COASTALENG.2010.08.001>.
- DOGAMI, 2010. Lidar Remote Sensing Data Collection: Southwest Washington.
- Dushaw, B.D., Egbert, G.D., Worcester, P.F., Cornuelle, B.D., Howe, B.M., Metzger, K., 1997. A TOPEX/POSEIDON global tidal model (TPXO.2) and barotropic tidal currents determined from long-range acoustic transmissions. *Prog. Oceanogr.* 40, 337–367. [https://doi.org/10.1016/S0079-6611\(98\)00008-1](https://doi.org/10.1016/S0079-6611(98)00008-1).
- Engle, V.D., Kurtz, J.C., Smith, L.M., Chancy, C., Bourgeois, P., 2007. A classification of U.S. Estuaries based on physical and hydrologic attributes. *Environ. Monit. Assess.* 129, 397–412. <https://doi.org/10.1007/s10661-006-9372-9>.
- Funakoshi, Y., Hagen, S.C., Bacopoulos, P., 2008. Coupling of hydrodynamic and wave models: case study for hurricane Floyd (1999) hindcast. *J. Waterw. Port, Coast. Ocean Eng.* 134, 321–335. [https://doi.org/10.1061/\(ASCE\)0733-950X\(2008\)134:6\(321\)](https://doi.org/10.1061/(ASCE)0733-950X(2008)134:6(321)).
- Ganju, N.K., Brush, M.J., Rashleigh, B., Aretxabaleta, A.L., Del Barrio, P., Grear, J.S., Harris, L.A., Lake, S.J., Mccardell, G., O'donnell, J., Ralston, D.K., Signell, R.P., Testa, J.M., Vaudrey, J.M.P., 2015. Progress and challenges in coupled hydrodynamic-ecological estuarine modeling. *Estuar. Coasts.* <https://doi.org/10.1007/s12237-015-0011-y>.
- Gano, S., Kim, H., Brown, D., 2006. Comparison of three surrogate modeling techniques: datascap, kriging, and second order regression. In: *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. American Institute of Aeronautics and Astronautics, Reston, Virginia. <https://doi.org/10.2514/6.2006-7048>.
- Gouldby, B., Méndez, F.J., Guaniche, Y., Rueda, A., Mínguez, R., 2014. A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis. *Coast. Eng.* 88, 15–26. <https://doi.org/10.1016/J.COASTALENG.2014.01.012>.
- Green, C., Viavattene, C., Thompson, P., 2011. Guidance for Assessing Flood Losses CONHAZ Report.
- Haigh, I.D., Wijeratne, E.M.S., MacPherson, L.R., Pattiaratchi, C.B., Mason, M.S., Crompton, R.P., George, S., 2014. Estimating present day extreme water level exceedance probabilities around the coastline of Australia: tides, extra-tropical storm surges and mean sea level. *Clim. Dyn.* 42, 121–138. <https://doi.org/10.1007/s00382-012-1652-1>.
- Hall, J.W., Manning, L.J., Hankin, R.K.S., 2011. Bayesian calibration of a flood inundation model using spatial data. *Water Resour. Res.* 47, W05529. <https://doi.org/10.1029/2009WR008541>.
- Higdon, D., Gattiker, J., Williams, B., Rightley, M., 2008. Computer model calibration using high-dimensional output. *J. Am. Stat. Assoc.* 103, 570–583. <https://doi.org/10.1198/016214507000000888>.
- Jia, G., Taflanidis, A.A., 2013. Kriging metamodelling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. *Comput. Methods Appl. Mech. Eng.* 261–262, 24–38. <https://doi.org/10.1016/J.CMA.2013.03.012>.
- Jia, G., Taflanidis, A.A., Nadal-Caraballo, N.C., Melby, J.A., Kennedy, A.B., Smith, J.M., 2016. 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, 909–938. <https://doi.org/10.1007/s11069-015-2111-1>.
- Jin, R., Chen, W., Simpson, T.W., 2001. Comparative studies of metamodelling techniques under multiple modelling criteria. *Struct. Multidiscip. Optim.* 23, 1–13. <https://doi.org/10.1007/s00158-001-0160-4>.
- Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin distance designs. *J. Stat. Plan. Inference* 26, 131–148. [https://doi.org/10.1016/0378-3758\(90\)90122-B](https://doi.org/10.1016/0378-3758(90)90122-B).
- Jones, B., Johnson, R.T., 2009. Design and analysis for the Gaussian process model. *Qual. Reliab. Eng. Int.* 25, 515–524. <https://doi.org/10.1002/qre.1044>.
- Kantha, L.H., Clayson, C.A., 2000. *Numerical Models of Oceans and Oceanic Processes*, vol. 66 Elsevier.
- Kennedy, M.C., Anderson, C.W., Conti, S., O'Hagan, A., 2006. Case studies in Gaussian process modelling of computer codes. *Reliab. Eng. Syst. Saf.* 91, 1301–1309. <https://doi.org/10.1016/J.RESS.2005.11.028>.
- Kim, S.W., Melby, J. a., Nadal-Caraballo, N.C., Ratcliff, J., 2015. A time-dependent surrogate model for storm surge prediction based on an artificial neural network using high-fidelity synthetic hurricane modeling. *Nat. Hazards* 76, 565–585. <https://doi.org/10.1007/s11069-014-1508-6>.
- Kohavi, R., 1995. A study of cross-validation and Bootstrap for accuracy estimation and model selection. In: *International Joint Conference on Artificial Intelligence*.
- Krien, Y., Dudon, B., Roger, J., Zahibo, N., 2015. Probabilistic hurricane-induced storm surge hazard assessment in Guadeloupe, Lesser Antilles. *Nat. Hazards Earth Syst. Sci.* 15, 1711–1720. <https://doi.org/10.5194/nhess-15-1711-2015>.
- Lakshmi, D.D., Murty, P.L.N., Bhaskaran, P.K., Sahoo, B., Kumar, T.S., Shenoi, S.S.C., Srikanth, A.S., 2017. Performance of WRF-ARW winds on computed storm surge using hydrodynamic model for Phailin and Hudhud cyclones. *Ocean Eng.* 131, 135–148. <https://doi.org/10.1016/J.OCEANENG.2017.01.005>.
- Lavery, S., Donovan, B., 2005. Flood risk management in the Thames Estuary looking ahead 100 years. *Philos. Trans. A. Math. Phys. Eng. Sci.* 363, 1455–1474. <https://doi.org/10.1098/rsta.2005.1579>.
- Le Provost, C., Genco, M.L., Lyard, F., Vincent, P., Canceil, P., 1994. Spectroscopy of the world ocean tides from a finite element hydrodynamic model. *J. Geophys. Res.* 99, 24777. <https://doi.org/10.1029/94JC01381>.
- Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., McInnes, K., Risbey, J., Schuster, S., Jakob, D., Stafford-Smith, M., 2014. A compound event framework for understanding extreme impacts. *Wiley Interdiscip. Rev. Clim. Chang.* 5, 113–128. <https://doi.org/10.1002/wcc.252>.
- Levy, S., Steinberg, D.M., 2010. Computer experiments: a review. *Adv. Stat. Anal.* 94, 311–324. <https://doi.org/10.1007/s10182-010-0147-9>.
- Lewis, M., Bates, P., Horsburgh, K., Neal, J., Schumann, G., 2013. A storm surge inundation model of the northern Bay of Bengal using publicly available data. *Q. J. R.*

- Meteorol. Soc. 139, 358–369. <https://doi.org/10.1002/qj.2040>.
- Lin, N., Emanuel, K.A., Smith, J.A., Vanmarcke, E., 2010. Risk assessment of hurricane storm surge for New York City. *J. Geophys. Res.* 115, D18121. <https://doi.org/10.1029/2009JD013630>.
- Lin, N., Emanuel, K.A., Oppenheimer, M., Vanmarcke, E., 2012. Physically-based assessment of hurricane surge Threat under climate change. *Nat. Clim. Change* 462–467 2.6.
- Liu, F., West, M., 2009. A dynamic modelling strategy for Bayesian computer model emulation. *Bayesian Anal* 4, 393–411. <https://doi.org/10.1214/09-BA415>.
- Liu, X., Guillas, S., 2017. Dimension reduction for Gaussian process emulation: an application to the influence of bathymetry on tsunami heights. *SIAM/ASA J. Uncertain. Quantification* 5, 787–812. <https://doi.org/10.1137/16M1090648>.
- Loepky, J.L., Sacks, J., Welch, W.J., 2009. Choosing the sample size of a computer experiment: a practical guide. *Technometrics* 51, 366–376. <https://doi.org/10.1198/TECH.2009.08040>.
- Love, M.R., Friday, D.Z., Grothe, P.R., Carignan, K.S., Eakins, B.W., Taylor, L.A., 2012. Digital Elevation Model of Astoria, Oregon: Procedures, Data Sources and Analysis.
- Luettich, R.A., Jr., Westerink, J.J., Scheffner, N.W., 1992. ADCIRC: an Advanced Three-Dimensional Circulation Model for Shelves, Coasts, and Estuaries. Report 1. Theory and Methodology of ADCIRC-2DDI and ADCIRC-3DL.
- Madsen, H., Jakobsen, F., 2004. Cyclone induced storm surge and flood forecasting in the northern Bay of Bengal. *Coast. Eng.* 51, 277–296. <https://doi.org/10.1016/J.COASTALENG.2004.03.001>.
- Malde, S., Oakley, J., Wyncott, D., 2016a. MUCM: Gaussian Process Emulator.
- Malde, S., Wyncoll, D., Oakley, J., Tozer, N., Gouldby, B., 2016b. Applying emulators for improved flood risk analysis. In: *E3S Web Conf*, vol. 7. pp. 04002. <https://doi.org/10.1051/e3sconf/20160704002>.
- Mass, C., Dotson, B., 2010. Major extratropical cyclones of the northwest United States: Historical review, climatology, and synoptic environment. *Mon. Weather Rev.* 138, 2499–2527. <https://doi.org/10.1175/2010MWR3213.1>.
- Mastrandrea, M.D., Field, C.B., Stocker, T.F., Edenhofer, O., Ebi, K.L., Frame, D.J., Held, H., Kriegler, E., Mach, K.J., Matschoss, P.R., Plattner, G.-K., Yohe, G.W., Zwiars, F.W., 2010. Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties IPCC Cross-Working Group Meeting on Consistent Treatment of Uncertainties.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239–245. <https://doi.org/10.1080/00401706.1979.10489755>.
- McKay, P., Blain, C.A., 2010. Toward Developing a Hydrodynamic Flow and Inundation Model of the Lower Pearl River.
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American regional reanalysis. *Bull. Am. Meteorol. Soc.* 87, 343–360. <https://doi.org/10.1175/BAMS-87-3-343>.
- Moel, H., Asselman, N., Aerts, J., 2012. Uncertainty and sensitivity analysis of coastal flood damage estimates in the west of The Netherlands. *Nat. Hazards Earth Syst. Sci.* 12, 1045–1058. <https://doi.org/10.5194/nhess-12-1045-2012>.
- Moftakhari, H.R., AghaKouchak, A., Sanders, B.F., Matthew, R.A., 2017. Cumulative hazard: the case of nuisance flooding. *Earth's Futur* 5, 214–223. <https://doi.org/10.1002/2016EF000494>.
- Montoya, R.D., Osorio Arias, A., Ortiz Royero, J.C., Ocampo-Torres, F.J., 2013. A wave parameters and directional spectrum analysis for extreme winds. *Ocean Eng.* 67, 100–118. <https://doi.org/10.1016/J.OCEANENG.2013.04.016>.
- Morris, M.D., Mitchell, T.J., 1995. Exploratory designs for computational experiments. *J. Stat. Plan. Inference* 43, 381–402. [https://doi.org/10.1016/0378-3758\(94\)00035-T](https://doi.org/10.1016/0378-3758(94)00035-T).
- NOAA National Centers for Environmental Information, 2003. Coastal Relief Model.
- OHagan, A., 2006. Bayesian analysis of computer code outputs: a tutorial. *Reliab. Eng. Syst. Saf.* 91, 1290–1300. <https://doi.org/10.1016/j.res.2005.11.025>.
- Oakley, J., 1999. Bayesian Uncertainty Analysis for Complex Computer Codes. Thesis.
- Orton, P.M., Hall, T.M., Talke, S.A., Blumberg, A.F., Georgas, N., Vinogradov, S., 2016. A validated tropical-extratropical flood hazard assessment for New York Harbor. *J. Geophys. Res. Ocean.* 121, 8904–8929. <https://doi.org/10.1002/2016JC011679>.
- Pendleton, L.H., 2010. The Economic and Market Value of Coasts and Estuaries: What's at Stake? Restore America's Estuaries, Arlington.
- Perez, J., Menendez, M., Losada, I.J., 2017. GOW2: a global wave hindcast for coastal applications. *Coast. Eng.* 124, 1–11. <https://doi.org/10.1016/J.COASTALENG.2017.03.005>.
- Pugh, D.T., 1996. Tides, Surges and Mean Sea-Level. John Wiley & Sons.
- Purvis, M.J., Bates, P.D., Hayes, C.M., 2008. A probabilistic methodology to estimate future coastal flood risk due to sea level rise. *Coast. Eng.* 55, 1062–1073. <https://doi.org/10.1016/j.coastaleng.2008.04.008>.
- Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. MIT press, Cambridge.
- Razavi, S., Tolson, B. a., Burn, D.H., 2012. Review of surrogate modeling in water resources. *Water Resour. Res.* 48. <https://doi.org/10.1029/2011WR011527>.
- Reichert, P., White, G., Bayarri, M.J., Pitman, E.B., 2011. Mechanism-based emulation of dynamic simulation models: concept and application in hydrology. *Comput. Stat. Data Anal.* 55, 1638–1655. <https://doi.org/10.1016/j.csda.2010.10.011>.
- Resio, D.T., Irish, J., Cialone, M., 2009. A surge response function approach to coastal hazard assessment – part 1: basic concepts. *Nat. Hazards* 51, 163–182. <https://doi.org/10.1007/s11069-009-9379-y>.
- Resio, D.T., Westerink, J.J., 2008. Modeling the physics of storm surges. *Phys. Today* 61. <https://doi.org/10.1063/1.2982120>.
- Roberts, S., Osborne, M., Ebdon, M., Reece, S., Gibson, N., Aigrain, S., 2013. Gaussian processes for time-series modelling. *Philos. Trans. A. Math. Phys. Eng. Sci.* 371, 20110550. <https://doi.org/10.1098/rsta.2011.0550>.
- Rogers, W.E., Kaihatu, J.M., Hsu, L., Jensen, R.E., Dykes, J.D., Holland, K.T., 2007. Forecasting and hindcasting waves with the SWAN model in the Southern California Bight. *Coast. Eng.* 54, 1–15. <https://doi.org/10.1016/J.COASTALENG.2006.06.011>.
- Rohmer, J., Idier, D., 2012. A meta-modelling strategy to identify the critical offshore conditions for coastal flooding. *Nat. Hazards Earth Syst. Sci.* 12, 2943–2955. <https://doi.org/10.5194/nhess-12-2943-2012>.
- Rueda, A., Gouldby, B., Mendez, F.J., Tomas, A., Losada, I.J., Lara, J.L., Diaz-Simal, P., 2016. The use of wave propagation and reduced complexity inundation models and metamodels for coastal flood risk assessment. *J. Flood Risk Manag.* 9, 390–401. <https://doi.org/10.1111/jfr3.12204>.
- Rusu, L., Pilar, P., 2008. Hindcast of the wave conditions along the west Iberian coast. *Coast. Eng.* 55, 906–919. <https://doi.org/10.1016/J.COASTALENG.2008.02.029>.
- Sacks, J., Schiller, S.B., Welch, W.J., 1989. Designs for computer experiments. *Technometrics* 31, 41–47. <https://doi.org/10.1080/00401706.1989.10488474>.
- Schulz, E., Speekenbrink, M., Krause, A., 2018. A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions. *J. Math. Psychol.* 85, 1–16. <https://doi.org/10.1016/J.JMP.2018.03.001>.
- Serafin, K.A., Ruggiero, P., 2014. Simulating extreme total water levels using a time-dependent, extreme value approach. *J. Geophys. Res. Ocean.* 119, 6305–6329. <https://doi.org/10.1002/2014JC010093>.
- Serafin, K.A., Ruggiero, P., Stockdon, H.F., 2017. The relative contribution of waves, tides, and non-tidal residuals to extreme total water levels on US West Coast sandy beaches. *Geophys. Res. Lett.* 44, 1839–1847. <https://doi.org/10.1002/2016GL071020>.
- Song, Y.K., Irish, J.L., Udoh, I.E., 2012. Regional attributes of hurricane surge response functions for hazard assessment. *Nat. Hazards* 64, 1475–1490. <https://doi.org/10.1007/s11069-012-0309-z>.
- Svensson, C., Jones, D.A., 2004. Dependence between sea surge, river flow and precipitation in south and west Britain. *Hydrol. Earth Syst. Sci.* 8, 973–992. <https://doi.org/10.5194/hess-8-973-2004>.
- Taylor, N.R., Irish, J.L., Udoh, I.E., Bilskie, M.V., Hagen, S.C., 2015. Development and uncertainty quantification of hurricane surge response functions for hazard assessment in coastal bays. *Nat. Hazards* 77, 1103–1123. <https://doi.org/10.1007/s11069-015-1646-5>.
- Timmermans, B., 2015. Uncertainty in Numerical Wind-Wave Models. University of Southampton.
- Tolman, H.L., 2009. User Manual and System Documentation of WAVEWATCH III TM Version 3.14 †.
- Wahl, T., Jain, S., Bender, J., Meyers, S.D., Luther, M.E., 2015. Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nat. Clim. Change* 5, 1093–1097. <https://doi.org/10.1038/nclimate2736>.
- Weaver, R.J., Luettich Jr., R.A., 2010. 2D vs. 3D storm surge sensitivity in ADCIRC: case study of hurricane isabel. In: *Estuarine and Coastal Modeling (2009)*. American Society of Civil Engineers, Reston, VA, pp. 762–779. [https://doi.org/10.1061/41121\(388\)44](https://doi.org/10.1061/41121(388)44).
- Weaver, R.J., Slinn, D.N., 2010. Influence of bathymetric fluctuations on coastal storm surge. *Coast. Eng.* 57, 62–70. <https://doi.org/10.1016/J.COASTALENG.2009.09.012>.
- Westerink, J.J., Luettich, R.A., Baptists, A.M., Scheffner, N.W., Farrar, P., 1992. Tide and storm surge predictions using finite element model. *J. Hydraul. Eng.* 118, 1373–1390. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1992\)118:10\(1373\)](https://doi.org/10.1061/(ASCE)0733-9429(1992)118:10(1373)).
- Zhang, K., Douglas, B.C., Leatherman, S.P., 1999. Twentieth-century storm activity along the U.S. East coast. *J. Clim.* 13.
- Zijlema, M., 2010. Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids. *Coast. Eng.* 57, 267–277. <https://doi.org/10.1016/J.COASTALENG.2009.10.011>.
- Zscheischler, J., Westra, S., van den Hurk, B.J.J.M., Seneviratne, S.I., Ward, P.J., Pitman, A., AghaKouchak, A., Bresch, D.N., Leonard, M., Wahl, T., Zhang, X., 2018. Future climate risk from compound events. *Nat. Clim. Change* 8, 469–477. <https://doi.org/10.1038/s41558-018-0156-3>.

