ABSTRACT

Title of thesis: INFLUENCES OF WAVE CLIMATE AND
SEA LEVEL ON SHORELINE EROSION
RATES IN THE MARYLAND CHESAPEAKE
BAY

Jia Gao, Master of Science, 2015

Thesis directed by: Professor Lawrence P. Sanford, University of
Maryland Center for Environmental Science
(UMCES)

SWAN and a parametric wave model implemented by the Chesapeake Bay Program

(CBP) were used to simulate wave climate from 1985 to 2005 in Chesapeake Bay (CB).

Calibrated sea level simulations from the CBP hydrodynamic model were acquired.

Spatial patterns of sea levels during high wave events were dominated by local north-

south winds in the upper Bay and by remote coastal forcing in the lower Bay. A dataset

comprising shoreline erosion rates and related characteristics was combined with the

wave and sea-level climates to explore the most influential factors affecting erosion. The

results show that wave power is the most significant factor for erosion in the Maryland

CB. Marsh shorelines present a nearly linear relationship between wave power and

erosion rates, whereas bank shorelines are less clear. The results of this study are

applicable at large scales. A more comprehensive data set is needed for building detailed

local predictive relationships.

INFLUENCES OF WAVE CLIMATE AND SEA LEVEL ON SHORELINE EROSION
RATES IN THE MARYLAND CHESAPEAKE BAY

By

Jia Gao

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2015

Advisory Committee:
Professor Lawrence P. Sanford, Chair
Professor William C. Boicourt, Co-Advisor
Associate Professor Cindy M. Palinkas

Dedication

To my parents, my boyfriend and my third aunt

Acknowledgements

My heartfelt thanks go to Professor Lawrence Sanford for his patience, support and guidance, which I am indebted to. It has been a great pleasure and wonderful experience working with him for the past three years. My sincere thanks go to Professor William Boicourt for his support and guidance through my graduate school. I am grateful of all my committee members for their time and help with this thesis. My thanks go to Dr. Dong Liang for his generous help with statistics.

I would like to give special thanks to my boyfriend Ian, who has been offering me love, technology and moral support. Thanks go to my dear friends, Alex Fries and Jacqueline Tay, who have been helping me with writing and statistics.

Thanks Horns Point Laboratory, Bay and River Fellowship and Maryland Sea Grant for their funding support through my graduate study.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Erosion is the process of the wearing away and removal of land by external forces. Shoreline erosion of banks or marshes can be described as iterations of a process in which waves undercut the cliff/marsh base, the cliff/marsh collapses, and then waves resuspend sediments at the cliff/marsh base. Finally, currents remove these materials. There are two components of shoreline erosion: fastland erosion, which occurs above the waterline; and nearshore erosion, which operates from the waterline to the base of wave action at water depths up to about 2.4 m in Chesapeake Bay (CB). The term 'erosion rate' used in this study refers primarily to fastland erosion rate, although both components must occur in tandem.

Erosion can lead to nutrient pollution, ecosystem degradation, and economic loss (U.S. Army Corp of Engineers and Maryland Department of Natural Resources, 2010; Leatherman et al., 1995). Erosion processes have been intensified by sea level rise, land subsidence, and increasing rates of shoreline development (Halka et al., 2005). Erosion is a complex process to study, not only because it involves various interacting factors but also because these factors can behave very differently in different geographic locations. Previous studies have explored relationships between erosion rates and potentially influential variables, such as wave energy, shoreline type, bank height, tidal range and sea level (Sunamura, 1992; Spoeri, 1985). The relationship between sea level and erosion on sandy shorelines has been described as a response of the equilibrium shoreline profile with wave activity as the hidden cause (Bruun, 1962; Schwartz, 1967; Dean, 1991). Several erosion models have been built for sandy beaches, such as Stormed-Induced

Beach Erosion (SEARCH) and Generalized Model for Simulating Shoreline Change (GENESIS).

Maryland CB shorelines consist of banks, with height ranges from 1 meter to over 30 m (such as Calvert Cliffs), and marshes, which are mostly located along the lower Eastern Shore of Maryland (Somerset, Wicomico, and Dorchester Counties). Year-round beaches only make up about 24 km of the entire CB shoreline (U.S. Army Corps of Engineers, 1971). Some regions, especially points and islands, are experiencing severe erosion (>2.4 m/year) on Western Shore Maryland (Pt. Lookout to St. Jerome, Holand Pt. and Thomas Pt.) and Eastern Shore Maryland (Kent Island, Lowes Pt. to Knapps, Mills Pt. to Hills Pt., James Island, Oyster Cove to Punch Island Creek and Barren Island)(Wang et al. 1982). There were 1.9×10^8 m^2 of land loss during 1850~1950 along Chesapeake shorelines (Slaughter 1967a).

Chesapeake Bay is the largest estuary in the United States with a shallow average depth of 8.5 m. Its length is about 315 km from the Susquehanna River to its outlet to the Atlantic Ocean and its width ranges from 5.6 km to 56 km (Langland and Cronin, 2003). The Bay's narrow dendritic geometry consists of about 18803 km of shoreline (Chesapeake Bay Program website http://www.chesapeakebay.net/discover/bay101/facts). CB receives ocean swell, (which seldom reaches mid-Bay), at its entrance between the Virginia Capes. The approximately north-south orientation of CB provides sufficient fetch for surface gravity waves, which dominate upper- and mid-Bay (Boon et al., 1996, Lin et al., 2002). Typical wave periods are about 3s and significant wave heights (Hsig) are less than 2m (Boon, 1998;Lin et al.,1998). Waves are typically fetch-limited and wind generated in CB, but they are still an important forcing for sediment transport

(Sanford, 1994; Boon et al., 1996) and shoreline erosion (US Army Corps of Engineers, 1990).

Shoreline elevation and orientation, shoreline type (vegetated, protected, bare, etc.), sediment type and availability, nearshore morphology, land subsidence, sea level rise, hydrodynamic and wave characteristics, and human activity can all be potential factors for shoreline erosion. Previous studies indicate that wave climate can be the most significant factor that influences erosion process in many places including CB (Skunda, 2000; Amin, 1997; Wang, 1982; Spoeri, 1985; Perry 2008; Kamphuis, 1987; Schwimmer, 2001).

In this study, we create a data set that contains a long-term wave climate (1985-2005) for the entire CB using Simulating Wave Nearshore (SWAN) and Chesapeake Bay Program (CBP) wave models, along with sea level data, which are simulated by CBP Hydrodynamic model developed by the U.S. Army Corp Engineers (USACE). All three models share the same grids, time duration and input wind fields. Thus, we can investigate the distributions and relationships between sea levels and wave height individually and jointly along CB shorelines. Sea levels during high-wave events are also examined for the purpose of investigating if the combination of high wave height at high or low sea level can influence erosion rates differently. Another derived dataset includes shoreline erosion rates, shoreline structure information, bank/marsh ratio, and mean bank height in 207 reaches in the Maryland portion of CB as assembled by Maryland Geological Survey (MGS). Also, a high resolution erosion database is accessible from the Coastal Atlas available at the website of Maryland Department of Natural Resources (DNR). We incorporate these data sets for implementing statistical analysis, such as

3

linear analysis, curve fitting, Generalized Additive Model (GAM), and Neural Network (NN). To date, this is the most comprehensive and high-resolution dataset used for analyzing relationships between erosion rates and wave climate, along with other shoreline characteristics, in Maryland CB. This study aims to gain a better understanding of different roles of controlling variables on erosion rates and attempts to establish a simple semi-empirical and statistical relationship between erosion rates and controlling variables, which could potentially be used by coastal managers.

Chapter 2 Wave climate and sea level in CB from 1985 to 2005

2.1 Introduction

Chesapeake Bay (CB) is the largest estuary in the United States. It has a relatively shallow average depth of 8.5 m, length of 315 km from the Susquehanna River to its outlet to the Atlantic Ocean, and width that ranges from 5.6 km to 56 km (Langland and Cronin, 2003). Such a narrow dendritic geometry gives CB 18803 km of shorelines (Chesapeake Bay Program website http://www.chesapeakebay.net/discover/bay101/facts). The axis of the Bay is mainly north-south orientated, though it is northeast-southwest in upper Bay, north-south in mid-Bay and northwest-southeast in the lower Bay. Generally, the mean tidal range is lowest near Annapolis (~0.3 m) and increases toward both ends to slightly less than 1m (Zhong and Ming, 2006). CB is a partially mixed estuary and has a classic two-layer estuarine circulation, with upper-layer fresh water going downstream and lower-layer oceanic water going upstream (Pritchard, 1956). The Susquehanna River provides about 60% of the freshwater to the Bay, with the rest from five major western tributaries.

Chesapeake Bay receives ocean swell, which seldom reaches mid-Bay, from between the Virginia Capes. The approximately north-south orientation of CB provides sufficient fetch for generation of wind-forced surface gravity waves, which dominate the upper- and mid-Bay region (Boon et al., 1996; Lin et al., 2002). For these waves, the typical period is about 3s and significant wave height (Hsig) is less than 2 m (Boon, 1998; Lin et al., 1998). Waves are likely the dominant forcing for sediment transport in shallow nearshore CB waters (Sanford, 1994; Boon et al., 1996) and for shoreline erosion (US Army Corps of Engineers, 1990). There were 1.9×10^8 m^2 of land loss in the century from

1850 to1950 along Chesapeake shoreline due to erosion (Slaughter, 1967a). Extra loading of sediments can lead to nutrient pollution, ecosystem degradation and economic loss. This process has been intensified by sea level rise, land subsidence and increasing rates of shoreline development (Halka et al., 2005).

Although there have been previous modeling studies of wind-waves in CB (e.g., Lin et al. 2002), these studies have either been short in duration or their predictions have not been thoroughly investigated. The motivation of this study is to build a long-term wave-climate data set for supporting shoreline erosion studies and to improve the understanding of the joint probability distribution between sea level and wave climate in CB. In this study, the SWAN and Chesapeake Bay Program (CBP) wave models were implemented in CB for long-term wave climate simulations from 1985 to 2005, and wave variability was compared to separately predicted sea level variability.

2.2 SWAN and CBP wave models

2.2.1 SWAN model

SWAN is a third generation wave model that is based on Eulerian formulation of the discrete spectral balance of wave action density. This model simulates random, short crested waves in coastal regions over arbitrary bathymetry, wind and current fields. SWAN is driven by boundary conditions and local winds. It accounts for triad wave-wave interaction, shoaling, refraction, white capping, bottom friction, depth-induced breaking, dissipation and diffraction. (Booij et al., 1999; Ris et al., 1999).

The version 40.91 of SWAN with 2D and third generation mode is used in this test. The configuration of the runs is as follows: Cartesian coordinates and curvilinear

grids are applied; wave period cutoff limits are 0.001 s and 25 s; the peak period is used

as characteristic wave period; the wave-growth term from Cavaleri and Malanotte-Rizzoli

(1981) and the surface drag coefficient from Wu (1982) are used; and the default

JONSWAP coefficient of 0.067 is adopted for bottom friction (Hasselmann et al., 1973).

The calculation time step is 10 minutes, a significant wave height of 0 m and peak period

of 0.1 s are used as ocean boundary conditions, and zero wave height and wave period

everywhere is used as the initial condition. Activated physical processes include white

capping, nonlinear quadruplet wave interactions, and triad wave-wave interactions.

2.2.2 CBP model

Young and Verhagen (1996) developed a semi-empirical method for calculating

fetch-limited wave growth that calculates wave heights and periods from water depths,

wind inputs and fetch. Noting the deep water asymptotic limits: JONSWAP relationship

(Hasselmann et al., 1973) and relationship of frequency and fetch (Kahma and Calkoen,

1992), Young and Verhagen (1996) proposed a fetch-limited and depth-limited shallow-

water waves relationship, of which the empirical parameters are based on measurements

at Lake George, which is approximately 20 km long and 10 km wide and has a relatively

uniform bathymetry of 2 m deep (Young and Verhagen, 1996). The expression of the

relationship is as follows:

$$\varepsilon = 3.64 \times 10^{-3} \left\{ \tanh A_1 \left[\frac{B_1}{\tanh A_1} \right] \right\}^n \tag{2.1}$$

$$\nu = 0.133 \left\{ \tanh A_2 \left[\frac{B_2}{\tanh A_2} \right] \right\}^m \tag{2.2}$$

$$A_1 = 0.292^{\frac{1}{n}} \delta^{\frac{1.3}{n}}; \quad B_1 = (4.396 \times 10^{-5})^{\frac{1}{n}} \chi^{\frac{1}{n}}; \quad A_2 = 1.505^{\frac{1}{m}} \delta^{\frac{3}{8m}}; \quad B_2 = 16.391^{\frac{1}{m}} \chi^{\frac{-0.27}{m}};$$

where ε is the non-dimensional wave energy; v is the non-dimensional wave frequency; χ is the non-dimensional fetch; δ is the non-dimensional water depth. A_1, B_1, A_2, B_2, m and n are all empirical parameters .Young and Verhagen (1996) calculated n = 1.74 and m = -0.37. From its expression, we can see that wave height and period can be calculated from fetch and bathymetry at each grid point. Solving this expression requires neither boundaries nor initial conditions, because it assumes instantaneous steady state at each point in time.

Following this method, the Engineer Research and Development Center (ERDC) of the U.S. Army Corp Engineers (USACE) developed a simple parametric wind-wave model (Harris et al., 2012). The most complex calculation in this model is determining smoothed effective overwater fetch at each grid point for each time step, which is accomplished by allowing for cosine-weighted spreading of the wind direction to avoid sudden changes in fetch due to slight changes in wind direction. We refer to this wave model as the CBP wave model for this study.

2.2.3 Model configuration

Hourly outputs of sea level for entire CB from 1985 to 2005 were available from the Waterways Experiment Station (CH3D-WES) model (Johnson et al., 1993), referred to as the CBP hydrodynamic model in this study. Sea level is among the best calibrated outputs of the hydrodynamic models in CB (Johnson et al., 1993; Cerco and Noel, 2004) SWAN is shown to be valid for wave simulations in CB (Lin et al., 2002). Wave climate from CBP wave model were reasonable estimates of observations and were used for

calculation of bed shear stress in the CBP Water Quality and Sediment Transport Model (Harris et al., 2012).

In order to be able to incorporate wave climate with sea level data, SWAN and CBP wave models were built with the same wind field, model grids, and bathymetry as the CBP hydrodynamic model, and were run over the same 21-year calibration period as that model. Input and output data are both hourly. The Bay model grid is curvilinear (178 × 282), with approximately 1-km resolution in the axial direction and 400-m resolution in the lateral direction throughout CB and its tributaries (Figure 2.1). The bathymetry and grids used in this study is from the current version of the CBP model (Harris et al., 2012), with 1.5 m as the shallowest shoreline model grid depth (Figure 2.2).

Wind input is interpolated from five wind stations to the Bay model grid: Thomas Point, Patuxent Air Base, Norfolk International Airport, Washington National Airport, and Richmond International Airport. Wind inputs of the latter four over-land stations are corrected by multiplying by 1.5 prior to interpolation to better match data at the over-water Thomas Point station (Johnson et al., 1993; Harris et al. 2012). This procedure approximately compensates for greater attenuation of overland wind velocity in the terrestrial atmospheric boundary layer, as compared to the marine atmospheric boundary layer (Goodrich, 1985; Xu, 2002).

2.2.4 Model validation

The SWAN Model was tested with 3 min, 10min, and 1 hour computational time-step simulation scenarios. From 1hr to 10 min, there is a significant improvement in model outputs, while from 10min to 3min the improvement is not obvious (Figure 2.3).

Time varying sea level was also tested as an input variable to the models, but the outputs of both SWAN and CBP wave models were almost the same with or without time varying sea level. For purposes of simplicity and to keep computational costs down, a computational time step of 10 minutes was used and time varying sea level was not included in the wave model runs.

We have observational data (Lin et al., 2002) from wave gauges that were deployed during 10-23 October 1995 at Calvert Cliffs and from 26 October 26 to 9 November 1995 at Poplar Island. The comparisons (Figure 2.4) show that both SWAN and the CBP wave model work reasonably well. Predictions at Poplar Island show a better match than at Calvert Cliffs in terms of both significant wave height (Hsig) and peak period. At Calvert Cliffs, predictions of Hsig match better with observational data than those of peak period from either model. Shown visually in Figure 2.4, SWAN is slightly better than the CBP wave model for both Hsig and peak period. Also, SWAN is able to provide direct estimates of quantities other than Hsig and peak period, such as maximum bottom orbital velocity, transport of wave energy, and steepness of waves. These variables are all known as important factors for shoreline erosion. Thus, we have chosen to use the SWAN output for all model data except for fetch, which is only calculated by the CBP wave model.

2.3 Results and discussion

In this section, using outputs from the wave and hydrodynamic models, we study the distribution of significant wave height and sea level statistically along the entire shoreline of CB, including the major tributaries. We quantify the joint probability between sea level and significant wave height (Hsig) using an empirical function (Section

2.3.1). We also discuss the spatial distribution of instantaneous sea level corresponding with high wave events in a statistical sense when dominated by different forcing (Section2.3.2). Finally, we examine time series of wave heights and sea level from three pairs of sites (lower, middle and upper Bay) during high-wave events (Section 2.3.3).

2.3.1 Joint probability distribution between sea level and wave height

The range of average Hisg from 1985 to 2005 varies between 0-0.3 m around CB (Figure 2.6). The smallest waves occur in tributaries, especially towards the head of tributaries, as would be expected due to limited fetch. Larger waves are concentrated in the mainstem regions, with highest wave heights located near the mouth of the Bay on both its western shore (Cape Henry) and eastern shore (Cape Charles) (Figure 2.6).

In order to study the joint probability of distribution between sea level and wave height, we explored the distribution of each variable separately first. A histogram of significant wave heights from the hourly SWAN output for the entire shoreline (2217 grids) of CB from 1985 to 2005 (Figure 2.5a) shows that 91% of all wave heights are located in the lowest bin of 0-0.27 m and that the amount of data falling into each bin decreases exponentially as Hsig increases. The dominance of Hsig within this low range may be partially due to the fact that we are only considering wind seas in this analysis. The probability of observing a particular Hsig (P(Hsig)) can be quantitatively modeled by an exponential function (Equation 2.3), with $R^2 \approx 0.99$ and SSE=1.6×10^{-3}.

$$P\left(\mathrm{Hsig}\right) = 0.25e^{-9.6Hsig} \tag{2.3}$$

A histogram of detrended sea level from hourly outputs of the CBP hydrodynamic model for the entire shoreline of CB from 1985 to 2005 (Figure 2.5b) shows that sea level is

close to normally distributed. Note that all sea-level data were locally detrended before calculation of this distribution. The probability distribution of sea level can be quantitatively described by a Gaussian function (Equation 2.4) with $R^2 = 0.99$ and SSE=1.9×10^{-5}, where w is detrended sea level. The red line in Figure 2.5b describes the best fitted normal distribution for the corresponding data presented by the histogram.

$$P(w) = 6.8 \times 10^{-2} e^{-(\frac{w - 2.9 \times 10^{-4}}{0.43})^2} \qquad (2.4)$$

Relationships between sea level and significant wave height can be further explored by examining the distribution of sea-level data within each bin of Hsig from Fig. 2.5a. Because most Hsig data fell into the first bin, bin sizes were expanded thereafter (0.27m-1.06m, 1.06m-1.86m, 1.86m-2.65m) to reduce the number of plots in Fig. 2.7. In Figure 2.7 a and b, all waves are relatively small, and sea level follows a normal distribution centered at zero (mean sea level). In Figure 2.7 c, the histogram of sea level still approximates a normal distribution but it is skewed to the right indicating a prevalence of high sea levels with high waves. There are not enough data points in the last plot (Hsig 1.86-2.65 m) to fit a normal distribution, but the center of the sea-level distribution is even more strongly skewed to the right. In other words, sea level during the highest wave events tends to be significantly elevated. These patterns are also evident in Figure 2.8a, which shows that most data are concentrated at low Hsig and centered at zero sea level; the joint probability (see below) decreases exponentially as Hsig increases and the most probably sea level increases as Hsig increases.

If Hsig and sea level are independent of each other, the joint probability should be exactly the product of P(w) (eq. 2.4) and P(Hsig) (eq. 2.3). However, if there are

correlations between Hsig and sea level, then a more complex joint probability distribution is expected. Assuming that this joint probability function would still have the form of the product of an exponential and a Gaussian function, and utilizing the Curve Fitting Tool in Matlab yields

$$P(w, \text{Hsig}) = 0.014e^{-7.04\,Hsig - 5.87 \times (w - 0.005)^2} \qquad (2.5)$$

Where Hsig is the significant wave height and w is the detrended sea level. Fitting a line between joint probability and the results from curve fitting shows that these two values match well when the joint probability is high (>0.005), which corresponds to low Hsig around mean sea level (Figure 2.8b). However, curve fit estimate of the joint probability deviates significantly from the 1:1 ratio line, with a higher probability of overestimation (Figure 2.8d) when the joint probability is low (<5×10^{-4}), which corresponds to high Hsig and extreme sea levels. Fitting the joint probability against P(w)×P(Hsig) tells a similar story except at high probability range, P(w)×P(Hsig) tends to underestimate (Figure 2.8c) and at low probability range, P(w)×P(Hsig) shows some scatter about the 1:1 line. Thus, the assumption of independence between the probabilities of sea level and Hsig is not as good for the most probable low waves, but is actually reasonable for less probable high waves.

Next, the spatial distributions of Hsig and sea level within different ranges of Hsig, especially in the high range, are discussed. Figure 2.9 shows that small waves occur nearly everywhere along the CB shoreline (Figure 2.9a), which also means that sea levels corresponding to low Hsig (see Figure 2.7a) are mostly evenly distributed along the whole CB shoreline. Most waves in the moderate Hsig range of 0.26 m-1.06 m are

located in the mainstem of the Bay and decrease toward the head of tributaries, where waves of these heights rarely occur. This indicates that most waves in the tributaries are smaller than 0.26 m high. Sea levels associated with medium high (1.06m-1.86m) waves are mostly located in the main stem of the Bay or near the mouth of tributaries. The highest values of Hsig (1.86m-2.65m) also occur near the mouth of the Bay and between the mouths of the major western shore tributaries, with the very highest waves occurring around Cape Henry (south side of the mouth; Figures 2.9c and d). It thus appears that the highest sea levels corresponding with highest Hsig occur primarily at the mouth of the Bay near Cape Henry.

Figure 2.10 demonstrates the count of occurrence of Hsig corresponding to the median of detrended sea levels in the ranges of Hsig: (a) 0-1.06 m and (b) 1.06-2.66 m. Medians of sea level are close to zero at all shoreline grids when Hsig is low to moderate (0-1.06 m). This means that during relatively low-wave events, the chances of having positive or negative sea level along the whole shoreline are statistically even. This phenomenon also agrees with Figure 2.9a and b and Figure 2.7a and b. Medians of sea level in the high range of Hsig (1.06-2.66 m) are mostly positive along the west side of the Bay with most of the negative points on the east side of the Bay. Interestingly, medians around Cape Charles are mostly negative, which means the positive sea levels in Figure 2.7c and d are mostly contributed by grid points around Cape Henry.

2.3.2 Spatial distribution of sea level during high wave events

The barotropically induced sea-level variability in a partially mixed estuary like CB can be influenced by tides, wind forcing, river runoff and forced damped seiche responses (Chuang and Boicourt, 1989). Tides in CB are forced by ocean tides at the

mouth, which usually have a semi-diurnal period that drives instantaneous sea level above/below the mean. Wang et. al (1978, 1979) studied sub-tidal sea-level variability and its relation to wind forcing. They found that the dominant sub-tidal sea-level fluctuation happened at a period >10 days induced by coastal sea-level fluctuations, which are driven by along-coast persistent winds. Water is driven out of the Bay by coastal set-down during northeastward winds and into the Bay by coastal set-up during southwestward winds, due to the coastal Ekman transport. The amplitude of this sub-tidal mode decreases from the mouth to the head of CB. For periods of fluctuation between 4 and 10 days, the fluctuation is dominated by eastward winds pumping water out of the Bay, while westward winds induce an inflow into the Bay. This east-west mechanism is more obvious in summer and fall than in winter and spring. For periods of fluctuation <4 days, local northward wind forcing induces an inflow, while local southward wind forcing drives water out of the Bay. Also, Chuang and Boicourt (1989) found that a lateral wind (east-west) can generate responses near the resonant frequency of CB. However, the local north-south wind can more easily induce a 2~3 day period seiche, which is characterized by a node at the mouth and an antinode at the head of CB. Thus, the fluctuations in sea level in the Bay we observe, if we only consider barotropic factors, are a combined action of astronomical tides, local longitudinal (north-south) and lateral (east-west) wind, remote coastal sea level, and a seiche response.

