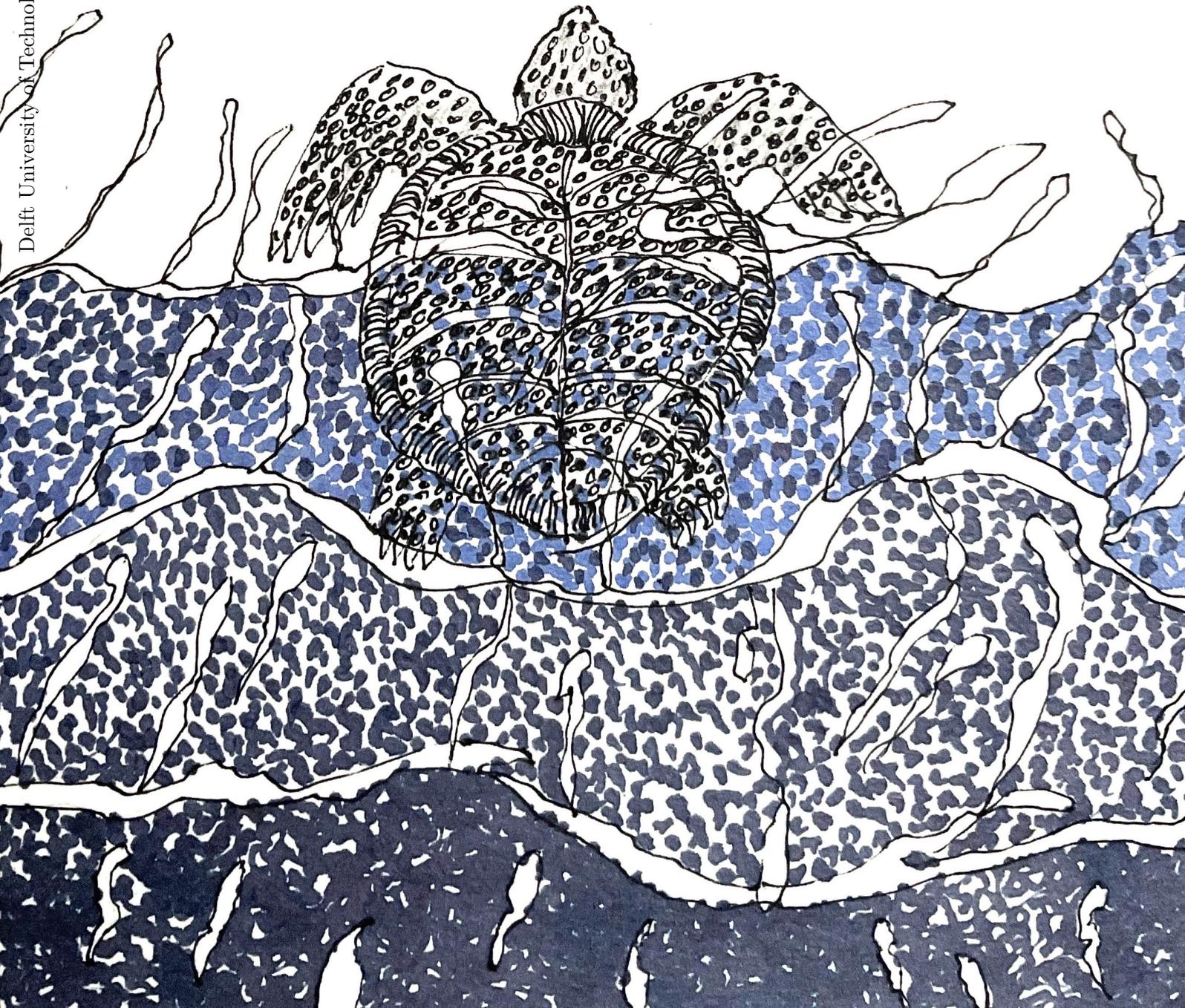# TUDelft

Enhancing Wave Runup Extraction on a Dissipative Beach through Machine Learning and Image Preprocessing Techniques

M. J. van der Grinten

Enhancing Wave Runup Extraction on a Dissipative Beach through Machine Learning and Image Preprocessing Techniques

by

M. J. van der Grinten

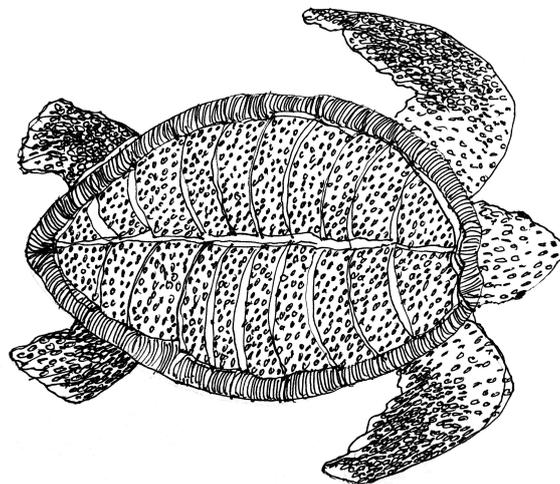|	
Students:	M. J. van der Grinten - 4760182
Supervisors:	Dr. Ir. J. A. Álvarez Antolínez
	Prof. Dr. Ir. A. J. H. M. Reniers
	Ir. J. C. Christiaanse
Institution:	Delft University of Technology
Faculty:	Civil Engineering and Geosciences
Place:	Delft, Netherlands
Date:	Sunday 23rd June, 2024

Illustrations by Marcus Aurelius van der Grinten, used with permission.

TUDelft

Abstract

Wave runup observations are key data for coastal management, as they help validate predictive models for the inundation frequencies and erosion rates. Efforts to develop automated algorithms that effectively identify the instantaneous water line from video imagery have led to a plurality of methods. However, under dissipative conditions, the presence of a seepage face often hinders proper extraction and requires time-intensive data quality control or manual digitization. In this study, we explore traditional color contrast (CC) preprocessing techniques and develop a novel method that incorporates a measure for texture roughness — specifically, local entropy — alongside saturation. The CC-model showed good agreement with the manually digitized water line (0.12 m root-mean-square-error (RMSE) and correlation coefficient (r) of 0.94), and with its runup statistics (0.08 m RMSE, and r of 0.97 for the 2% runup exceedance, $R_{2\%}$). The timing aspect of the method also showed good agreement (3.88 s RMSE, and r of 0.70 for $T_{m-1,0}$). Concurrently, a convolutional neural network (CNN) informed by CC-preprocessed images was cross-validated using nine manually labeled video time series, each lasting 1 hour and 30 minutes. The CNN model demonstrated good agreement during cross-validation with manually labeled time series (0.10 m RMSE and r of 0.96 for the full-time series, and 0.09 m RMSE and r of 0.97 for $R_{2\%}$). The temporal dimension of the CNN estimate was also satisfactory (3.51 s RMSE, and r of 0.79 for $T_{m-1,0}$). The observed $R_{2\%}$ values showed the best agreement with the formula for extremely dissipative conditions from Stockdon et al. (2006), with RMSE-values lower than 0.13 m and r-values that exceeded 0.70 for all three methods. When applied to other datasets, the CNN method occasionally failed to accurately capture the water line due to specific characteristics of the new timestack images. These results validate our ML method as a viable proof of concept and challenge us to enhance its adaptability and accuracy across varied environmental conditions. Despite these limitations, the CNN method can be effectively implemented for long-term runup analysis. Additionally, the CC method is anticipated to be applicable across similar beaches along the northern Gulf of Mexico for long-term extreme value analysis and wave-by-wave analysis. Both methods demonstrate potential in reducing the time required to extract the instantaneous runup from video imagery under dissipative conditions and enhance real-time monitoring, enabling better predictive modeling of coastal processes.