Appendix B

The following is a reproduction of the publication:

Serafin, K., P. Ruggiero, K. Parker, and D. Hill. 2019. “What's Streamflow Got To Do With It? A Probabilistic Simulation of the Competing Oceanographic and Fluvial Processes Driving Extreme Along-River Water Levels.” *Nat. Hazards Earth Syst. Sci. Discuss.*, in review.
<https://doi.org/10.5194/nhess-2018-347>.

What's streamflow got to do with it? A probabilistic simulation of the competing oceanographic and fluvial processes driving extreme along-river water levels

Katherine A. Serafin^{1,2}, Peter Ruggiero¹, Kai A. Parker³, and David F. Hill³

¹College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA

²Department of Geophysics, Stanford University, Stanford, CA, USA

³College of Engineering, Oregon State University, Corvallis, OR, USA

Correspondence: Katherine A. Serafin (kserafin@stanford.edu)

Abstract. Extreme water levels generating flooding in estuarine and coastal environments are often driven by compound events, where many individual processes such as waves, storm surge, streamflow, and tides coincide. Despite this, extreme water levels are typically modeled in isolated open coast or estuarine environments, potentially mischaracterizing the true risk of flooding facing coastal communities. This manuscript explores the variability of extreme water levels near the tribal community of La Push, within the Quileute Indian Reservation on the Washington state coast where a river signal is apparent in tide gauge measurements during high discharge events. To estimate the influence of multiple forcings on high water levels a hybrid modeling framework is developed, where probabilistic simulations of joint still water level and river discharge occurrences are merged with a hydraulic model that simulates along-river water levels. This methodology produces along-river water levels from thousands of combinations of events not necessarily captured in the observational records. We show that the 100-yr still water level event and the 100-yr discharge event do not always produce the 100-yr along-river water level. Furthermore, along specific sections of river, both still water level and discharge are necessary for producing the 100-yr along-river water level. Understanding the relative forcing driving extreme water levels along an ocean-to-river gradient will help communities within inlets better understand their risk to the compounding impacts of various environmental forcing, important for increasing their resilience to future flooding events.

15 1 Introduction

Coincident or compound events are a combination of physical processes in which the individual variables may or may not be extreme, however the result is an extreme event with a significant impact (Zscheischler et al., 2018; Bevacqua et al., 2017; Wahl et al., 2015; Leonard et al., 2014). Flooding is often caused by compound events, where multiple factors impact both open coast and estuarine environments. Storm events, for example, often generate concurrently large waves, heavy precipitation driving increased streamflow, and high storm surges, making the relative contribution of the actual drivers of extreme water levels difficult to interpret. Studies at the global (e.g., Ward et al. (2018)), national (e.g., Wahl et al. (2015); Svensson and Jones (2002); Zheng et al. (2013)) and regional scale (e.g., Odigie and Warrick (2017); Moftakhari et al. (2017)) have evaluated the

likelihood for variables such as high river flow and precipitation to occur during high coastal water levels, demonstrating that dependencies often exist between these individual processes.

Around river mouths, the elevation of the water level measured by tide gauges, or the still water level (SWL), varies depending on the mean sea level, tidal stage, and the non-tidal residual contributors which may include the following forcings; storm surge, seasonally-induced thermal expansion (Tsimplis and Woodworth, 1994), the geostrophic effects of currents (Chelton and Enfield, 1986), wave setup (Sweet et al., 2015; Vetter et al., 2010), and river discharge. Most commonly, estimates of non-tidal residuals are determined by subtracting predicted tides from the measured water levels. However, residuals computed in this way often contain artifacts of the subtraction process from phase shifts in the tidal signal and/or timing errors (Horsburgh and Wilson, 2007). Another approach for extracting the non-tidal residual is through the skew surge, which is the absolute difference between the maximum observed water level and the predicted tidal high water (de Vries et al., 1995; Williams et al., 2016; Mawdsley and Haigh, 2016). While this methodology removes the influence of tide-surge interaction from the non-tidal residual magnitude, it does not differentiate between the many factors contributing to the water level, an important step for distinguishing when and why high water, and thus flooding, is likely to occur.

Hydrodynamic and hydraulic models have recently been used in attempts to quantify the relative importance of river and ocean-forced water levels to flooding. The nonlinear coupling of wind and pressure driven storm surge, tides, wave-driven setup, and riverine flows has been found to be a vital contributor to overall water level elevation (Bunya et al., 2010). Furthermore, river discharge is often found to interact nonlinearly with storm surge (Bilskie and Hagen, 2018), exacerbating the impacts of coastal flooding (Olbert et al., 2017), which suggests that the extent or magnitude of flooding is often underpredicted when both river and oceanic processes are not modeled (Bilskie and Hagen, 2018; Kumbier et al., 2018; Chen and Liu, 2014). The computational demand of two and three-dimensional hydrodynamic models, however, typically precludes a large amount of events to be examined. Therefore, while accurately modeling the physics of the combined forcings, researchers taking this approach are often limited to modeling only a select number of boundary conditions. On the other hand, statistical models allow for the investigation of compound water levels through the simulation of combinations of dependent events which may not have been physically realized in observational records (Bevacqua et al., 2017; van den Hurk et al., 2015). In addition, researchers have recently begun to generate hybrid models that link statistical and physical modeling approaches for understanding compound flood events (Moftakhari et al., 2019; Couasnon et al., 2018). Similar to the results solely from hydrodynamic and hydraulic models, statistical and hybrid modeling strategies show that simplifications of the dependence between multiple forcings may lead to an underestimation of flood risk.

This study explores the influence of oceanographic and riverine processes on extreme water levels along a coastal river where there is a substantial river signal recorded in the tide gauge. In order to better understand the river- and ocean-forced water levels at this location, a hybrid methodology is developed for linking statistical simulations of ocean and river boundary conditions with a hydraulic model that simulates along-river water levels. First, river-influenced water levels are defined and removed from SWLs. Then, both river discharge and river-influenced water levels are incorporated into a non-stationary, probabilistic total water level model, which allows for multiple synthetic representations of joint ocean and riverine processes that may not have occurred in the relatively short observational records. Next, a 1-dimensional hydraulic model is used to simulate water

surface elevations along a 10 km stretch of river. Surrogate models are generated from the hydraulic model simulations and used to extract along-river water levels for each probabilistic joint-occurrence of SWL and river discharge in a computational efficient manner. Rather than determining the along-river return level from an equivalent return level forcing (e.g., the 100-yr discharge event drives the 100-yr water level), spatially-varying along-river return levels are extracted and matched to the driving boundary conditions. This technique allows for a spatially explicit analysis of the ocean and river conditions generating extreme water levels.

The following sections describe the study area, present the hybrid modeling framework linking oceanographic and riverine systems, and evaluate the compounding drivers of along-river extreme water levels.

2 Study Area

The Quillayute River is located in Washington state along the US West coast and drains approximately 1630 km² of the north-western Olympic Peninsula into the Pacific Ocean (Czuba et al., 2010). The Quillayute River is approximately 10 km long, is formed by the confluence of the Bogachiel and Sol Duc Rivers (Figure 1), and enters the Pacific Ocean at La Push, Washington, home to the Quileute Tribe. The Quileute Indian Reservation is approximately 4 km² and the majority of community infrastructure sits at the river mouth, bordering the river and open coast. The Quileute Harbor Marina is also situated just inside the river mouth, and is the only port between Neah Bay and Westport, Washington. Rialto spit, which connects Rialto Beach to Little James Island, contains a rocky revetment which protects the marina and the community from ocean and storm wave impact.

The Quillayute River is a natural, unstablized river that is relatively straight at the confluence of the Bogachiel and Sol Duc rivers and increases in sinuosity moving towards the river mouth. Channel-bed materials are coarse (gravel and cobble) in the free-flowing channels and dominated by sand in the small estuary (Czuba et al., 2010). Upstream of river km 3 there are numerous point bars and bends in the river. Between river km 1.5 and 3, the Quillayute is braided with several side channels, usually containing woody debris (Czuba et al., 2010). The channel is straight near the river mouth and is confined by the Rialto spit revetment before draining into the Pacific Ocean.

The oceanic climate of the coastal Pacific Northwest (PNW) is cool and wet with a small range in temperature variation and the majority of rainfall between October and May. Four river basins, the Sol Duc, Bogachiel, Calawah, and Dickey rivers, feed into the Quillayute River and comprise the majority of the watershed. Streamflow in the region is primarily from storm-derived rainfall in the winter and snowmelt during the spring and summer (WRCC, 2017).

Oceanographically-driven SWLs are generally comprised of mean sea level, tides, and non-tidal residuals, which include storm surge. Regional variations in shelf bathymetry, shoreline orientation, storm tracks (Graham and Diaz, 2001), seasonality (Komar et al., 2011), and winds drive differences in storm surge along the US West coast. However, the US West coast's narrow continental shelf, in relation to broad-shelved systems, controls the magnitude of storm surge, which is rarely larger than 1 m (Bromirski et al., 2017; Allan et al., 2011). The PNW is also influenced by a unique interannual climate variability due to the El Niño Southern Oscillation. During El Niño years, the PNW experiences increased water levels for months at a time, along

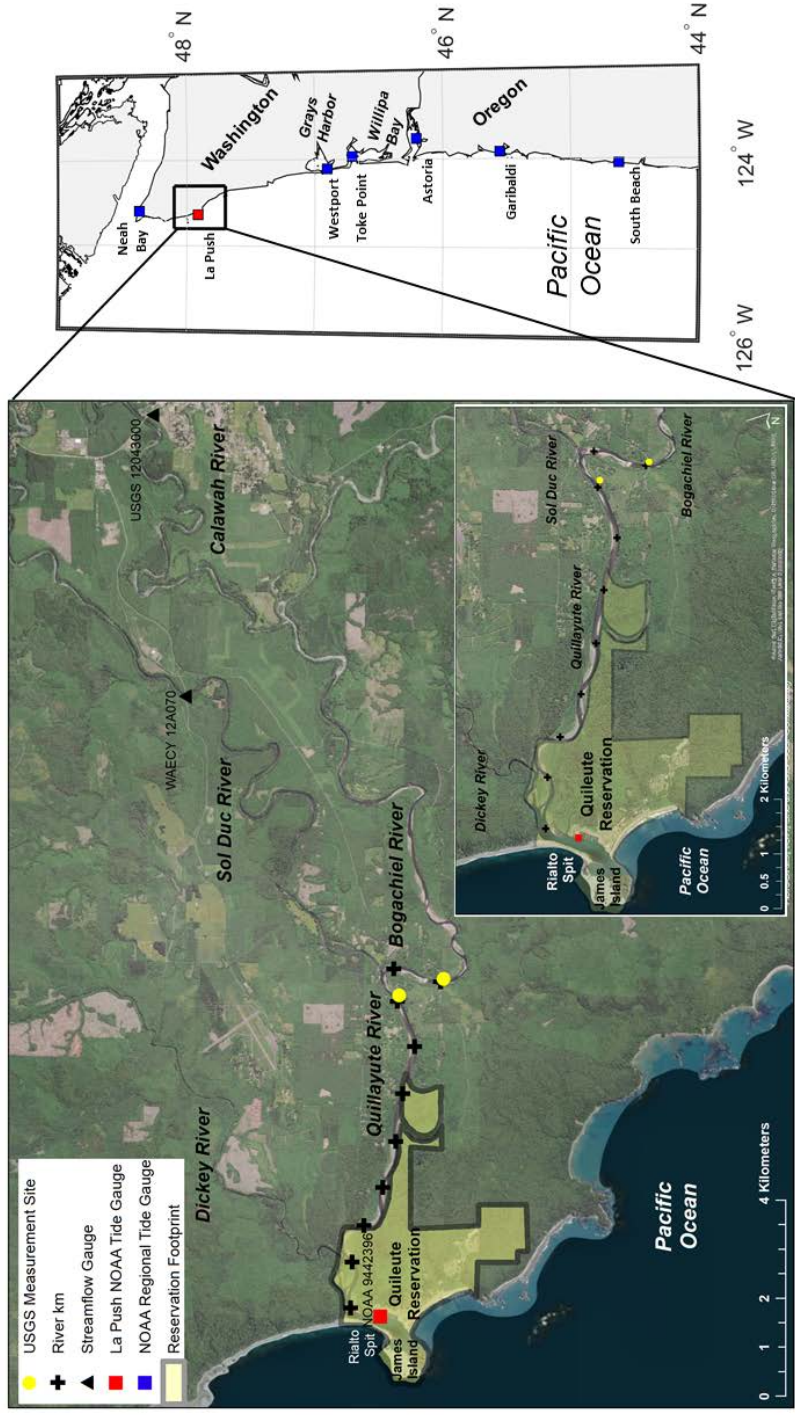


Figure 1. Map of study area (left), which is denoted on the regional map (right) in the black box. The La Push tide gauge is represented as a red square while other regional tide gauges are represented as blue squares. The Calawah and Sol Duc river gauges are represented as black triangles and USGS measurement sites from the May 2010 survey (see supplemental information) are depicted as yellow circles. Approximate river kilometers are denoted as black crosses on the study area map.

with changes in the frequency and intensity of storm systems (Komar et al., 2011; Allan and Komar, 2002). In the PNW, tides are micro- and mesotidal, and at La Push the tidal range is mixed, predominantly semidiurnal, with a mean range of 1.95 m and a great diurnal range of 2.58 m (<https://tidesandcurrents.noaa.gov/datums.html?id=9442396>).

Global rise in sea level and local changes in vertical land motions result in significant longshore variations of relative sea level along the Washington coastline. The northern Washington coast is experiencing relative sea level rates of -1.85 ± 0.42 mm/yr due to a rising coastline, while relative sea level in Willapa Bay in southern Washington is 0.94 ± 2.14 mm/yr (Komar et al., 2011). Tide gauge records at La Push are too short to calculate robust trends in sea level, however, sea level is likely rising in this location, rather than falling, partly due to local land subsidence (Miller et al., 2018).

3 Data

Observational records in the region are generally sparse; one tide gauge exists in the marina near the river mouth and only two of the four rivers which feed into the Quillayute watershed are gauged (Figure 1). The Sol Duc River gauge (WA Dept of Ecology 12A070) is located 7 miles upriver from the Quillayute River and measures hourly discharge and stage observations from 2005-2014. The second river gauge is located on the Calawah River (USGS 12043000), approximately 15 miles upriver from the Quillayute River. The Calawah River flows into the Bogachiel River, and has hourly discharge and stage measurements from 1989 - 2016. The hourly record of discharge measurements from the Sol Duc River is 100% complete, while the Calawah River is 99% complete. An area scaling watershed analysis (Gianfagna et al., 2015) is undertaken to rectify the discharge by the amount of ungauged watershed. The watershed delineation shows that the Bogachiel, Calawah, Sol Duc, and Dickey rivers account for 24%, 22%, 37%, and 17% of the total Quillayute River watershed area, respectively. Noting the similar watershed characteristics and proportional watershed areas, the contribution of the Bogachiel River is estimated by scaling the Calawah River discharge measurements by a factor of 2.09. This scaling factor for estimating Bogachiel River discharge is validated by comparing to 8 discharge point measurements taken during a U.S. Geological Survey (USGS) survey in 2010 (see supplemental information). Discharge for the Quillayute River is estimated by adding together discharge from the Sol Duc and Boagachiel rivers.