Another view of the spatial distribution of median sea level during high-wave events (defined here as wave heights above the 60% of maximal Hsig) is shown in Figure 2.11. For positive medians of sea level, a value of 1 is assigned; -1 is assigned for negative medians of sea level. If counts of positive and negative sea level are

approximately the same (the difference in counts is smaller than 1% of the total number of data), 0 is assigned. On this map, positive medians of sea level are mostly located on the western shore in the mainstem Bay, the southern side of tributaries in the lower Bay, and the northern side of tributaries in the upper Bay, while negative medians have generally the opposite pattern.

Elliott (1978) and Wang et al. (1978) noted that half of the sea-level fluctuations in the Potomac River originate from the sea-level changes in the Bay; the other half is generated by local wind. In smaller tributaries, local forcing within tributaries should be less effective than in bigger tributaries like the Potomac River. Thus, it is assumed that sea level in small tributaries is mostly influenced by sea level in the Bay in our discussion. We also assume that when strong wind events occur, most areas in the Bay share a similar wind pattern.

When north winds prevail, especially northeast winds, sea level in the Bay may rise due to coastal sea-level setup (remote forcing) through Ekman transport. However, the local effect, which directly blows water out of the Bay, causes a sea-level set-down. In contrast, south winds can lower sea level from the influence of remote forcing, but can also generate sea level set-up through local effects (Wang, 1979).

During high-wave events (wave heights above the 60% of maximal Hsig at each grid point), medians of sea level on the western shore of the main stem, the eastern side of islands, and the southern side of tributaries in the lower Bay are positive (Figure 2.11). This means that northeasters cause these high wave events and the effects of remote forcing on sea-level fluctuations dominate over local forcing in the lower CB. During

high wave events in upper CB, positive medians of sea level occur on the western shore, at the head of CB, and on the north sides of tributaries; negative medians occur on the south sides of tributaries. Negative medians on southern sides of tributaries show that north winds lower sea level in the upper Bay instead of causing sea-level increase. This indicates that local effects overcome remote forcing in the upper Bay. Moreover, the positive medians on northern side of tributaries and at the head of CB indicate that strong south winds cause a rise in sea level in the upper Bay. So we can speculate that positive high wave sea levels in the upper mainstem western shore are caused by south winds instead of northeasters. Thus, the effects of local forcing on sea-level fluctuations dominate over remote forcing in upper CB.

Even though most shoreline points agree with the former statements, there are exceptions at complicated curved shorelines, and in tributaries, especially the heads of tributaries. Note that defining the geographic boundaries for different dominance of remote forcing and local forcing is beyond the scope of this study. Due to complicated shoreline orientations, freshwater input from land, and other site-specific factors, detailed small-scale, site-specific analysis requires finer resolution and a more comprehensive local datasets.

2.3.3 Specific examples of paired sites

In this section, sites on both eastern and western Shore in the lower, middle, and upper Bay are examined during high-wind events. These sites are all located in regions that are only weakly influenced by local river runoff. Two points at roughly the same latitude should have a similar wind field due to our method of interpolation. So, only one wind vector time series for each pair of sites is shown in the plots.

Grid 14 and 2074 are on opposite sides of the mouth of the Bay (see Figure 2.11 for locations). Wind coming from the north gives grid 14 a longer fetch, while wind coming from the south yields a longer fetch for grid 2074. Thus, Figure 2.12 shows that north winds increase Hsig dramatically at grid 14 but not grid 2074, while south winds result in higher Hsig at grid 2074 than at grid 14. Correspondingly, north winds result in sea levels above the mean, while south winds result in sea levels below the mean. These patterns result from coastal Ekman transport, as discussed previously. Figure 2.12 shows the dominance of Ekman transport driven by north-south winds along the coast (i.e., remote forcing) on sea level fluctuation over local forcing at the mouth of CB. Thus, north wind causes high (and mostly positive) sea level and high Hsig at grid 14, while south wind causes low (and mostly negative) sea level and high Hsig at grid 2074. Grid points on the lower western shore should follow a similar pattern as grid 14, while those on lower eastern shore should be similar to grid 2074.

At the head of CB, the responses of Hsig for grid 1353 and grid 1222 are very similar to each other except during strong northwest winds, due to differences in the fetch for each grid. Grid 1222 is sheltered when wind comes from the northwest, which leads to a small fetch. Regardless of the strength of northwest wind, waves cannot grow due to the lack of fetch. Both grids have high Hsig when strong south winds blow, since they have a similar fetch length in this direction. Thus, both strong northwest wind and wind with south component lead to a high Hsig at grid 1353, while only the latter result in a high Hsig at grid 1222 (Figure 2.13).

Increasing sea level corresponds to winds with a strong south component, while decreasing sea level corresponds to winds with a north component in Figure 2.13. This is

exactly the opposite pattern compared to two grids (14 and 2074) at the mouth of the Bay, where the effects of remote forcing on sea level fluctuation dominates. The response here follows the local forcing mechanism: north winds should result in a volume flux toward the mouth of CB, while south winds result in surface-water transport toward the head of the Bay. Figure 2.13 shows that during significant local wind event, sea level setup/setdown induced by local north-south wind can dominate over remote forcing at the head of CB. Thus, strong northwest winds might cause a decrease in sea level and high Hsig at grid 1353, while winds with a strong south component cause an increase in sea level and high Hsig at both grids.

At grid 1222, high Hsig corresponds to increased (mostly positive) sea level due to strong local south winds, and so the median of sea level at high Hsig should be positive (see Figure 2.11 for locations). On the contrary, high Hsig can correspond to both an increase (mostly positive) and decrease in sea level (more negative) at grid 1353. High Hsig corresponding to sea level increase happens when local winds with a strong south component occur; the high Hsig corresponding to sea level decrease happens with strong local northwest winds. Figure 2.11 shows that the median of sea level at high Hsig is negative at this grid, which means a low sea level with northwest wind has a higher frequency of occurrence during high-wave events. Grid points on the upper western shore should follow a similar pattern as grid 1222, while those on the upper eastern shore should be similar to grid 1353 in Figure 2.11.

In Mid-Bay, grid 1015 (Calvert Cliffs) and 1708 are located on the west and east sides of the Bay respectively, where the local Bay axis is oriented northwest-southeast (see Figure 2.11 for locations). Because of this change in orientation, strong wind coming

from south only generate high Hsig at grid 1708, but there is not sufficient fetch for high waves at Calvert Cliffs. Meanwhile, northwest winds can result in high Hsig at both grids, but northeast wind creates larger Hsig at Calvert Cliffs (Figure 2.14). South winds cause an increase in sea level while north winds cause a decrease, indicating that sea-level fluctuations caused by local longitudinal winds can still overcome the remote forcing generated from the mouth of the Bay at this latitude (Figure 2.14. High Hsig corresponds to low (mostly negative) sea level at Calvert Cliffs due to strong northeast and northwest winds, resulting in a negative median of sea level at high Hsig. However, at grid 1708, high Hsig corresponds to both positive and negative sea level is due to south or northwest winds, respectively. The negative median of sea level at high wave height means that there is a higher frequency of northwest winds than south winds during high-wave events at grid 1708. Grid points on the mid-eastern shore should be similar to grid 1708. However, only grids close to or north of Calvert Cliff (grid 1015) on the mid-western shore should follow a similar pattern. Most grids points south of Calvert Cliff should follow the pattern of grid 1222 due to the change of shoreline orientation from northeast-southwest to northwest-southeast.

During high-wave events, even though the local forcing dominates over remote forcing in the upper Bay and remote forcing dominates over local forcing in lower Bay, the behavior during any single event is not predictable due to the combined action of local and remote forcing and dynamic seiche response.

2.4 Conclusions

Considering only locally generated wind waves along the shorelines of CB (no ocean swell), the frequency of occurrence of Hsig decreases exponentially with increasing Hsig. The frequency of occurrence of sea level follows a normal distribution.

Along the entire CB shoreline, only considering wind seas, relatively large waves (Hsig>0.27 m) are mostly found along the mainstem and lower reaches of tributaries. Higher waves (Hsig>1.06 m) occur more frequently at the mouth of the Bay, especially on the western shore, due to the long open fetch to the north.

At lower wave heights (Hsig<1.06 m), the frequency of occurrences of positive or negative sea level at each shoreline grid point are approximately the same. However, at high wave heights (Hsig>1.06m), the difference between the frequency of occurrence of positive and negative sea level at each shoreline grid point differs spatially.