Acknowledgements

And yes, as if I were receiving an Oscar, I take this stage to express my gratitude to the people who supported me through this Thesis. Because let's face it, making moments bigger than they really are is part of the fun. A special thanks goes to my Thesis committee. To Ir. Jakob Christiaanse, the turtle-loving guy who patiently listened to me ramble at lightning speed every week. I'll always cherish the epic surfing trips. To Dr. Jose Antolínez, who gave me all the space I needed and taught me some valuable lessons. You helped me see when there were things to be proud of, even if I didn't realize it at the time. And to Professor Ad Reniers, who always knew the right moments to subtly steer me in the right direction, even though I'm not exactly known for following advice. I also want to thank Falco Taal, my fellow Thesis companion. I enjoyed the two sporty runs we took on the beach, just like the real athletes we are. And I don't think I would have laughed so much without you in Texas. My heartfelt appreciation goes to the team in Texas that welcomed us so warmly in Galveston. To Professor Jens Figlus, who made everything possible and even lent us a hand during fieldwork days. To Professor Timothy Dellapenna, who drove us around the island and enriched us with fascinating facts about soil layers and sediment. Thank you as well for the greatest shrimp boil of all time. And to Professor Christopher Marshall, for taking us on his boat for turtle tagging. What a shame we caught none, but at least we got to see some of these beautiful creatures in the hospital you set up. Of course, I am deeply grateful to my parents for giving me a balanced set of interests. To my mom, for her passion about the great challenges we face, her voracious reading, and her endless ability to inspire and captivate. To my dad, who drew the cover of this Thesis and always checked in. These pages feel even more reflective of myself thanks to his beautiful artwork. To the three human beings who share the same building blocks as me, I want to say that everything we share makes us one hell of a team. To Julius, who always promised to read my Thesis but only got to see the cover, I truly appreciate you always being there for me. And finally, to all my friends, I can't say how much I love y'all.
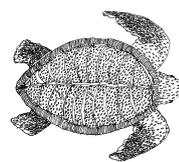
Contents

Introduction

1

The Gulf of Mexico is a serene nesting ground for many sea turtles. Yet, beneath the tranquil sands lies a critical concern: rising groundwater tables threaten the sanctuaries where these ancient mariners entrust their eggs. Delving into the depths of coastal hydrology reveals a breathing system, where ebb and flow, waves, storm surges, and sea level rise shape the landscape and dictate the fate of sea turtle nests. In this fragile equilibrium, understanding shoreline processes becomes vital in safeguarding these delicate nesting grounds against the pressing threat of climate change.

1.1. Endangered species

The International Union for Conservation of Nature (2021) lists two of the seven species of sea turtles worldwide as critically endangered, one as endangered, and three as vulnerable. Sea turtles face a myriad of anthropogenic stressors. These challenges encompass habitat loss and degradation (Patrício et al., 2021; Caillouet Jr et al., 2018), indigenous harvesting practices (Jensen et al., 2022), inadvertent by-catch (Cheng et al., 2019; Darquea et al., 2020), illicit trade activities (Nooren and Claridge, 2002; Pheasey et al., 2021; Senko et al., 2022), vessel strikes (Fuentes et al., 2021), and the pervasive impact of pollution (Sinaei et al., 2021; Arienzo, 2023). Globally, the majority of sandy coastlines in marine protected areas have been subject to erosion (Luijendijk et al., 2018), raising concern over sea turtle nesting grounds even more. Coastal interventions and continual river manipulation have disrupted sediment supplies, resulting in the depletion of beaches and habitat loss for sea turtles (Lutcavage, 2017). Along with these pressing concerns, the intricate interplay of rising temperatures, shifting ocean currents, and sea level rise could amplify the challenges faced by these marine species. These effects of climate change will impact their nesting grounds (Fuentes et al., 2011), reshape their migration trajectories (Hays, 2017), and alter their reproductive success.

Sea turtle eggs are susceptible to temperature variations. For the incubation to be successful, the thermal range should be between 25°C and 32°C (Yntema and Mrosovsky, 1982). Exposure to higher or lower temperatures reduces the hatchlings' survival and increases the probability of morphological abnormalities (Matsuzawa et al., 2002; Segura and Cajade, 2010). Besides, sea turtle eggs show temperature-dependent sex determination (TSD). The global warming of the oceans results in a bias towards more females being born (Tomillo et al., 2014; Tanner et al., 2019), which will decrease genetic variability among sea turtle species. The rapid pace of anthropogenic climate change and the slow maturation to adulthood of sea turtles introduce uncertainty regarding the timely occurrence of phenotype changes through natural selection or adaptation in nesting behavior (Dalleau et al., 2012).

Christiaanse et al. (2024) showed that potential suitable nesting regions are available for sea turtles to migrate to. The question remains whether sea turtles will make this migration in time.

Besides seasonal disturbances, climate change also impacts sea turtle populations through more extreme weather events (Dewald and Pike, 2014), rising groundwater tables by increased precipitation (Rivas et al., 2018), and sea level rise (Fish et al., 2005; Veelenturf et al., 2020; Sönmez et al., 2021). These changing environmental conditions impact the inundation frequency of nests, hindering the proper exchange of gases for sea turtle eggs. Pike et al. (2015) compared the hatching rate of green turtles (*Chelonia mydas*) eggs when submerged in saltwater for a certain period. They found that submergence for 1 or 3h reduced the viability of the eggs by less than 10% but that submergence for 6h reduced the viability of the eggs by as much as 30%. Besides, eggs are most vulnerable during the two first weeks after being laid and just before hatching. During that period an inundation of 30 minutes can cause 100% mortality of the clutch (Nooren and Claridge, 2002).

1.2. No standard for nest relocation

As a precaution, sea turtle nests are often relocated when prone to inundation. However, there seem to be no standard criteria to determine whether a nest should be moved (Ware et al., 2019). And, while nest relocation can be effective, unnecessary nest manipulation can have detrimental consequences on the hatchlings' survival rate (Grand and Beissinger, 1997; Limpus et al., 1979; Parmenter, 1980; Wood and Bjorndal, 2000), and alter the sex ratio and morphological development (Glen et al., 2003; Maulany et al., 2012; Mrosovsky, 2006; Whitmore and Dutton, 1985) of the embryos. Besides, it has been shown that relocation of the nests should happen quickly after the eggs have been laid. Ahles and Milton (2016) demonstrated that nests moved after 12 hours from deposition exhibited higher rates of embryonic death compared to eggs that were relocated within the initial 12 hours. Nest relocation should thus preferably be a preventive method based on historical data and long-term predictions and not on last-moment forecasts. However, hazards related to rapid climate change and other anthropogenic stressors such as oil spills may not always allow for nests to be relocated under ideal conditions.