Hourly measured SWLs at the La Push tide gauge (NOAA station 9442396, 2004 - 2016) relative to Mean Lower Low Water (MLLW) are downloaded, transformed into NAVD88, and decomposed into mean sea level (η_{MSL}), tide (η_A), and non-tidal residual (η_{NTR}). The η_{NTR} is further decomposed into monthly mean sea level anomalies (η_{MMSLA}), seasonality (η_{SE}), and storm surge (η_{SS}), using methods described in Serafin et al. (2017). Peak η_{SS} events at La Push are found to be the highest on record compared to all US West coast tide gauge stations (Serafin et al., 2017). Upon further investigation of the η_{SS} record, a large portion of extreme η_{SS} events occur during low wave events (Figure 2a) and high river discharge events (Figure 2b). This is inconsistent with η_{SS} in Westport, Washington (Figure 2a and Figure 2c), just south of La Push, and with other tide gauges along the US West coast (not shown). It is therefore hypothesized that the anomalously large signal seen in the η_{SS} is river-induced.

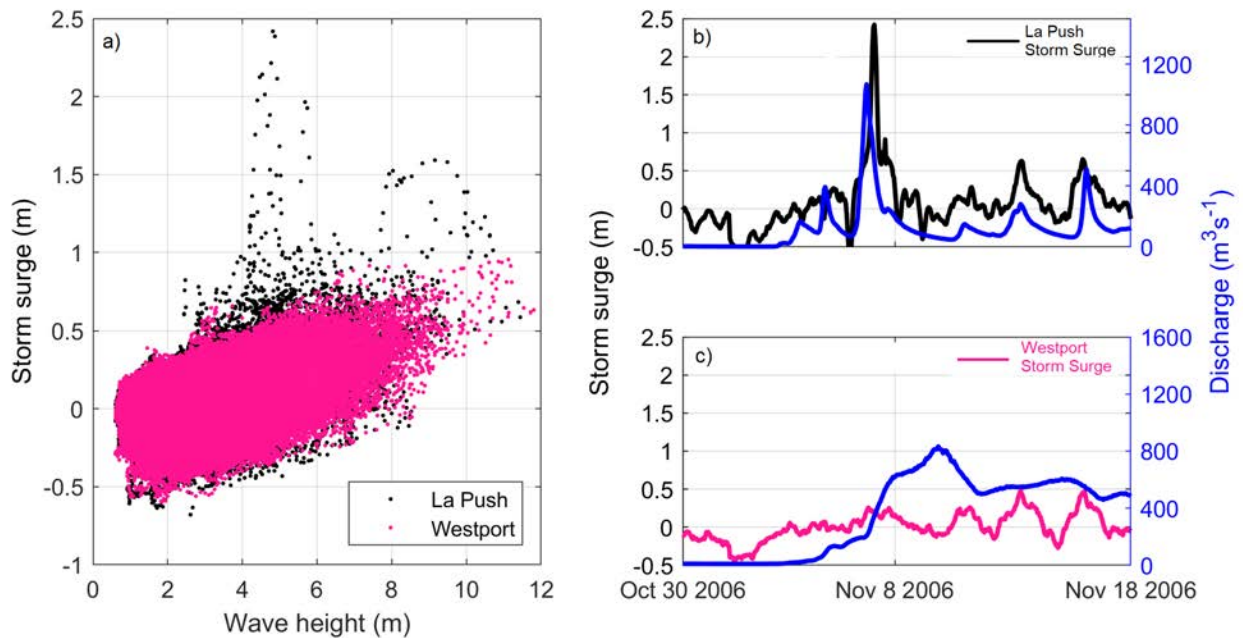


Figure 2. a) The joint relationship between storm surge and wave height for La Push, Washington (black) and Westport, Washington (pink). Example storm surge and discharge relationship at b) La Push and c) Westport, Washington.

To further investigate the anomalously large η_{SS} at the La Push tide gauge, the hydrodynamic model ADvanced CIRcculation (ADCIRC, Luetlich Jr et al. (1992)) and Simulating Waves nearshore (SWAN, Zijlema (2010)) model (ADCSWAN; Dietrich et al. (2011)) is used to simulate water levels at the tide gauge during a storm event corresponding with the peak river discharge on record occurring on January 8, 2009. ADCIRC is run in 2D depth-integrated barotropic mode which performs well for calculating water surface elevations during storm events (Weaver and Luetlich, 2010). SWAN is run in non-stationary mode on an unstructured grid, allowing for tight coupling to ADCIRC. The model is run with two forcing implementations: one including a full forcing of waves, wind, pressure, streamflow, sea level anomalies, seasonality, and tides and one including only streamflow and tides. Once the river-influenced water level is validated, it is removed from the η_{SS} signal and saved as a 6th geophysical variable (η_{Ri} , see supplemental information for removal technique).

Because of the short length of the La Push tide gauge record, decomposed water levels from the La Push tide gauge are merged with decomposed water levels from the Toke Point tide gauge (NOAA station 9440910) to create a combined water level record with a length of 36 years. Details of this methodology are explained in the corresponding supplemental information, as well as in Serafin et al. (2019). Once the two tide gauges are merged, the combined hourly tide gauge record extends from 1980 - 2016 and is 97% complete.

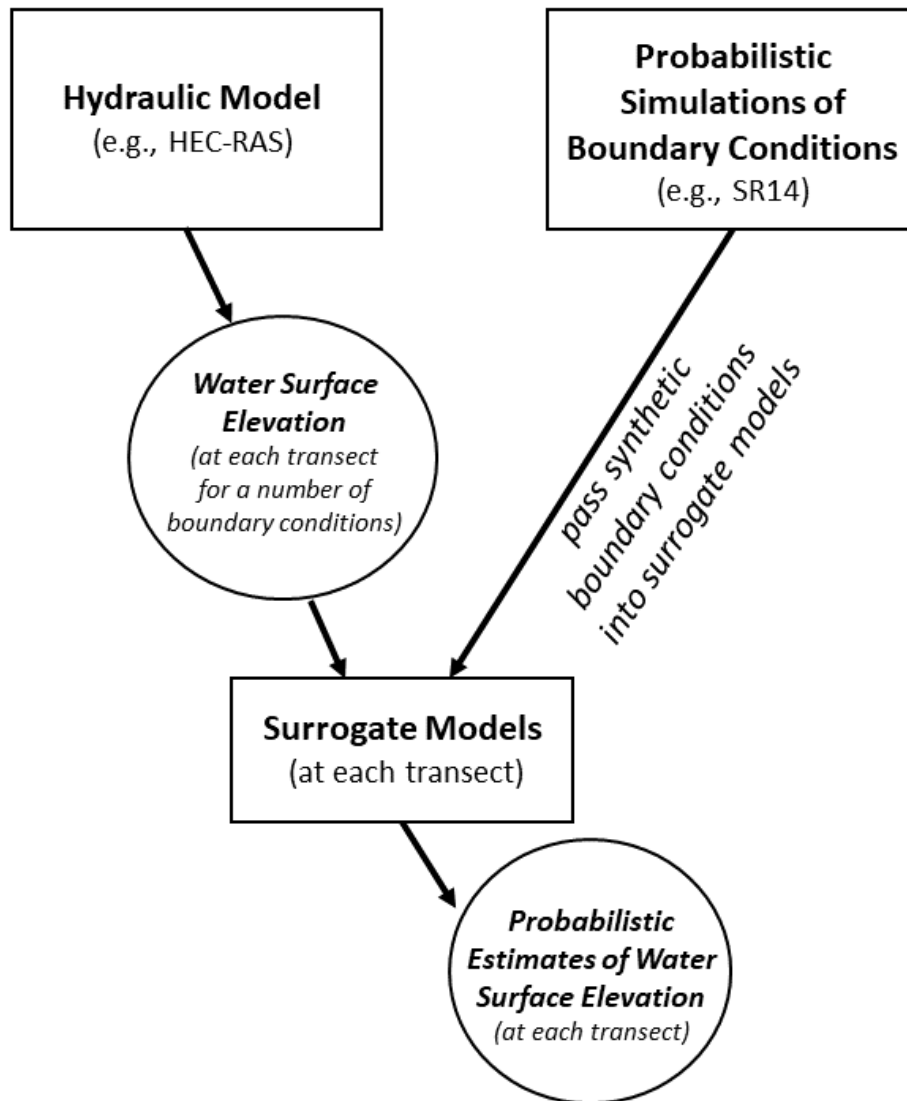


Figure 3. Schematic of hybrid statistical-physical modeling technique. Models are portrayed as squares, while circles portray model outputs.

4 Methods

Return level flood magnitudes, such as the 100-yr event, are typically assumed to be driven by a specific forcing event, such as the 100-yr rainfall or storm surge. However, for processes driven by multiple dimensions, different sizes and combinations of forcing conditions could potentially generate extreme flood magnitudes. To explore the role of compounding forcings in generating extreme water levels, a hybrid modeling framework is developed by merging a hydraulic model simulating river flow with probabilistic simulations of jointly occurring boundary conditions, in this case SWL and river discharge (Figure 3). Statistical simulations allow for long, synthetic records of joint forcings that may not have occurred in the short observational records but are physically capable of co-occurring. Modeling all of the statistically simulated boundary conditions in a hydraulic model to output along-river water levels would be prohibitively expensive. As an alternative to time consuming simulations, surrogate models (Razavi et al., 2012) are developed to approximate the response of a hydraulic model simulation at each along-river location. This technique allows for the analysis of along-river water levels driven by a variety of boundary conditions. Long synthetic records on the order of 500 years allows for the direct empirical extraction of water level return levels rather than an extrapolation from historic observational forcing conditions. In addition, the large sample space of simulated variables permits a comparison of event-based return levels, where the 100-yr water level is determined by the 100-yr forcing, to response-based return levels, where the 100-yr water level is derived and then mapped to its respective forcing conditions. This novel framework is flexible for input of any statistical or hydraulic model. In this application, we use the Serafin and Ruggiero (2014) full simulation total water level model and the US Army Corps of Engineers' (USACE) Hydrologic Engineering Center's River Analysis System (HEC-RAS; Brunner (2016)), which are described in more detail below.

4.1 Probabilistic simulations of boundary conditions

The non-stationary, probabilistic simulation model of Serafin and Ruggiero (2014) (hereinafter SR14) was developed to produce synthetic time series of daily maximum total water levels (TWLs), the combination of waves, tides, and non-tidal residuals, on open-coast sandy beaches. SR14 simulates the individual components of the TWL in a Monte Carlo sense, while appropriately accounting for any dependencies existing between the variables. This modeling technique is able to include non-stationary processes influencing extreme and non-extreme events, such as seasonality, climate variability, and trends in wave heights and water levels. SR14 outputs a number of synthetic records of all variables driving TWLs that produce alternate, but physically plausible, combinations of waves and water levels along an identified stretch of coastline. This technique is flexible to allow for both the simulation of the present-day climate for computing robust statistics on extreme TWL events, as well as the simulation of future climates and their impact on extreme TWLs. Because SR14 was developed for use in open-coast environments, it does not include a procedure for simulating estimates of river discharge, which is present in the local tide gauge at the La Push study site. SR14 is therefore modified to produce synthetic time series of river discharge as well as a river-induced water level.

High discharge events on the two gauged rivers in the watershed, the Sol Duc and Calawah rivers, tend to occur within hours of peak wave events recorded in offshore wave buoy records and water level events recorded in the tide gauge data. Due to the

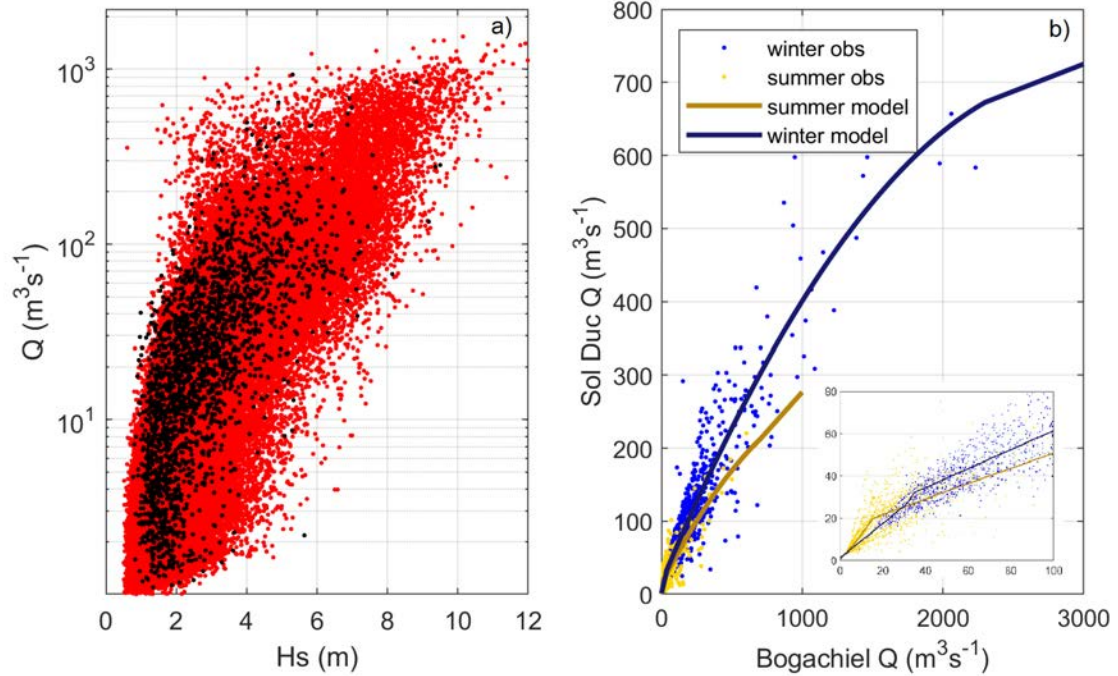


Figure 4. a) Joint relationship between wave height (Hs) and discharge (Q) for the observational record (black) and one example 500 year simulation (red). b) Seasonal model fit for the probabilistic simulation of the Sol Duc River Q in relation to the Bogachiel River Q. The inset displays the model fits for discharge less than $100 \text{ m}^3\text{s}^{-1}$.

interrelated nature of these forcings, daily maximum estimates of Calawah River discharge (Q_C) are compared to all variables simulated in the SR14 model (e.g., wave height, η_{SS} , η_{NTR} , η_{MMSLA}) to capture any dependencies between these processes. The variable with the highest monthly correlation to Q_C is wave height (Hs). Extreme Q_C events are simulated using a bivariate logistic model, which is the same technique used to simulate η_{SS} . The bivariate logistic model preserves the dependency and frequency of occurrence of joint Hs- Q_C events in extreme and non-extreme space. This technique generates a synthetic record of Q_C that is seasonally varying, related to larger-scale climate variability through wave height (essentially as a proxy for storms), and carries the same dependency between variables as the observational record (Figure 4a). Q_C is then multiplied by 2.09 to represent inflow from both the Bogachiel and Calawah rivers.

Discharge measurements at the Sol Duc River are highly correlated with the discharge measurements at the Calawah River ($\rho = 0.9$, $\tau = 0.83$), thus Sol Duc River discharge (Q_{SD}) is modeled based on a relationship with the scaled Q_C , representing the Bogachiel River (Q_B). Estimates of Q_{SD} are related to Q_B during the summer and winter seasons. First, daily maximum Q is split into summer (May, June, July, August, September, and October) and winter (January, February, March, April, November,

December) seasons. Next, models are fit to the joint relationship between the Q_{SD} and Q_B each season, such that for the summer season,

$$Q_{SD} = 1.186Q_B + 0.226. \quad (1)$$

is used when Q_B falls between 0-10 m^3s^{-1} , and

$$5 \quad Q_{SD} = -1.0 \times 10^{-4}Q_B^2 + 0.38Q_B + 14.07. \quad (2)$$

is used when Q_B falls between 10 - 700 m^3s^{-1} (Figure 4b). When Q_B is greater than 700 m^3s^{-1} , Q_{SD} is determined using

$$Q_{SD} = 0.216Q_B + 61.25. \quad (3)$$

For the winter season,

$$Q_{SD} = 0.816Q_B + 1.168. \quad (4)$$

10 is used when Q_B falls between 0-25 m^3s^{-1} , and

$$Q_{SD} = -1.0 \times 10^{-4}Q_B^2 + 0.46Q_B + 16.11. \quad (5)$$

when Q_B falls between 25 - 2300 m^3s^{-1} (Figure 4b). When Q_B is greater than 2300 m^3s^{-1} , Q_{SD} is determined using

$$Q_{SD} = 0.075Q_B + 500.42. \quad (6)$$

Summer and winter Q_B is binned and residuals of Q_{SD} from the above model fits are generated. Normal distributions are fit to
 15 Q_{SD} residuals in each bin, except for low bins (less than 25 m^3s^{-1}) where residuals are fit to exponential distributions. Q_{SD} is then directly related to simulated estimates of Q_B ; Q_{SD} is first determined by fitting the prescribed model to each estimate of Q_B , and then a random sample is taken from the residuals per binned Q_B and added to the model. This technique captures the joint-peaks of the river systems visible in the observed dataset, while allowing for variability between the simulated estimates (Figure 4b).