During high-wave events the effects of local forcing on sea-level fluctuations dominate over remote forcing (coastal sea level) in the upper Bay; the effects of remote forcing on sea level fluctuations dominate over local forcing in the lower Bay. Whether high-wave events more likely correspond to high (positive) or low (negative) sea level depends on the geographic location in CB. The general patterns are that high-wave events more likely correspond to high sea level on the western shore of the mainstem Bay, on the southern side of tributaries in the lower Bay and on the northern side of tributaries in the upper Bay.   High-wave events are more likely to correspond to low sea level on the eastern shore of the mainstem Bay, the northern side of tributaries in the lower Bay and the southern side of tributaries in the upper Bay.
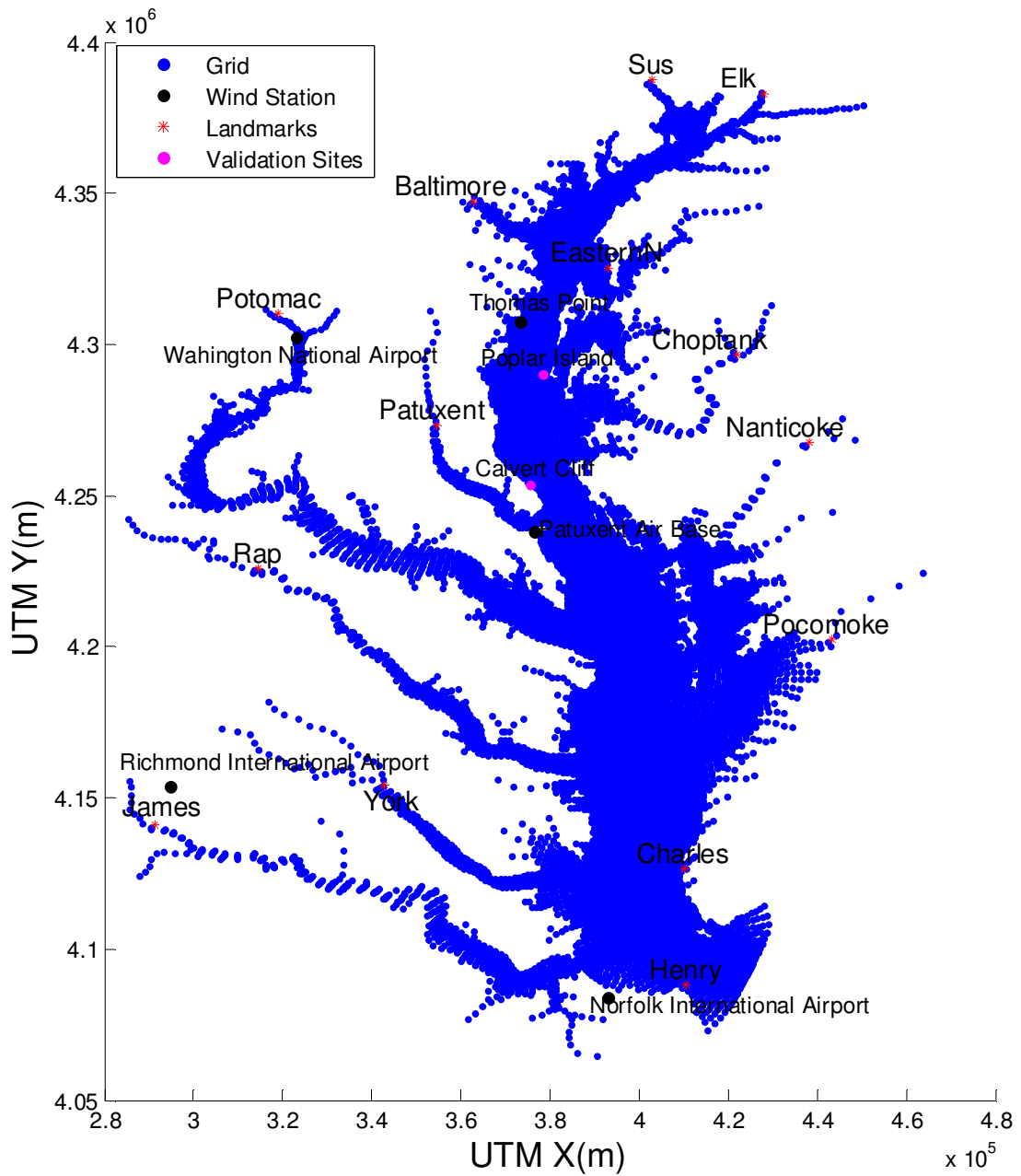
Figure 2.1. Model grid locations, wind stations, landmarks and two sites with observational data for model validation (Calvert Cliffs and Poplar Island)
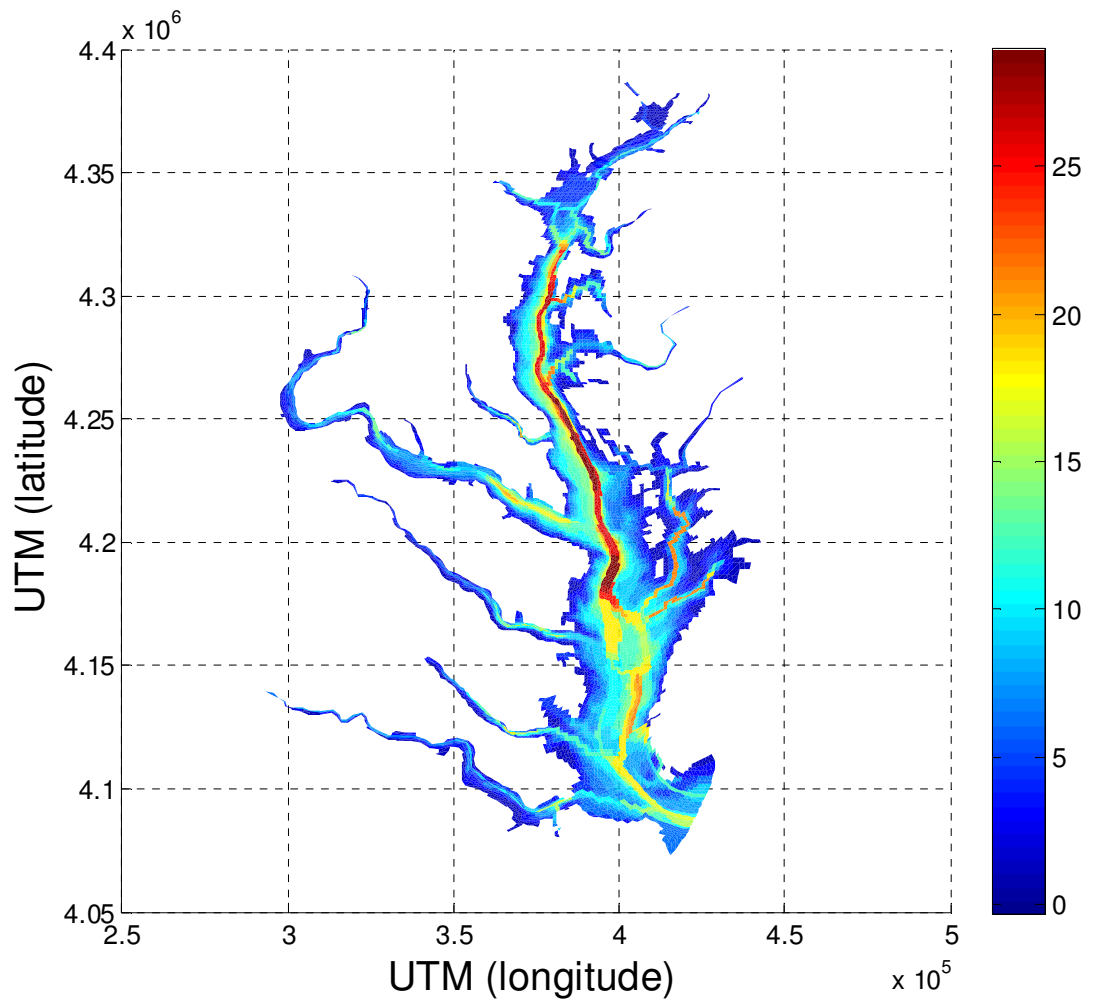
Figure 2.2. Bathymetry of Chesapeake Bay used in model (m).
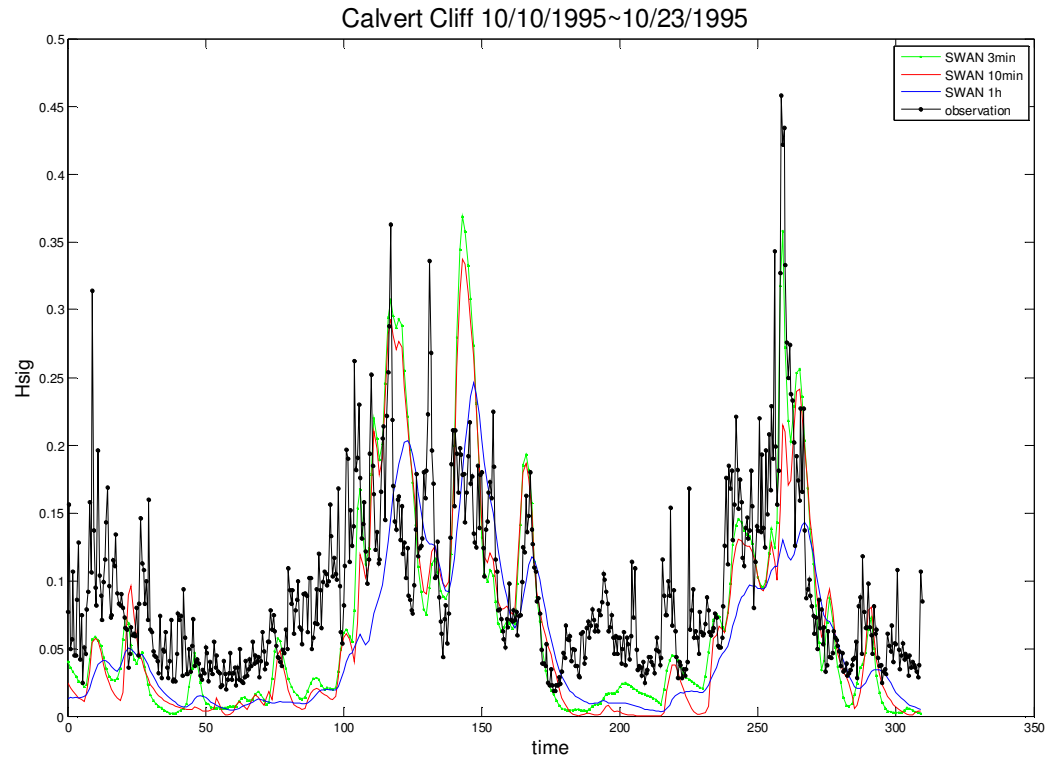
Figure 2.3. Comparison of different computational time step (3min, 10min, 1h) in SWAN
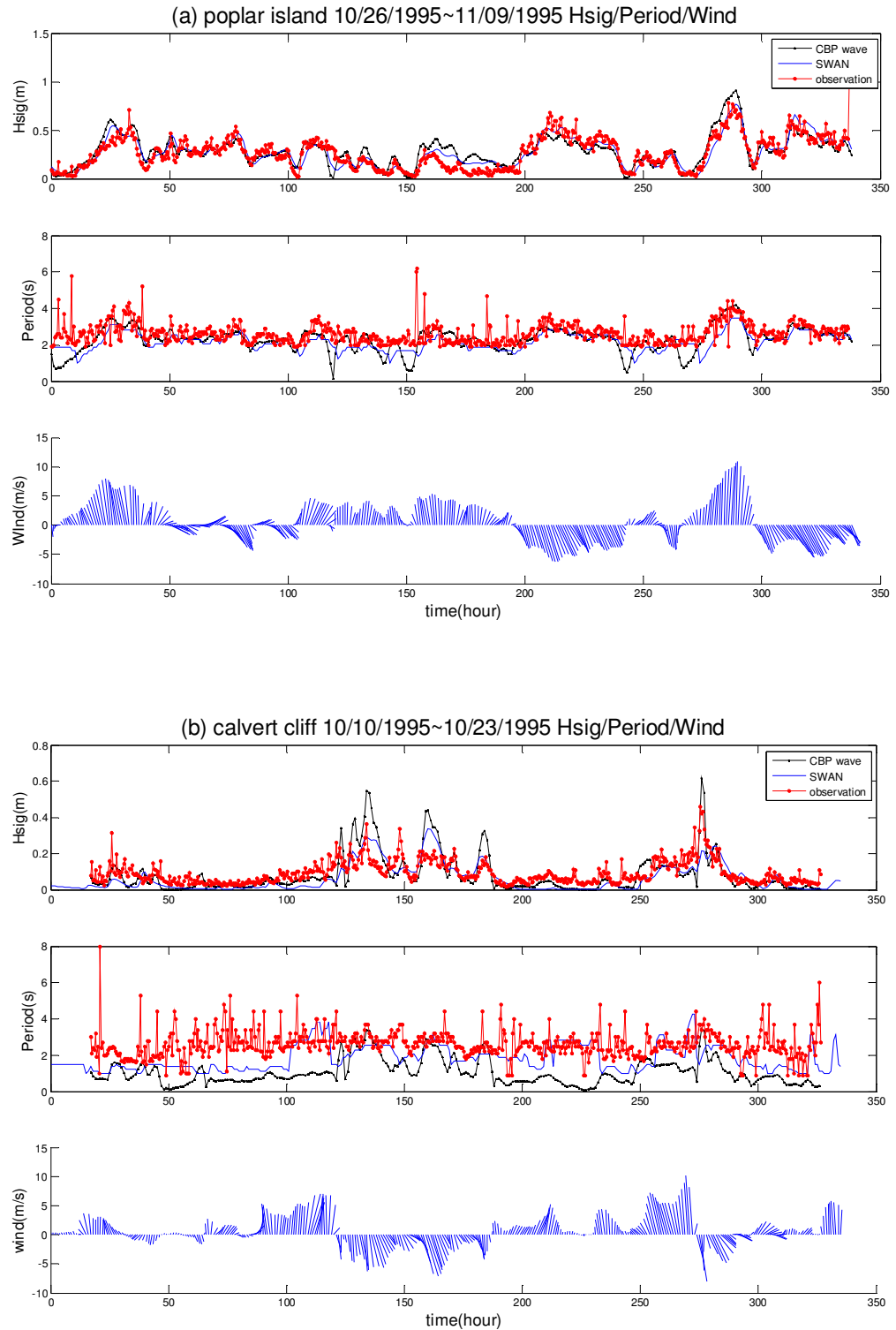
Figure 2.4. Comparisons of time series between model outputs and observations at (a) Poplar Island and (b) Calvert Cliff. For each location, panels show (upper to lower) significant wave height (Hsig), peak period, and wind interpolated from 5 stations (wind direction is shown by the direction of lines) from the two wave models.
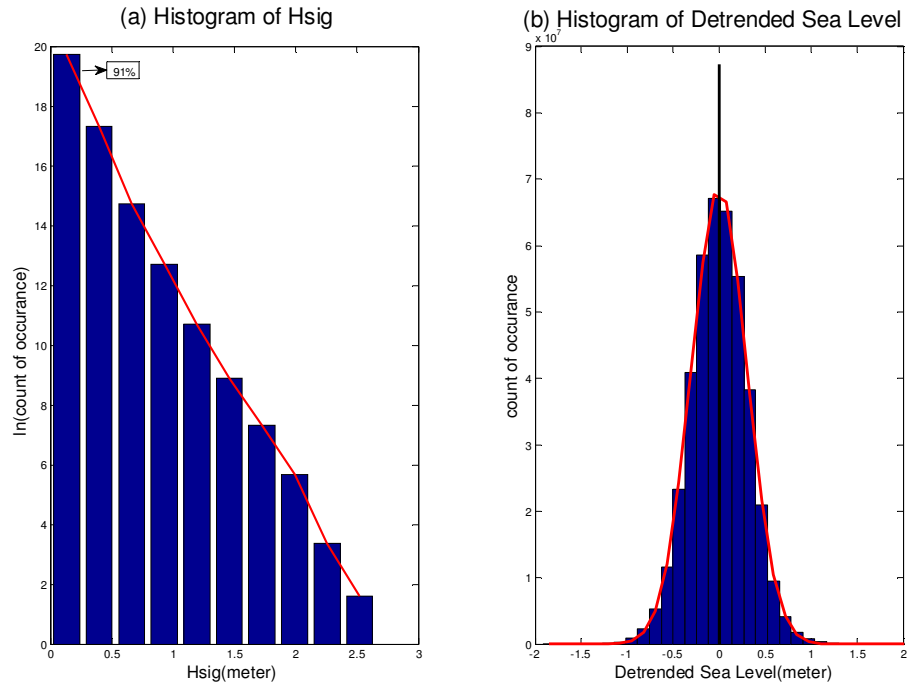
Figure 2.5. Histograms of (a) Hsig (y-axis is natural log) and (b) detrended sea level from the hourly output of SWAN and CBP hydrodynamic models for the entire Chesapeake Bay shoreline from 1985-2005.
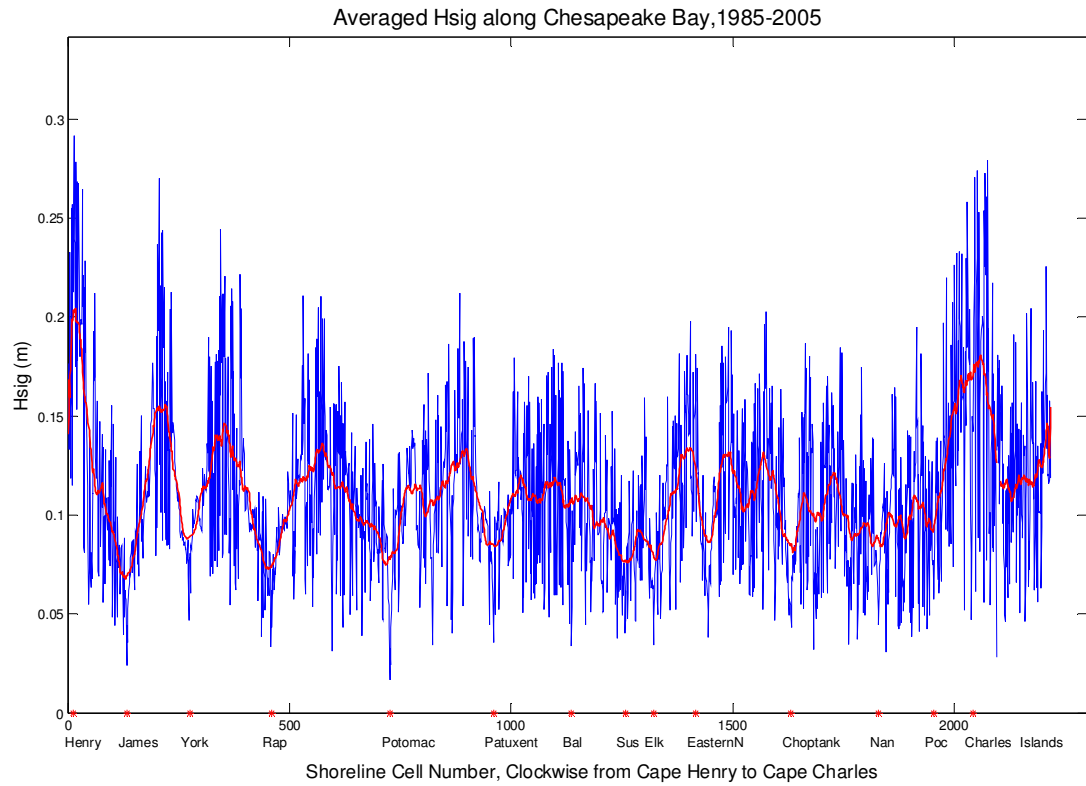
Figure 2.6. Averaged Hsig along the CB shoreline, clockwise from Cape Henry to Cape Charles; the red line is a smooth function of averaged Hsig. Red stars along x-axis are head of tributaries and landmarks, details in Table 2.1.
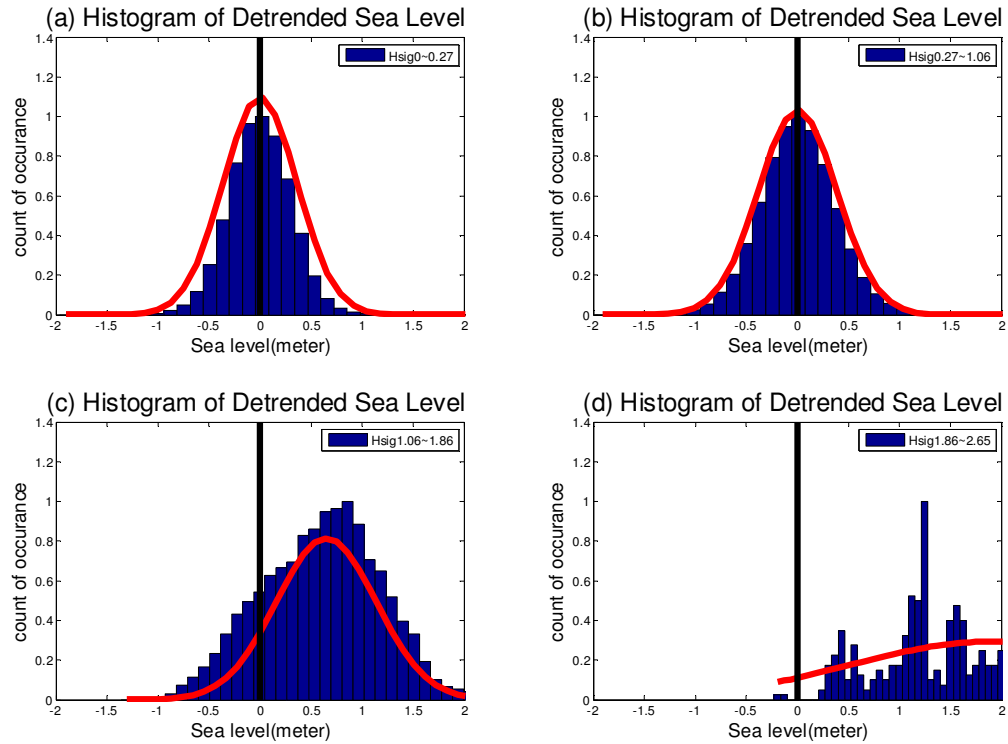
Figure 2.7. Histogram of sea level corresponding to 4 different ranges of Hsig: (a) 0-0.27 m (b) 0.27-1.06 m (c)1.06-1.86 m (d) 1.86-2.65 m. Sea level data are detrended and from the hourly output of the CBP hydrodynamic model for the entire Chesapeake Bay shoreline from 1985-2005, a scaled normal distribution for reference only in each subplot.
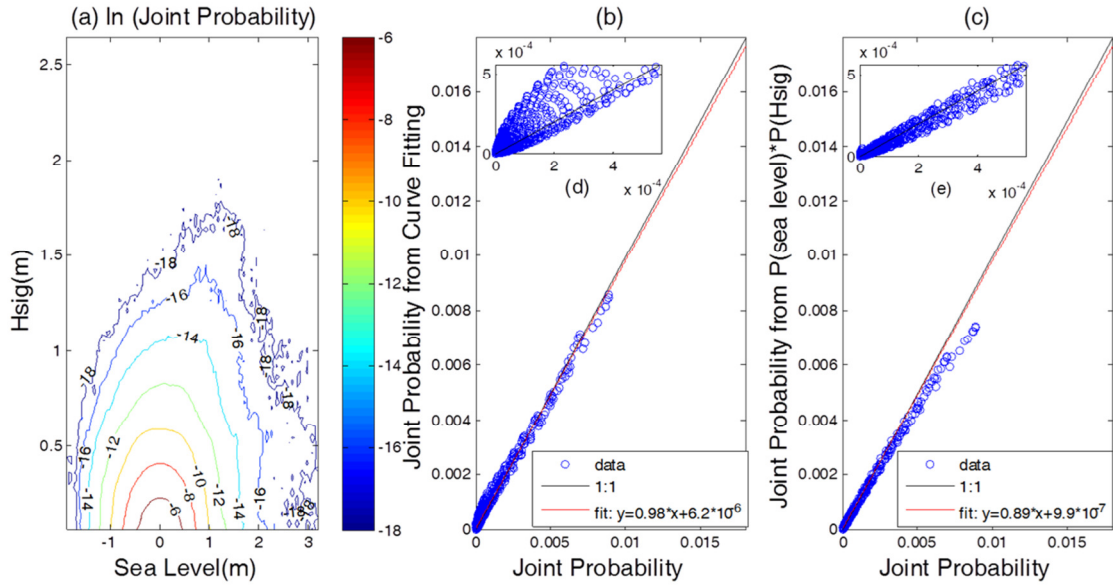
Figure 2.8. (a) Natural log of the joint probability calculated from wave and hydrodynamic model outputs; (b) joint probability vs. joint probability from curve fitting; (c) joint probability vs. the product of the individual probability of sea level and Hsig (P(w)×P(Hsig)). (d) and (e)give an expanded view of the data near low joint probability.
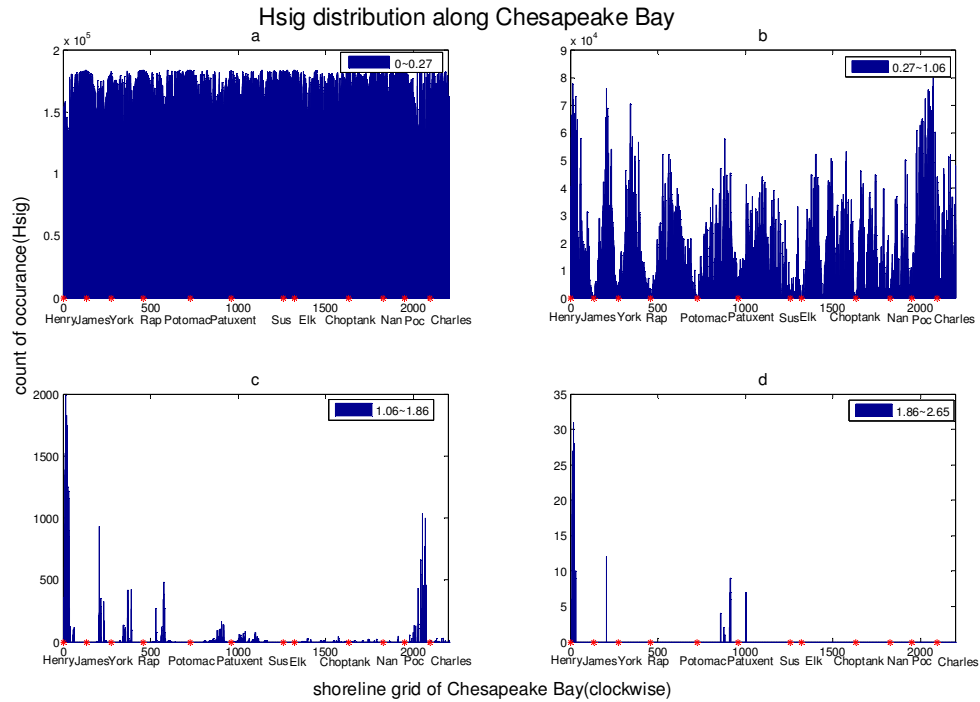
Figure 2.9. Distribution of Hsig within its 4 ranges: (a) 0-0.27 m, (b) 0.27-1.06 m, (c)1.06-1.86 m, (d) 1.86-2.65 m. X-axis: grid point along CB shoreline from Cape Henry to Cape Charles (clockwise) and around large islands for grid points > 2095 (see also Table2.1), with the heads of major tributaries indicated by the red stars (Table 2.1); Y-axis: number of occurrences of Hsig in each range at each grid.
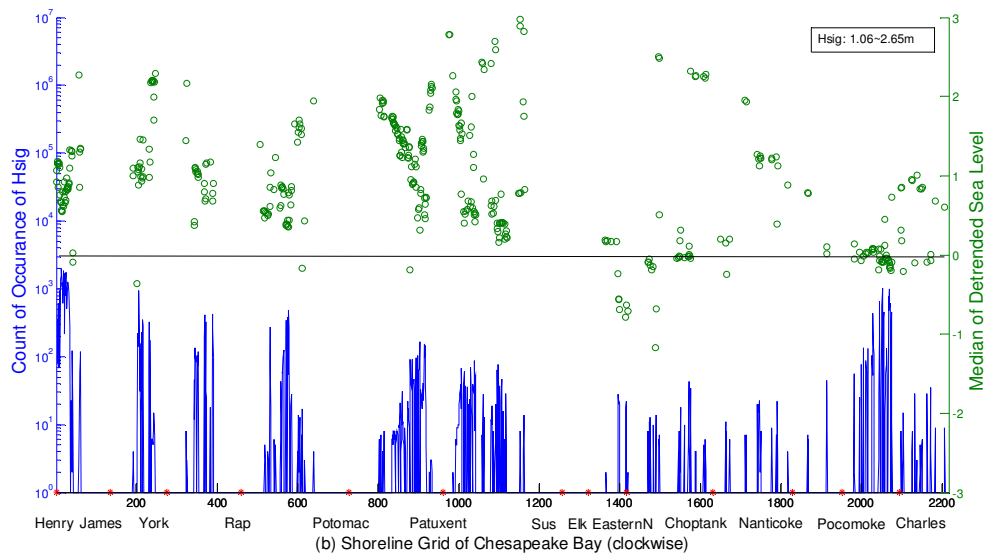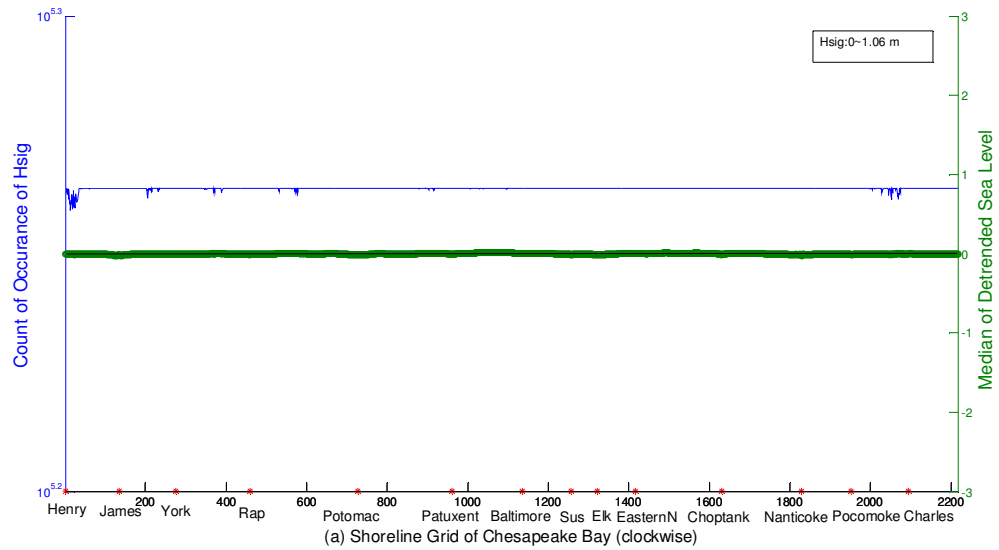
Figure 2.10. Count of occurrence of Hsig corresponding to the median of detrended sea levels in the ranges of Hsig: (a) 0-1.06m and (b) 1.06-2.65 m. Green markers represent the median of detrended sea level in each range; blue line shows the count of occurrence at each grid point from Cape Henry to Cape Charles (clockwise) and around large islands for grid points > 2095.
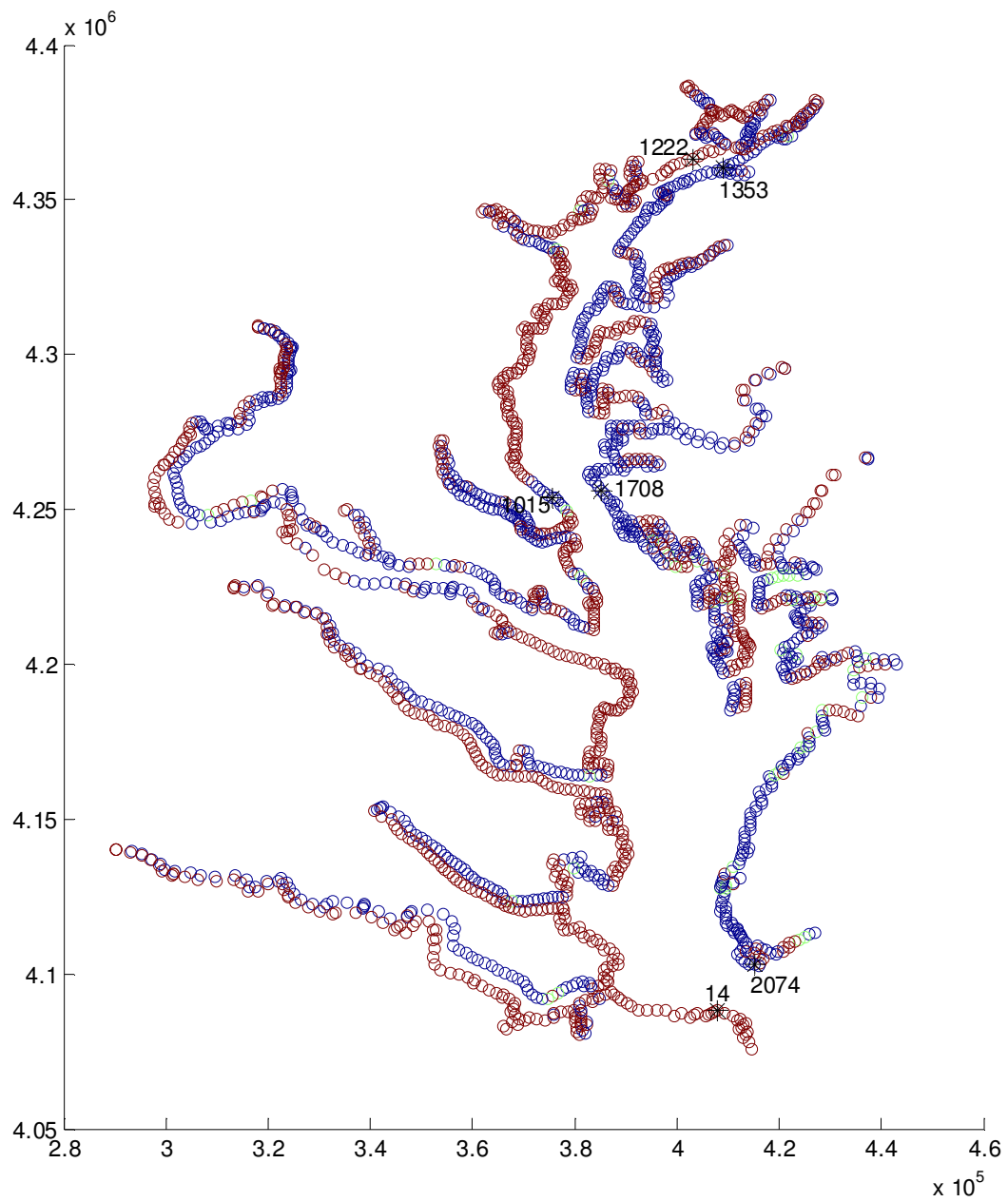
Figure 2.11. Medians of detrended sea levels that correspond with the top 60% of maximal Hsig at each grid point in Chesapeake Bay. Red represents positive medians, blue represents negative medians, and green represents zero median.
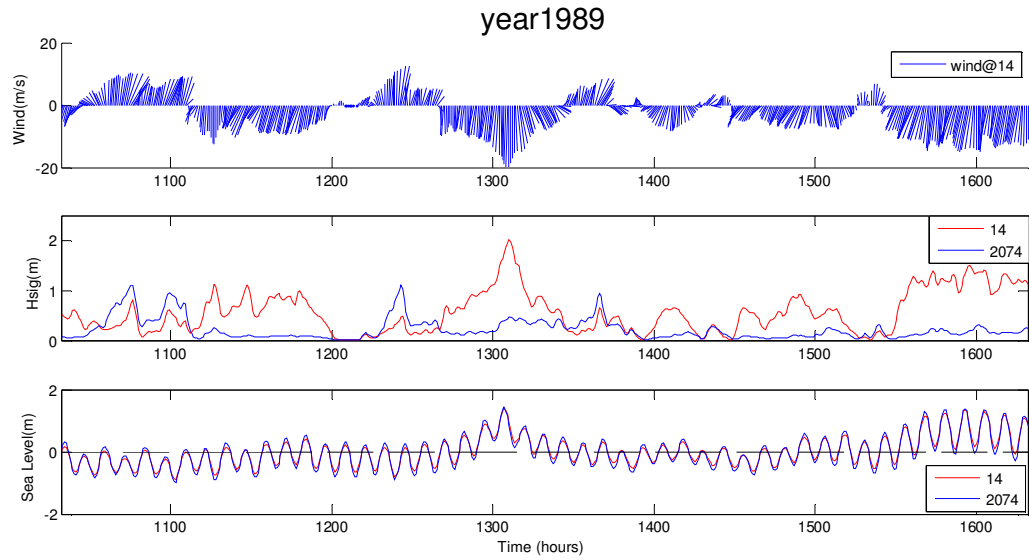
Figure 2.12. Time series of two sites (grid 14 and grid 2074) at mouth of the Bay. Upper panel: wind interpolated from 5 stations, the wind direction is shown by the directions of the lines; Middle panel: Hsig from the SWAN model; Lower panel: sea level from CBP hydrodynamic model.
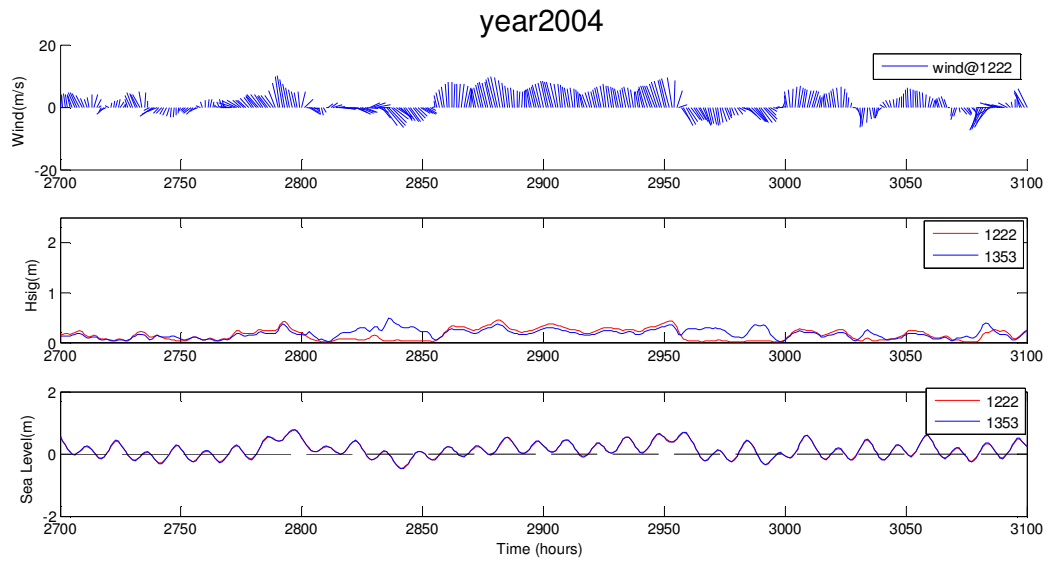
Figure 2.13. Time series of two sites (grid 1222 and grid 1353) at the mouth of the Bay. Upper panel: wind interpolated from 5 stations, the wind direction is shown by the directions of the lines; Middle panel: Hsig from the SWAN model; Lower panel: sea level from the CBP hydrodynamic model.
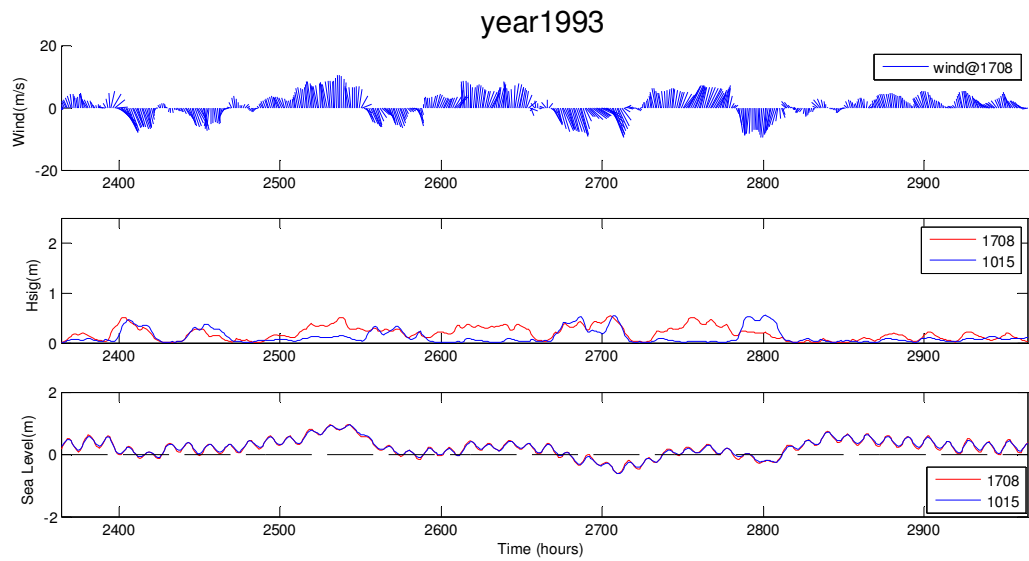
Figure 2.14. Time series of two sites (grid 1708 and grid 1015) at the mouth of the Bay. Upper panel: wind interpolated from 5 stations, the wind direction is shown by the direction of the lines; Middle panel: Hsig from the SWAN model; Lower panel: sea level from the CBP hydrodynamic model.