Christiaanse et al. (2024) analyzed the global nesting behavior of sea turtles and suggested sea surface temperature, tidal range, extreme surges, and proximity to coral and seagrass habitats significantly influence nesting preferences. Using these coastal indicators they found that sea turtles are only nesting in 30% of suitable global coastal regions between -39° and 48° latitude and indicating the possibility for sea turtles to broaden their nesting habitat. This research can provide information on eventual migration trends, and on where to relocate sea turtle nests but does not predict the flooding likelihood of single nests.

Out of the twenty-two coastal indicators employed by Christiaanse et al. (2024), there is no parameterization for the conditions beneath the sand, where the eggs are laid. According to Nooren and Claridge (2002) 'The only valid reason for moving a nest from its original location is that the hatch success is threatened in some way. Typically [...] because the nest is located too low and will be inundated by high tides, either directly or by seawater seeping in below the surface'. To date, nest inundation has been related to waterline predictions only, and has not yet been translated to groundwater variations. The extent of wave-induced groundwater fluctuations on sea turtle nests remains uncertain, and the circumstances under which these fluctuations, if substantial, become relevant in predicting inundation are not clearly defined.

To successfully determine whether a sea turtle nest should be relocated or not, a better understanding of both the hydrodynamics and the eggs' resilience to environmental influences is required. Presently, our knowledge of species-specific tolerance to hypoxic events is limited, and, factors such as frequency, duration, and timing relative to embryonic development remain understudied (Caut et al., 2010; Foley et al., 2006; Pike et al., 2015; Ware and Fuentes, 2018) and require in-depth investigation. On the other hand, the hydrodynamics influencing sea turtle nest inundation remain relatively unexplored and call for thorough investigations.

1.3. Prediction of nest inundation

Until now most nest relocations are based on in situ observations of the high tide water line (Ware et al., 2019). However, some attempts have been made to predict nest inundation from offshore wave conditions.

Wave runup models have been used to point out locations susceptible to inundation. The Stockdon et al. (2006) formula has already been used by Ware et al. (2019, 2021) to improve the identification of nests prone to inundation. In a study along the coast of Fort Morgan, Alabama, USA different definitions of beach slope were compared (Ware et al., 2019). The wash-over state was correctly identified in 83.3% of sites when using the nest slope while the dune-to-water slope performed better when predicting the wash-over frequency of the nests. In another study conducted along 40 nesting beaches along the northern Gulf of Mexico, 34% of nesting locations per beach were found to face significant risk to wave exposure (Ware et al., 2021). In these studies, the wash-over state was defined according to the high tide water level identified at the most recent wrack, at the wet-dry boundary, by signs of erosion or deposition of sand, or by direct observation of waves interacting with the nests.

As an alternative to empirical predictions Dédina (2023) employed two numerical metamodels (HyCReWW and BEWARE 2.0) to forecast wave exposure for nests within a data-scarce environment on a fringing reef environment at Ras Baridi, Saudi Arabia. Due to the limited data, validation of the metamodels against local runup time series was compromised, and the research methodology was only assessed through comparison with validated Xbeach models elsewhere.

Although there has been a shift towards predicting the exposure of species nesting in coastal areas to wave wash-over using wave runup models, validation of these models against runup time series has not been conducted. While empirical predictions based on in situ observations of the high tide water line have been the predominant method for nest relocation, wave runup models' potential to enhance nest inundation predictions remains largely untapped. The limited validation of these models against runup time series impedes their reliability and effectiveness in informing management decisions. Therefore, there is a clear need for further research to validate wave runup models through comparison with runup time series data, thereby enhancing their utility in coastal ecosystem management strategies.

1.4. The importance of reliable runup data

To effectively protect sea turtle nests from inundation, it is essential to better understand the shoreline dynamics. Our study focuses on improving the accuracy of runup extraction from video imagery, which is a critical component for the validation of runup predictors and subsequent nest inundation. But improving runup observation, and herewith runup predictions not only aids in protecting sea turtle nests but contributes to a broader context of coastal management. Precise runup data is

essential for design purposes, and for predicting and mitigating coastal erosion, sea-level rise, and storm surges.

From previous studies, many preprocessing techniques have proved themselves to efficiently enhance specific nearshore features from timestack images, which are created by stacking sequential video frames over time, capturing the temporal changes in wave runup. These methods encompass the analysis of grayscale image Bailey and Shand (1994), Color Channel Difference (CCD) operations Turner et al. (2001), Pixel Intensity Clustering (PIC) (Aarninkhof et al., 2005; Uunk et al., 2010), pixel variance (Simarro et al., 2015), or a measure for image roughness (Zhang and Zhang, 2009).

However, on dissipative beaches, extracting runup can be challenging due to the low contrast between the swash and the sand in low-energy conditions. Additionally, the presence of a seepage face keeps the beach covered with a thin layer of water, making it difficult to distinguish the swash from the sand. Color contrast (CC) methods often fall short in capturing these subtle processes at play (Huisman et al., 2011; Simarro et al., 2015). Process-based methods, such as the Radon method developed by Almar et al. (2017), require certain energetic conditions to operate properly. Given the shortcomings of traditional methods and the prohibitively slow process of manual runup digitization, which leaves a growing set of data unexplored, the interest in machine-learning implementations in the field of runup extraction has grown substantially over the past two decades (Marti-Puig et al., 2024; Kang et al., 2024a,b).

Recently, Collins et al. (2023) demonstrated that machine-learning (ML) algorithms, informed by multi-channel LiDAR - elevation, reflectance, and spatial-temporal elevation variance - can delineate the water line up to high levels of precision. While the high cost of LiDAR equipment limits its widespread applicability, this method prompts a new question in video imagery: Can an ML algorithm, informed by image preprocessing techniques for feature extraction, accurately extract the instantaneous water line from timestack images? Traditionally, video images are limited to RGB color channels and do not provide additional attributes such as real-world coordinates, reflectance, or surface orientation, available with LiDAR data. However, the insights gained from CC methods could be beneficial here. Applying preprocessing techniques to timestack images can create new input channels for a multi-channel machine learning algorithm, thereby enhancing its ability to extract meaningful features from video imagery.

1.5. Enhancing runup extraction on a dissipative beach

This study addresses the broad challenges sea turtle nesting habitats face, particularly the threats posed by varying groundwater levels induced at the sea boundary. These environmental factors are critical in understanding the sustainability of these habitats. Given the importance of groundwater level monitoring, this research specifically focuses on developing and evaluating a detection algorithm designed to improve the precision of runup measurements which are key data in connecting the nearshore hydrodynamics to the beach hydrology. By honing in on the detection algorithm, this study seeks to provide a robust tool that can be used in the broader context of coastal management and conservation efforts.