20 4.1.1 Modeling the river-induced water level

At tide gauges along the US West coast, the maximum daily SWL generally occurs during, or close to, the daily high tide (Serafin and Ruggiero, 2014; Serafin et al., 2017). Modeling peaks in η_{Ri} that occur during low tide would therefore erroneously increase simulated estimates of the SWL occurring during high tide. Thus, instances of η_{Ri} occurring approximately during high tide are retained and all other η_{Ri} peaks are discarded, resulting in 155 η_{Ri} events.

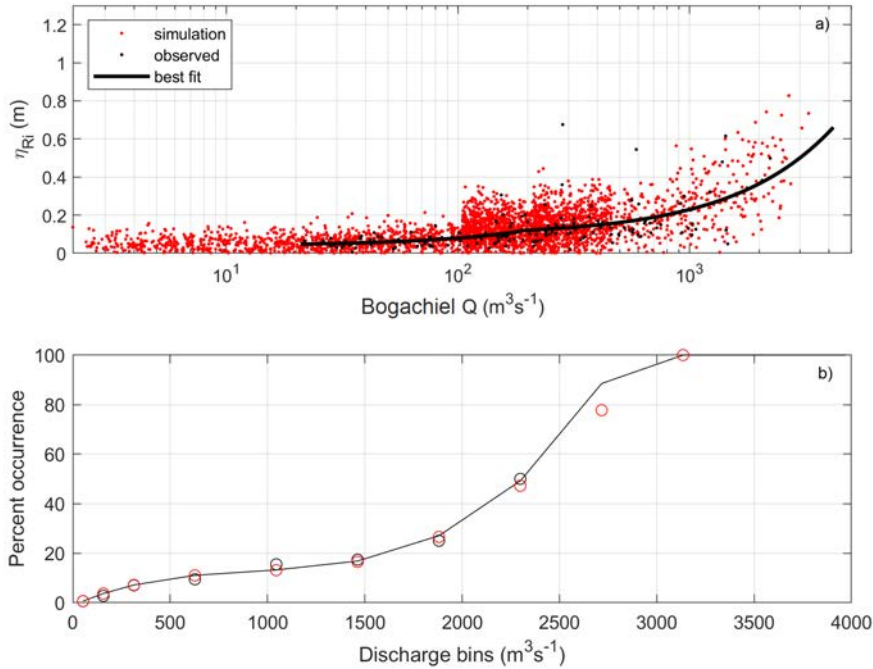


Figure 5. a) The relationship between the river-influenced water level (η_{Ri}) and Bogachiel River discharge on a log-linear scale. The solid black line represents the model fit to the observational records (black dots). b) The percentage of time η_{Ri} occurs in the record during a specific Q_B . In both panels, black represents the observational record and red represents one example 500 year simulation.

Synthetic estimates of η_{Ri} are developed by relating Q_B and η_{Ri} . This relationship is modeled using

$$\eta_{Ri} = 0.039Q_B + 0.854 \times 10^{-3}. \quad (7)$$

when Q_B is below $190 \text{ m}^3\text{s}^{-1}$ and

$$\eta_{Ri} = 0.093Q_B + 0.284 \times 10^{-3}. \quad (8)$$

- 5 when Q_B is above $190 \text{ m}^3\text{s}^{-1}$ (Figure 5a). Next, coarse bins ranging from 100 to $4000 \text{ m}^3\text{s}^{-1}$ are created and the standard deviation (σ) of η_{Ri} within each bin is saved. For bins that contain less than 10 observations, observations from the previous bins are included until there are more than 10 observations per bin for σ calculations. Finally, a 2-point running average is used to smooth σ from each bin to ensure continuous transitions and to avoid the edge-effects from binning a sparse dataset.

There are times of high Q_B without a distinguishable η_{Ri} in the tide gauge record, thus a model is also developed to
 10 simulate the frequency of occurrence of η_{Ri} during daily maximum SWLs. The frequency of occurrence of η_{Ri} is defined as the percentage of time η_{Ri} occurs in the observational record, which is less than 10% of the time when Q_B is less than $210 \text{ m}^3/\text{s}$, and 15 - 25% of the time when Q_B is between 840 and $2090 \text{ m}^3\text{s}^{-1}$ (Figure 5b). For Q_B greater than $2090 \text{ m}^3\text{s}^{-1}$,

η_{Ri} occurs approximately 50% of the time. The frequency of occurrence of η_{Ri} is modeled using a best-fit cubic function, where the frequency of occurrence is a function of Q_B based on the percentage of time the values have occurred in the record. Because there are no events greater than $2500 \text{ m}^3\text{s}^{-1}$ on record, we represent the percentage of occurrence over this value as 100% (Figure 5b).

- 5 Once Q_B is simulated using SR14, η_{Ri} is simulated for every day in time by selecting the synthetic daily estimate of Q_B and randomly sampling from a normal distribution for each Q_B bin, where μ is the regression model and σ is the standard deviation from each bin (Figure 5a). The frequency of occurrence model is then used to select the correct proportion of η_{Ri} events to retain for each synthetic simulation. These techniques capture both the spread of η_{Ri} related to Q_B as well as the percentage of time of occurrence (Figure 5).

10 4.2 Hydraulic model for along-river water levels

While a variety of hydraulic models can be used for determining the elevation of along-river water levels, we employ the Hydraulic Engineering Center's River Analysis System (HEC-RAS; Brunner (2016)). HEC-RAS is used to estimate water surface elevations in rivers and streams in both steady and unsteady flow and under subcritical, supercritical, and mixed flow regimes (Goodell, 2014). HEC-RAS has been previously used to model water surfaces for a range of applications including, but not limited to, floodplain mapping (Yang et al., 2006), flood forecasting (Saleh et al., 2017), dam breaching (Butt et al., 2013), and flood inundation (Horritt and Bates, 2002). HEC-RAS computes water levels by solving the 1D energy equation with an iterative procedure, termed the step method, from one cross-section to the next (Brunner, 2016). For subcritical flows, the step procedure is carried out moving upstream; computations begin at the downstream boundary of the river and the water surface elevation at an upstream cross-section is iteratively estimated until a balanced water surface is obtained. Energy losses between cross-sections are comprised of a frictional loss via the Manning's Equation and a contraction/expansion loss via a coefficient multiplied by the change in velocity head (see Brunner (2016) for more details).

In this application, HEC-RAS is used to model 1D water levels under gradually varied, steady flow conditions at transects along the Quillayute River. While a simplification of flood processes, the 1D application is commonly used to create flood hazard maps. A detailed Digital Elevation Model (DEM) is developed for the river network, including bathymetry and topography for the floodplains of interest (Figure 6). Model domain boundary conditions are chosen as the SWL at the tide gauge (m; downstream boundary) and river discharge from the Sol Duc and Bogachiel rivers (m^3s^{-1} ; upstream boundary). The HEC-RAS model is validated using water surface measurements from a 2010 survey. Details of the HEC-RAS model validation and calibration procedures are documented in supplemental information.

4.3 Hybrid statistical-physical modeling

- 30 The modified simulation technique of SR14 is used to produce 70 500 year long synthetic records representing present-day climate for the time period of 1980-2016 of daily maximum SWL and Q for both the Sol Duc and Bogachiel rivers. Rather than run the ~ 13 million conditions simulated through a numerical model, a limited set of joint boundary conditions of SWL and Q (at the Bogachiel and Sol Duc rivers) are run through HEC-RAS, outputting the elevation of the along-river water

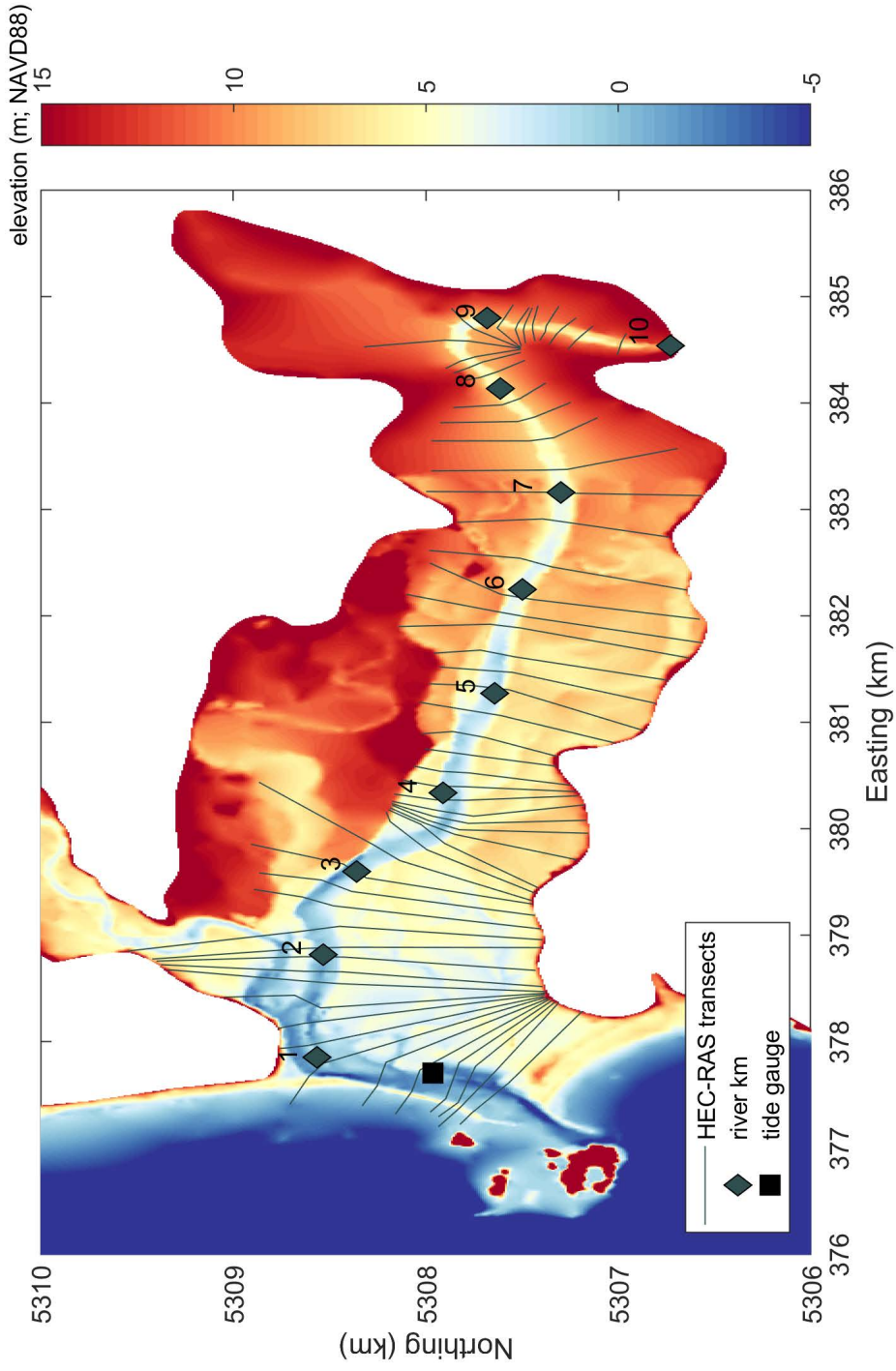


Figure 6. Digital Elevation Model (DEM) used for the HEC-RAS simulations of the Quillayute River. HEC-RAS cross sections are depicted as grey lines. Approximate river kilometer and the location of the tide gauge are depicted as diamonds and a square, respectively.

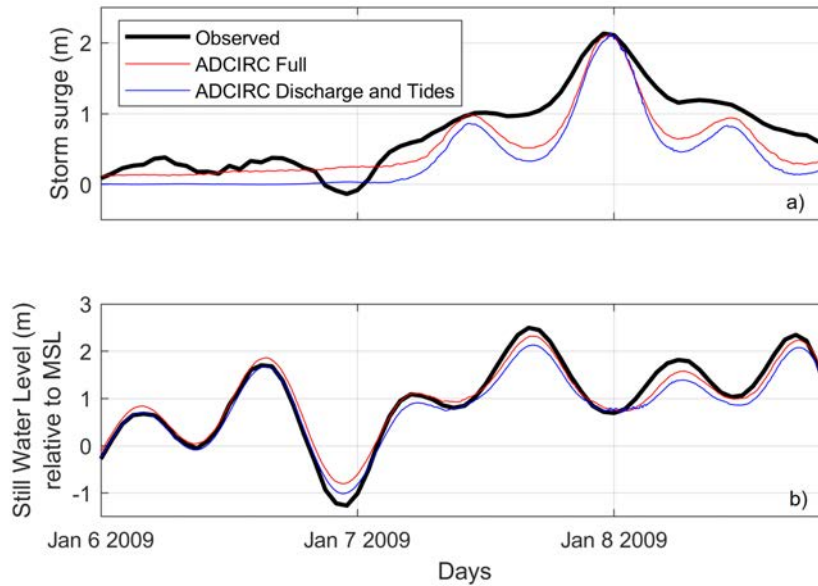


Figure 7. Resulting storm surge (a) and still water level (b) at the La Push tide gauge modeled using ADCIRC for a simulation including full forcing (red) and a simulation including only discharge and tides (blue) compared to the observed storm surge (black). The ADCIRC simulation was run for the maximum discharge event on record occurring on January 8, 2009.

level at each HEC-RAS transect. Surrogate models are generated from the HEC-RAS runs for each transect using a scattered linear interpolation of the 3D surface of boundary conditions. The number of combinations of SWL and Q used to develop the surrogate models are chosen to minimize interpolation errors during validation runs. A daily estimate of water level elevation at each transect is produced by inputting all daily maximum SWL and Q conditions into the surrogate models, which efficiently
 5 extract along-river water levels for any set of SWL and Q inputs. Using the countback method, where for example, the 5th largest event for each synthetic record would be the 100-yr event, water level return levels are extracted for all 70 500 year synthetic records for the 1) along-river water levels at each transect, 2) SWLs, and 3) Q. This methodology provides both an estimate of the return level magnitude (e.g., the average of the 70 100-yr events), as well as the uncertainty around that magnitude (e.g., the distribution of the 70, 100-yr events). It also provides a technique to compare the response-based return
 10 level (e.g., the 100-yr water level) to the event-based return level (e.g., water level driven by the 100-yr SWL or 100-yr Q event).

5 Results

The following section first validates the presence of a river-induced water level within the tide gauge signal, then demonstrates the effectiveness of the surrogate models in representing along-river water levels for unmodeled HEC-RAS boundary condi-

tions. Next, the spatial and temporal variability of the magnitude of along-river water levels and their driving conditions are examined. Finally, low probability water levels, like the 100-yr event, are extracted from the simulated records of along-river water levels and the dominant drivers are evaluated.

5.1 River-induced water level validation

5 Results from ADCSWAN modeling of the January 8, 2009 storm event show that the simulation including only river discharge and tides is nearly able to recreate the measured peak η_{SS} signal at the tide gauge (Figure 7a). The addition of wind, pressure, waves, sea level anomalies, and seasonality is found to have minimal impact on the peak observed η_{SS} . Furthermore, the maximum η_{SS} occurs during low tide (Figure 7b), which indicates a potential relationship between water surface elevation, tidal level, and river discharge. While the ADCSWAN runs only explore one instance of this phenomenon, it provides physics-
10 based evidence that anomalously high η_{SS} at the La Push tide gauge is likely being driven by large discharge events.

5.2 Surrogate model validation

A number of validation scenarios are modeled in HEC-RAS to determine whether the combinations of Q and SWL boundary conditions used to develop the surrogate models represent a large enough sample space of forcing conditions for the interpolation of along-river water levels. The validation scenarios are chosen to cross through both HEC-RAS modeled and unmodeled
15 conditions (Figure 8a). Across all validation scenarios, the average root mean square error (RMSE) between the HEC-RAS directly-modeled and the surrogate model-interpolated water levels is 1 cm. Only 1.5% of the validation scenarios have a bias greater than 10 cm, and the largest RMSE at any transect is 20 cm across all scenarios (Figure 9). The validation scenario with the worst performance occurs during high Q_B and low Q_{SD} paired with low SWL events. However, even during this case, the differences between the HEC-RAS directly-modeled the surrogate model-interpolated water level is small (Figure
20 8b). The main research focus here is extreme water levels, and the conditions driving low probability return level events rarely fall around the scenarios with the highest bias.