Name	Abbreviation	Index of Model Grid	Name	Abbreviation	Index of Model Grid
Cape Henry	Henry	1	Elk River	Elk	1321
James River	James	134	Eastern Neck	EasternN	1416
York River	York	275	Choptank River	Choptank	1631
Rappahannock River	Rap	459	Nanticoke River	Nan	1829
Potomac River	Potomac	727	Pocomoke sound	Poc	1953
Patuxent River	Patuxent	960	Cape Charles	Charles	2095
Baltimore	Bal	1136	Islands(Pooles, Blood Worth, Smith, and Tangier)[a]	Islands	2096-2217
Susquehanna River	Sus	1258			

a: Other islands (Kent Island, Poplar Island, Hoopers Island, Barren Island, Taylor Island, etc.) are counted as part of the major shoreline due to the limited resolution of shoreline grids.

Table 2.1. Land marks

Chapter 3 Shoreline erosion rates in Maryland CB

3.1 Introduction

In the Maryland portion of Chesapeake Bay (CB), shorelines consist of banks, with heights ranging from 1 meter to over 30 m (at Calvert Cliffs), and marshes that are mostly found along the lower eastern shore (Somerset, Wicomico, and Dorchester Counties). Year-round beaches only exist on about 24 km of the entire CB shoreline (U.S. Army Corps of Engineers, 1971). Some regions, points and islands are experiencing severe erosion (>2.4 meter/year) on the western shore (Pt. Lookout to St. Jerome, Holland Pt. and Thomas Pt.) and eastern shore (Kent Island, Lowes Pt. to Knapps, Mills Pt. to Hills Pt., James Island, Oyster Cove to Punch Island Creek and Barren Island)(Wang et al., 1982). There were 1.9×10^8 m^2 of land loss during 1850-1950 along Chesapeake shorelines (Slaughter, 1967a). Erosion can lead to nutrient pollution, ecosystem degradation and huge economic loss (USACE and MDNR, 2010; Leatherman et al., 1995). Erosion process has been intensified by sea-level rise, land subsidence and increasing rates of shoreline development (Halka et al., 2005).

Erosion is a highly complicated process to study not only because it involves various interacting factors but factors that can behave very differently in different geographic locations. The relationship between sea level and erosion on sandy shorelines has been described as a response of the equilibrium shoreline profile, with wave activity as the hidden cause (Bruun, 1962; Schwartz, 1967; Dean, 1991), which forms the basis for several erosion models for sandy beaches, such as the Stormed-Induced Beach Erosion (SEARCH) and Generalized Model for Simulating Shoreline Change (GENESIS) models. The role of wave activity (parametrized as wave power) is less clear for sea cliffs.

For example, Benumof et al. (2000) found that material strength appears to largely determine sea-cliff-retreat rates, with wave power as a secondary effect, but experimental results of soft cliffs found that erosion was correlated with oblique wave power (Damgaard and Dong 2004). Wave power was also found to correlate with erosion rates for glacial till bluff in Lake Erie (Kamphuis, 1987), marsh shorelines in Rehoboth Bay (Schwimmer 2001), and uniform cohesive bluffs in Lake Ontario (Amin, 1997). A recent study in Hog Island Bay (in Virginia) reinforces the important role of waves in driving erosion along marsh edges (Mcloughlin et al., 2015). Thus, wave power is likely one of the most important factors in predicting erosion rates.

Chesapeake Bay shorelines are mostly marshes and banks. Previous studies in CB have considered the ratio of silt to sand, bluff height, and cohesive soil strength as predictive variables for erosion rates (Dalrymple, 1986; Wilcock et al., 1998), with less attention to the effect of wave actions. Meanwhile, other studies have found waves as the primary factor for erosion process in CB (Wang et al., 1982; Spoeri, 1985; Skunda, 2000; Perry, 2008). Skunda (2000) modeled shoreline instability along the western shore of CB in Virginia and found wave power as a significant factor for erosion. Spoeri et al. (1985) and Wang et al. (1982) used data of 107 reaches (2-5km in length) in Maryland CB to analyze relationships between variables, including wave power sediment types, tidal range, rainfall, and 100-year storm surge, and erosion rates. Using traditional regression and discriminant analysis, Spoeri et al. (1985) were not able to provide an adequate erosion rate prediction model, but they concluded that wave energy still seems to be primarily responsible for the changes in shoreline erosion. Perry (2008) applied discriminant analysis (CART) and linear models to Maryland CB and found that fetch

seems to be the most important factor affecting shoreline erosion, again highlighting the role of waves, while geography is the second-most important predictor, complicating efforts to develop a single model for state-wide predictions.

In this study, we compared the predictions of the wave and sea level climatology developed in the previous chapter to a dataset that includes historical shoreline erosion rates, shoreline structure information, bank/marsh ratio, and mean bank height in 207 reaches for the Maryland part of CB, which has been assembled by Maryland Geological Survey (MGS) (Hennessee et al., 2006). Our study focuses on the Maryland portion of CB because an equivalent detailed shoreline inventory (combining erosion rates with shoreline characteristics) is not available for the Virginia portion of CB. First, we combine the climatological forcing and shoreline characteristics datasets to create a comprehensive and high-resolution dataset for analyzing relationships between erosion rates and wave climate, sea level, along with other shoreline characteristics, in Maryland CB. Next, we implement linear analysis, curve fitting, Generalized Additive Model (GAM), and Neural Network (NN) analyses on these datasets. Comparisons of erosion rates and wave characteristics are made between relatively large (reach and grid cell) and local scales.

Previous attempts to develop straightforward relationships between physical forcing and shoreline characteristics in CB have met with limited success. The work presented here utilizes a more comprehensive and accurate dataset to attempt to improve on those previous attempts. The datasets built in this study are so far the most comprehensive dataset available for studying shoreline erosion in the Maryland CB. They cover longer shorelines (e.g., major tributaries) and the climatological forcing is from

more advanced numerical models that allow more various and accurate environmental variables, such as bottom orbital velocity and wave power calculated from a full spectral wave model. With the information on marsh/bank ratio, erosion of marsh and bank shorelines can be analyzed separately. Data at different scales allows us to compare results and seek improvements for future data collections. Considering physical dynamics when selecting environmental variables and exploring innovative non-linear statistical methods for erosion predictions is also a step forward from previous studies. This study aims to gain a better understanding of the different contributions of a variety of controlling variables to shoreline erosion rates and attempts to establish relatively simple semi-empirical and statistical relationships between erosion rates and controlling variables, which could potentially improve estimates of erosion rates in the CBP sediment transport model and be helpful for coastal managers.

3.2 Methods

3.2.1 Data

The study area includes the Maryland portion of CB, including its major tributaries and islands. Some interior ponds, creeks and heads of tributaries are not included due to limited spatial resolution. The Maryland Geological Survey assembled data on shoreline erosion, shoreline structure percentage, bank percentage and mean bank height into one dataset at the resolution of 'reach' (Hennessee et al., 2006) (Figure 3.1). First, a reach was defined from one point of land to another. Then the reach was further subdivided if the rates of shoreline change shown on the regional map varied widely within the reach. The mouths of tributaries, county boundaries and marked changes in shoreline orientations all influence reach scopes (Hennessee et al., 2006). MGS

demarcated Maryland CB shorelines into 207 reaches, which were divided almost equally between the eastern shore (100 reaches) and the western shore (107 reaches).

The Virginia Institute of Marine Sciences (VIMS) identified man-made structures along the tidal shorelines of navigable waterways in Maryland. Appearances of structures were observed from a slow-moving boat traveling parallel to the shorelines and organized into a geographically referenced set of shoreline data (Hennessee et al., 2006). For each reach, there is a value for the corresponding percentage of protected shoreline length. Of the 207 reaches in Maryland, marshes (19 reaches) are mostly unprotected. Fifteen of them are entirely unprotected and 4 reaches have 3% or less protection along their length.

The Maryland Geological Survey cooperated with the U.S. Geological Survey (USGS) and Towson University's Center for Geographic Information Sciences (CGIS) to determine erosion rates for the coastal and estuarine shorelines in Maryland. They used digital shorelines dating from 1841-1995 as inputs into a computer program, the Digital Shoreline Analysis System (DSAS; Danforth and Thieler, 1992). In DSAS, a 50-m inland baseline was constructed, which was parallel to shorelines, as well as transects that were 20 m apart and perpendicular to the baseline. Then, rates of change were determined along each transect. DSAS produced nearly 250,000 transects with associated rates of change, including the Atlantic coast, the coastal bays, and the CB and its tributaries (Hennessee et al., 2002; 2003a,b). This database is available as a product called Coastal Atlas on Maryland Department of Natural Resources (DNR) website (http://gisapps.dnr.state.md.us/coastalatlas/iMap-master/basicviewer/index.html). In order to acquire shoreline erosion rates for each reach, MGS averaged the rates of change at all transects that were located in each shoreline reach. Erosion rates of transects that

intersected protected shorelines were excluded from the calculations, such that the reach averaged erosion rates only include unprotected shoreline. In the Coastal Atlas convention, negative erosion rates represent shoreline erosion, while positive erosion rates indicate accretion.

Maryland Geological Survey used the 7.5-minute USGS topographic quadrangles to identify two features: marshes and topographic contours along the shoreline (at contour intervals of 3 or 6 m). These two features are used to estimate the ratio of bank versus marsh (bank percentage), which differ in their physical characteristics, and the average bank height of a reach (Hennessee et al., 2006).

Erosion data are available at three resolutions: reach (1.9km to 87km in length), grid cells (approximately 1km in length) and transects (20m resolution). We incorporated output variables (Table 3.1) from the SWAN model (Booij et al., 1999; Ris et al., 1999), fetch data from the CBP wave model (Young and Verhagen, 1996; McLoughlin et al., 2015) and sea-level data from the CBP hydrodynamic model (Johnson et al., 1993) , all available at the grid scale. These models have 1316 shoreline grid cells located in Maryland. A map of each reach is overlaid with a map of grid cells and grid cells are manually assigned into each corresponding reach (Figure 3.1). These models have hourly outputs from 1985 to 2005. After averaging throughout the 21-year period, each model grid has a single average value for each output variable. All the values within each reach can then be averaged to obtain the average wave parameters and sea level at each reach. Thus, wave parameters and sea levels are derived at the same resolution (at reach resolution) as shoreline characteristics assembled by MGS.

To obtain erosion rates at the resolution of model grid cells, erosion rates calculated in DSAS were averaged over the size of model grid cells. Each shoreline transect was simply assigned to the nearest model grid. If a model grid was not assigned, the average of nearest two grid cells was used instead. The resolution of grid cells cannot resolve the Bay side (high erosion rates) or the sheltered side (low erosion rates) of a few islands. Thus, erosion rates at grid-cell resolution are underestimated at the Bay sides of Taylors Island, Hoopers Island, Barren Island, and Bloodworth Island, as well as the upper part of the Bay side of Smith Island (including Martin National Wildlife Refuge). Grid-scale estimates of erosion rate are biased by non-eroding or accreting hardened shoreline segments because the distinction between hardened and naturally eroding shoreline was not available at the grid scale. Since marsh grids are mostly unprotected, only marshy shorelines were selected for quantitative analysis at the model grid resolution in this study.

We obtained two datasets as a result: one consisting of only marsh dominated shorelines at the resolution of model grids, including all wave variables and erosion rates; and another at the resolution of reach, including all wave variables, erosion rates, bank percentages, bank heights and structure percentages. Wave predictions and sea-level simulations from SWAN, the CBP wave model and the CBP hydrodynamic model were available at both resolutions.

We also added some derived variables from model outputs into our data set for analysis, such as the onshore wave power and weighted fetch. α is the angle between incoming waves direction and the orthogonal line of shoreline orientation (Figure 3.2). It was calculated as $\alpha = \beta - \theta + 90^{o}$, where β is the direction of incoming wave-energy

propagation calculated from the transports of wave energy along x/y-axis of Cartesian coordinate (east as x-axis; transp_x and transp_y), and θ is the shoreline orientation. These calculations follow the geometric convention that east is 0 degrees, with angle values increasing counterclockwise. $\cos\alpha > 0$ represents offshore directed waves, $\cos\alpha < 0$ represents onshore directed waves, and $\cos\alpha = 0$ represents along-shore directed waves. The total wave-energy flux was calculated as

$transpall = \sqrt{transp\_x^2 + trasnp\_y^2}$. The average onshore wave energy flux was calculated as

$$transp\_onshore = \frac{\sum Transport\ of\ Wave\ Energy\ Onshore}{Number\ of\ All\ Wave\ Energy\ Estimates} = \frac{\sum transpall \times \cos\alpha}{Number\ of\ All\ Wave\ Energy\ Estimates}; \text{any}$$

calculations of offshore energy flux were excluded. Transp_onshore is all negative due to $\cos\alpha < 0$ representing the onshore direction, but its absolute value (positive) is used for analysis in this study for simplicity. Fetch and weighted fetch were acquired from the CBP wave model; the average wind weighted fetch at reach i was calculated by

$$WeightedFetch_i = \frac{\sum_j wind_{ij} \times fetch_{ij}}{\sum_j wind_{ij}}. \text{'Tidal Range' is not the exact tidal range, but rather}$$

the standard deviation of sea level, which is proportional to tidal range – it is referred to as 'Tidal Range' in this study for simplicity. Depth is from the model bathymetry, which is relative to the Mean Sea Level (MSL) in 1983. Sea levels are outputs from the CBP hydrodynamic model, which is relative to 1983 MSL as well. The choice of 1983 as a reference year is to be consistent with the CBP model. Thus, the steepness can be calculated as:

$BathSteepness = \frac{Depth+SeaLevel}{Half\ Width\ of\ Shoreline\ Grid\ Cell}$, where the denominator is the distance

offshore at the center of each shoreline model grid. Elevation is the height of the cliff

with respect to the tidal flat bottom, and the Volumetric Erosion Rate (VER) = Erosion

rate× Elevation.

3.2.2 Region divisions

For the dataset at reach resolution, shoreline types can be identified as 'marsh',

'bank' or 'mixed' type through the variable 'bank percentage'. We defined a reach as

type 'marsh' if 'bank percentage' was ≤ 10% and a reach as type 'bank' if 'bank

percentage' was ≥ 90%. Because of distinctive erosion processes due to the different

sediment properties (e.g., particle size, vegetation type, etc.) between marsh and bank

(shown below), the mixed type adds unnecessary complexity to the issue. Thus, this study

will only focus on discussing reaches that fall into bank (117 reaches) or marsh 27

reaches) categories. We further divided each of these types into the sub-regions of

mainstem, tributary, eastern shore and western shore for our analysis.

Wave heights are much higher in the mainstem of CB due to much longer fetches

and stronger fetch-aligned winds than in most tributaries (see previous chapter). Thus we

divided each type of shoreline into tributary and stem for analysis. Also, because the

median sea level during high-wave events is the opposite on eastern and western shores,

each type of shoreline is divided into eastern shore and western shore for exploration as

well. 'Bank' type ends up with 17 reaches in the mainstem, 99 reaches in tributaries, 43

reaches on the eastern shore, and 73 reaches on the western shore. Ideally, only reaches in

the mainstem would be included in 'eastern' and 'western'. However, all 117 reaches of

bank data are used because further subdivision of 17 reaches in the mainstem would not allow statistically meaningful analysis.

Marsh data at the reach scale also have too few points for further subdivision. To solve this issue, we employed the marsh dataset at the resolution of model grids. Utilizing the information of category (marsh or bank) from reach dataset, we could assign 'marsh' or 'bank' type to each model grid. Grids located on Taylors Island, Hoopers Island, Barren Island and a portion of the Bay side of Smith Island are excluded from the dataset due to the incongruous resolution when averaging erosion rates from transects to each model grid (Figure 3.3). As a result, 163 grid cells were identified as marsh, and they are further divided into 'stem', 'tributary', ' stemEastern' and 'stemWestern' for analysis. The Coastal Atlas transects of erosion rates include hardened shorelines, but since marsh shorelines are mostly unprotected, only marsh data at the resolution of model grid were selected for further analysis. In summary, we acquired eleven groups of data: 'Bank', 'Bank Stem', 'Bank Tributary', 'BankEastern', 'BankWestern', 'Marsh', 'MarshHD', 'MarshHDstem', 'MarshHDtributary', 'MarshHDstemEastern', 'MarshHDstemWestern'. The first six groups are at the resolution of reach, and the latter five groups represent marsh data at the resolution of grid cells.

3.2.3 Quantitative approaches

3.2.3.1 Statistical methods

Linear correlation analyses, including Pearson Correlation Coefficient and Multiple Linear Regression (MLR), were performed for the purpose of preliminary quantitative data screening. However, the effects of predictor variables on erosion rates

are clearly not uniform across CB due to geographical variations, which make Bay-wide relationships non-linear and/or non-uniform. Non-parametric and non-linear GAM and NN analyses were used selectively to characterize statistical relationships when data sets were too complex or too non-linear for simple linear techniques. Thus, non-parametric and non-linear GAM and NN analysis were also performed tentatively in this study.

NN (Beale et al., 2014) imitates the mechanics of neural systems, which consist of multiple layers and interconnected neurons within each layer. NN has the ability of representing both linear and non-linear relationships and learning these relationships directly from the data being modeled. GAM estimates non-parametric functions of predictor variables and connects dependent variables with a link function, which is more flexible than linear regressions (Hastie and Tibshirani, 1990). We applied GAM using 4, 6 and 12 as the degrees of freedom for smoothing terms and the identity function as the link function for this study. NN analysis was achieved using NN toolbox in Matlab; pairwise linear correlation analysis was performed in Matlab; and GAM and MLR were done in R.

'BankEastern', 'BankWestern', 'MarshHDstemEastern', and 'MarshHDstemWestern' are excluded from these analyses for simplicity. In NN analysis, there are two layers in total, the hidden layer and the output layer, which usually has one neuron. The hidden layer is configured with one neuron for data groups 'bank', 'bankstem', 'bankTributary' and 'Marsh', and 3 neurons for data groups 'MarshHD', 'MarshHDstem' and 'MarshHDTributary', to avoid overfitting due to small sample sizes.

The Akaike Information Criterion (AIC) is a measure of the relative quality of a statistical model for a given set of data, but AIC cannot evaluate the quality of a model in

an absolute sense (Bozdogan, 1987). AICc is AIC with a strict penalty for introducing too many variables or a correction for small sample sizes, so using AICc instead of AIC can reduce the probability of choosing models with too many parameters. The formula is as follows:

$$AIC = 2k + n\log(\frac{SSE}{n}) \text{ and } AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

where n is the sample size and k is the number of estimated parameters.

The flow of statistical analysis is as follows. First, all predictor variables (Table 3.1) were preprocessed by eliminating similar variables and scaling. For example, 'fetch' and 'weighted_fetch' represent a very similar physical factor, although their values were occasionally quite different. Thus, the 'fetch' variable with a lower correlation coefficient was eliminated from each group before applying statistical models. All variables were then scaled using a 'z-score'. This process is linear and thus should not influence our statistical models. NN analysis used unscaled data because the NN toolbox in Matlab pre-scales input data automatically. Next, non-metric multidimensional scaling was used to map each predictor variable onto a 2-dimensional space (Borg and Groenen, 2005). The more similar variables will be closer to each other in the 2-dimensional space. Many groups of clustered variables will be detected and only the variable with the highest correlation coefficient within each clustered group will be included in statistical analysis. Using 'bank' data as an example, 'URMS', 'UBOT'and 'UBOTsq' were tightly clustered; 'Hsig90' and 'Hsig95' were tightly clustered; and 'Tps', 'TM01', 'WLEN', 'LWAVP', and 'transpall' (see table3.1) were tightly clustered. Of these, only 'UBOTsq', 'Hsig95' and 'LWAVP' were chosen from each group, along

with all other unclustered variables, for statistical analysis. These selected variables were then used in stepwise MLR, GAM and NN analysis. R^2, adjusted R^2, P-value, AIC, AICc and RMSE (Root Mean Square Error) were employed for the comparison of statistical methods.

3.2.3.2 Curve fitting of erosion rates versus onshore transport of wave energy

Relationships between erosion rates and onshore wave energy flux were explored using the Curve Fitting toolbox in Matlab to compare with the regression equations found in Rehoboth Bay and Lake Erie (Gelinas and Qidgley, 1973; Kamphuis, 1987; Schwimmer, 2001).

3.2.4 Comparisons of wave climate and erosion rates among different scales (reach, grid cells, transects)

Wave height and wave period were measured in the Bohemia River (5/20-27/2014), Broad Creek (9/4-11/2012), Elk River (8/17-22/2011), Honga River (9/2-5/2010), St Mary River (8/9-15/2012) and Severn River (5/8-13/2014) by E. Koch and D. Booth (unpublished data). Wave gauges were deployed at about 1m depth at each site, one in front of the natural shoreline and one in front of a directly adjacent rip-rapped shoreline. The data were all collected during the potential growing season for submerged aquatic vegetation (SAV) for a study focusing on the effects of shoreline hardening on SAV distributions, and were intended to characterize seasonal local waves. For our purposes, the observations of wave height and wave period were averaged over the duration of each deployment and then compared with the simulated wave climate (averaged from 1985 to 2005) for the corresponding reach or grid cell from SWAN. Wave period was compared with predicted bottom wave period (TMBOT), because wave

gauges were pressure sensors deployed just above the bottom. Neither the reach dataset nor grid-cell dataset can resolve the shoreline morphology at Broad Creek, which is a sheltered cove. The Severn River is truncated in our model grid cells but can be resolved by the reach dataset. Thus, Broad Creek was excluded for analysis and the output of the closest model cell was used for the Severn River (Figure 3.3).