This research aims to develop accurate algorithms for extracting the instantaneous water line from video images. The following questions are addressed to achieve this: What CC method can be implemented on a dissipative beach? Can image preprocessing techniques create effective new input channels for a convolutional neural network (CNN) algorithm to extract wave runup from video

imagery? Compared to manual digitization and the CC-based method, how does this enhanced ML algorithm extract the water line on a dissipative beach? And, how do the observed water lines compare to empirical estimates of runup?

Video imagery was collected from a beach along the Gulf of Mexico from which timestack images were extracted. A range of preprocessing techniques was applied to these timestack images, and from this, a CC method based on saturation and local entropy was developed following earlier work from Zhang and Zhang (2009). Simultaneously, a multi-channel ML model was developed and optimized using these preprocessed images as input channels and manually labeled timestacks for training. The performance was then evaluated using runup statistics, with the $R_{2\%}$ compared to both the general formula and the formula for extremely dissipative beaches as presented by Stockdon et al. (2006). The aim is to significantly improve the precision and reliability of runup extraction from video imagery by providing a cost-effective alternative to LiDAR-based methods. This study does not address long-term trends in shoreline dynamics, limiting its applicability, and the findings of this research are based on a proof-of-concept implementation, which requires further validation and refinement before practical deployment in diverse real-world scenarios.

The report is organized as follows: Chapter 2 provides a detailed review of wave runup, groundwater processes, and extraction techniques of the water line from video imagery, Chapter 3 presents the procedure for the collection of timestack images and outlines the CC-based and ML method as well as the optimization process, in Chapter 4 and 5 we describe and discuss the results, and suggest future perspectives, and Chapter 6 summarizes the conclusions of this study, along with the main recommendations for future research.

<div align="right">
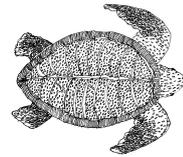
2

</div>

State of the Art

This chapter provides a comprehensive review of the current state of the art in shoreline dynamics, groundwater processes, and runup extraction techniques. It is divided into three main sections. Section 2.1, explores the dynamic nature of the water line, focusing on key terminologies and conceptual frameworks essential for understanding wave runup. Section 2.2 addresses the interactions between groundwater processes and coastal dynamics. Various responses of groundwater systems to different types of wave forcing are examined. Finally, Section 2.3 discusses the techniques used for extracting runup data from video imagery. Topics include general camera setups for monitoring the water line, georeferencing techniques, and various methods for runup and shoreline extraction. Further considerations and challenges in extracting accurate runup data from video images are also presented.

2.1. A moving water line

2.1.1. Lexicon of shoreline extrusion

Shoreline extrusion is defined as the motion of the waterline in the cross-shore direction up and down the beach profile. It is characterized by a broad range of frequencies, going from incident waves to tides, and storm surges. Figure 2.1 proposes a visualization of the shoreline extrusion caused by incident waves, also referred to as runup.
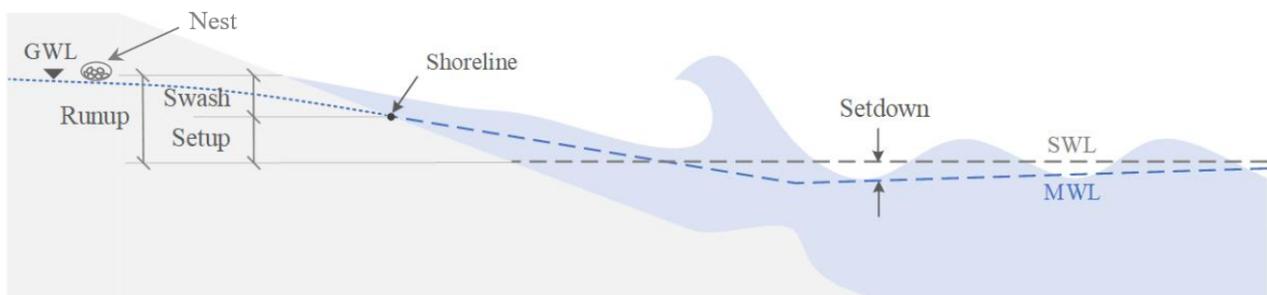
Figure 2.1: Diagram illustrating wave runup, setup, setdown, groundwater level (GWL), still water level (SWL), mean water level (MWL), runup as a sum of swash and setup, the SWL on a single barred beach profile and a sea turtle nest.

When incident waves travel to shore and start to shoal, the wave force is directed offshore resulting in a lowering of the mean water level (MWL) below the still water level (SWL), also referred to as wave-induced setdown. Then as they carry on and meet even shallower water waves start breaking,

and the wave force changes direction toward the shore, this time increasing the MWL above the SWL. A phenomenon referred to as wave-induced setup. The remaining energy from the incident waves, or swash, travels to shore as bores raising the water line even further. Setup and swash are the two main components of wave runup.

As the name already gives away, the still water level corresponds to the hypothetical water level without the action of any waves. Let's consider a steady state over some time shorter than the tidal fluctuation and longer than the wave actions, in the order of 30 minutes. The mean water level corresponds to the time-averaged water level over this period. The still water level is a function of sea level rise, the tide, and water level anomalies such as storm surges, climatic cycles (e.g. El Niño), seasonal effects, or steric variations. Adding the time-averaged action of waves results in the mean water level.

An important aspect of wave runup is the relative impact of infragravity waves versus incident wind-generated waves. Generally, infragravity waves range from 0.004 Hz to 0.04 Hz, and incident waves from 0.04 Hz to 1 Hz, or periods of 250 s to 25 s, and 25 s to 1 s respectively (Bertin et al., 2018). However, the exact definition of infragravity versus incident waves is case-dependent. Under reflective conditions, wind-generated waves dominate the nearshore dynamics of the swash but infragravity waves can influence the mean water level near the shoreline - or setup. Under dissipative conditions, the relative influence of infragravity waves is greater as incident waves break over a longer distance and barely reach the shoreline. Therefore 'dissipative beaches tend to have long swash periods and infrequent swash excursions above the effluent line (i.e., transition between saturated and unsaturated sand) across a wide intertidal zone compared to reflective beaches with short swash periods and frequent excursions above the effluent line in a narrow intertidal zone' (Ware et al., 2019).

2.1.2. Empirical formulations for runup

The accurate prediction of runup holds great significance in the design of infrastructure, as well as in the assessment of coastal safety and environmental impact during extreme events. Yet, due to the complexity of non-linear processes along the surf and swash zone and the perpetual changes in bathymetry, runup prediction through numerical modeling remains challenging. Consequently, runup is often predicted through empirical formulas (Gomes da Silva et al., 2020).