5.3 Hybrid modeling of along-river water levels

5.3.1 Temporal variability

Seasonal variability exists in the elevation of along-river water levels. The highest elevation water level occurs during the winter
25 (here defined as December, January, February), while the lowest elevation water level occurs during the spring (March, April, May) (Figure 10a). The spring along-river water level is on average 0.50 m lower than the winter along-river water level, 0.33 m lower than the fall (September, October, November) along-river water level, and 0.03 m lower than the summer (June, July, August) along-river water level (Figure 10b). The difference between seasonal along-river water levels is nonlinear upstream, and certain sections of the river have larger changes in elevation between months (Figure 10b). However, this variation becomes
30 relatively linear downstream of river km 3.

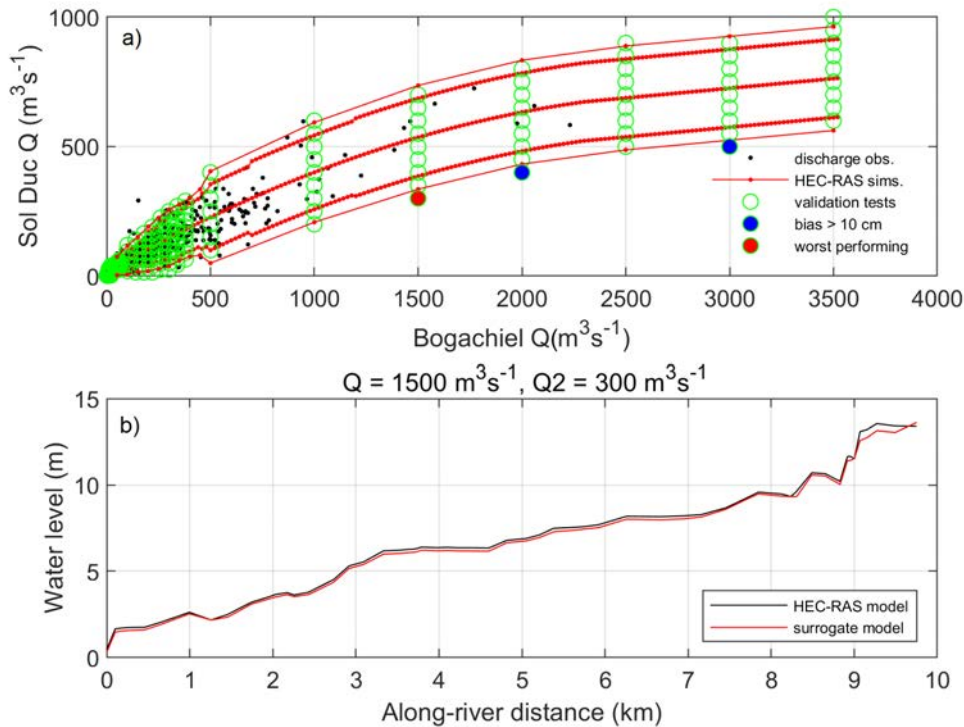


Figure 8. a) Modeled HEC-RAS Q boundary conditions used to generate the surrogate models (red-dotted lines) compared to the simulated conditions used for surrogate model validation (green dots). The black dots represent the observational daily max conditions, while the colored circles represent the worst-performing of the validation tests. The red and blue colored circles represent the scenarios where the interpolated water surface had a bias of over 10 cm lower than the model. b) Example along-river water level for the worst performing condition in the validation tests.

The seasonal variability of the along-river water level is driven by the seasonality of the forcings, which are well represented in the simulations compared to the observations (Figure 11). The monthly median SWLs and η_{NTRs} are higher in the winter than in the summer (Figure 11a and Figure 11b). This cyclical variability is also depicted in the monthly median river discharge from the Quillayute River (combined Sol Duc and Bogachiel Q), and is approximately $200 \text{ m}^3\text{s}^{-1}$ higher in winter months than summer months (Figure 11c). The 98th percentile of SWL, η_{NTR} , and Q have a similar seasonal variability as the median conditions (Figure 11d, Figure 11e, and Figure 11f).

5.3.2 Spatial variability

The large number of joint SWL and Q conditions allows for the direct extraction of water level return levels and the corresponding univariate or multivariate drivers along each HEC-RAS transect. The magnitude of the 100, 25, 10, and annual

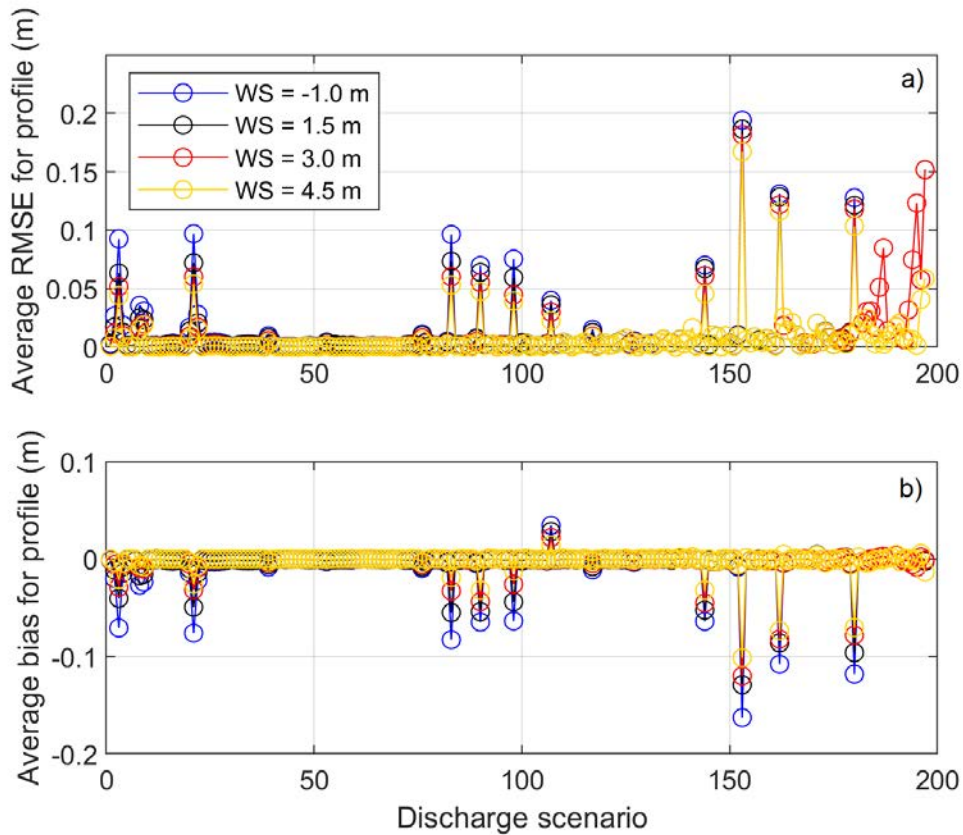


Figure 9. a) Average root mean square error (RMSE) and b) bias for all 197 discharge validation scenarios across 4 out of the 15 SWL scenarios. The worst-performing model is discharge scenario 153.

return level water levels is between 3 and 17 m (NAVD88, Figure 12a). While the peaks in return level events occur at similar locations, the difference between return level events varies spatially moving upriver. For example, at river km 1, the difference between the average (of all simulations) annual and 100-yr event is approximately 0.9 m, whereas at river km 8 and upstream, the difference between these two events is closer to 2 m (Figure 12b).

- 5 The dominant forcing conditions driving water level return levels varies along-river. At the river mouth, the annual water level event (e.g., the event that is expected every year) in each simulation occurs during Q ranging from 40 - 2600 m^3s^{-1} and SWLs around 3.3 m, which corresponds with the annual SWL event (Figure 13a). Moving upstream to river km 1 and 2, the annual water level event is driven by both high SWL occurring during low Q and low SWL occurring during high Q. At river km 4, the annual water level event occurs during the annual Q event coincident with SWLs that range from 1.8 - 3.9 m (Figure
- 10 13a). These results are similar, albeit events are larger magnitude, for the 100-yr water level event. Downstream 100-yr water levels are driven by SWLs, upstream 100-yr water levels are driven by Q, and the 100-yr water level between km 1 and 2 is driven by different combinations of high and low SWL and Q events (Figure 13b).

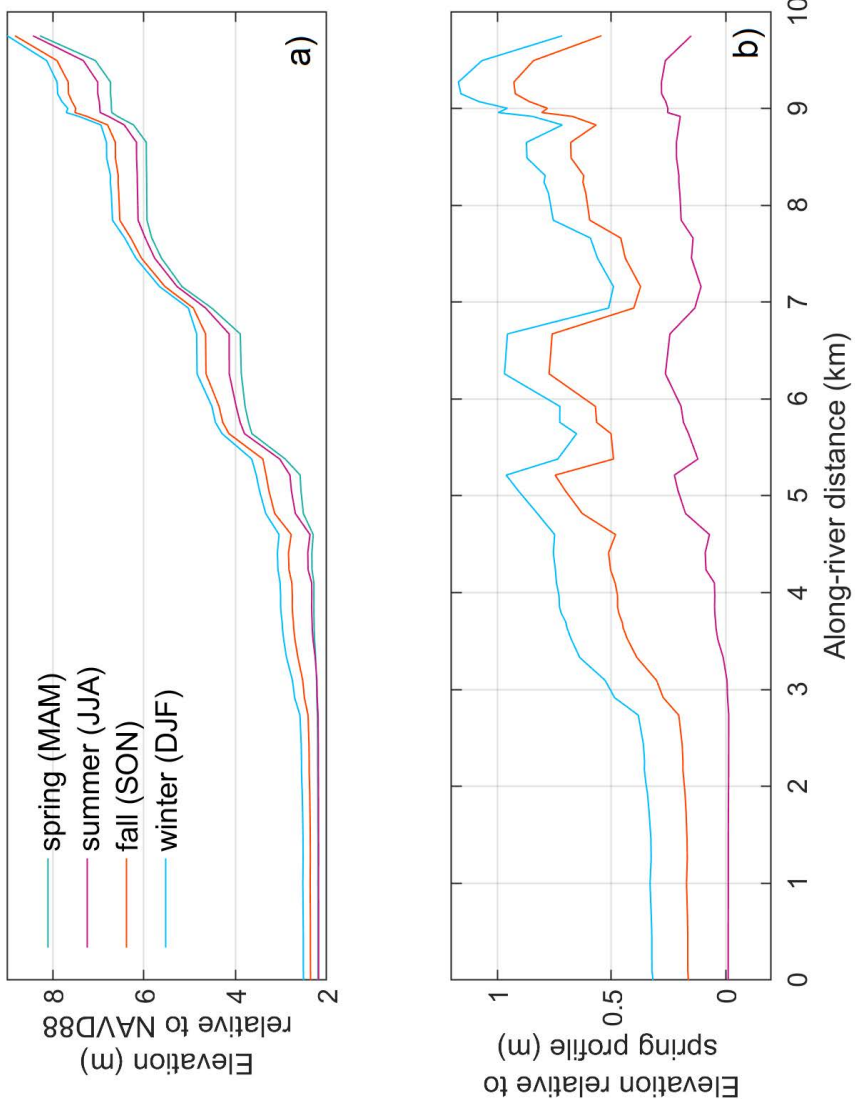


Figure 10. a) Variability of along-river water levels averaged over spring (MAM), summer (JJA), fall (SON), and winter (DJF). b) The difference between the spring and summer, fall, and winter along-river water levels.

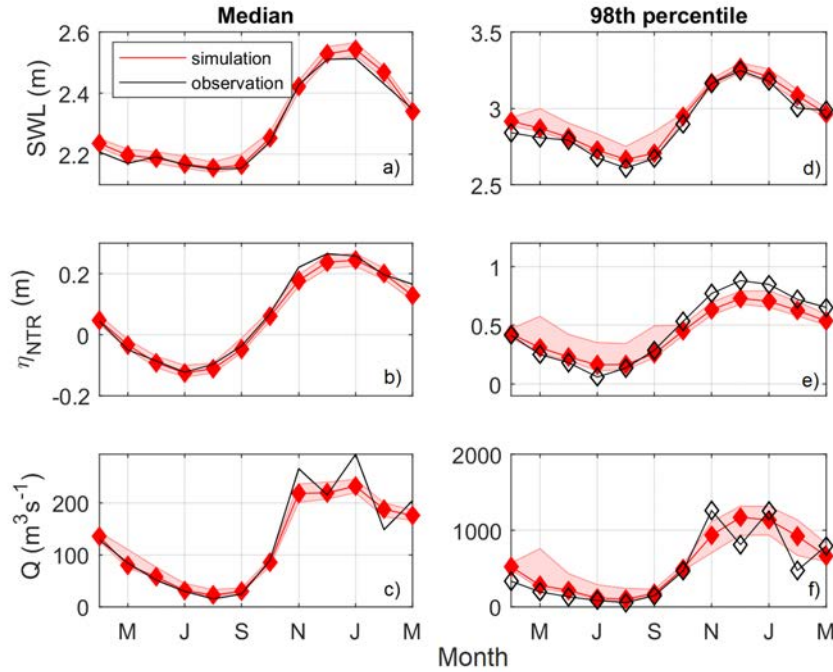


Figure 11. Left) Observational (black) and simulated (red) monthly median still water level (SWL), discharge (Q), and non-tidal residual (η_{NTR}). Right) Observational (black) and simulated (red) monthly 98th percentile of the SWL, and η_{NTR} Q . Red shading indicates the bounds value from each simulation.

The relative importance of both oceanic and riverine forcing to extreme water levels emerges when averaging the magnitude of the drivers of the water level return levels at each transect from all 70 500 year long simulations (Figure 14). The magnitude of the average Q driving water level return levels gradually increases by approximately $1000 m^3s^{-1}$ over river km 0 - 2 and then is consistent from river km 2 to 10 (Figure 14a). Downstream, between river km 0 and 0.25, the magnitude of the average SWL driving water level return levels is consistent and then gradually decreases over a 1 km zone (Figure 14b).

When comparing to water level return levels driven by a univariate forcing or event return level (e.g., along-river water levels modeled from the 100-yr Q or SWL event), we find that the stretches of river driven by a consistent SWL or Q forcing approximates the univariate return level event. Therefore, the 100-yr SWL does indeed cause the 100-yr water level downstream, between river km 0 and 0.25, while the 100-yr Q event drives the 100-yr water level upstream, between river km 2 - 10 (dashed lines, Figure 14). However, between river km 0.25 - 1.75 a flood transition zone is present, where neither the SWL return level or the Q return level events drive the water level return level. This is consistent across all return level events.

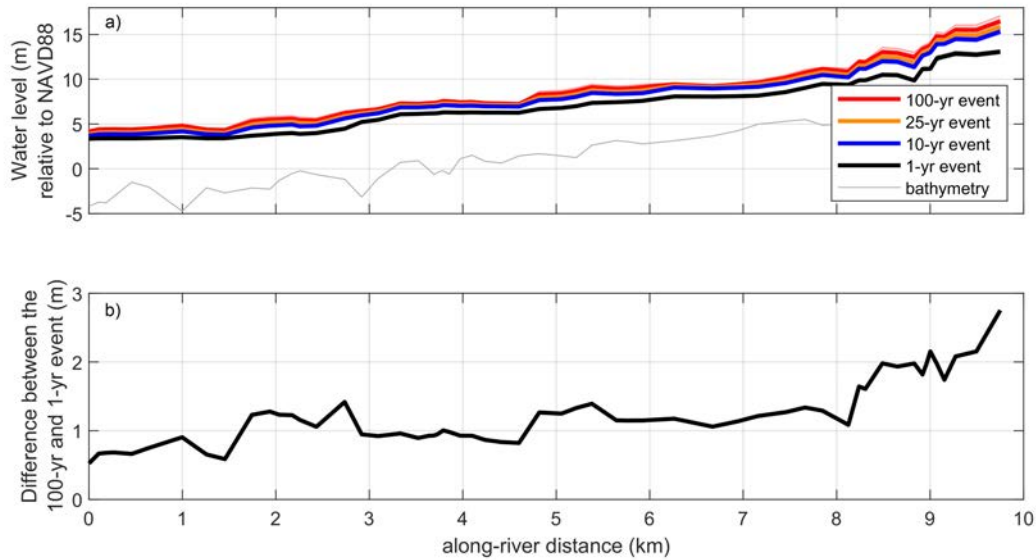


Figure 12. a) The water level return level at each transect for all 70 probabilistic simulations. Each return level event displays the average of the simulations (solid line) as well as the range around the average (shaded). b) The along-river difference between the annual and 100-yr event, averaged over 70 simulations.

6 Discussion

The hybrid model developed in this study, which combines statistical simulations with a physics-based model, provides an approach for probabilistically evaluating the conditions that drive extreme water levels, not only in an open-coast setting, but also miles upriver. The ability to simulate millions of combinations of Q and SWL events allows for a robust estimate of result-
 5 ing along-river water levels, which numerical models alone are unable to consider due to large computational expenses. While some of our modeling techniques are specific to this location, the overall framework for combining statistical and physics-based models is general enough for use in coastal locations throughout the globe where flooding arises from compounding processes.