3.3 Results

3.3.1 Erosion rates

Erosion rates at the resolution of model grids include all types of shorelines, so the value for each grid cell reveals the actual yearly average erosion rates (blue line in Figure 3.4a) instead of only unprotected shorelines. Erosion rates at the reach resolution only contain unprotected shorelines (red line in Figure 3.4a). They agree well with erosion rates at the resolution of grid cells except the two high peaks of accretion. These peaks are located in Baltimore Harbor and Hart-Miller Island (created from dredge materials), which are both heavily impacted by humans.

Erosion rates vary widely. Using grid cell 1273 as a division of the eastern and western shore, shorelines on the eastern shore generally have higher erosion rates than on the western shore. The Bay side of islands undergoes the most severe erosion. High erosion (negative values in Fig. 3.4a) occurs on the Bay sides of Taylor Island, Hoopers Island, Barren Island, Pooles Island, Bloodsworth Island, and the upper part of the Bay side of Smith Island. Most islands have marshy shorelines, which appear to be more vulnerable to erosion than banks in our data, except the Bay side of Hoopers Island and Barren Island. The onshore component of wave-energy flux appears minimal in

tributaries, especially at the head of tributaries, but significant at islands. Most high

erosion rates correspond to high wave-energy flux but high wave energy does not

necessarily lead to high erosion (Figure 3.4a).

In Figure 3.4b, most of the average bank height at the reach resolution is less than

5 m, except the banks near Calvert Cliffs, which rise up to 30 m. Banks occupy a much

larger percentage of shoreline than marshes in CB; marshes are mostly located on the

lower eastern shore and islands.

Because the transects of erosion rates in the Coastal Atlas include unidentified

hardened shorelines, it is inappropriate to apply these erosion rates averaged over grid

cells for most analyses. However, marshy shorelines are almost all unprotected, so the

erosion rates for grid cells identified as marsh qualify as unprotected erosion rates. Reach

erosion rate estimates, on the other hand, excluded hardened shorelines. Grid cells from

1741 to 1952 are mostly marsh on the lower eastern shore, allowing the use of their

erosion rates at both the resolution of grid cells and reaches. Even with scatter, these

areas exhibit a general trend of high wave-energy flux corresponding with high erosion

(Figure 3.5a). For the lower western shore (grid cells 797-1054), where banks dominate,

erosion rates can be compared to the wave-energy flux only at the reach resolution,

though no obvious trend is detected (Figure 3.5b).

3.3.2 Statistical analysis

3.3.2.1 Linear correlation

Outliers are excluded using quantile ranges. Data outside the range of Q1 – 3*IQ

and Q3+3*IQ are identified as outliers, where Q1 is the 25th percentile, Q3 is the 75th

percentiles and IQ equals Q3-Q1. For bank data, the upper bound is relaxed to -0.96

instead of Q3+3*IQ, which equals to -0.79, in order to preserve more reasonable data for

analysis. As a result, one data point of bank (the Bay side of Hoopers Island and Barren

Island), and one data point of marsh (entire Taylors Island) are detected as outliers. For

marsh data at the resolution of grid cells, three outliers are detected on the Bay side of

Smith Island.

We calculated the Pearson correlation coefficients among all 24 potential

influential variables (Table 3.1) for 'marsh' and 'bank' respectively, but only those

relatively significant (R>0.5, P<0.05 for marsh; R<0.2, P<0.05 for bank) variables are

listed in Table 3.2 and Table 3.3. The correlation matrix shows most variables correlate

with each other (Figure 3.6), with stronger correlations generally occurring for marshes

than for banks. The strongest correlations occur between wave characteristics such as

significant wave height, wave period, wave length, etc. (Figure 3.6).

We refer to the data of all reaches as the 'All Data' group, which covers all 203

reaches (with the 4 outliers excluded) along the Maryland shorelines, including marshes,

banks and a mixture of different percentages of these two types. Correlations between

wave variables and erosion for 'All Data' are not very significant, perhaps because there

are so many other distinctive geological properties affecting erosion of this

heterogeneous data set. For example, erosion rates between marshes and low elevation

banks (0-1.5m) behave differently when compared with increasing onshore wave-energy

flux (Figure 3.7). Marshes undergo more severe erosion than low banks in general.

Erosion rates of marsh become much higher as the onshore wave-energy flux increases,

while erosion rates of low banks remain relatively constant. In the range of low onshore

wave-energy flux ($<7\times10^{-4}$ kw/m), both low banks and marshes have a background

erosion rate about -0.2 m/year. Thus, marshes need to be treated differently than low

banks. This is also evident by the correspondence of higher bank percentages to lower

erosion rates (r^2=0.44; p=5.48×10^{-11}; Figure 3.8).The mean of erosion rates decreases as

bank percentages increase, and the mean of different groups are indeed statistically

significantly different (ANOVA; 99.9%). The difference of mean erosion rates between

0-10% and 70-90%/90-100%; and, the counterparts between 90-100% and 10-40%/40-70%

exceed 95% statistical significance in Tukey's test. In other words, if a reach has higher

percentage of bank, it is more resistant to erosion. The category '0-10%' and category

'90%-100%' are defined as marsh and bank, respectively, for most of the analyses

presented here. It is shown that banks experience both erosion and accretion, while

marshy shorelines only undergo erosion; and, banks undergo less severe erosion than

marshy shorelines.

In the marsh data, wave power (transpall) is the most linearly correlated variable

with erosion rate (Table 3.2). MGS assumed a constant 0.5 m as marsh elevation, thus

the correlation coefficients of volumetric erosion with all influential variables are exactly

the same as linear erosion rate. In the bank data, wind weighted fetch and bottom shear

stress (UBOTsq) are almost equally important. Volumetric erosion is highly correlated

with erosion itself, but no significant correlation was found between the volumetric

erosion and other likely influential variables. In the 'BankStem' data group, which only

includes 16 reaches, many bottom-stress variables, including URMS, UBOT and

UBOTsq, and the steepness of bathymetry show the highest correlations. In 'Bank

Tributaries' and 'Bank Eastern' data groups, 'fetch' has the most significant linear

correlation, and 'MedianWater60Hsig' had the most significant correlation with erosion. Only 4 variables are significantly correlated in 4 or more shoreline categories: 'FSPR', 'UBOTsq', 'Transport Onshore', and 'Tidal Range'.

In general, marshy shorelines have much stronger correlations than banks between erosion rates and other variables. Wave-power related variables (such as Hsig, transp_normal, fetch, weighted_fetch etc.) are most closely correlated with erosion rates. Wave-period and wave-length variables are all highly correlated with wave height and wave power (see Figure 3.6), thus they are also highly correlated with erosion rates. Bottom-stress related variables show high correlation with erosion rates as well, especially for 'Bank', 'Bank Stem' and 'Bank Western'. The correlation between tidal range and erosion is a geographical effect - it is higher in tributaries than the mainstem, increasing upstream in tributaries, except at the mouth of CB, while wave power has just the opposite pattern (Figure 3.9 and Figure 3.4). Thus, reaches with lower tidal ranges would likely have high erosion rates due to higher wave power. In 'Marsh', 'Bank' and 'Bank Tributary' group of data, correlations also increase with higher percentiles of wave height (Hsig90, Hsig95).

The correlation of 'MedianWater60Hsig' with erosion rate is negative on the eastern shore but positive on the western shore (Table 3.2). This statistic is not significant on the eastern shore. A positive correlation means that higher medians of sea level correspond to less erosion on the western shore; this makes sense recalling that most median sea levels are negative on the western shore during high-wave events (top $60^{\%}$ of maximum Hsig). Long-shore drift only shows a weak correlation with erosion rate on tributary banks. Onshore wave power shows a high correlation with erosion rates in most

data groups. Surprisingly, its correlation is lower than the unidirectional total magnitude of wave energy flux ('transpall') in 'Marsh' group (though only slightly so). 'FSPR' has relative high positive correlations with erosion rate, because younger waves with wider frequency spectrum have smaller wave power, which lead to less erosion.

Bank height is not significantly correlated with erosion rates, and so therefore it is not included in Table 3.2. However, mean erosion rates do decrease as bank height increases for bank heights between 0-6 m (Figure 3.10). ANOVA analysis shows that there is a 92% significance that the mean among the three categories in the 0-6 m range are different. Furthermore, Tukey's test identified a 99% statistical significance of difference in the mean erosion rate of bank heights 0-1.5 m and 3-6 m, but the difference of the mean erosion rate among the other two combinations of groups (0-1.5 m and 1.5-3 m; 1.5-3 m and 3-6 m) was not significant. Two data points of bank height higher than 10 m show moderate erosion and are treated as outliers in both the ANOVA and Tukey's test, simply because there are far fewer data points than in the other categories.

On a finer scale, the correlation coefficients decrease, but their significance (P-value) increases due to an increase of sample size and thus scatter. The correlations in marsh data at grid cell resolution (Table 3.3), which is a finer scale, generally agree with the marsh data at reach resolution (Table 3.2), except that bottom-stress related variables and wave-energy flux show relatively less correlation.

3.3.2.2 MLR, GAM and Neural Network

With outliers excluded, the histograms of erosion rates for the three groups of bank data ('Bank', 'BankStem' and 'BankTributary') are symmetrical and nearly

normally distributed (Figure 3.11). Thus, GAM with an identity link function is applied

to the bank data. However, the four groups of marsh data ('Marsh', 'MarshHD',

'MarshHDStem', and 'MarshHDtributary') have skewed distributions and are not

included in the GAM analysis (Figure 3.11). NN and MLR were applied to both the bank

and marsh data. Table 3.4 shows the variables that were included or excluded for MLR,

GAM and NN analysis. Every run of NN analysis is different due to different initial

weight assignments. Results of NN shown in Table 3.5 are the best results among 10 test

runs for each group of data.

In the data group 'BankStem', all three multi-variant statistical methods are over-

parametrized and thus over-fitted due to the small sample size. In all three groups of bank

data, GAM shows higher correlations (R^2), less residuals (RMSE) and lower AIC, which

make GAM a better predictable model than MLR and NN within these data groups. MLR

surprisingly shows a better performance than NN for all three groups of bank data. For

all marsh data at reach and grid cell resolutions, there is no evidence that NN simulates

erosion rates better than linear MLR (Table 3.5). The number of estimated parameters in

GAM and NN analysis are usually higher than MLR using the same dataset, and adjusted

R^2, AIC and AICc all penalize the model performance for having a high number of

estimated parameters. Also, no obvious advantages of NN and GAM are shown over

MLR when comparing simulated erosion rates with measured erosion rates (Figure 3.12).

3.3.3 Curve fitting of erosion rates versus onshore transport of wave energy

We applied the Curve Fitting toolbox in Matlab to fit erosion rates to the onshore

wave-energy flux for marsh data at both resolutions and bank data at the reach resolution.

Only marsh data at the reach resolution yield statistically significant empirical functions. The nonlinear least-squares method yields a power function (Equation 3.1, r^2=0.55):

$$y = -33.71 \times x^{0.86} \tag{3.1}$$

$$y = -54.06 \times x - 7.55 \times 10^{-2} \tag{3.2}$$

where x is onshore wave power and y is erosion rate. However, a linear relationship (Equation 3.2, $R^2 = 0.55, P = 1.3 \times 10^{-5}$) can also describe this quantitative relationship with equivalent statistical robustness (Figure 3.13). Nevertheless, the power function is preferred, since the linear fit cannot capture the slight deceleration of the increase in erosion rate as the wave-energy flux increases.

3.3.4 Comparisons of wave climate and erosion rates among different scales (reach, grid cells, transects)

Using erosion rates averaged over a large scale could under- or overestimate local conditions, depending on the details of local variability. This is also why the average erosion rate of a reach can differ significantly from the erosion rate of the corresponding grid cells. The smallest scale local erosion rates quoted here are from the nearest individual transect to a site. Averaged erosion rates at the scale of either a reach or a grid cell seem to poorly estimate local erosion rates for the Bohemia River, Elk River and St Marys River sites, but are more representative at the Honga River site (Figure 3.14). For the Severn River site, the closest model grid cell was used instead of the actual location, so there is little agreement between local and grid cell erosion rates as expected.

'Grid cells' and 'reach' are all spatially and temporally averaged representations of the wave climate. Spatially, they are averaged along the length of either a reach or the size of a grid cell. Temporally, they are averaged over 1985-2005 from hourly SWAN model outputs. Local 'natural' and 'riprap' wave measurements were only temporally averaged over the duration of deployment (4 to 8 days). Thus, it is somewhat surprising that the average wave climate seems to reasonably estimate local wave conditions (Figure 3.15). Long-term average significant wave height (Hsig) tends to be a slight underestimation of short-term average Hsig, while the long-term top 5% of Hsig tends to be higher than its short-term counterparts. This might be due to errors in the SWAN estimation under different wind conditions, but it could also be that the short-term local observations were mostly collected during summer when strong winter storms were absent. It is not possible to compare modeled waves to observations directly because the model period ended before any of the observations were collected.

3.4 Discussion

3.4.1 Erosion rates

There are two components of shoreline erosion: fastland erosion, which happens above the waterline, and nearshore erosion, which occurs from the waterline to the base of wave action. The term 'erosion rate' used in this study refers to the fastland erosion rate. Shoreline elevation and orientation, shoreline type (e.g., vegetated, protected, bare), sediment type and availability, nearshore morphology, land subsidence, sea-level rise, hydrodynamic and wave characteristics, and human activity can all be potential factors for shoreline erosion of both types.

The actual process of shoreline erosion usually proceeds through a sequence of events: waves undercut the cliff/marsh base; the cliff/marsh collapses; waves resuspend sediments at the cliff/marsh base; and currents remove these materials. Thus, using erosion rates that count volume or mass might be more beneficial than simple lateral erosion rates since the former quantifies the erosion process more comprehensively. Marani et al. (2011) found that wave power is proportional to volumetric erosion, the product of erosion rates and the corresponding height, on marsh edges, instead of simply the lateral shoreline erosion rates. In the present study, marsh elevation was not measured directly but was assumed to be 0.5 m, while bank height is given for each reach by MGS. With an assumed constant marsh elevation, there is no difference between shoreline erosion rates compared to volumetric erosion rates. If comparing among locations with various bulk densities, the mass erosion rate should be the product of erosion rates, elevation, and bulk density. So, variable bulk density along the banks of CB might explain the lack of correlation of bank data with volumetric erosion rates. Another reason could be that the product of average bank height and average erosion rate is different from the average of the product of bank height and erosion rate. We could calculate the former, while the latter is what we really expect.

Sediment type and availability for each reach is unknown, but there are some data on sediment characteristics for banks and marshes. Analysis of 76 sediment samples from 21 bank sites shows 44% sand and gravel, 56% silt and clay, and negligible organic matter. 20 sediment samples from 4 marsh sites show 22% sand and gravel, 44% silt and clay and 34% organic matter (Hennessee et al., 2006). The largest difference between banks and marshes then is that the marsh sediments contain much more organic matter,

presumably in the form of roots and decaying above ground biomass. We hypothesize that this large organic matrix makes the marshes more erodible than their bank counterparts (Figure 3.7).

During extreme storms, strong winds can cause sea-level to rise above the marsh elevation, dissipating most of the wave energy due to friction instead of eroding the marsh base. Excluding sea levels above marsh elevation decrease the scatter in the relationship between erosion rates and wave-power density (Marani et al., 2011), but we did not exclude events when sea levels exceeded marsh elevation due to lack of local data on marsh elevation. Moreover, measurements of relatively permanent submergence due to sea-level rise or land subsidence along the entire Maryland part of CB are not available and not considered in this study.

3.4.2 Statistical analysis

Linear analysis assumes that the effects of each predictor variable on erosion rates are similar among different sites. However, in a spatial dataset as large as the one examined here, different predictor variables may play different roles in different regions. Thus, linear methods are not favorable for quantifying statistical relationships unless the studied region is small enough to avoid large-scale variability in erosional processes. In this study, data were separated as 'marsh' and 'bank' and then subdivided into 'tributary' and 'stem'. However, this breakdown by region did not show improvements of performance on the applied statistical methods, which implies that finer divisions with higher resolution data might help reduce the complexity and improve statistical model results.

NN and GAM are very sensitive to outliers and training data sets. Thus, data usually need to be preprocessed (e.g. excluding outliers, scaling) before they are used for training in statistical models. When data-adaptive non-linear models are applied to highly complex issues, especially with a small sample size, the resulting predictions may have reasonable outputs but tend to be overfitted. Pruning model complexity usually leads to poorer model outputs. With increasing parameters, sample size will become relatively smaller. Thus, if redundant variables were included before carrying out GAM or NN in this study, our results might be much closer to the target data, but the statistical models would be more over-parametrized and thus more overfitted due to our relatively small sample size. Because our predictor variables are correlated with each other to different extents and our sample size is relatively small, the results of GAM and NN analysis in this study tend to be over-parameterized and should be used with great caution for prediction.

Colinearity among variables, relatively small sample size (or over-parameterization), absence of geological characterization, and poorly estimated outputs can all lead to invalidation of the predicting ability of all three applied statistical models. Using NN, GAM, or other similar machine-learning techniques, our data might be too generic and incomplete to build an accurate, quantitative predictive model for erosion rates. However, given a more complete and comprehensive dataset, either temporally averaged high-resolution spatial data or time-series data at a local site, Generalized Additive Model, Neural Network analysis, or other similar machine-learning techniques might be much more powerful for building quantitative predicting models between erosion rates and its influential variables.

3.4.3 Wave power versus erosion rates

3.4.3.1 Wave power versus erosion rates

Wave-energy flux is the most often investigated factor affecting erosion rates and found to be significantly related with erosion rate in many cases (Dalrymple, 1986; Gelinas and Qidgley, 1973; Kamphuis, 1987; Marani et al., 2011; Mariotti et al., 2010; Mcloughlin et al., 2015; Ronald and Douglas, 2005; Schwimmer, 2001), although other studies have observed a lack of significant relationships between the wave-energy flux (or wave power) and shoreline erosion rates at local and low-wave-energy shorelines (Cowart et al., 2010; Ravens et al., 2009), The regression equation that Schwimmer (2001) found is

$$y = 0.35x^{1.1} \; , \tag{3.3}$$

where y is erosion rate and x is onshore wave power. Equation 3.3 is close to a linear relationship using nine marsh shoreline sites in Rehoboth Bay, where the analyzed wave power ranged from 0.66 Kw/m to 9.21 Kw/m. It is similar to what Kamphuis (1987) found,

$$y = 1.06x^{1.37} \tag{3.4}$$

for erosion rates of glacial bluffs along the north shoreline of Lake Erie, where wave power was observed to be in the same range as Rehoboth Beach. Our relationships $y = -33.71 \times x^{0.86}$ and $y = -54.06 \times x - 7.6 \times 10^{-2}$ reveal that the relationship between wave power and erosion rates in CB is also nearly linear but the multiplicative coefficient of the wave term is two orders of magnitude higher than Equation 3.3. Note that negative

represents erosion in this study, while positive represents erosion in the other studies. In other words, if applying the Schwimmer (2001) empirical relationship to our data, the highest 0.03 Kw/m wave power would result in $7.4×10^{-3}$ m/year erosion rate, while the observed erosion rate was about 1.53 m/year. Average wave power ranges from $1.6×10^{-3}$ Kw/m to $2.4×10^{-2}$ Kw/m for the marsh shorelines in our study. This order-of-magnitude smaller wave-power environment of CB, compared to Rehoboth Beach, leads to the orders-of-magnitude difference in the multiplicative coefficients of wave power term.

A study in Galveston Bay, Texas, where wave climate is also much smaller than Rehoboth Beach, found that the Schwimmer (2001) relationship leads to substantial underestimation of erosion rates (Ravens, 2009). And, a study in Hog Island Bay, Virginia, used linear regression fits to acquire coefficients around 130 (units converted from w/s in the original paper to Kw/m) for its slopes between wave power and erosion rate (Mcloughlin et al., 2015), which is about two times of 54.06 from this study, implying that applying the Schwimmer (2001) equation to Hog Island would end up with underestimation as well (figure 3.13). This suggests that it might not be appropriate to apply empirical relationships of erosion rates with wave power to coasts/estuaries with significantly different wave climates.

Equation 3.1 reveals that, even using power function, the relationship is quite close to linear. Also, the linear regression has a similar R^2 compared to the power-function curve fitting. Marani et al. (2011) demonstrate a linear relationship for marsh edge erosion using dimensional analysis:

$$\frac{Rhc}{\overline{P}} = f\left(\frac{h}{d}\right),$$
(3.5)

where R is the erosion rate; h is cliff face height with respect to the tidal-flat bottom; c is the sediment effective cohesion; $\overline{P}$ is the average wave power; and d is the tidal-flat bottom depth with respect to mean sea level. If $\dfrac{f\left(\frac{h}{d}\right)}{c}$ doesn't appreciably depend on $\dfrac{h}{d}$ in the range of values covered by the data, and the erodibility of soil isn't dramatically different, the volumetric erosion rate (V) is a linear function of the average wave power,

$V = Rh = a\overline{P}$.

3.4.3.2 Wave power from SWAN versus from linear wave theory

In this study, the wave-energy flux is calculated as

$transp\_onshore = \dfrac{\sum Transport\ of\ Wave\ Energy\ Onshore}{Counts\ of\ All\ Wave\ Energy\ Occurances}$ instead of divided only by the counts of

onshore events. In this way, the potential bias of only including onshore events in the wave climate is avoided, as discussed in great detail by Mcloughlin et al. (2015). In previous studies (Marani etc., 2011; Schwimmer, 2001; Mcloughlin et al., 2015), the wave-energy flux is mostly calculated using linear wave theory

$$P = \frac{1}{8}\rho g H_s^2 C_g \cos\alpha \qquad\qquad (3.6)$$

where ρ is water density; g is gravitational acceleration; H_s is significant wave height; C_g is group velocity.

We also calculated the wave-energy flux as above (Equation 3.6) to compare with the output wave power from SWAN used in this study, using $\rho = 1008 kg/m^3$, g= $9.81 m/s^2$, and H_s is Hsig from SWAN. The same α is used as described in the methods

section 3.2. The definition of α can be slightly different, thus $\cos\alpha > 0$ might end up as onshore or offshore component in different studies, but it makes no change in absolute value. C_g is the group velocity calculated by the minimum value between $\sqrt{gd}$ (shallow-water) and $\frac{gT}{4\pi}$ (deep-water), d is the water depth, and T is peak period (variable 'Tps') from SWAN. Applying $C_g = \sqrt{gd}$ on deep-water waves results in overestimating group velocity (Koch et al., 2006). Even though CB is a shallow-water environment, wind generated waves can still behave as deep-water waves due to their small wavelength.

Linear wave power calculated this way is approximately 2.3 times the wave-power output from SWAN, if the minimal intercept in the equation is ignored, expressed as:

$$\text{LinearWavePower}(\text{onshore}) = 2.33 \times \text{TranspOnshore} - 5.28 \times 10^{-4} \qquad (3.7)$$

In SWAN, Hsig is calculated as

$$H_s = 4\sqrt{\iint E(\sigma,\theta)d\sigma d\theta} \qquad (3.8)$$

Thus, the wave power calculated using linear wave theory should be calculated as

$$P = \frac{1}{8}\rho g H_s^2 C_g \cos\alpha = 2\rho g C_g \iint E(\sigma,\theta)d\sigma d\theta \cos\alpha \qquad (3.9)$$

SWAN calculates wave power using

$$Transp = \sqrt{Transp_x^2 + Transp_y^2} = \rho g \sqrt{\left(\iint C_{gx}E(\sigma,\theta)d\sigma d\theta\right)^2 + \left(\iint C_{gy}E(\sigma,\theta)d\sigma d\theta\right)^2} \qquad (3.10)$$

Where $E(\sigma,\theta)$ is the variance density spectrum; σ is the absolute radian frequency

determined by the Doppler shifted dispersion relation; and θ is the wave direction.

If Cg is not a function of σ or θ, Equation 3.10 can be rewritten as

$$Transp = \rho g \iint E(\sigma,\theta)d\sigma d\theta \sqrt{\left(C_{gx}\right)^2 + \left(C_{gy}\right)^2} = \rho g C_g \iint E(\sigma,\theta)d\sigma d\theta \quad (3.11)$$

Thus,

$$TranspOnshore = \rho g C_g \iint E(\sigma,\theta)d\sigma d\theta cos\alpha \quad (3.12)$$

gives us P=2×TranspOnshore. In reality, C_g is a function of σ, so the factor between

Equation 3.9 and Equation 3.12 cannot be exactly 2. This explains the 2.33 slope and

negligible intercept in the linear relationship between 'Linear Wave Power Onshore

Component' and 'Transp Onshore' (Figure 3.16). The reason why there is a factor of two

between the wave power from the integration of the wave spectrum and the wave power

calculated from linear wave theory is not clear to us, but we are more confident in the

former rather than the latter, because it is the actual spectral definition of wave power

rather than a linear equation for monochromatic waves applied to random waves.