These formulas use simplified characteristics such as the beach slope, and the deep water wave height and periods as input parameters to compute statistical runup values such as the maximum runup (R_{max}) or the runup exceeded by 2% of the waves ($R_{2\%}$). Due to the frequent bimodality of wave spectra, empirical formulations commonly address infragravity and incident waves separately. Also is it important to acknowledge that every formula comes with its own range of applicability, based on its respective training and validation datasets. Certain equations are to be used solely for gravel beaches (Poate et al., 2016; Didier et al., 2016; Dodet et al., 2018), others for sandy beaches (Stockdon et al., 2006; Senechal et al., 2011; Guedes et al., 2011, 2012; da Silva et al., 2018; Power et al., 2019), and some were built for estuaries and fetch limited areas only (Didier et al., 2020).

In the present case of a mildly sloping sandy beach, the widely used formula of Stockdon et al. (2006) would be the choice of reference since it 'outperformed previous works, in terms of generality of the datasets considered and accuracy of the predictor' (Gomes da Silva et al., 2020). By using a dataset of 10 locations in the United States and the Netherlands they built an empirical formula for

$R_{2\%}$ (Eq. 2.1) that included wave setup (Eq. 2.2), infragravity (Eq. 2.3), and swash (Eq. 2.4):

$$R_{2\%} = \alpha \left[\bar{\eta} + \frac{\sqrt{S_{ig}^2 + S_{inc}^2}}{2} \right] \tag{2.1}$$

$$\bar{\eta} = 0.35 \beta_s \left(H_0 L_0 \right)^{0.5} \tag{2.2}$$

$$S_{ig} = 0.06 \left(H_0 L_0 \right)^{0.5} \tag{2.3}$$

$$S_{inc} = 0.75 \beta_s \left(H_0 L_0 \right)^{0.5} \tag{2.4}$$

where $\alpha = 1.1$, $\bar{\eta}$ is the wave-induced setup, S_{ig} infragravity wave-height, S_{inc} the incident swash, and β_s the beach slope. In addition to this, they developed a runup formula for dissipative beaches (Eq. 2.5):

$$R_{2\%d} = 0.043 \left(H_0 L_0 \right)^{0.5} \tag{2.5}$$

This formula shares some similarities with the formula for S_{ig} (Eq. 2.3) of Stockdon et al. (2006). During dissipative conditions, the beach slope is milder, and as all incident waves break they saturate a wider surf zone. Therefore, setup is of less importance as it spreads over a longer distance, and mainly infragravity waves reach the shoreline. Hence the resemblance. However, one should be aware of the eventuality of saturation extending towards the lower frequencies. Phenomena that was observed by Ruessink et al. (1998) and later by Ruggiero et al. (2004).

2.1.3. Measurement methods for runup

As mentioned before the interest in runup prediction has grown exponentially over the past decades. A large set of empirical formulations have entered the scientific world (Gomes da Silva et al., 2020) simultaneously resulting in the development of wave runup measurement techniques.

Resistance wires have been used on the beach face by Guza and Thornton (1982). The method consists of a wire under high-frequency alternating voltage deployed in the swash zone. Voltage changes due to water coverage of the wire can be translated to runup values. The method provides a high sampling frequency and spatial resolution but is restricted to mild environments due to its fragility. Besides, Holland et al. (1995) found that the method underestimated runup values and, that the height at which the wire is deployed influences the measurements. A higher deployment above the bed results in a smaller runup and horizontal extrusion.

The human eye stands as the oldest instrument for measuring wave runup that remains in use to this day. Nielsen and Hanslow (1991) installed a series of stakes over the beach face with known locations and visually retrieved runup values as singular waves passed the markers. The primary drawbacks of this straightforward method are its restricted spatial and temporal accuracy and the labor intensity that it requires in the field. But, if coupled with a video monitoring system the temporal accuracy and the labor intensity in the field can be easily overcome. Still, another limitation of this method is its range of applicability as only mild wave conditions allow the proper installation of stakes over the beach face.

Turner et al. (2008) first measured bed elevation and wave runup using ultrasonic sensors and pressure sensors have been deployed to measure runup on steep rocky cliffs by Dodet et al. (2018). These techniques both allow for high sampling frequencies but the effectiveness of these methods is constrained by the availability of sensors, resulting in limited spatial resolution. The latter is even more true on mildly sloping beaches where the required sensors are higher due to the longer beach face. Moreover, deploying sensors in the surf zone proves challenging for high-energy environments, and installing sensors on non-cohesive material beaches can be intricate due to the time scale of bed level changes.

Runup can also be measured manually using a GPS with high spatial accuracy. A possibility here is to search for spatial markers such as plant wracks, the limit between wet and dry sand, or beach scarps, which are all indicators of past runup events. This technique was applied by Ware et al. (2019) in their research on 'wave runup modeling to inform coastal species management'. The primary drawbacks are that GPS antennas often have slow performance and that the method demands intensive and redundant efforts in the field, limiting the output to single values such as R_{max}.

The application of video imagery to runup measurement has significantly influenced scientific interest in wave runup. Since the early 2000s, the volume of publications related to runup has grown exponentially, coinciding with the onset of deploying video monitoring systems for nearshore applications (Mendes et al., 2022). Video imagery allows for a spatial resolution refined to 0.01 m and a typical temporal resolution of 30 Hz. It also enables continuous monitoring while requiring less manual labor. Further details on this technique are available in Section 2.3.

X-band radars have also been deployed to assess wave runup (Hasan and Takewaka, 2014; Lyddon et al., 2021). Radars are top-notch technology as they allow for both very high spatial and temporal resolution but, they are also very expensive instruments. And, as higher resolution comes hand in hand with higher computational costs they should only be applied when needed.

2.2. A pervious boundary

2.2.1. Lexicon of groundwater processes

Beach groundwater can be considered as the interface between land and sea, where the coastal aquifer meets the ocean. In this unconfined aquifer flows are driven by tides, storm surges, wave-induced setup, swash, and water levels at the land boundary, but eventually also by rainfall and evaporation. All contribute to a continuous exchange of water, sediment, gases, and organic matter across the shoreline (Horn, 2006). Typically, the land boundary shows low-frequency variations at the seasonal rate while the ocean boundary exhibits higher frequencies, going from tides to incident waves (Bakhtyar et al., 2011). Figure 2.2 provides a visualization of the time-averaged groundwater processes during falling and rising tides.

Subsequent discourse will delve into the key terminologies of groundwater for beaches defined in Horn (2002). The groundwater table corresponds to the depth at which the pore pressure equals the atmospheric pressure. Beneath the groundwater table, is the fully saturated phreatic zone where pore pressure is higher than the atmospheric pressure. Above the groundwater table, the soil can eventually be fully saturated as well due to capillary effects. In this so-called capillary fringe or tension-saturated zone, pore pressure is lower than the atmospheric pressure. Above the capillary fringe and up to the ground surface lays the intermediary zone where the pores are not fully saturated and the pore pressure varies from place to place. Together, the capillary fringe and the intermediary

Figure 2.2: Diagram illustrating the groundwater terminologies over a beach profile. In (a) the exit point and the water line are decoupled during a falling tide and, in (b) the groundwater during a rising tide is depicted. Adapted from Horn (2002).

zone form the vadose. Beware that these classical definitions hold for homogeneous soils only.