The decomposition of the SWL into low and high frequency signals, including a river-influenced component, helps identify the importance of physical processes for generating high water levels across various regional settings. This is especially
 10 important in locations like the US West coast, where the steep, narrow continental shelf prevents wind and pressure driven storm surge from being overwhelmingly large (Allan et al., 2011). The influence of the river signal in the tide gauge is directly related to the setting of our study site. The estuary is relatively small and narrow with the river discharging directly into the ocean. This is dissimilar to other tide gauges in the region which are located in larger estuaries, situated away from river input. Estuaries typically exhibit wave, tide, or river-dominant morphology, based on the relative energy of each process (Dalrymple
 15 et al., 1992). The Quillayute River outlets directly to a high wave energy environment and has a small estuary volume compared to its river input volume. The steep catchment of the mountainous environment means a short response time for rainfall,

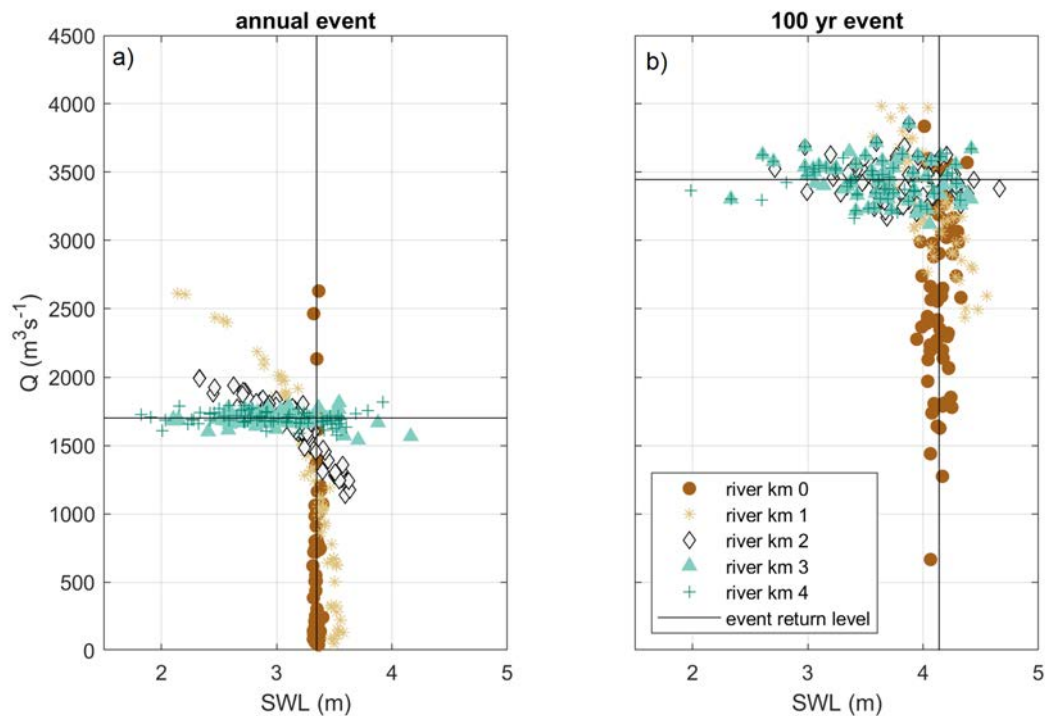


Figure 13. The individual Q or SWL condition driving the a) annual and b) 100-yr water level event at specific along-river locations for each 70 500 year simulation. In both figures, the black lines represent the annual and 100-yr return level magnitude for Q and SWL.

therefore producing peak discharges temporally similar to peak storm-induced still water levels, allowing for interaction between the two. In contrast, water level elevations with large estuary volume compared to river discharge are less influenced by fluvial processes. Furthermore, a larger estuary may experience variability in the water surface elevation due to wave-induced setup and/or other local storm-induced processes (Cheng et al., 2014; Olabarrieta et al., 2011), which may further dampen the influence of a river signal.

This research confirms the presence of an oceanographic-fluvial transition zone, where traditional, univariate methodologies for defining return level events are insufficient for defining water level return levels. Between river km 1 and 2, we find that a range of SWL and Q conditions drive all return level events, and water levels are driven by neither the univariate SWL or Q return level event. A similar flood zone transition was recently modeled numerically, and albeit for a single event, physically demonstrates the importance of including multiple variables to reproduce accurate flooding (Bilskie and Hagen, 2018). Thus, flood hazard assessments on systems with multivariate forcings may misrepresent water level elevations for low probability events if only univariate variables are modeled. This has large implications for characterizing the risk to flooding, especially in the context of mapping flooding hazards. Furthermore, we show that return level water levels can occur over a range of

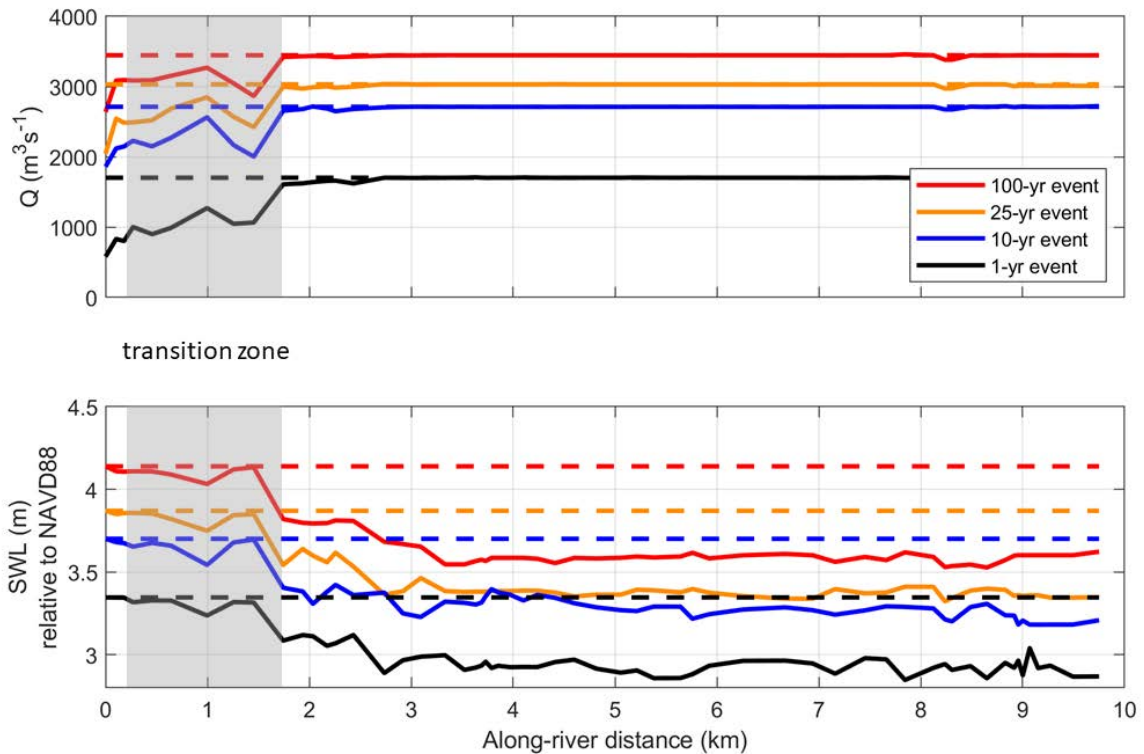


Figure 14. The average forcing condition driving along-river return levels at each transect where a) displays the Quillayute Q conditions and b) displays the SWL conditions. The dashed lines depict the univariate forcing conditions, where the along-river return level is assumed to be driven by either Q or SWL. Red, orange, blue, and black lines represent the 100, 25, 10, and annual return level event. The grey shaded area represents a transition zone, where the water level is driven by a combination of SWL and Q events.

combined extreme and non-extreme forcing in the flood transition zone. This illustrates that in order to properly understand the impacts of compounding flooding, more than just design scenarios need to be considered for the proper assessment of risk.

Many of our results can be explained by dynamics that occur during interacting ocean and river flows. For example, a coincidence of high SWL and peak river discharge may induce blocking, where river-induced water levels are trapped upstream and either flood overbank or outlet to the ocean as the tide recedes (Kumbier et al., 2018; Chen and Liu, 2014). The outletting to the ocean as the tide recedes artificially inflates SWLs at the tide gauge, increasing water levels for days at a time and prolonging exposure to flooding. When subtracting a tide time series from this signal, storm surge would appear to be elevated at low tide. While the ADCSWAN simulation confirms the presence of this effect by matching the peak storm surge at low tide, our hybrid methodology only models steady flow scenarios. Thus, with co-occurring daily maximum SWL and discharge, our model may miss certain dynamics important for flooding over unsteady conditions. Furthermore, interactions between storm

surge and river discharge may increase the overall elevation of the residual (Maskell et al., 2013). While beyond the scope of our present study, this unsteady characteristics are important to consider in future research.

Because sea level rise, along with other changes to the climate, will exacerbate the compounding effects of flood drivers (Moftakhari et al., 2017; Wahl et al., 2015), it is also important to consider the impact of changes to processes driving flooding events in the future (Zscheischler et al., 2018). By 2100, the likely range of relative sea level rise in the La Push area is projected to be between 18 and 80 cm, considering vertical land motion and various emissions scenarios (Miller et al., 2018). The western Olympic Peninsula is projected to experience increased winter precipitation (Mote et al., 2013; Halofsky et al., 2011) which could subsequently increase either the frequency or intensity of high Q events along the Quillayute River. While we have characterized the spatial variability in extreme water levels in the present-day, there is a high likelihood changes in the future climate will shift the importance of these interacting processes.

7 Conclusions

This research illustrates the importance of considering a large number of forcing conditions to model compounding processes when evaluating extreme water levels. Here we find that in coastal settings, river discharge can be an important driver of high water levels measured in a tide gauge. We also find that the univariate, event-based return level event, like the 100-yr discharge, does not always match the response-based return level, like the 100-yr water level. Furthermore, when processes compound, along-river return levels may be driven by events that are not considered extreme themselves. Probabilistic techniques allowing for the analysis of thousands to millions of combinations of events not captured in the observational record provides a characterization of where river, ocean, or the combination of the two, may be important for generating extreme events.

Overall, the hybrid merging of a statistical and numerical model provides a methodology for better understanding the drivers of flooding along the length of a river. While our model does not actively resolve the physical interaction of river and oceanographic flow, it develops an approach for characterizing and extracting river-influenced water levels measured at tide gauges while robustly modeling the drivers of extreme along-river water levels. Understanding the dominant, spatially variable drivers of flooding events will help coastal communities better understand their risks, which is important for increasing resilience to future events.

8 Data availability

Data can be made available by the authors upon request.

9 Author contribution

The study and methodology were conceived by KAS and PR. KAS carried out the analyses, produced the results, and wrote the manuscript under the supervision of PR. KAP carried out the analyses and produced the results of the ADCIRC simulations.

KAP also developed the topography/bathymetry DEM as well as the geometric files for use in HEC-RAS. KAS, PR, KAP, and DFH all contributed by generating ideas, discussing results, and manuscript editing.

10 Competing interests

The authors declare that they have no conflict of interest.

- 5 *Acknowledgements.* Tide gauge records are available through the National Oceanic and Atmospheric Administration (NOAA) National Ocean Service (NOS) website and river discharge is available through the U.S Geological Survey (USGS) National Water Information System (<https://waterdata.usgs.gov/wa/nwis/rt>). Bathymetric and topographic data for DEM creation were obtained from NOAA's Elevation Data viewer. Thank you to Michael Rossotto and Garrett Rasmussen for providing updated shapefiles of the Quileute Reservation boundaries. Thank you also to two anonymous reviewers whose comments improved the quality of this manuscript. This work was funded by the NOAA
- 10 Regional Integrated Sciences and Assessments Program (NA15OAR4310145) and a contracted grant with the Quinault Treaty Area (QTA) tribal governments (Quinault Indian Nation, Hoh Indian Tribe, and Quileute Tribe).

References

- Allan, J. C. and Komar, P. D.: Extreme storms on the Pacific Northwest coast during the 1997-98 El Niño and 1998-99 La Niña, *Journal of Coastal Research*, pp. 175–193, 2002.
- Allan, J. C., Komar, P. D., and Ruggiero, P.: Storm Surge Magnitudes and Frequency on the Central Oregon Coast, in: Proc. Solutions to Coastal Disasters Conf, 2011.
- Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., and Vrac, M.: Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy), *Hydrology and Earth System Sciences*, 21, 2701–2723, 2017.
- Bilskie, M. and Hagen, S.: Defining Flood Zone Transitions in Low-Gradient Coastal Regions, *Geophysical Research Letters*, 45, 2761–2770, 2018.
- Bromirski, P. D., Flick, R. E., and Miller, A. J.: Storm surge along the Pacific coast of North America, *Journal of Geophysical Research: Oceans*, 122, 441–457, 2017.
- Brunner, G. W.: HEC-RAS River Analysis System Hydraulic Reference Manual, Version 5.0, Tech. rep., 2016.
- Bunya, S., Dietrich, J. C., Westerink, J., Ebersole, B., Smith, J., Atkinson, J., Jensen, R., Resio, D., Luettich, R., Dawson, C., et al.: A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part I: Model development and validation, *Monthly Weather Review*, 138, 345–377, 2010.
- Butt, M. J., Umar, M., and Qamar, R.: Landslide dam and subsequent dam-break flood estimation using HEC-RAS model in Northern Pakistan, *Natural Hazards*, 65, 241–254, 2013.
- Chelton, D. B. and Enfield, D. B.: Ocean signals in tide gauge records, *Journal of Geophysical Research: Solid Earth*, 91, 9081–9098, 1986.
- Chen, W.-B. and Liu, W.-C.: Modeling flood inundation induced by river flow and storm surges over a river basin, *Water*, 6, 3182–3199, 2014.
- Cheng, T., Hill, D., and Read, W.: The Contributions to Storm Tides in Pacific Northwest Estuaries: Tillamook Bay, Oregon, and the December 2007 Storm, *Journal of Coastal Research*, 31, 723–734, 2014.
- Couasnon, A., Sebastian, A., and Morales-Nápoles, O.: A Copula-Based Bayesian Network for Modeling Compound Flood Hazard from Riverine and Coastal Interactions at the Catchment Scale: An Application to the Houston Ship Channel, Texas, *Water*, 10, 1190, 2018.
- Czuba, J. A., Barnas, C. R., McKenna, T. E., Justin, G. B., and Payne, K. L.: Bathymetric and streamflow data for the Quillayute, Dickey, and Bogachiel Rivers, Clallam County, Washington, April–May 2010, vol. 537, US Department of the Interior, US Geological Survey, 2010.
- Dalrymple, R. W., Zaitlin, B. A., and Boyd, R.: Estuarine facies models: conceptual basis and stratigraphic implications: perspective, *Journal of Sedimentary Research*, 62, 1992.
- de Vries, H., Breton, M., de Mulder, T., Krestenitis, Y., Proctor, R., Ruddick, K., Salomon, J. C., Voorrips, A., et al.: A comparison of 2D storm surge models applied to three shallow European seas, *Environmental Software*, 10, 23–42, 1995.
- Dietrich, J., Zijlema, M., Westerink, J., Holthuijsen, L., Dawson, C., Luettich Jr, R., Jensen, R., Smith, J., Stelling, G., and Stone, G.: Modeling hurricane waves and storm surge using integrally-coupled, scalable computations, *Coastal Engineering*, 58, 45–65, 2011.
- Gianfagna, C. C., Johnson, C. E., Chandler, D. G., and Hofmann, C.: Watershed area ratio accurately predicts daily streamflow in nested catchments in the Catskills, New York, *Journal of Hydrology: Regional Studies*, 4, 583–594, 2015.
- Goodell, C.: *Breaking the HEC-RAS Code: A User’s Guide to Automating HEC-RAS*, h2Is, Portland, OR, 2014.
- Graham, N. E. and Diaz, H. F.: Evidence for intensification of North Pacific winter cyclones since 1948, *Bulletin of the American Meteorological Society*, 82, 1869–1893, 2001.