3.4.4 Comparisons of wave climate and erosion rates among different scales

(reach, grid cells, transects)

We were able to establish the quantitative relationship (Equation 3.1; Equation

3.2) using averaged marsh data at reach resolution but not with the higher resolution

marsh data in grid cells. At sites fringing Hog Island Bay, VA, the relationship between

wave power and erosion rates was significant when variables were averaged over the

entire length of a marsh edge, but no significant correlations were found for individual

segments of the marsh shorelines (Mcloughlin et al., 2015). Higher resolution data

deliver more information while increasing the scatter and noise, and thus decrease the

correlation. Averaged data result in higher correlation but resolve less information. In the

linear analysis of this study, the correlation between potential variables and erosion rates

of 'marsh' (in reach resolution) is higher but generally agrees well with counterparts of

group 'marshHD' (in grid cell resolution). Data at the reach resolution were able to

reasonably demonstrate the relationship between wave characteristics and erosion rates

(Table 3.2, Figure 3.13). However, the average of erosion rates along a certain length of

shoreline can differ significantly from the erosion rates at some of its local sites,

especially small tributaries and creeks (Figure 3.14).

As a result, quantitative results of this study are not necessarily applicable to local

scales. When comparing sites at highly local scales, data should be acquired in a

reasonably higher resolution. At each local site, the factors affecting erosion might affect

the erosion process quite differently. Thus, conclusions and results analyzed from one

local site need to be verified cautiously before applying to other locations.

3.4.5 Climate change

Increase in storm intensity and frequency due to climate change will accelerate

erosion process (Finkelstein and Hardaway, 1988). If the local wave regime is sensitive

to water depth, then sea-level rise can potentially lead to higher wave power impinging

on shorelines due to longer fetches and deeper estuaries/coasts.

Storms are episodic and can cause cliff failure and severe shoreline retreats

(Adams et al., 2005; Brain et al., 2014; Earlie et al., 2015). Extreme wave events (top 5%

67

and top 10% of Hsig) and erosion rates tend to have a higher correlation than average wave climate (Table 3.2). To quantify the influence of storm events on erosion process, local measurements of shoreline retreat before and after storms and wave observations during storms are required. However, this is beyond the scope of our study.

3.5 Conclusions

In Maryland CB, marsh shorelines generally erode faster than bank shorelines, and the Bay sides of islands undergo the most severe erosion. Most marsh shorelines are unprotected and are located on the lower eastern shore. Most high erosion corresponds to high wave power but high wave power does not necessarily lead to high erosion, consistent with previous studies (Wang et al., 1982). Marsh shorelines have strong linear correlations between erosion rates and wave climate, while bank shorelines show weaker correlations. A correlation coefficient of -0.75 between onshore wave-energy flux and erosion rates strongly suggests that wave power can be the dominant factor for marsh shoreline erosion in CB. We speculate that wave power can be a dominant factor for bank shorelines as well, but other confounding factors hindered it from arising in our statistical analysis. A nearly linear, statistically significant relationship between wave-energy flux and erosion rates was found using marsh data in CB, but its coefficient is orders of magnitude different than some of the literature from other locations (Schwimmer, 2001; Kamphuis, 1987) due to the significantly smaller wave climate in CB. This suggests that it might be inappropriate to apply empirical relationships of erosion rates versus wave power to coasts/estuaries with significantly different wave climates.

Marsh has been assigned high economic and ecological values (Groot et al., 2012) for buffering eutrophication and serving ecosystems (Bricker and Stevenson 1996).

Knowing marshes are more erodible than banks, which is counter intuitive, and experiencing high wave power exposure in Maryland CB, management strategies should favor marsh protection over bank protection by building 'living shorelines' (http://dnr.maryland.gov/ccs/livingshorelines.asp). Meanwhile, evaluation of the feasibility of non-structural shoreline protection practices should consider the degree of wave power exposure of marsh shorelines.

The number of influential factors, their interactions, and the spatial and temporal variability of their effects on erosion rates ensure that understanding how different influential factors impact erosion rate in CB is a complex problem. If carrying out a comprehensive study, all potential influential factors should be considered: whether inundation due to sea level rise and land subsidence is significant in measured erosion rate, shoreline elevation and orientation, sediment type and stability, soil mechanics (e.g., yield stress, ground water content), shoreline type (protected vs unprotected), effects of different vegetation and frequency of inundation, currents and longshore transport, human activity (including boat wakes), and last but not least, wave climate.

Neither linear analysis nor non-linear analysis was successful in building a single comprehensive empirical model for erosion rate prediction for the Maryland portion of CB in this study. A higher resolution and more comprehensive data set, which includes more geological and geographical information besides wave characteristics, will be needed for building reliable erosion rate prediction models, especially at local scales. Starting from building local models for relatively simple environments will be more straightforward. Then developing models for more complicated environments and finally

generalizing them into a comprehensive model for a reasonable size of domain seems like
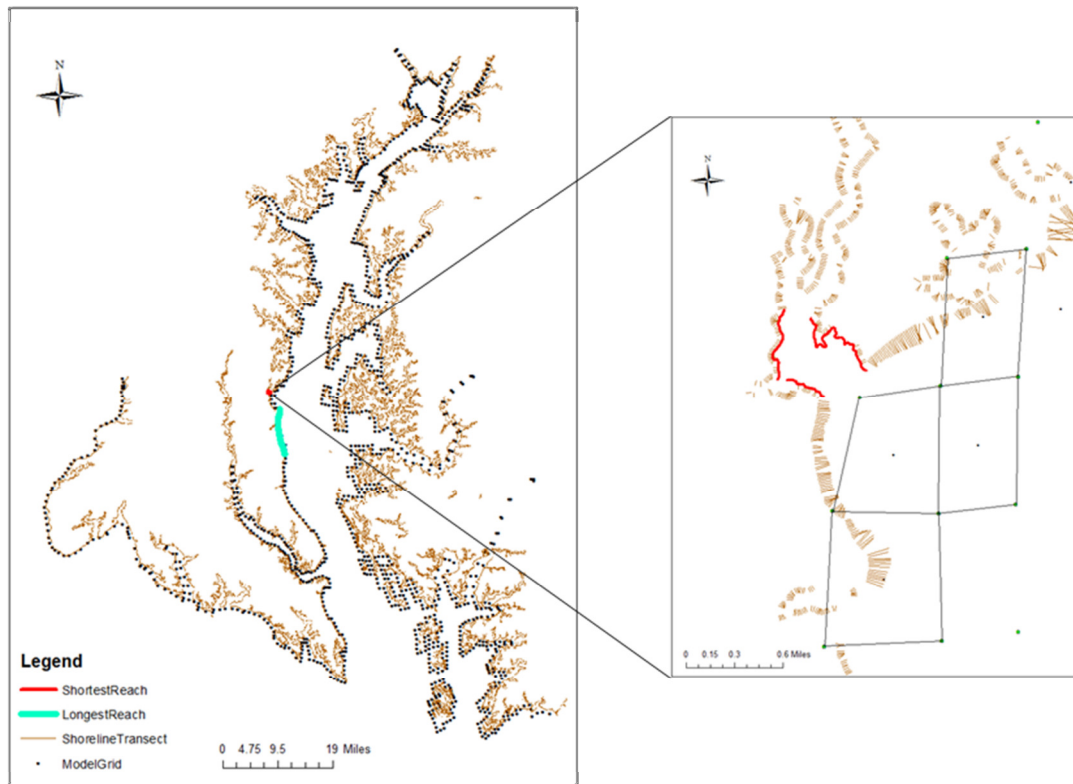
a good approach.

Figure 3.1. Illustration of the length of a transect, model grid and reach. Left: transects (brown lines) and model grids (black dots), with shortest/longest reach labeled in red/blue; right: enlarged local area from the shortest reach. Irregular rectangles represent the sizes of model grids.
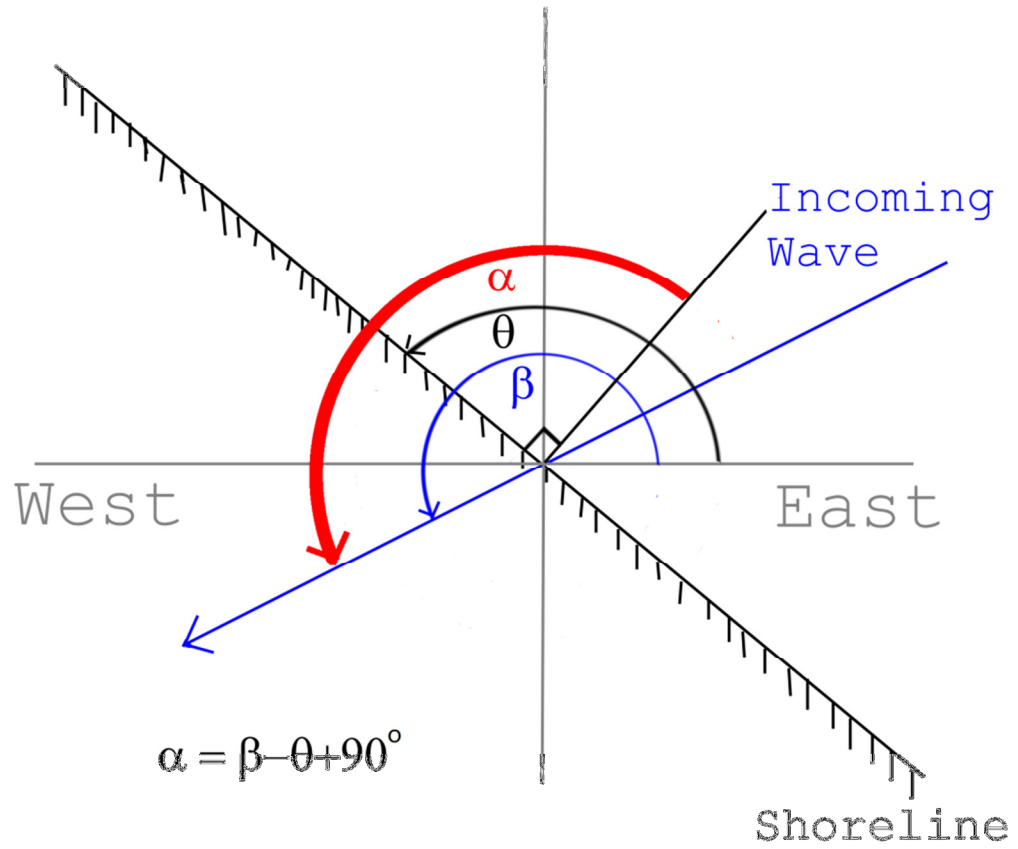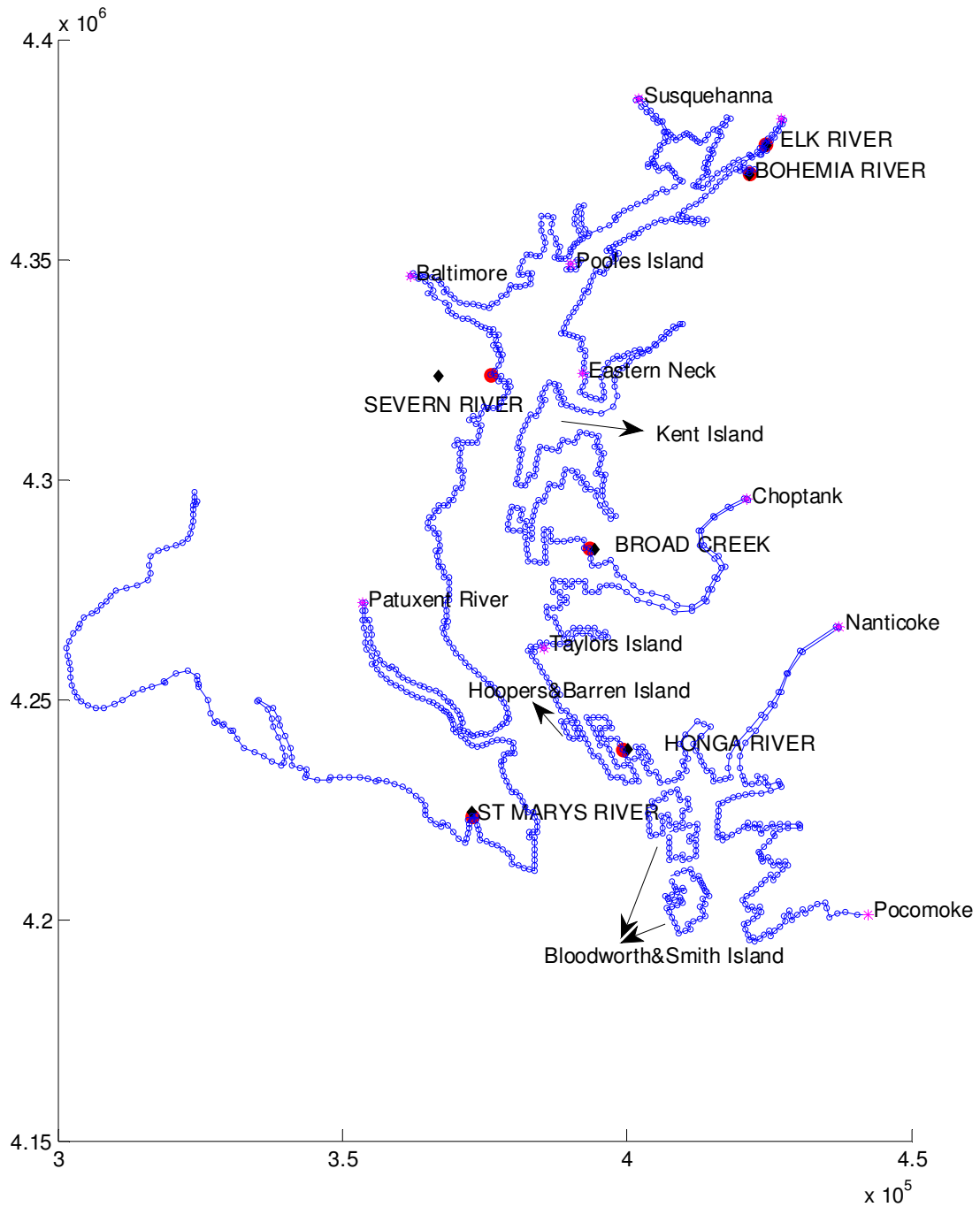
Figure 3.2. Demonstration of the calculation of α.

Figure 3.3. Map of Maryland CB shorelines with landmarks. Red dots label six local sites: Bohemia River, Broad Creek, Elk River, Honga River, St Mary River and Seven River, where model outputs were obtained. Black dots are where observations of these 6 sites were taken. Magenta star labels landmarks and Islands.
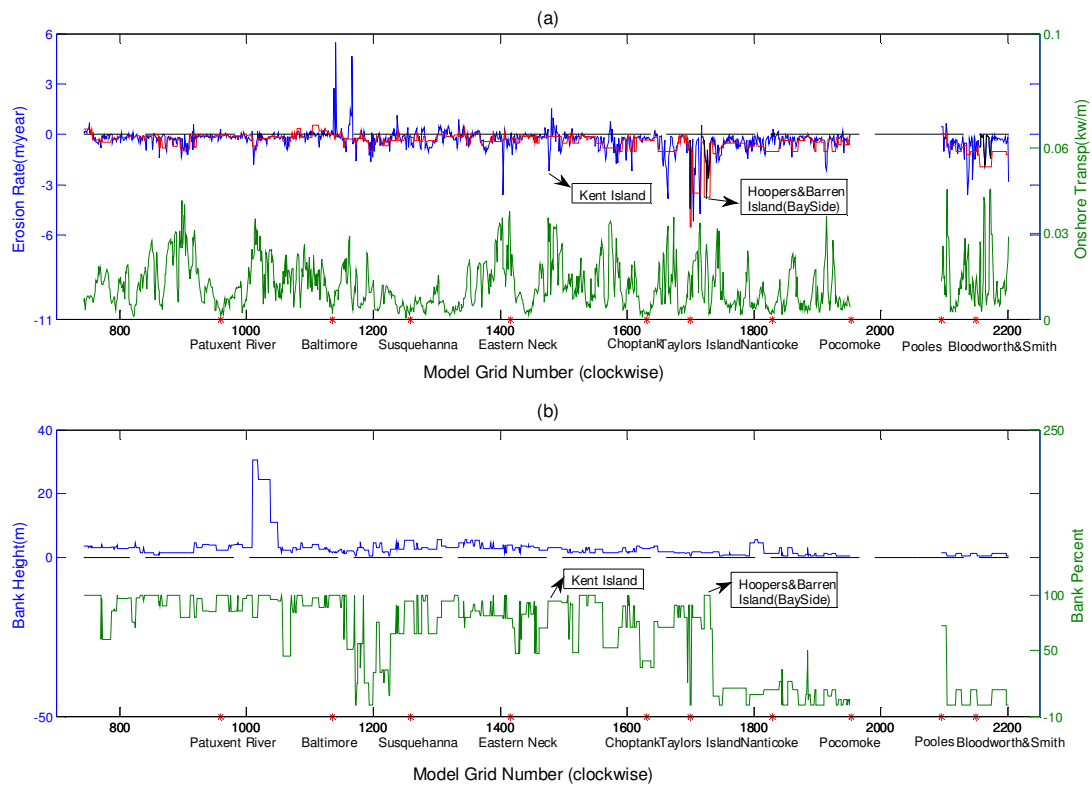
Figure 3.4. Distributions of erosion rates, onshore wave power, bank height and bank percentage along CB shorelines in Maryland. Black horizontal dash line in both figures identifies zero. Landmarks are labelled for the head of tributaries and islands. (a) Erosion and onshore wave power: erosion rates (blue line) and wave power (green line) are at model grid resolution; the red line shows erosion rates at reach resolution. Grid cells numbers in Maryland range from 743 to 1952 and from 2096 to 2201. The absence near cell 2000 is the excluded shoreline of Virginia. The erosion rate at grid cell resolution is underestimated at the Bay sides of Taylors Island, Hoopers Island, Barren Island, Bloodworth Island, and a portion of the Bay side of Smith Island and indicated by the black solid line. (b) Bank height and bank percentage: both are at reach resolution, but the value in each reach is assigned to the corresponding grid cells; bank height is assigned 0.5m for marsh type.
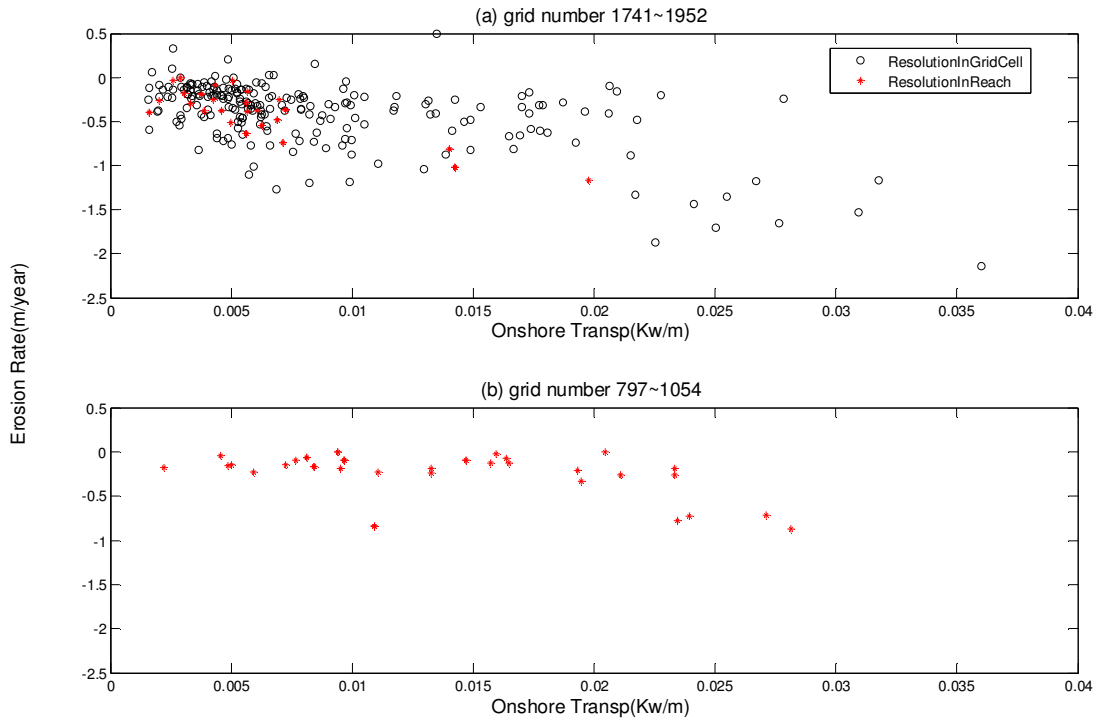
Figure 3.5. Relationship between erosion rates and the onshore wave power on the (a) lower eastern shore, mostly composed of marsh, with the grid-cell and reach resolutions indicated by black circles and red stars, respectively; (b) lower western shore, mostly composed of bank, with only the reach resolution shown.