According to Horn (2002), 'the slope of the water table changes with the tide, sloping seaward on a falling tide and landward on a rising tide'. When the tide falls faster than the beach can drain (Figure 2.2.a), the water table can become decoupled from the water line. The groundwater table and the ground surface then coincide at an exit point above the instantaneous water line. In between these two points, a seepage face develops from which groundwater exfiltrates the beach matrix. Under a rising tide (Figure 2.2.b), the exit point and the water line are joined in one single point above which water can infiltrate the beach matrix.

At first sight, the concept of a groundwater table appears to be a simple notion as it marks the upper boundary of the saturated zone. Yet, further inspection of this concept has shown some discrepancies between the theory and practice. Many issues arise from the heterogeneity of the soil complexion. For instance, differences in permeability can result in perched and inverted water tables. Another phenomenon is the encapsulation of air and the formation of biogenic gas bubbles below the groundwater table. Also, strong bi-modality of the pore size distribution can result in the smaller pores not filling up with water. Last but not least, rapid fluctuations in the groundwater table can also occur when the capillary fringe reaches the ground surface. For a thick capillary fringe close to the ground surface, 'only a small depth of water is required to convert the water from the tension-saturated zone to one of positive pore-water pressures, causing the water table to rise to the ground surface' (Baird and Low, 2022).

Beach groundwater processes are mainly driven by hydraulic gradients and hydraulic conductivity (K), also referred to as coefficient of permeability. Many empirical formulas exist to estimate the hydraulic conductivity from single soil characteristics such as the mean grain diameter (D) and the degree of sediment sorting (σ). The formula from Krumbein and Monk (1943) is still broadly used to estimate hydraulic conductivity (Eq. 2.6):

$$K = \frac{kg}{\nu} \qquad (2.6)$$

where g is the gravitational constant (ms^{-2}), and k the permeability (in units of Darcy, where 1 Darcy = $9,87 \cdot 10^{-13} m^2$), that can be estimated as (Eq. 2.7):

$$k = 760D^2 e^{-1.3\sigma} \tag{2.7}$$

Although widely used, empirical formulations for hydraulic conductivity should always be considered with caution due to the variability in time and space. The soil's heterogeneity, temperature changes, and saturation can all affect the hydraulic conductivity. Baird et al. (1998) showed that measured hydraulic conductivity can diverge from the estimated values by an order of magnitude. And, in Benson et al. (1997) different measurement methods resulted in large differences in hydraulic conductivity. Besides, most empirical formulations are appropriate for sandy soils only (Horn, 2002).

2.2.2. Response to tidal forcing

Early research on beach groundwater dynamics was oriented toward the effect of higher-frequency signals. Hence, the tide. Studies including Emery and Foster (1948), Ericksen (1970), Waddell (1973, 1976, 1980) Lewandowski and Zeidler (1978), Nielsen (1990), Hegge and Masselink (1991), Kang et al. (1994), Nielsen and Kang (1995), Baird et al. (1998), Raubenheimer et al. (1999) Cartwright and Nielsen (2001), and Robinson et al. (2005), have shown the tidal signal to decay landwards, and to be asymmetrical (fast rising and slow falling tide), and skewed (positively) in time. The phase lag increases linearly in the landward direction while the amplitude decays exponentially (Cartwright et al., 2004). Asymmetry of the groundwater response to the tidal forcing also increases in the landward direction.

The work of Raubenheimer et al. (1999) showed that the tidal signal decays landwards, but, that the lower frequency signal decays to a lesser extent than the higher frequency signals. In their specific case, diurnal and semidiurnal signals were damped out completely 100m land inwards from the shoreline while the spring-neep signal was still clearly visible at that same location.

Two regions can be distinguished when considering beach groundwater. In the upper intertidal region, the beach matrix is successively saturated and unsaturated whereas in the lower section of the intertidal zone, the beach matrix is permanently saturated (Horn, 2002). When the tide rises above the beach groundwater table, seawater infiltrates the beach matrix landwards. As the tide falls again, water exfiltrates the beach toward the sea. The beach groundwater table is sometimes considered as the prolongation of the mean water level, which in turn relates closely to the tidal elevation. But as mentioned before, when the tide drops faster than the beach matrix can drain the mean water level becomes decoupled from the groundwater level (see Fig. 2.2.a). Between the two then, a seepage face appears where water exfiltrates the beach matrix (Horn, 2006). This phenomenon is thought to be of great importance for sediment transport in the intertidal zone.

The beach groundwater table was often measured higher than the tidal elevation (e.g. Nielsen et al., 1988; Kang et al., 1994b; Turner et al., 1997, as cited in Horn, 2006). This so-called superelevation of the beach groundwater table is because of the non-linearity of the horizontal groundwater flow (Philip, 1973; Knight, 1981; Nielsen, 1990, as cited in Horn, 2006). Overheight of the beach groundwater table is thought to increase with decreasing beachface slope and sediment size and with decreasing tidal range and wave infiltration (Turner et al., 1997).

Cartwright et al. (2004) showed that the sinusoidal tidal signal generates higher harmonics when interacting with the sloping boundary of the beach aquifer. The generation of these harmonics was found to be largest in the upper part of the aquifer (sand surface close to the shoreline) and smallest in the lower parts of the aquifer (sand surface seaward of the shoreline). This nonlinear filtering of

the sloping boundary is the largest where flow magnitudes are greatest. Hence, in the upper part of the aquifer, where seawater infiltrates the beach matrix.

2.2.3. Response to incident wave forcing

Until now, the focus was set on a time-averaged situation over a period longer than individual waves but shorter than the tidal cycle. In doing so, only the lower frequency signals have been assessed, ignoring the response of the groundwater table to incident waves.

The tide results in the beach groundwater table's slope changing over time. With the slope being landward for a rising tide and seaward for a falling tide. However, other shapes of beach groundwater tables have been measured. Namely, with a bulge landward of the shoreline and close to the runup limit. Which, can be attributed to the infiltration of incident waves Horn (2006). Near the shoreline, the wave-induced setup increases the mean water level, and a part of the swash infiltrates the beach matrix raising the groundwater table even further. Besides, wave-induced setup drives large circulation cells across the beach Bakhtyar et al. (2011).

Hegge and Masselink (1991) presented a thorough analysis of individual wave runup during a falling tide. They distinguished three categories of runup events. The first concerns low-amplitude runup events that did not affect the water table. The second was moderate runup events that did transmit a pressure force to the saturated sand and, stabilized the water table. The third category was the high-amplitude swashes that extended beyond the groundwater table and could contribute to a rise in the groundwater table. Spectral analysis showed how the beach matrix functions as a band-pass filter, where some frequencies (in the specific case, 0.013 Hz) are less attenuated than others. The mechanisms behind the selective signal transmission of the beach matrix remain unclear (Horn, 2006).