- Halofsky, J. E., Peterson, D. L., O'Halloran, K. A., and Hoffman, C. H.: Adapting to climate change at Olympic National Forest and Olympic National Park, Gen. Tech. Rep. PNW-GTR-844. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 130 p, 844, 2011.
- Horritt, M. and Bates, P.: Evaluation of 1D and 2D numerical models for predicting river flood inundation, *Journal of Hydrology*, 268, 87–99, 5 2002.
- Horsburgh, K. and Wilson, C.: Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, *Journal of Geophysical Research: Oceans*, 112, 2007.
- Komar, P. D., Allan, J. C., and Ruggiero, P.: Sea level variations along the US Pacific Northwest coast: Tectonic and climate controls, *Journal of Coastal Research*, 27, 808–823, 2011.
- 10 Kumbier, K., Cabral Carvalho, R., Vafeidis, A. T., and Woodroffe, C. D.: Investigating compound flooding in an estuary using hydrodynamic modelling: a case study from the Shoalhaven River, Australia, *Natural Hazards and Earth System Sciences*, 18, 463–477, 2018.
- Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., McInnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M.: A compound event framework for understanding extreme impacts, *Wiley Interdisciplinary Reviews: Climate Change*, 5, 113–128, 2014.
- Luettich Jr, R. A., Westerink, J. J., and Scheffner, N. W.: ADCIRC: An Advanced Three-Dimensional Circulation Model for Shelves, Coasts, 15 and Estuaries. Report 1. Theory and Methodology of ADCIRC-2DDI and ADCIRC-3DL., Tech. rep., Coastal Engineering Research Center, Vicksburg MS, 1992.
- Maskell, J., Horsburgh, K., Lewis, M., and Bates, P.: Investigating River–Surge Interaction in Idealised Estuaries, *Journal of Coastal Research*, 30, 248–259, 2013.
- Mawdsley, R. J. and Haigh, I. D.: Spatial and temporal variability and long-term trends in skew surges globally, *Frontiers in Marine Science*, 20 3, 29, 2016.
- Miller, I., Morgan, H., Mauger, G., T., N., R., W., D., S., M., W., and E., G.: Projected sea level rise for Washington state: A 2018 assessment, Tech. rep., A collaboration of Washington Sea Grant, University of Washington Climate Impacts Group, Oregon State University, University of Washington, and US Geological Survey. Prepared for the Washington Coastal Resilience Project, 2018.
- Moftakhari, H., Schubert, J. E., AghaKouchak, A., Matthew, R., and Sanders, B. F.: Linking Statistical and Hydrodynamic Modeling for 25 Compound Flood Hazard Assessment in Tidal Channels and Estuaries, *Advances in Water Resources*, 2019.
- Moftakhari, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., and Matthew, R. A.: Compounding effects of sea level rise and fluvial flooding, *Proceedings of the National Academy of Sciences*, 114, 9785–9790, 2017.
- Mote, P. et al.: *Climate: Variability and Change in the Past and the Future*. Chapter 2, 25-40, 2013.
- Odigie, K. O. and Warrick, J. A.: Coherence between Coastal and River Flooding along the California Coast, *Journal of Coastal Research*, 30 2017.
- Olabarrieta, M., Warner, J. C., and Kumar, N.: Wave-current interaction in Willapa Bay, *Journal of Geophysical Research: Oceans*, 116, 2011.
- Olbert, A. I., Comer, J., Nash, S., and Hartnett, M.: High-resolution multi-scale modelling of coastal flooding due to tides, storm surges and rivers inflows. A Cork City example, *Coastal Engineering*, 121, 278–296, 2017.
- 35 Razavi, S., Tolson, B. A., and Burn, D. H.: Review of surrogate modeling in water resources, *Water Resources Research*, 48, 2012.
- Saleh, F., Ramaswamy, V., Wang, Y., Georgas, N., Blumberg, A., and Pullen, J.: A Multi-Scale Ensemble-based Framework for Forecasting Compound Coastal-Riverine Flooding: The Hackensack-Passaic Watershed and Newark Bay, *Advances in Water Resources*, 2017.

- Serafin, K. A. and Ruggiero, P.: Simulating extreme total water levels using a time-dependent, extreme value approach, *Journal of Geophysical Research: Oceans*, 119, 6305–6329, 2014.
- Serafin, K. A., Ruggiero, P., and Stockdon, H. F.: The relative contribution of waves, tides, and nontidal residuals to extreme total water levels on US West Coast sandy beaches, *Geophysical Research Letters*, 44, 1839–1847, 2017.
- 5 Serafin, K. A., Ruggiero, P., Barnard, P., and Stockdon, H. F.: The influence of shelf bathymetry and beach topography on extreme total water levels: Linking large-scale changes of the wave climate to local coastal hazard, *Coastal Engineering*, 2019.
- Svensson, C. and Jones, D. A.: Dependence between extreme sea surge, river flow and precipitation in eastern Britain, *International Journal of Climatology*, 22, 1149–1168, 2002.
- Sweet, W. V., Park, J., Gill, S., and Marra, J.: New ways to measure waves and their effects at NOAA tide gauges: A Hawaiian-network perspective, *Geophysical Research Letters*, 42, 9355–9361, 2015.
- 10 Tsimplis, M. and Woodworth, P.: The global distribution of the seasonal sea level cycle calculated from coastal tide gauge data, *Journal of Geophysical Research: Oceans*, 99, 16 031–16 039, 1994.
- van den Hurk, B., van Meijgaard, E., de Valk, P., van Heeringen, K.-J., and Gooijer, J.: Analysis of a compounding surge and precipitation event in the Netherlands, *Environmental Research Letters*, 10, 035 001, 2015.
- 15 Vetter, O., Becker, J. M., Merrifield, M. A., Pequignet, A.-C., Aucan, J., Boc, S. J., and Pollock, C. E.: Wave setup over a Pacific Island fringing reef, *Journal of Geophysical Research: Oceans*, 115, 2010.
- Wahl, T., Jain, S., Bender, J., Meyers, S. D., and Luther, M. E.: Increasing risk of compound flooding from storm surge and rainfall for major US cities, *Nature Climate Change*, 5, 1093–1097, 2015.
- Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., Veldkamp, T. I., Winsemius, H. C., and Wahl, T.: Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries, *Environmental Research Letters*, 13, 084 012, 2018.
- 20 Weaver, R. and Luettich, Jr, R.: 2D vs. 3D storm surge sensitivity in ADCIRC: Case study of hurricane isabel, in: *Estuarine and Coastal Modeling (2009)*, pp. 762–779, ASCE, 2010.
- Williams, J., Horsburgh, K. J., Williams, J. A., and Proctor, R. N.: Tide and skew surge independence: New insights for flood risk, *Geophysical Research Letters*, 43, 6410–6417, 2016.
- 25 WRCC: Climate of Washington: Western Regional Climate Center website, https://wrcc.dri.edu/Climate/narrative_wa.php, 2017.
- Yang, J., Townsend, R. D., and Daneshfar, B.: Applying the HEC-RAS model and GIS techniques in river network floodplain delineation, *Canadian Journal of Civil Engineering*, 33, 19–28, 2006.
- Zheng, F., Westra, S., and Sisson, S. A.: Quantifying the dependence between extreme rainfall and storm surge in the coastal zone, *Journal of Hydrology*, 505, 172–187, 2013.
- 30 Zijlema, M.: Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids, *Coastal Engineering*, 57, 267–277, 2010.
- Zscheischler, J., Westra, S., Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M., Wahl, T., et al.: Future climate risk from compound events, *Nature Climate Change*, p. 1, 2018.

Supplemental Information

Katherine A. Serafin^{1,2}, Peter Ruggiero¹, Kai A. Parker³, and David F. Hill³

¹College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA

²Department of Geophysics, Stanford University, Stanford, CA, USA

³College of Engineering, Oregon State University, Corvallis, OR, USA

Correspondence: Katherine A. Serafin (kserafin@stanford.edu)

1 Hydraulic model domain and setup

HEC-RAS model runs require detailed terrain information for the river network, including bathymetry and topography for the floodplains of interest. Topography data is sourced from a 2014 U.S Army Corps of Engineers (USACE) lidar survey (USACE, 2014). Bathymetry data is developed by blending two NOAA digital elevation models (DEM): National Geophysical Data Center's (NGDC) La Push, WA tsunami DEM (1/3 arc second; NGDC (2007)) and the coastal relief model (3 arc seconds; NGDC (2003)). These datasets, however, do not accurately resolve the channel depths of the Quillayute River inland of the coast, so a 2010 US Geological Survey (USGS)-conducted bathymetric survey of the river is also blended into the DEM (Czuba et al., 2010).

In 2010, depths of along-river cross sections and an 11 km long longitudinal profile from the Bogachiel River to the mouth of the Quillayute River were surveyed (Czuba et al., 2010). The survey of the longitudinal river profile also recorded the elevation of the water surface. Ideally, the collected bathymetry dataset would be merged directly into the existing DEM. The Quillayute River, however, is uncontrolled and meanders over time, producing a variation in the location of the main river channel between the DEM and the high-resolution USGS-collected bathymetric data. Therefore, the USGS bathymetric profiles are adjusted to match the location of the DEM channel. While a product of multiple datasets and processing steps, the final DEM provides bathymetric/topographic data with the most up-to-date channel depths for the Quillayute River (Figure 6, main text).

A series of 58 transects are extracted from the DEM using HEC-GeoRas (Ackerman, 2009) and written into a geometric data file for input into HEC-RAS. Each river transect extends across the floodplain to the 10 m contour, where applicable. Otherwise, each transect terminates at the highest point landward of the river. Because HEC-RAS computes energy loss at each transect via a frictional loss based on the Manning's equation, Manning's coefficients, an empirically derived coefficient representing resistance of flow through roughness and river sinuosity, are selected for the river channel and the floodbanks. In-channel Manning's coefficients are tuned to calibrate the model's resulting water surface elevations with that of the observed water surface data. Manning's coefficients for the rest of the computational domain (e.g., anything overbank) are estimated using 2011 Land Cover data from the Western Washington Land Cover Change Analysis project (NOAA, 2012) and visual inspection of aerial imagery and range from 0.04 (cleared land with tree stumps) - 0.1 (heavy stands of timber/medium to dense brush). These values are extracted from the HEC-RAS Hydraulic Reference Manual, Table 3-1 (Brunner, 2016). Model

domain boundary conditions are chosen as the water surface elevation at the tide gauge (m; downstream boundary) and river discharge from a combination of records representing the Quillayute River watershed (m^3s^{-1} ; upstream boundary).

1.1 HEC-RAS model validation

In order to determine the dominant inputs to Quillayute River discharge, combined estimates of the Sol Duc and Calawah Rivers are compared to measurements taken on the Quillayute River in May 2010 (Czuba et al., 2010). Combined discharge estimates from the Sol Duc and Calawah rivers underpredict streamflow in the Quillayute River by approximately 33%. An area scaling watershed analysis (Gianfagna et al., 2015), described in the main text, found that the Bogachiel and Calawah Rivers had similar contributions. Thus the Calawah river is scaled by a factor of 2.09 to represent the Bogachiel River. Combined discharge estimates from the Sol Duc River and Bogachiel River, representing the Quillayute River, are also compared to the Quillayute discharge measurements taken during the 2010 survey. Using this methodology, the discharge estimates of the Quillayute River fall within the uncertainty of the discrete USGS measurements in most cases (Table 1).

The longitudinal measured water surface profile allows for the verification and calibration of HEC-RAS modeled water surface elevations on the day of the survey (Figure 1). HEC-RAS is run using discharge of the watershed-scaled Bogachiel River as the upstream boundary condition during the hour of the field survey and this discharge is combined with a lateral inflow from the Sol Duc River around river km 8.5. Manning's coefficients within the main channel of the Quillayute River are calibrated to best represent the water surface elevation on the day of the USGS longitudinal survey. Final Manning's coefficients range from to 0.005 to 0.1, and are on average 0.025.

The final calibrated HEC-RAS model produces a water surface elevation with an average bias less than 1% (less than 1 cm) and an average standard deviation of approximately 5% (7.5 cm). The maximum difference between the two water surfaces is approximately 14 cm (20%). The percent difference between the depth of the observed and modeled water surface is almost always less than 10% (Figure 1).

2 Tide gauge processing

The continuous La Push tide gauge record begins in 2004, recording 12 years of water levels. This record, however, does not capture the extreme water levels occurring during the 1982/83 and 1997/98 El Niños. Therefore, water levels from the La Push tide gauge are merged with water levels from the Toke Point tide gauge (beginning in 1980, NOAA station 9440910) to create a combined water level record representing a larger range of extreme conditions. η_A and η_{SE} , water level components deterministic to the La Push tide gauge, are extended to 1980. Water level components influenced by regional or local forcings like η_{MMSLA} and η_{SS} , are compared before combining. η_{MMSLA} between the Toke Point and La Push tide gauges are similar, so Toke Point η_{MMSLA} are appended to the beginning of the La Push η_{MMSLA} . Toke Point, however, has slightly higher magnitude η_{SS} than La Push and there is a noticeable offset in the highest η_{SS} peaks. A correction is thus applied to the Toke Point η_{SS} before appending it to the beginning of the La Push η_{SS} . η_{MSL} is extended back to 1980 using relative sea level rise trends for the region. Once the two tide gauges are merged, the combined hourly tide gauge record extends from

Table 1. Quillayute River discharge measurements from the USGS survey (Czuba et al., 2010) compared to the Quillayute River discharge estimates computed by adding the Sol Duc USGS gauge measurements with the Bogachiel River discharge, estimated via scaling of the Calawah River gauge measurements. The parenthesis in the last column is the standard deviation of USGS survey measurements ($\text{m}^3 \text{s}^{-1}$).

Date of Survey	Sol Duc ($\text{m}^3 \text{s}^{-1}$)	Calawah ($\text{m}^3 \text{s}^{-1}$)	Bogachiel ($\text{m}^3 \text{s}^{-1}$) (estimated)	Quillayute ($\text{m}^3 \text{s}^{-1}$) (estimated)	Quillayute ($\text{m}^3 \text{s}^{-1}$) (measured)
4/20/2010	52	28	58	110	116 (7)
4/21/2010a	48	25	53	101	108 (1)
4/21/2010b	48	25	52	100	103 (3)
4/21/2010c	46	24	50	96	107 (1)
5/4/2010a	73	69	144	217	220 (5)
5/4/2010b	70	66	137	207	207 (4)
5/5/2010	59	51	107	166	170 (3)
5/6/2010	50	40	84	134	136 (3)

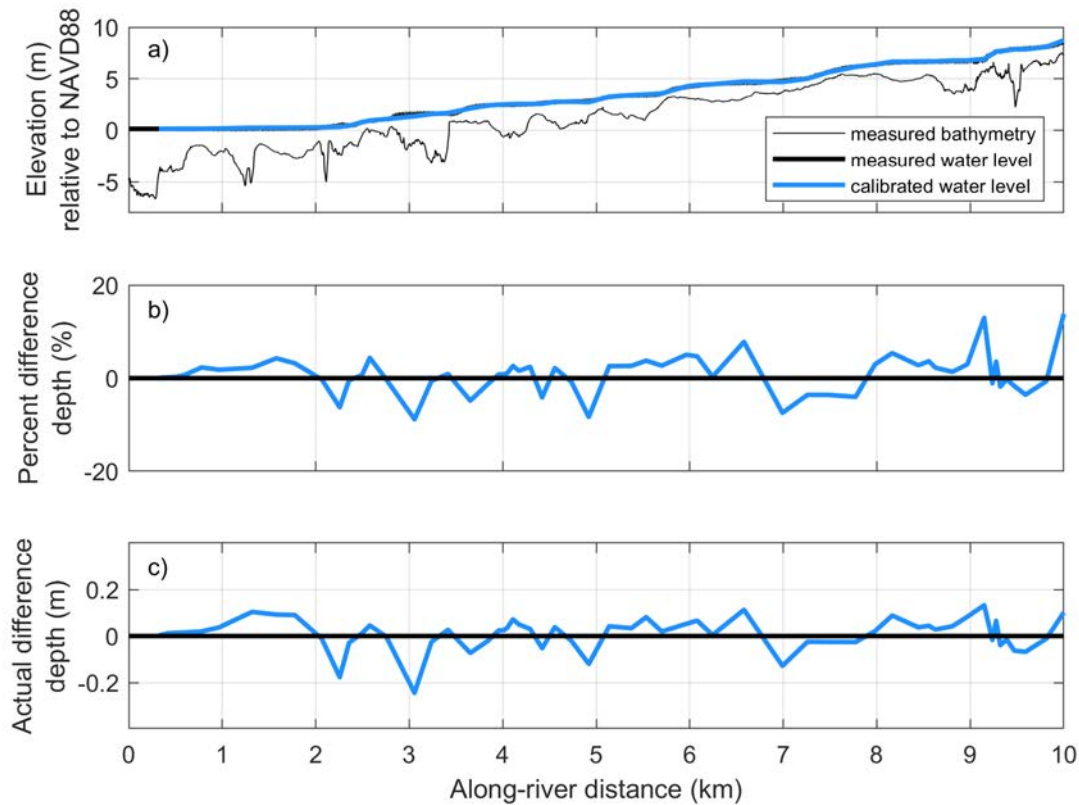


Figure 1. a) Bathymetry and longitudinal profile from the Bogachiel River to the mouth of the Quillayute River surveyed by the USGS in May of 2010 (black). The longitudinal water level for the calibrated HEC-RAS model is depicted in blue. b) Percent difference between the measured (black) and HEC-RAS modeled (blue) water level. c) Actual difference between the measured (black) and HEC-RAS modeled (blue) water level.

1980 - 2016 and is 97% complete. Discharge measurements sampled at 15 minute intervals for the Calawah and Sol Duc rivers are interpolated to hourly increments to match the timing of the SWL measurements.

2.1 Removal of river-influence from the oceanographic signal

Storms tend to influence large stretches of coastline at once, and while site-specific variations in the coastline or distance from storm can drive local variations in the amplitude of η_{SS} , the overall η_{SS} signal is fairly coherent across regional tide gauges across the PNW. The river-influenced water levels are therefore isolated and removed from the La Push η_{SS} record by developing a relationship between the La Push η_{SS} and a regionally-averaged η_{SS} .