Bank:Pearson Correlation Coefficients Matrix (21 * 21) **Marsh:Pearson Correlation Coefficients Matrix (21 * 21)**
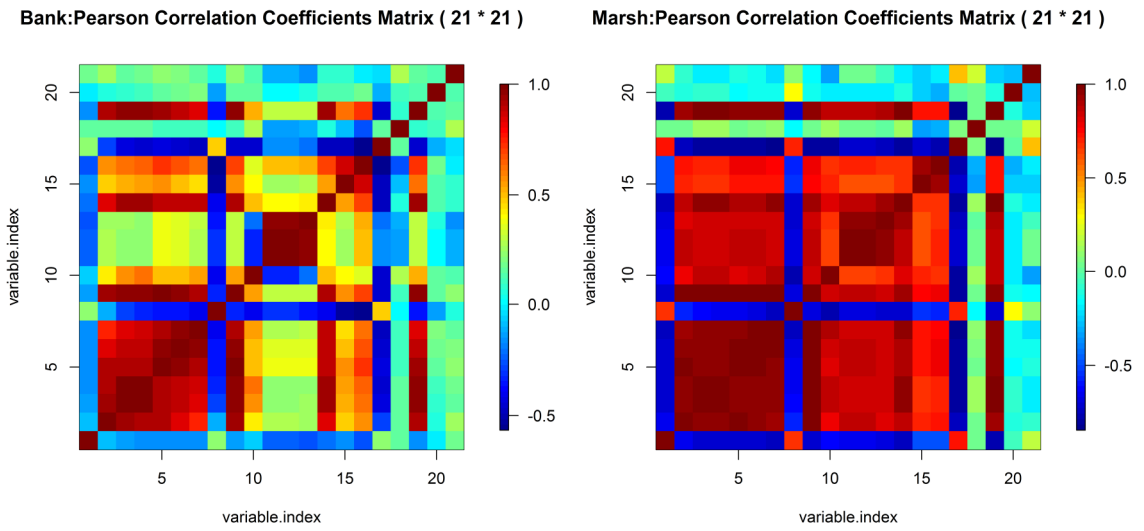
Figure 3.6. Pearson correlation coefficients matrix of erosion rates and all potential influential variables; index of variable is the same as the index in Table 3.3
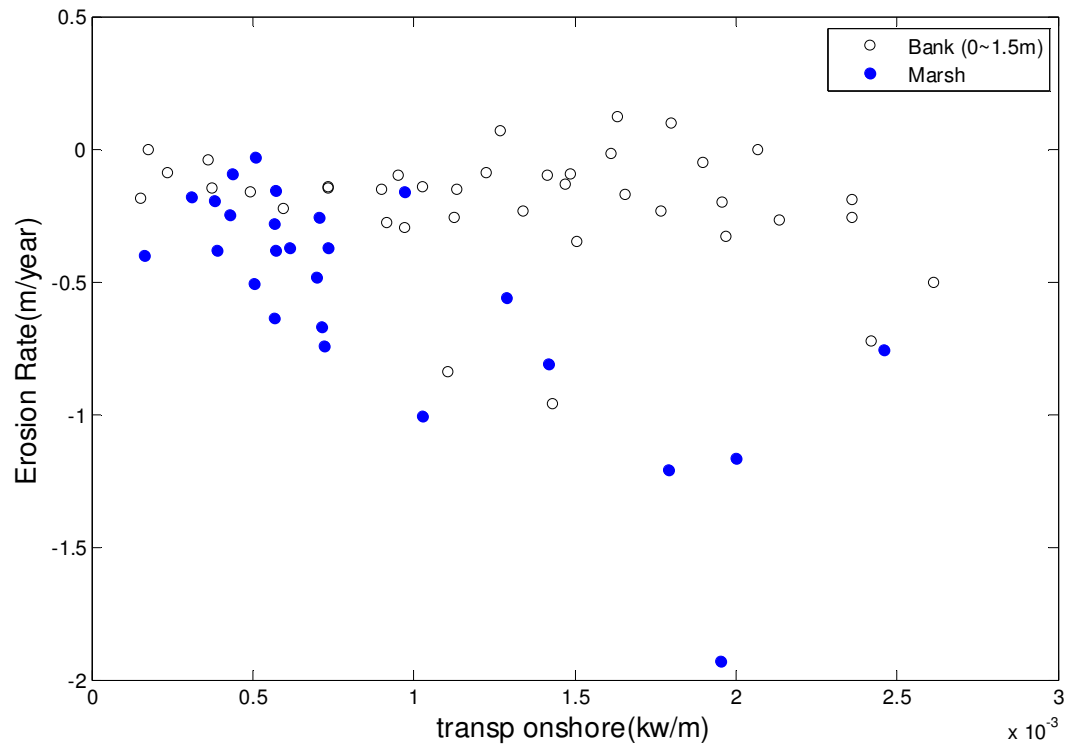
Figure 3.7. Comparison of erosion rates between marshy shorelines and low-elevation banks (0-1.5m) versus onshore wave power.
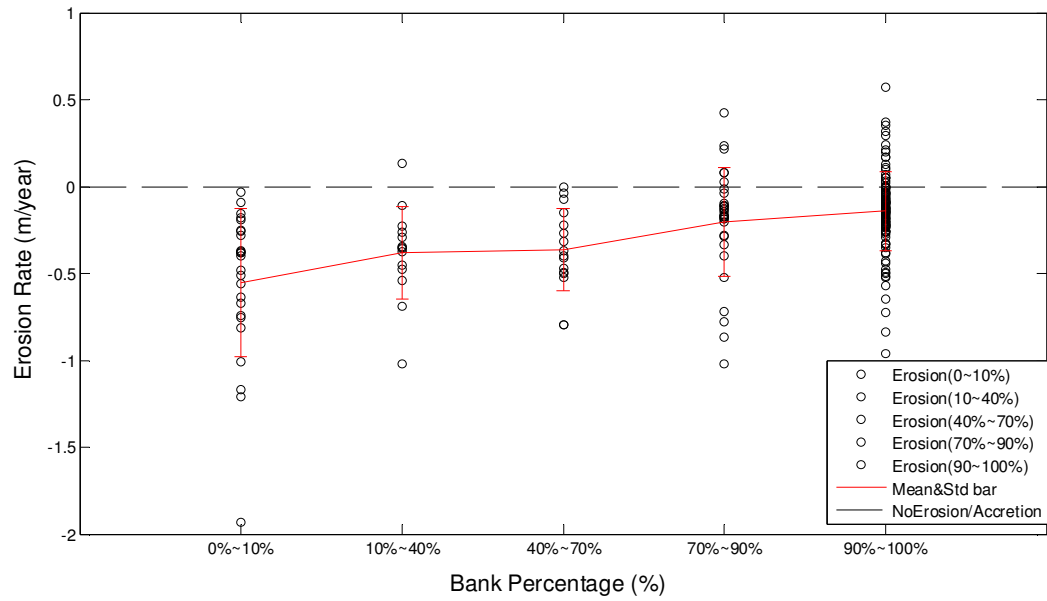
Figure 3.8. Distribution of erosion rates categorized by bank percentage. Marsh is defined as 0~10%, while 90%~100% is defined as bank. The red line connects the mean erosion rate of each category; the red bar shows the standard deviation of erosion rates within each category.
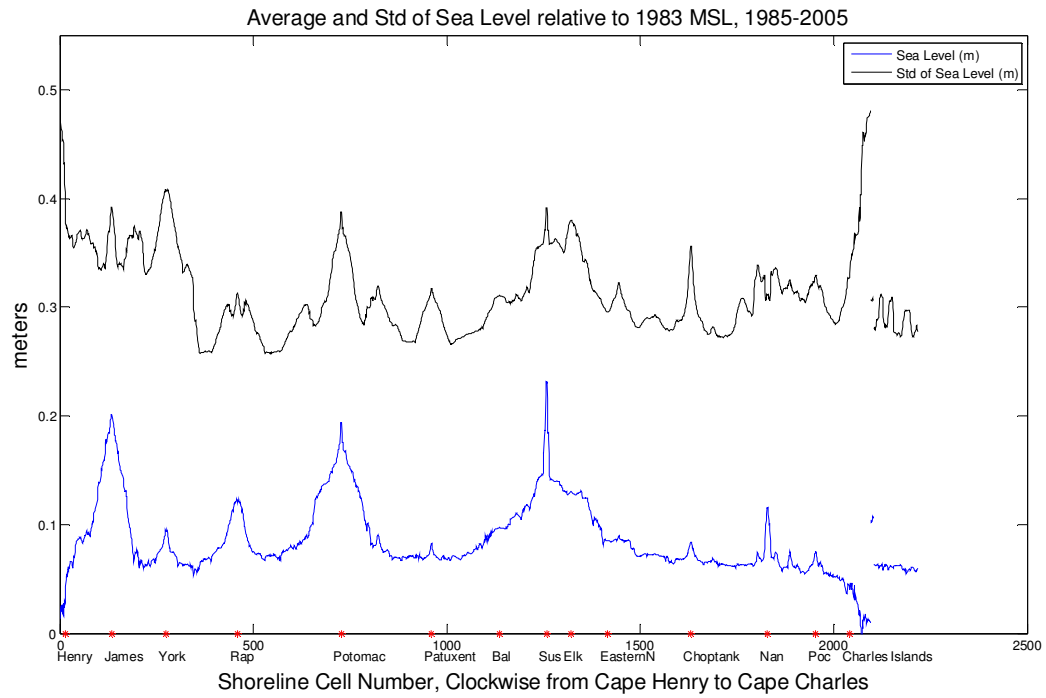
Figure 3.9. Average sea levels and the standard deviation (Std) of sea level from 1985 to 2005, clockwise along Chesapeake Bay shoreline.
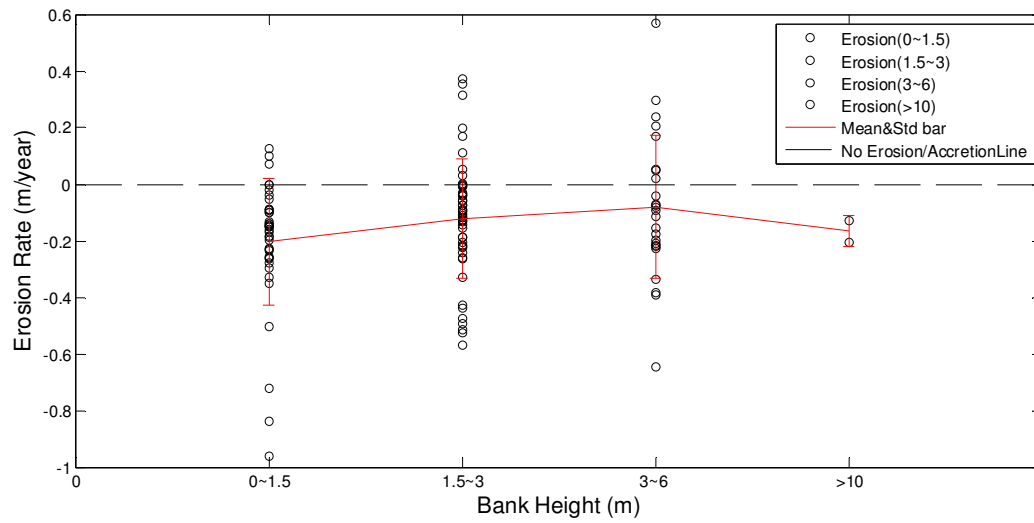
Figure 3.10. Erosion rate distributions categorized by bank height. The red line connects the mean erosion rate of each category; the red bar shows the standard deviation of erosion rates within each category.
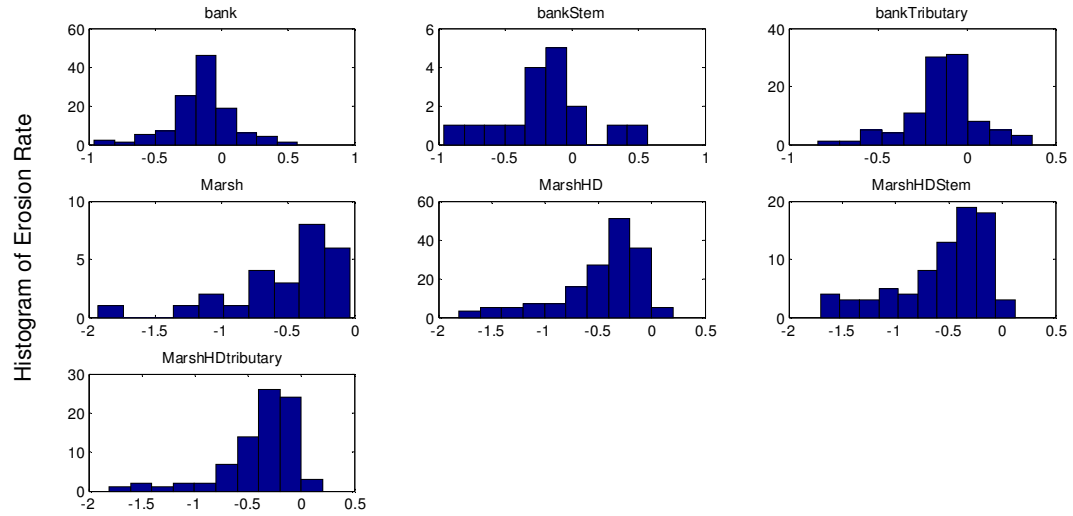
Figure 3.11. Histograms of erosion rates for different groups of data: bank, marsh, stem of banks (bankStem), tributaries of banks (bankTributary) in reach resolution; marsh (MarshHD), stem of marsh (MarshHDStem), tributaries of marsh (MarshHDtirbutary) in grid cells resolution.
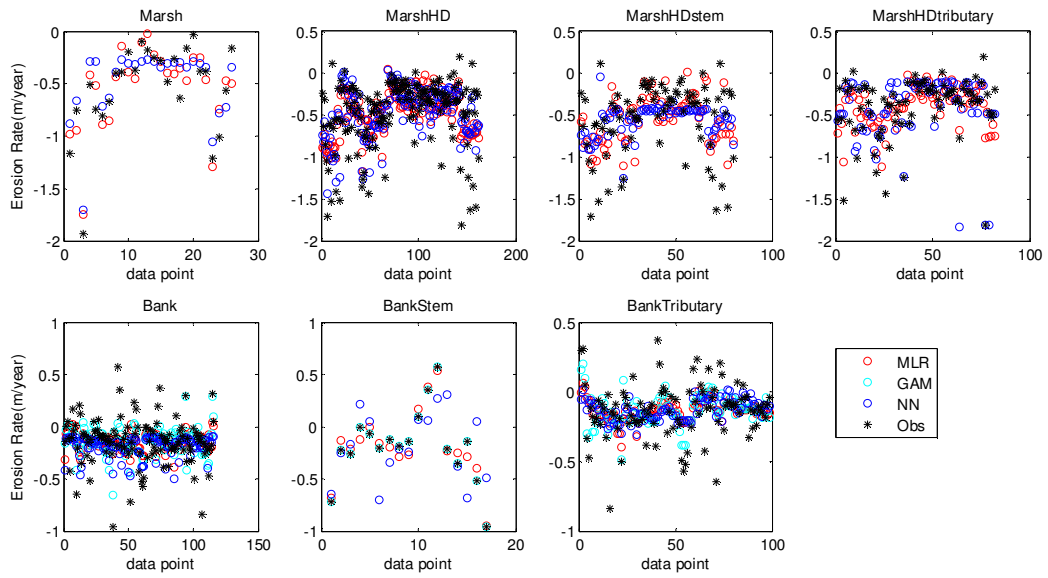
Figure 3.12. Plot of observed erosion rates and simulated erosion rates from statistical models: Multiple Linear Regression (MLR), Generalized Additive Model (GAM), and Neural Network (NN) analysis.
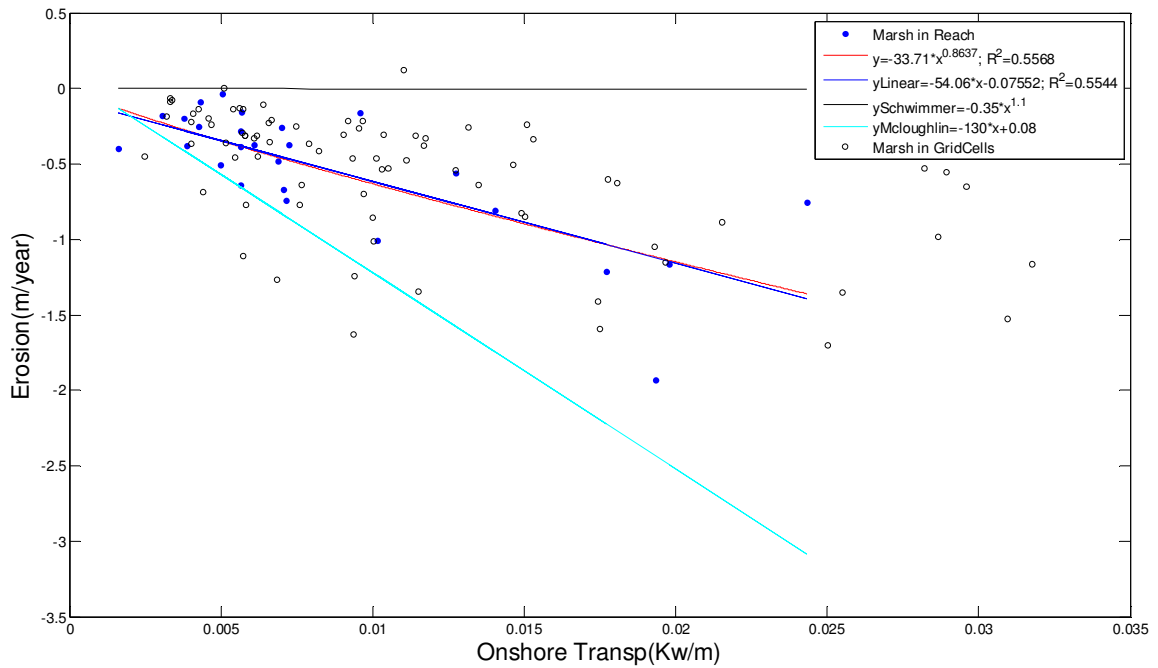
Figure 3.13. Curve fitting of marsh data for erosion versus the onshore component of the wave-energy flux. Blue dots are data at reach resolution; the red line is the best power function curve fitting, the dark blue line is the best linear curve fitting, black line is Schwimmer's equation, light blue line is Mcloughlin's linear relationship, and the black dots are data at grid-cell resolution. All fittings and equations are applied on marsh data of Maryland CB in reach resolution.
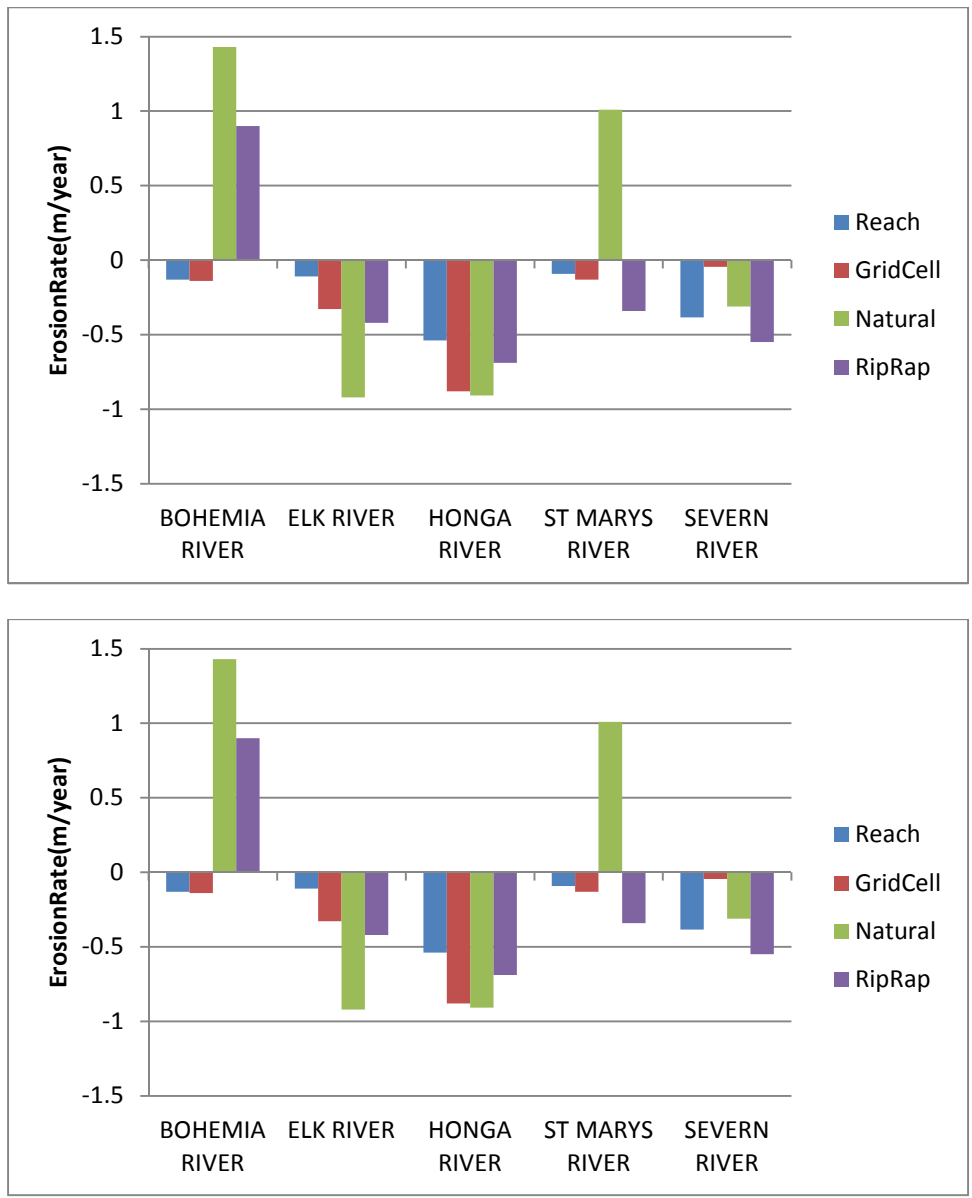
Figure 3.14. Erosion rates at three different scales (reach, grid cell and transect). 'Reach' and 'Grid Cell' are the averaged erosion rates over the length of the corresponding reach or grid cell that include the site. 'Natural' (unprotected) and 'Rip Rap' (protected) is the erosion rate from its closest transect.
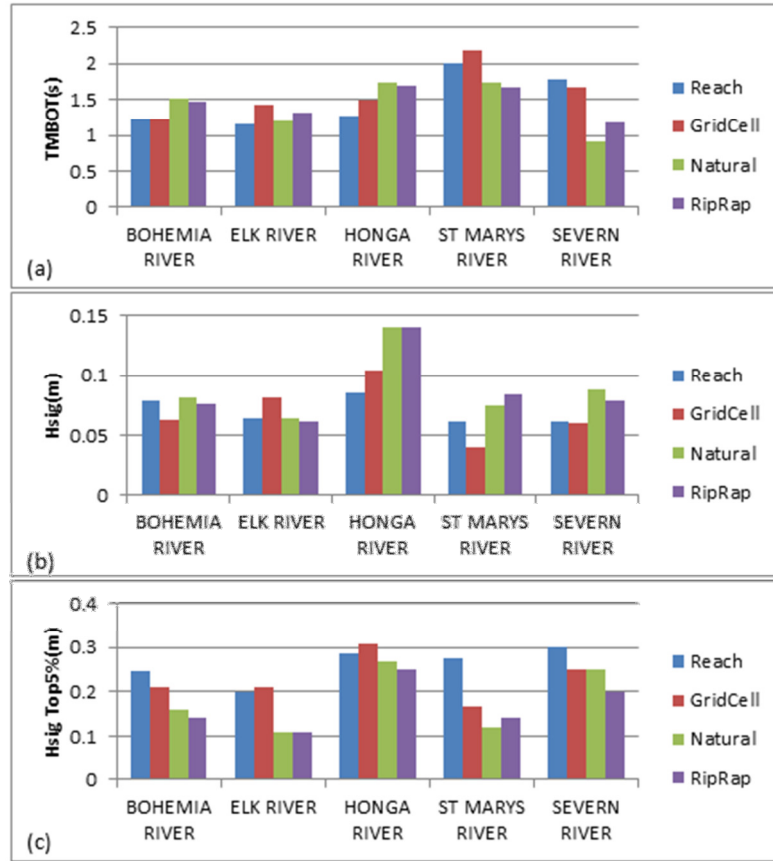
Figure 3.15. Plots of (a) bottom peak period (TMBOT), (b) significant wave height (Hsig), and the (c) top 5% of Hsig for three different scales (reach, grid cell and local site). 'Reach' and 'Grid Cell' represent the long term-averaged (1985-2005) value for the site. Natural (unprotected) and Rip Rap (protected) represent the measurements from local short-term deployments.
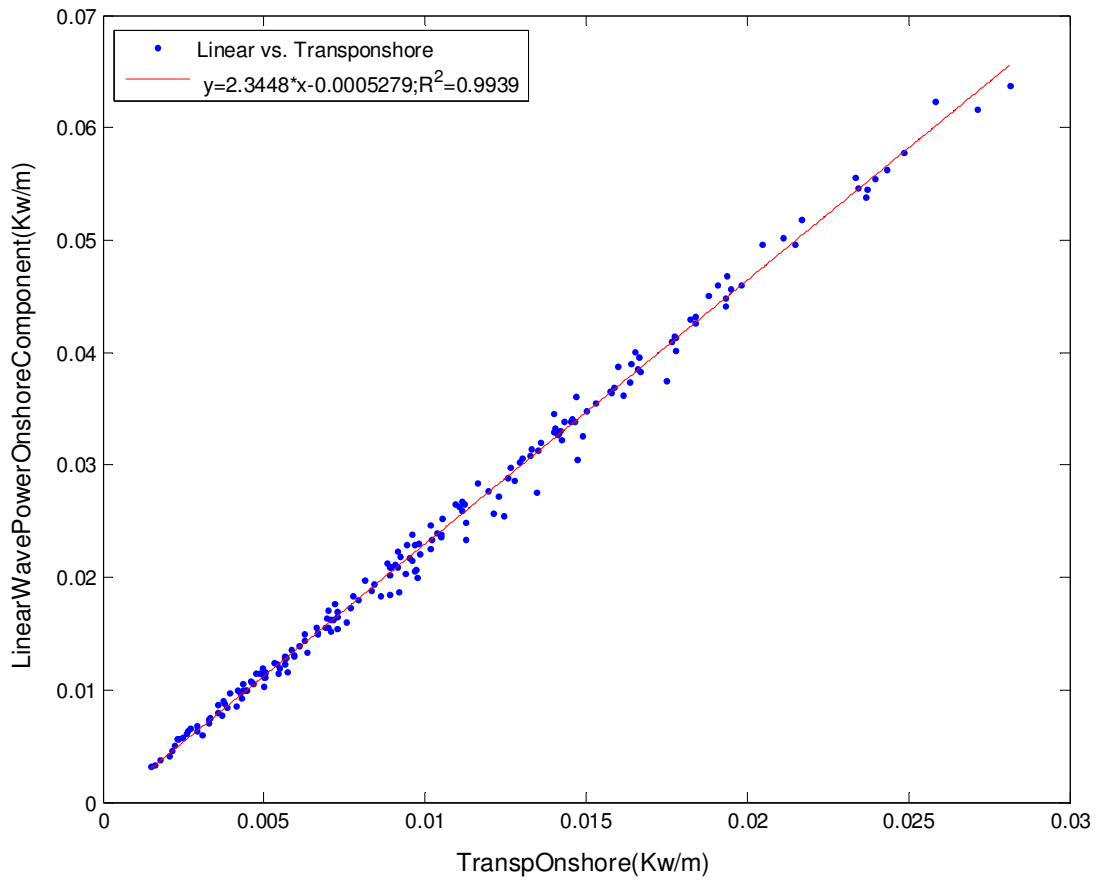
Figure 3.16. Comparison of onshore transport of wave-energy flux (wave power) calculated using linear wave theory with significant wave height and as output by SWAN.