Hegge and Masselink (1991) also showed that the signal from the incident waves shifts towards the lower frequencies when penetrating the beach matrix. This was confirmed by Turner et al. (1997) and, Cartwright et al. (2005) suggested this to be related to the overtopping of the exit point. When a wave overtops the exit point it infiltrates the beach matrix whereas when a wave does not overtop the exit point it transmits momentum directly to the saturated beach aquifer. It is only logical that the infiltration process affects the frequency of the incidence signal.

The higher frequency signals from incident waves have been shown to behave in an analog manner to the tidal-induced perturbations. The amplitude of perturbations of the groundwater table decays exponentially and the phase lag increases linearly (Cartwright et al., 2005). It is worth emphasizing once more that lower frequencies are damped out less than higher frequencies. Hence, the further landwards, the narrower the groundwater spectrum becomes and, the more it shifts towards the lower frequencies. In the latest research (Guest and Hay, 2017; Stark et al., 2022; Kranenborg et al., 2023), additional insights were provided, indicating that the pore pressure signal diminishes and undergoes a shift as the depth increases. Lower frequencies tend to penetrate the bed deeper and shift less than higher frequencies.

2.2.4. Response to infragravity waves

Darcy's law states that flow through porous media is driven by pressure head variations. When traveling to shore through the surf zone, waves apply differential pressure on the sea bed. This

wave-induced pressure head is attenuated with depth, but the longer the wave, the more gradual the attenuation. Infragravity waves are characterized by relatively long wavelengths and can thus induce flow through the beach matrix. However, these flows remain relatively insignificant as the resulting vertical flow is approximately 0.01 mm/s, and the corresponding horizontal flow is two orders of magnitude smaller (Bertin et al., 2018).

More important is the impact of infragravity waves on the beachface. This is even more true for dissipative sandy beaches, where infragravity waves dominate over incident waves and the related setup. In dissipative conditions, the lower part of the swash zone is saturated. However, large infragravity-uprushes can reach the unsaturated part of the beachface leading to infiltration of the beach matrix. Sous et al. (2016) studied the groundwater processes under dissipative conditions. They showed that the groundwater pressure field in the swash zone fluctuated in the same frequency range as the overlying infragravity-driven swash.

2.2.5. Effect of the capillary fringe

A capillary fringe largely influences the dynamics of the groundwater table. Coarser sand and gravel are not concerned, as pores in their soil matrices are too large for capillary forces to prevail over gravity. However, soils are seldom uniformly distributed in space. And since smaller pores induce greater tension, the capillary fringe will be thicker where the sediment is smaller and well-graded. This complicates modeling of the capillary fringe since in-situ measurements are always limited. Yet some empirical formulations have been developed to predict the thickness of the capillary fringe (Turner and Nielsen, 1997; Atherton et al., 2001).

Apart from empirical formulations, numerical models for groundwater flow including the capillary fringe have also grown substantially over the past few decades. Geng et al. (2017) compared numerical and field results and showed that a thicker capillary fringe reduced the infiltration of waves into the unsaturated zone. A thicker capillary fringe results in higher moisture content near the soil surface and reduces the available space for infiltration. Besides, a thicker capillary fringe was found to enhance horizontal flow induced by tidal forcing.

Earlier work from Nielsen and Turner (2000), and Cartwright et al. (2003, 2004), as cited in Horn (2006), stated the influence of a capillary fringe to be even more important for higher frequency forcing. Observations have revealed that groundwater waves propagated faster and decayed slower with the presence of a capillary fringe.

The reversed Wieringermeer effect corresponds to an excessive enhancement of the water table when a small amount of water rises the capillary fringe to the soil surface causing the pore tension to be released all at once. A small amount of water then causes a jump in the groundwater level in the order of magnitude of the capillary fringe, even though it is triggered by only a few millimeters of additional water content. A schematic of the phenomenon is given in Figure 2.3. This phenomenon also occurs near the shoreline due to the infiltration of swash into the beach matrix. Swash oscillations cause the sequential destruction and formation of meniscuses at the ground surface. This results in pressure head variations in the order of decimeters even tho the effective infiltration of swash is a few millimeters of water (Turner and Nielsen, 1997).

However, this phenomenon is mostly restricted to the swash zone where the capillary fringe lays close to the ground surface and infiltration by waves occurs. The work of Li et al. (1997, 1999), as cited in Horn (2006), suggested that the location of the top of the capillary fringe away from the
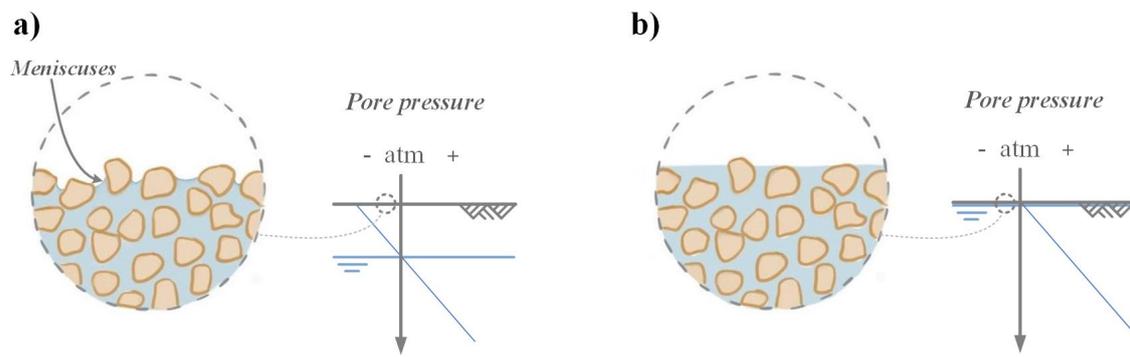
Figure 2.3: The reversed Wieringermeer effect is illustrated through a schematic representation. In (a), the capillary fringe nearly reaches the ground surface, displaying meniscuses. In (b), the system's response to a small water influx is depicted. Adapted from Turner and Nielsen (1997).

swash zone changes slowly in time. For the capillary fringe to change, water has to flow through the unsaturated zone and this process is significantly slower than flow through saturated media.

2.2.6. Fluidization of the bed

The earlier discussed infiltration and exfiltration of the beach matrix are thought to be the main drivers of the morphological changes in the swash zone. However important these may be, it is essential to address the potential for fluidization within the bed as a driver for morphological changes (Horn, 2002).

Fluidization takes place when water flows upward and, the induced drag force is greater than the specific weight of the submerged sediment. For fluidization to occur hydraulic gradient must be large enough and 'ground water outflow from a beach in response to a falling tide is unlikely to be sufficient' (Baird et al., 1998). Hence, fluidization is presumed to happen preferably under swash-induced hydraulic gradients over the seepage face (Baird et al., 1998). The exact mechanisms behind fluidization in the swash zone remain intricate and are still under investigation.

2.2.7. Three-dimensionality of beach groundwater flow

To date, predominant scholarly attention on groundwater dynamics has been directed toward two-dimensional flow and transport along a shore-perpendicular plane. The heterogeneous character of soils was already mentioned. Soil heterogeneity can be either natural due to slow geological processes or, anthropomorphic due to beach nourishments or engineered structures. Either way, the along-shore variability of the soil matrix can influence the groundwater dynamics significantly.