η_{SS} decomposed from the Neah Bay, Westport, Astoria, Garibaldi, and South Beach tide gauges are averaged each hour to create a regional η_{SS} record (black line; Figure 2). The standard deviation (σ) of the available η_{SS} records at each hour is used

to represent the variability of η_{SS} due to local effects at each station. η_{SS} at La Push that are larger than the regional average + 2.5σ are considered anomalous to the region, and defined as river-influenced water levels (η_{Ri}). Observations flagged as larger than the regional average + 2.5σ (dashed line; Figure 2) were replaced with the regional average + σ . A value of + σ was chosen to minimize jumps in time series when substituting in a smoother dataset. While this methodology does not remove all the effects of η_{Ri} in the η_{SS} signal, it captures the majority of anomalous water levels driven by high discharge events.

η_{Ri} is produced from the difference between the original La Push η_{SS} and the η_{SS} modified described above which removes η_{SS} anomalous events. η_{Ri} occurring during low discharge events (here low is defined as less than $10 \text{ m}^3\text{s}^{-1}$, the approximate summer average discharge) is added back into the La Push η_{SS} , as it is likely not driven by river forcing. After η_{Ri} was removed from the η_{SS} signal, it is saved as a time series of river-forced water level events.

Extreme Hs and Q events at the Calawah River are determined using the Peak Over Threshold approach, where all independent daily maximum events over a defined threshold are selected. Threshold excesses are fit to non-stationary Generalized Pareto distributions, which include seasonality as a covariate. Both variables are transformed to approximately Fréchet margins. A bivariate logistic model is then used to model the dependency between the variables. To simulate, random numbers are sampled from a uniform distribution and mapped to each variable's prescribed Fréchet cumulative probability distribution function. Based on the probability of occurrence of the transformed value, the estimate is transformed back to the physical scale using the Generalized Pareto distribution if extreme, dependent on the variable's threshold. If not extreme, the estimate is transformed back to the physical scale using monthly-varying Gaussian copulas. This technique generates a synthetic record of Q at the Calawah River gauge that is seasonally varying, related to larger-scale climate variability through wave height (essentially as a proxy for storms), and carries the same dependency between variables as the observational record. Q is then multiplied by 2.09 to represent inflow from both the Bogachiel and Calawah rivers.

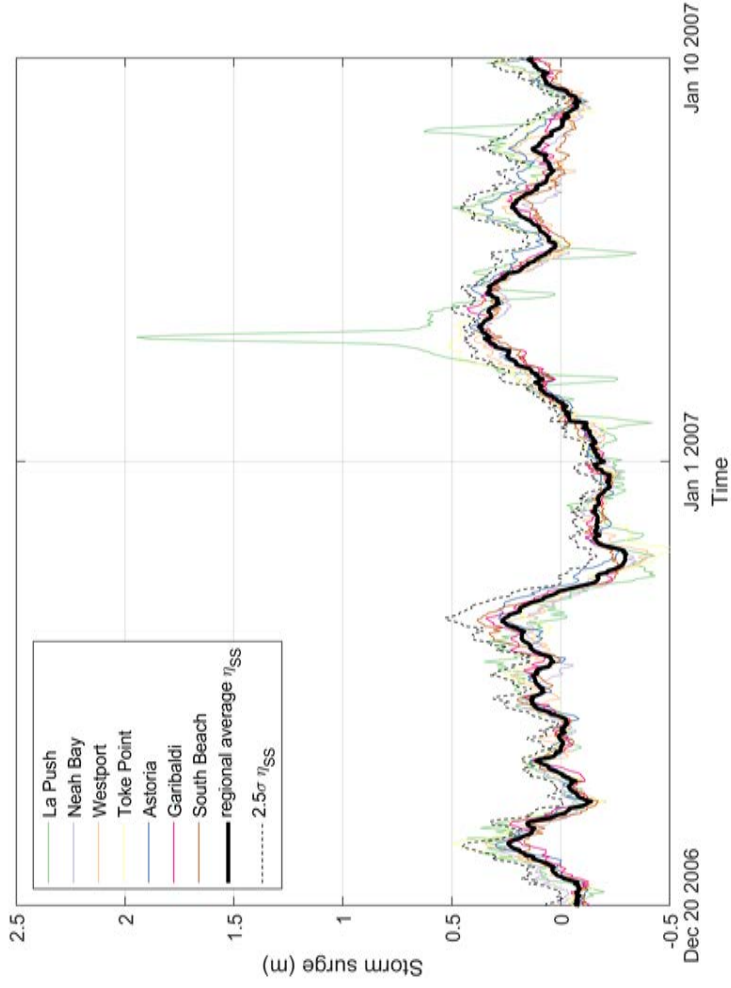


Figure 2. A comparison of storm surge (η_{SS}) decomposed from all tide gauges along the northern Washington to central Oregon coastline. The solid, black line depicts the regional average of all of the η_{SS} signals, while the dashed black line represents the regional average $\eta_{SS} + 2.5*\sigma$ of all η_{SS} in the region. When the La Push η_{SS} exceeds the regional average $\eta_{SS} + 2.5*\sigma$ it is removed from the record and considered river influence.

References

- Ackerman, C. T.: HEC-GeoRAS; GIS Tools for support of HEC-RAS using ArcGIS, Tech. rep., United States Army Corps of Engineers, Davis, 2009.
- Brunner, G. W.: HEC-RAS River Analysis System Hydraulic Reference Manual, Version 5.0, Tech. rep., 2016.
- 5 Czuba, J. A., Barnas, C. R., McKenna, T. E., Justin, G. B., and Payne, K. L.: Bathymetric and streamflow data for the Quillayute, Dickey, and Bogachiel Rivers, Clallam County, Washington, April–May 2010, vol. 537, US Department of the Interior, US Geological Survey, 2010.
- Gianfagna, C. C., Johnson, C. E., Chandler, D. G., and Hofmann, C.: Watershed area ratio accurately predicts daily streamflow in nested catchments in the Catskills, New York, *Journal of Hydrology: Regional Studies*, 4, 583–594, 2015.
- 10 NGDC: U.S. Coastal Relief Model - Northwest Pacific. National Geophysical Data Center, NOAA, <https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/gov.noaa.ngdc.mgg.dem:288/html>, <https://doi.org/doi:10.7289/V5H12ZXJ>, 2003.
- NGDC: La Push, Washington 1/3 arc-second MHW Coastal Digital Elevation Model, <https://data.noaa.gov//metaview/page?xml=NOAA/NESDIS/NGDC/MGG/DEM/iso/xml/247.xml&view=getDataView&header=none>, 2007.
- NOAA: C-CAP Washington 2011-Era Land Cover Metadata, <https://coast.noaa.gov/ccapatlas/>, 2012.
- USACE: 2014 USACE NCMP Topobathy Lidar: Washington, <https://coast.noaa.gov/dataviewer/#/lidar/search/where:ID=6263>, 2014.

Appendix C

The following is a reproduction of the Executive Summary for Chapter 5: Coastal Hazards in “Climate Change Vulnerability Assessment for the Treaty of Olympia Tribes” (Dalton et al., 2016).

Coastal ecosystems and communities **along the outer coast** of the western Olympic Peninsula are currently at risk of erosion and flooding hazards driven by extreme total water levels (TWLs)—mean sea level combined with storm surge, high tide, seasonal and interannual variability in sea level, and ocean wave characteristics. Understanding the magnitude and frequency of these extreme water level events will better prepare coastal communities for dealing with both present day and future coastal hazards which effect critical shoreline habitats and infrastructure. TWLs along the shoreline of the Treaty of Olympia area are modeled for both present day conditions and future conditions under rising sea levels and increasing wave heights.

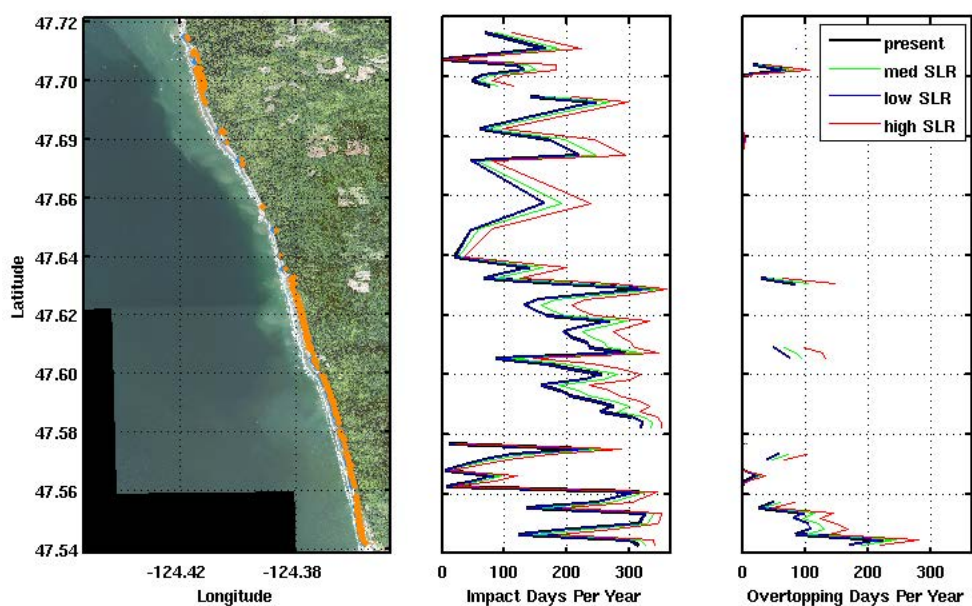
Along the Pacific Northwest coast, sea level is projected to rise by 4-56” by 2100, but local projections vary depending on the local tectonic setting. The central and southern Washington coast is uplifting at a slower rate than the northwest Olympic Peninsula, even subsiding in some areas, and thus is poised to experience the effects of sea level rise sooner than the northwest coast. In this analysis, sea level projections for 2050 range from -0.10 to 0.5 meters (-3.9” to 19.7”) and the observed increasing wave height trend is allowed to continue to mid-century. Two types of high water events are analyzed: nuisance events (everyday hazards characterized by the average amount of days per year that the coastline experiences either overtopping or collision) and extreme events (such as the annual maximum event or the 100-year return level event).

The type of backshore¹ is important for characterizing TWLs and their impacts. By comparing elevations of TWLs to the elevation of the foremost backshore feature, we can estimate the risk of overtopping and inundating the backshore area and the risk of colliding with and eroding the backshore feature be it a bluff, cliff, dune, or engineered structure. The Hoh, Quileute, and Quinault coastlines, and surrounding beaches, consist of highly variable morphology including both low-sloping and steep beaches comprised of either sand or gravel and are backed by dune, bluffs, or cliffs. In general, however, the most characteristic morphology includes steep cliffs and bluffs, occasionally fronted by ephemeral beach berms. The number of days per year that these smaller fronting features experience impact (erosion) and overtopping (flooding) varies along the coastline (Table C1), as do extreme TWL return level events. These ephemeral features most likely act as a buffer to backing cliff or bluff erosion and critical habitat.

¹ The backshore is the region of the beach extending from the high water line to the landward extent of the beach. The most foremost backshore feature is therefore the first feature (dune, cliff, bluff, etc.) in which high water levels may impact.

The beach segments from south of Ozette Lake to Rialto Beach, along the Quileute reservation, and Ruby Beach to the Queets River experience impact (erosion) regime on the leading backshore feature for at least half the year largely due to steep beach slopes and low dune toes. Most other areas along the Treaty of Olympia coastline experience impact days about a third of the year. In general, overtopping occurs much less often than the collision regime, but is most common near Kalaloch and some parts of the Taholah area where it occurs about 100 days per year. Under all sea level rise scenarios water levels increase along the coast, driving increases in number of impact or overtopping days per year (Table C1). The largest increases in impact days per year occur along the beach segment from Ruby Beach to the Queets River (Figure C1) and areas within the Cape Alava to Rialto Beach segment. The overtopping regime remains infrequent, due to the large amount of cliffs and bluffs backing the majority of the coastlines, but increases by a few more weeks per year along the Cape Alava to Rialto Beach segment and the Hoh Reservation coastline.

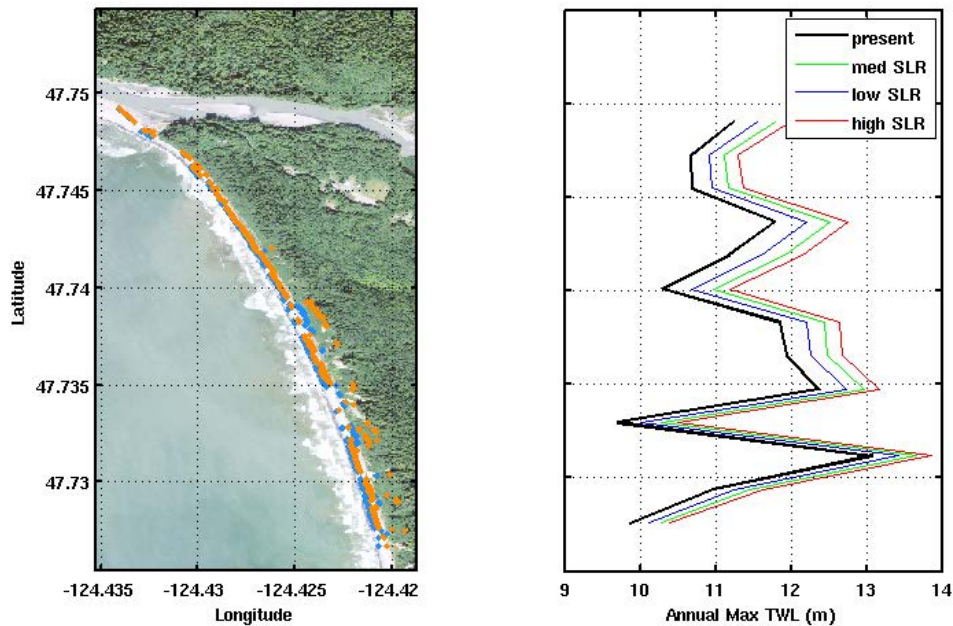
Figure C1 Impact days per year (IDPY; middle) and overtopping days per year (ODPY; right) for Ruby Beach to the Queets River. Bolded black lines indicate the average (across 35 simulations) IDPY/ODPY computed using the present-day simulations while blue, green, and red lines indicate future simulations for low, medium, and high sea level rise, respectively.



The annual maximum TWL event increases under all SLR scenarios for the Hoh, Quileute, and Quinault coastlines, but the Hoh coastline may receive the largest changes (Figure C2) likely due to the overall steeper beaches. Likewise, the steeper beaches from Ozette South may result in larger increases in the annual maximum TWL event compared with First Beach (Table C1). The

southern Quinault coast segment experiences the smallest increase in the annual maximum event. However, even slight increases in water levels may matter more in locations with critical habitat or infrastructure (e.g., the Taholah area) rather than locations where little infrastructure or high backing cliff morphology exists.

Figure C2 The annual maximum TWL for the Hoh reservation. Bolded black lines indicate the average (across 35 simulations) annual maximum TWL computed using the present-day simulations while blue, green, and red lines indicate the average annual maxima of future simulations (12 each scenario) for low, medium, and high sea level rise, respectively.



Certain species with economic and cultural significance to the Treaty of Olympia tribes (e.g., surf smelt, razor clams, shorebirds) may be impacted by the projected future changes in extreme total water levels. While the projected changes in TWLs by 2050 are likely not severe enough to significantly threaten coastal habitat, some intertidal species may shift landward. For example, across all of the locations, the 3 m contour is inundated every day of the year during the maximum daily TWL under a high SLR scenario. The largest amount of change, on average, is in the southern extent of the Quinault area. While intertidal species, like razor clams and surf smelt, may have the ability to move vertically up the beach, snowy plovers and other back-barrier nesting species may face habitat loss as sea levels continue to rise.

Table C1 Percent change (and standard deviation) in impact (erosion) and overtopping (inundation) days per year for present conditions and mid-21st century high sea level rise projections and increase in the average annual maximum TWL event for segments of the Treaty of Olympia coastline. The coastline segment is ranked by the average amount of Impact Days Per Year (e.g., on average, Quileute presently experiences the most IDPY and Southern Quinault experiences the least IDPY).

Coastline Segment	% Change in IDPY	% Change in ODPY	Average Increase in Annual Maximum TWL Event
Quileute (Rialto Beach and First Beach)	+18% ($\pm 11\%$)	+35% ($\pm 31\%$)	+50 cm
Kalaloch (Ruby Beach to the Queets River)	+30% ($\pm 25\%$)	+55% ($\pm 34\%$)	+50 cm
Ozette (Cape Alava to Rialto Beach)	+25% ($\pm 36\%$)	+50% ($\pm 45\%$)	+60 cm
North Quinault (Northern Taholah to Queets River)	+18% ($\pm 15\%$)	+35% ($\pm 60\%$)	+50 cm
Middle Quinault (Point Grenville to Northern Taholah)	+17% ($\pm 40\%$)	+25% ($\pm 55\%$)	+50 cm
Second Beach	+54% ($\pm 10\%$)	+90% ¹	+43 cm
Hoh	+16% ($\pm 13\%$)	+27% ($\pm 12\%$)	+77 cm
Third Beach	+47% ($\pm 6\%$)	+43% ¹	+63 cm
Southern Quinault (Moclips to Point Grenville)	+65% ($\pm 33\%$)	+95% ($\pm 30\%$)	+30 cm

¹ Only one point was included in the analysis, so no standard deviation could be calculated.