Erosion	Erosion Rate (in meter/year)
Bank Height	Elevation of Bank (in meter)
Bank Percentage	Percentage of bank (versus Marsh) at each Reach (dimensionless)
Hsig	Significant Wave Height (in meter)
Hsig90	Top 10% of Hsig (in meter)
Hsig95	Top 5% of Hsig (in meter)
Tps	Smoothed Peak Period (in second)
TM01	Mean absolute wave period (in second)
WLEN	Average wave length (in meter)
FSPR	The normalized width of the frequency spectrum (dimensionless)
LWAVP	Peak wave length (in meter)
TMBOT	The bottom wave period (in second)
URMS	The Root Mean Square of the orbital velocity near the bottom (in m/s)
UBOT	The Root Mean Square of the maxima of the orbital velocity near the bottom (in m/s)
UBOTsq[a]	$UBOT^2$;proportional to bottom shear stress
transp_onshore	the onshore component of transport of energy (in m^3/s); equals to transpall*cosα
Fetch	Fetch at each reach(in meter)
Weighted fetch	Wind weighted fetch(in meter)
Tidal Range	Tidal range calculated from sea level (in meter)
Bath_steepness	Steepness of bathymetry at each reach (dimensionless)
Transpall	Transport of all wave energy (scalar, in m^3/s)
drift	Longshore drift =tranpall*sinα*cosα (in m^3/s);
Sea Level	Relative sea level to the mean sea level
MedianWater60Hsig	The median of sea level, of which corresponding Hsig is greater than 60% maximum of Hsig

Table 3.1. Definition of variables

	Type	Marsh		Bank		Bank Stem	Bank Tributary	Bank Eastern	Bank Western
	Correlated Variable	Erosion	VER[a]	Erosion	VER[a]	Erosion			
		R>0.5, P<0.05		R>0.2, P<0.05					
	number of data/reach	26		116	115	17	99	43	73
1	Erosion	1		1	0.88	1	1	1	1
2	Hisg	-0.66		*-0.09[b]*					
3	Hsig90	-0.71		*-0.14[b]*			-0.20		
4	Hsig95	-0.72		*-0.15[b]*			-0.21		
5	Tps	-0.70							
6	TM01	-0.71							
7	WLEN	-0.74					-0.21		
8	FSPR	0.67		0.23			0.27	0.34	
9	LWAVP	-0.72							
10	TMBOT	-0.64							
11	URMS	-0.65		-0.23		-0.51			
12	UBOT	-0.65		-0.23		-0.51			
13	UBOTsq	-0.68		-0.26		-0.52			-0.29
14	transp_onshore	-0.75		-0.21			-0.24		-0.24
15	fetch						-0.31	-0.46	
16	Weighted fetch			-0.27			-0.28	-0.39	
17	Tidal Range	0.73		0.23			0.27		0.24
18	Bath_steepness					0.56			
19	transpall	-0.77							
20	drift						-0.23		
21	MedianWater60Hsig							*-0.22[b] (P~0.16)*	0.34

a: VER(Volumetric Erosion): VER(bank)= ErosionRate*BankHeight; VER(marsh)=ErosionRate*0.5

b: doesn't fit in corresponding R and P category but included in table for additional information.

Table 3.2. Pearson correlation coefficients between erosion rates (or volumetric erosion) and selected potential influential variables at reach resolution.

Type	MarshHD[a]	MarshHD Stem	MarshHD Tributary	MarshHDStem Eastern	MarshHDStem Western
	R>0.35, P<0.01				R>0.35,P<0.05
number of data/reach	162	80	67	82	13
Hisg	-0.49	-0.43	-0.49	-0.53	
Hsig90	-0.53	-0.49	-0.53	-0.58	
Hsig95	-0.53	-0.49	-0.53	-0.57	
Tps	-0.56	-0.51	-0.56	-0.62	-0.55
TM01	-0.54	-0.51	-0.53	-0.61	-0.56
TM02	-0.50	-0.48	-0.47	-0.57	
WLEN	-0.49	-0.46	-0.46	-0.56	
LWAVP	-0.55	-0.50	-0.58	-0.61	-0.57
TMBOT	-0.49	-0.46	-0.48	-0.54	
URMS	-0.38			-0.44	
UBOT	-0.38			-0.44	
UBOTsq	-0.40	-0.37	-0.36	-0.47	
trasnp_along			-0.37		
transp_onshore	-0.45	-0.41	-0.46	-0.48	
fetch	-0.46	-0.36	-0.61	-0.65	
Weighted fetch	-0.45		-0.56	-0.66	
Tidal Range	0.45	0.41	0.45	0.55	
transpall	-0.50	-0.45	-0.54	-0.55	-0.58

a: HD is appended as a symbol for data in the resolution of model grid to differentiate from data in Table 2

Table 3.3. Pearson correlation coefficients between erosion and selected potential influential variables for marsh data at the resolution of model grid cells (outliers excluded).

	MLR (explanatory variables)	NN and GAM (excluded variables)
Bank	Erosion ~ UBOTsq + TidalRange + MedianWater60Hsig	"fetch" "URMS" "UBOT" "Hsig90" "Tps" "TM01" "WLEN" "transpall"
Bank stem	Erosion ~ Bank.height + Hsig95 + FSPR + TMBOT + UBOT + UBOTsq +transp_normal + Weighted.fetch + transpall + drift +MedianWater60Hsig	"fetch" "URMS" "Hsig90" "Tps" "TM01" "WLEN" "LWAVP"
Bank tributary	Erosion ~ fetch_coast + Tidal.Range + MedianWater60Hsig	"Weighted_fetch" "URMS" "UBOT" "Hsig95" "Hsig90" "Tps" "TM01" "WLEN" "LWAVP"
Marsh	Erosion ~ FSPR + TMBOT + UBOTsq + fetch_coast + TidalRange +Bath_steepnees + transpall + sealevel	"Weighted_fetch" "URMS" "UBOT", "Hisg" "Hsig90" "Hsig95" "Tps" "TM01" "WLEN" "LWAVP"
MarshHD	Erosion ~ Tps + transp_onshore + fetch_coast + TidalRange + sealevel	"Weighted_fetch" "URMS" "UBOT" "Hsig95" "TM01" "WLEN" "LWAVP" "transpall"
MarshHDstem	Erosion ~ Hsig90 + Tps + TMBOT + drift + sealevel + MedianWater60Hsig	"Weighted_fetch" "URMS" "UBOT" "Hsig95" "TM01" "WLEN" "LWAVP" "transpall"
MarshHD tributary	Erosion ~ LWAVP + fetch_coast	"Weighted_fetch" "URMS" "UBOT" "Hsig90" "Tps" "TM01" "WLEN" "transpall"

Table 3.4. Variables that are included in stepwise Multiple Linear Regression and variables that are excluded using non-metric multidimensional scaling before GAM and Neural Network Analysis.

	Bank (116 samples)			BankStem (17 samples)			BankTributary (99 samples)		
	GAM	NN[c]	MLR	GAM[b]	NN[b,c]	MLR[b]	GAM	NN[c]	MLR
Dof[a]	89	N/A	112	-4	N/A	5	84	N/A	95
AIC	-362	-320	-353	N/A	N/A	-61	-339	-304	-329
AICc	-345	N/A	-353	N/A	N/A	17	-333	N/A	-328
R^2	0.47	0.12	0.12	1	0.31	0.94	0.38	0.13	0.14
AdjustedR^2	0.30	N/A	0.10	N/A	N/A	0.81	0.28	N/A	0.11
P-value	2.9×10^{-17}	1.3×10^{-4}	10^{-4}	3×10^{-101}	1.9×10^{-2}	1.1×10^{-10}	8.3×10^{-12}	2.2×10^{-4}	1.4×10^{-4}
RMSE	0.028	0.047	0.045	0	0.098	0.007	0.024	0.034	0.033

	Marsh (26 samples)		MarshHD (162 samples)		MarshHDstem (80 samples)		MarshHDtributary (82 sample)	
	NN[c]	MLR	NN[c]	MLR	NN[c]	MLR	NN[c]	MLR
Dof[a]	N/A	17	N/A	156	N/A	73	N/A	79
AIC	-60	-72	-264	-366	-82	-165	-116	-212
AICc	N/A	-61	N/A	-366	N/A	-164	N/A	-211
R^2	0.84	0.82	0.46	0.40	0.37	0.42	0.46	0.44
AdjustedR^2	N/A	0.74	N/A	0.38	N/A	0.37	N/A	0.42
P-value	5.9×10^{-11}	1.2×10^{-10}	3.7×10^{-23}	3.1×10^{-19}	2.2×10^{-9}	7.1×10^{-11}	1.9×10^{-12}	1.4×10^{-11}
RMSE	0.034	0.031	0.087	0.097	0.123	0.107	0.085	0.070

a: Dof is degree of freedom

b: overfitted due to relatively small sample size. N/A is used to fill unvailable variables due to this matter.

c: degree of freedom is unavalible (N/A) for Neural Network analysis. Thus, AICc and Adjusted R^2 are absent (N/A).

Table 3.5. Degree of freedom, AIC, AICc, R^2, adjusted R^2, P value and Root Mean Square Error (RMSE) of statistical models: Multiple Linear Regression (MLR), Generalized Additive Model (GAM), and Neural Network (NN) analysis.

References

Adams, P.N., Storlazzi, C.D., and Anderson, R.S., 2005. Nearshore wave induced cyclical flexing of sea cliffs. *Journal of Geophysical Research*, **110:** F02002

Amin, S. M. N. and R. G. D. Davidson-Arnott, 1997. A Statistical Analysis of the Controls on Shoreline Erosion Rates, Lake Ontario. *Journal of Coastal Research*, **13**: 1093-1101.

Beale, M., M. Hagan and H. Demuth., 2014a. Neural Network Toolbox, MathWorks Cooperation.
Benumof, B. T., Sorlazzi, C. D., Seymour, R. J., and Griggs, G. B., 2000. The relationship between incident wave energy and seacliff erosion rates: San Diego County, California. *Journal of Coastal Research*, **16**: 1162-1178.

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third generation wave model for coastal regions. 1. Model description and validation. *Journal of Geophysics Research,* **104**: 7649–7666.

Boon, J.D., 1998. Wave climate in Chesapeake Bay. Proceedings of the Third International Conference on Ocean Wave Measurement and Analysis—ASCE, Virginia Beach, VA, pp. 1076–1087.

Boon, J.D., Green, M.O., Suh, K.D., 1996. Bimodal wave spectra in lower Chesapeake Bay, sea bed energetics and sediment transport during winter storms. *Continental Shelf Research,* **16**: 1965–1988.

Borg , I. and P. Groenen, 2005. Modern Multidimensional Scaling: Theory and Applications. Springer Verlag.

Bozdogan, H. , Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions. *Psychometrika*, **52**: 345–370.

Brain, M.J., Rosser, N.J., Norman, E.C., and Petley, D.N., 2014. Are microseismic ground displacements a significant geomorphic agent? *Geomorphology*, **207**: 161-173.

Bricker, S. B., Stevenson, J. C., 1996. Nutrients in coastal waters: a chronology and synopsis of research. *Estuaries,* **19**,337-341.

Bruun, P., 1962. Sea Level Rise as a Cause of Shore Erosion. Journal of Waterway, Port, Coastal and Ocean Engineering. *American Society of Civil Engineers*, **88**: 117-130.

Cavaleri, L., and P. M. Rizzoli, 1981. Wind wave prediction in shallow water: Theory and applications*. Journal of Geophysical Research*, **86**: 10961–10973.

Cerco, C. and M. Noel, 2004. The 2002 Chesapeake Bay Eutrophication Model. US Environmental Protection Agency Chesapeake Bay Program Office, Annapolis, Maryland(EPA 903-R-04-004), 375p.

http://www.chesapeakebay.net/publications/title/the_2002_chesa
peake_bay_eutrophication_model, accessed July 2013.

Chuang, W. S. and W. C. Boicourt, 1989. Resonant seiche motion in the Chesapeake Bay. *Journal of Geophysical Research,* **94**:2105–2110.

Cowart, Lisa, J. P. W., and D. Reide Corbett, 2010. Analyzing Estuarine Shoreline Change: A Case Study of Cedar Island, North Carolina. *Journal of Coastal Research*, **26**: 817–830.

Dalrymple, R. A., et al. , 1986. Bluff recession rates in Chesapeake Bay. *Journal of Waterway Port Coastal and Ocean Engineering*, **112**: 164-168.

Damgaard, J. S. and P. Dong, 2004. Soft cliff recession under oblique waves: Physical model tests. *Journal of Waterway Port Coastal and Ocean Engineering,* **130**: 234-242.

Danforth, W.W., and Thieler, E.R., 1992. Digital Shoreline Analysis System (DSAS) User's Guide, Version 1.0: U.S. Geological Survey Open-File Report 92-355, 18 p.

Dean, R. G., 1991. Equilibrium Beach Profiles: Characteristics and Applications. *Journal of Coastal Research*, **7**: 53-84.

Earlie, S. Claire, Adam, P. Young, Gerd Masselink and Paul E. Russell, 2015. Coastal cliff ground motions and response to extreme storm waves. *Geophysical Research Letters*, DOI: 10.1002/2014GL062534

Elliott, A. J., 1978. Observations of the meteorologically induced circulation in the Potomac estuary. *Estuarine Coastal Marine Science,* **6**: 285-299.

Finkelstein, K. and Hardaway, C.S., 1988. Late Holocene sedimentation and erosion of estuarine fringing marshes, York River, Virginia. *Journal of Coastal Research*, **4**: 447-456.

Gelinas, P.J. and Qidgley, R. M., 1973. The influence of geology on erosion rates along the north shore of Lake Erie. Proceedings 16thConference of Great Lake Research, 421-430.

Goodrich, D. M., 1985. On stratification and wind-induced mixing in the Chesapeake Bay, Ph.D. thesis, 134 pp., State Univ. of New York at Stony Brook, Stony Brook, N. Y.

Groot, Rudolf de, Luke Brander, Sander van der Ploeg, Robert Costanza, Florence Bernard, Leon Braat, Mike Christie, Neville Crossman, Andrea Ghermandi, Lars Hein, Salman Hussain, Pushpam Kumar, Alistair McVittie, Rosimeiry Portela, Luis C. Rodriguez, Patrickten Brinkm, Pieter van Beukering, 2012. Global estimates of the value of ecosystems and their services in monetary units. *Ecosystem Services,* **1**, 50-61.

Halka, J. P., Bergstrom, P., Buchanan, C., Cronin, Thomas M., Currey, L., Gregg, A., Herman, J., Hill, L., Kopecky, S., Larsen, C., Luscher, A., Linker, Lewis, Murphy, D. and Stewart, S., May 2005, Sediment in the Chesapeake Bay and management issues:

tidal erosion processes, Chesapeake Bay Program Nutrient Subcommittee Sediment Workgroup's Tidal Sediment Task Force

Harris, C. K., J. P. Rinehimer and S.-C. Kim, 2012. Estimates of Bed Stresses within a Model of Chesapeake Bay. Proceedings of the Twelfth International Conference on Estuarine and Coastal Modeling, American Society of Civil Engineers, Washington, D.C.

Hasselmann, K. et al., 1973. Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). Dtsch. Hydrogr. Z. Suppl. A, 8(12), 95 p

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall.

Hennessee, L., Offerman, K.A., and Halka, J., 2006. Determining Shoreline Erosion Rates for the Coastal Regions of Maryland (Part 2), Coastal and Estuarine Geology File Report NO. 06-03

Hennessee, L., Valentino, M.J., and Lesh, A.M., 2003a. Updating shore erosion rates in Maryland: Baltimore, Md., Maryland Geological Survey, Coastal and Estuarine Geology File Report No. 03-05, 26 p.

Hennessee, L., Valentino, M.J., and Lesh, A.M., 2003b. Determining shoreline erosion rates for the coastal regions of Maryland (Part 2): Baltimore, Md., Maryland Geological Survey, Coastal and Estuarine Geology File Report No. 03-01, 53 p.

Hennessee, L., Valentino, M.J., Lesh, A.M., and Myers, L., 2002, Determining shoreline erosion rates for the coastal regions of Maryland (Part 1): Baltimore, Md., Maryland Geological Survey, Coastal and Estuarine Geology File Report No. 02-04, 32 p.

Johnson, B., Kim, K., Heath, R., Hsieh, B., and Butler, L., 1993. Validation of a three-dimensional hydrodynamic model of Chesapeake Bay. *Journal of Hydraulic Engineering*, **199**: 2-20.

Kahma, K.K. and Calkoen, C.J., 1992. Reconciling discrepancies in the observed growth of wind-generated waves. *Journal of Physical Oceanography*, **22**: 1389- 1405.
Kamphuis, J.W., 1987. Recession rate of glacial till bluffs. *Journal of Waterway, Port, Coastal and Ocean Engineering*, **113**: 60-73.

Kamphuis, J.W., 1987. Recession rate of glacial till bluffs. *Journal of Waterway, Port, Coastal and Ocean Engineering*, **113**: 60-73.

Koch, E. W., L. P. S., Shih-Nan Chen, Deborah J. Shafer, and Jane McKee Smith ,2006, Waves in Seagrass Systems: Review and Technical Recommendations, p10, System-Wide Water Resources Program and Submerged Aquatic Vegetation Restoration Research Program, Army Corp of Engineers(ERDC TR-06-15).

Langland, M. and T. Cronin, 2003. A Summary Report of Sediment Processes in Chesapeake Bay and Watershed. Water-Resources Investigations Report 03-4123, U.S. Department of the Interior & U.S. Geological Survey.

Leatherman , S. P., Ruth Chalfont, Edward C. Pendleton, Tamara L. McCandless and Steve Funderburk, 1995. Vanishing Lands: Sea Level, Society and Chesapeake Bay Laboratory for Coastal Research, University of Maryland; Chesapeake Bay Field Office U.S. Fish and Wildlife Service.

Lin, W., Sanford, L.P., Alleva, B.J., and Schwab, D.J., 1998.Surface wind wave modeling in Chesapeake Bay. Proceedings of the Third International Conference on Ocean Wave Measurement and Analysis—ASCE, Virginia Beach, VA, pp. 1048–1062.

Lin, W., Sanford, L.P., Suttles, S.E., 2002. Wave measurement and modeling in Chesapeake Bay. *Continental Shelf Research*, **22**: 2673-2686.

Marani, M., A. D. A., Lanzoni, and M. Santalucia, 2011. Understanding and predicting wave erosion of marsh edges. *Geophysical Research Letters*, **38**:L21401.

Mariotti, G. , S. F., P. L. Wiberg, K. J. McGlathery, L. Carniello, and A. Defina, 2010. Influence of storm surges and sea level on shallow tidal basin erosive processes. *Journal of Geophysical Research*, **115**: C11012.

McLoughlin, S.M., Wiberg, P.L., Ilgar Safak, McGlathery, K.J., 2015. Rates and Forcing of Marsh Edge Erosion in a Shallow Coastal Bay. *Esturaries and Coasts*, **38**: 620–638.

Perry, E. S., 2008. Assessing the Relation of Shoreline Erosion Rates to Shoreline Features and Wave Action in Maryland Chesapeake Bay, U.S. Army Corps of Engineers: 53, W912DR-07-P-0224.

Pritchard, D., 1956. The dynamic structure of a coastal plain estuary. *Journal of Marine Research*, **15**: 33-42.

Ravens, T. M., R. C. T., Kimberly A. Roberts, and Peter H. Santschi, 2009, Causes of Salt Marsh Erosion in Galveston Bay, Texas. *Journal of Coastal Research*, **25**: 265–272.

Ris, R.C., Holthuijsen, L.H., Booij, N., 1999. A third generation wave model for coastal regions. 2. Verification. *Journal of Geophysics Research,* **104**: 7667–7681.

Roland, M. E. and Douglass, L. S., 2005. Estimating Wave Tolerance of Spartina alterniflora in Coastal Alabama. *Journal of Coastal Research,* **21**: 453–463.

Sanford, Lawrence P., 1994. Wave-forced resuspension of upper Chesapeake Bay muds. *Esturaies*, **17**: 148-165.

Schwartz, M. L., 1967. The Bruun Theory of Sea-Level Rise as a Cause of Shore Erosion. *The Journal of Geology*, **75**: 76-92.

Schwimmer, R. A., 2001. Rates and Processes of Marsh Shoreline Erosion in Rehoboth Bay, Delaware, U.S.A.. *Journal of Coastal Research*, **17**: 672-683.

Skunda, K. G., 2000. Application of the shoreline instability model along the western side of the Chesapeake Bay, VA. 85 pp.

Slaughter, T. H., 1967a, Shore erosion control in tidewater Maryland. *Journal of the Washington Academy of Sciences*, **57**: 117-129.

Spoeri, R. K., Zabawa, C. F., Coulombe, B., 1985. Statistical Modeling of Historic Shore Erosion Rates on the Chesapeake Bay in Maryland. *Environmental Geology and Water Science*, **7**: 171-187.

Sunamura, T., 1992. Geomorphology of Rocky Coasts. New York: John Wiley & Sons

U.S. Army Corp of Engineers and Maryland Department of Natural Resources, 2010. Chesapeake Bay Shoreline Erosion in Maryland: A Management Guide.

U.S. Army Corps of Engineers, 1971. National shoreline study, North Atlantic region, vol. I: New York, U.S. Army Corps North Atlantic Division.

US Army Corps of Engineers, 1990. Chesapeake Bay shoreline erosion study. Department of the Army, US Army Corps of Engineers, Baltimore, MD, USA, 111pp.

Wang, Dong-Ping, 1979. Subtidal sea level variations in the Chesapeake Bay and relations to atmospheric forcing. *Journal of Physical Oceanography,* **9**: 413-421.

Wang, D-P., A. J. Elliott, 1978. Nontidal variability in the Chesapeake Bay and Potomac River: Evidence for nonlocal forcing. *Journal of Physical Oceanography*, **8**: 225-232.

Wang, H., et al., 1982. Relationships of Coastal Processes to Historic Erosion Rates. An Assessment of Shore Erosion in Northern Chesapeake Bay and of the Performance of Erosion Control Structures ,Annapolis, Maryland Department of Natural Resources**:** 5,1-5,75.

Wilcock, P. R., D. S. Miller, R. H. Shea, and R. T. Kerkin, 1998. Frequency of effective wave activity and the recession of coastal bluffs: Calvert Cliffs, Maryland. *Journal of Coastal Research*, **14**:256-268.

Wu, J., 1982. Wind-stress coefficients over sea surface from breeze to hurricane. *Journal of Geophysical Research*, **87**: 9704-9706.

Xu, J., S.-Y. Chao, R. R. Hood, H. Wang, and W. C. Boicourt, 2002. Assimilating high-resolution salinity data into a model of a partially mixed estuary. *Journal of Geophysical Research*, **107**, C7, 3074.

Young, I., Verhagen, L., 1996. The growth of fetch limited waves in water of finite depth. Part 1. Total energy and peak frequency. *Coastal Engineering*, **29**: 47-78.

Zhong, L. and Li, M., 2006. Tidal Energy Fluxes and Dissipation in the Chesapeake Bay. *Continental Shelf Research*, **26**: 752-770.