Geng and Michael (2021) simulated groundwater flow through into different geologic structures. They found that the along-shore component accounted for 40% to 50% of the total flow path for the considered volcanic and deltaic soil structures. As anticipated, preferential passages with high permeability significantly influence the trajectory of water particles. An extra parameter should be added to the intricate relationship between beach characteristics and the beach reaction to hydrodynamic forcing.

2.3. Runup extraction from video imagery

2.3.1. Georeferencing of video images

Previously, the complexity of nearshore dynamics and the importance of reliable runup data for validating predictive models were highlighted. Various runup measurement techniques were discussed, leading to the selection of video imagery. The implementation of this method will be described here.
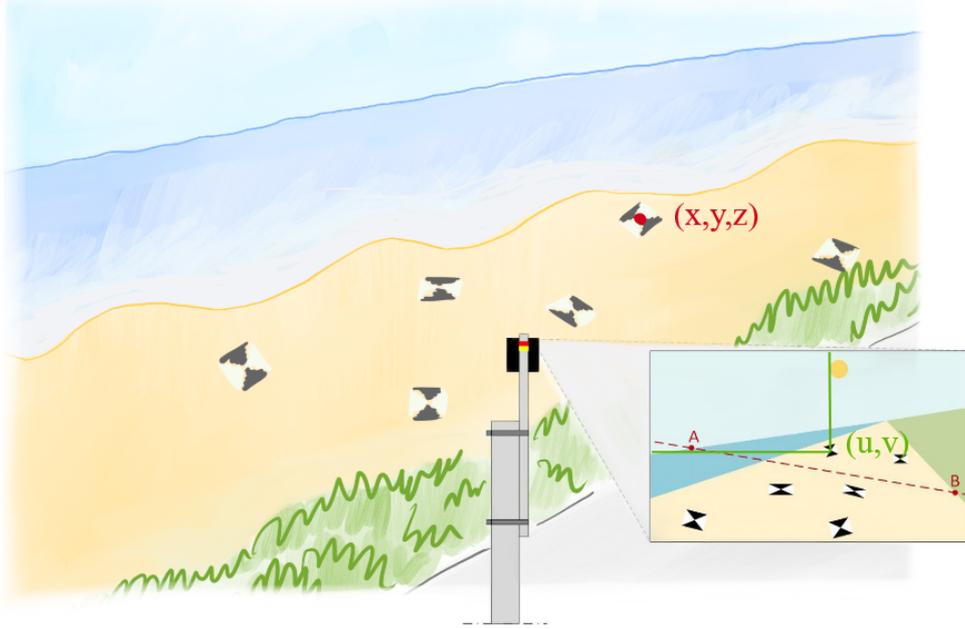

Figure 2.4: Artist impression of the relation between real-world and pixel coordinates in a setup for runup extraction.

Runup extraction from video images requires a relation between the 3-dimensional real-world coordinates x, y, z and the 2-dimensional pixel coordinates u, v (see Figure 2.4). For this purpose, the homogeneous photogrammetric equation described in (Hartley and Zisserman, 2003) can be followed (Eq. 2.8):

$$
\begin{bmatrix} U_D \\ V_D \\ 1 \end{bmatrix} \leftrightarrow \begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & U_0 \\ 0 & -f_y & V_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -C_{XW} \\ 0 & 1 & 0 & -C_{YW} \\ 0 & 0 & 1 & -C_{ZW} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}
$$

$$
= \quad K \quad \cdot \quad R \quad \cdot \quad C \quad \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}
$$

$$(2.8)$$

where on the left hand side U_D, V_D represent the distorted pixel coordinates and U, V denote the undistorted pixel coordinates. On the right-hand side, the calibration matrix, denoted by K, incorporates the focal length f_x, f_y along the x- and y-axis and, U_0, V_0 correspond to the coordinates of the principal point of the image. The rotation matrix, denoted by R, hides the three unknown rotation angles: azimuth, tilt, and swing, defined in Figure 2.5. The camera matrix C contains the
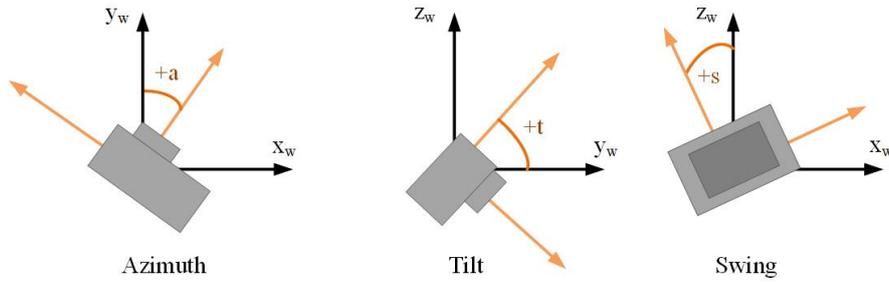
Figure 2.5: Angle definition for ZXZ rotation matrix.

camera's positional coordinates. Finally, the X_W, Y_W, and Z_W correspond to the real-world system.

Two types of camera distortion can be distinguished. Namely, radial and tangential distortion. Radial distortion can be described by Equation 2.9:

$$U_{distorted} = u(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$
$$V_{distorted} = v(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

(2.9)

where k_1, k_2, and k_3 are the radial distortion coefficients, and r the distance from the center of the image such that $r^2 = (u - u_0)^2 + (v - v_0)^2$. Tangential distortion is described by Equation 2.10:

$$U_{distorted} = u + 2t_1 uv + t_2(r^2 + 2u^2)$$
$$V_{distorted} = v + t_1(r^2 + 2v^2) + 2t_2 uv$$

(2.10)

where t_1 and t_2 are the tangential distortion coefficients and r is the distance to the center of the image. Following the ZXZ rotation (Bruder and Brodie, 2020), the rotation matrix can be described by Eq. 2.11:

$$R = \begin{bmatrix} -\sin(a)\sin(s)\cos(t) - \cos(s)\cos(a) & \cos(s)\sin(a) - \sin(s)\cos(t)\cos(a) & -\sin(t)\sin(s) \\ \sin(a)\cos(t)\cos(s) - \cos(a)\sin(s) & \sin(s)\sin(a) + \cos(a)\cos(t)\cos(s) & \cos(s)\sin(t) \\ \sin(a)\sin(t) & \sin(t)\cos(a) & -\cos(t) \end{bmatrix}$$

(2.11)

with the three rotation angles azimuth (a), tilt (t) and swing (s) defined in Figure 2.5.

Solving this system requires intrinsic and then extrinsic camera calibration. The intrinsic calibration comes back to solving all the parameters inherent to the camera settings, namely, lens distortion coefficients, focal lengths, and the coordinates of the principal point of the image, specific to every camera setting combination. The distortion coefficients and the parameters for the camera calibration matrix can be estimated by shooting multiple photos of a checkerboard pattern with known dimensions under different orientations. Then, a least square adjustment of the parameters is made relative to the known relative real-world coordinates of the corners of the checkerboard. The *OpenCV* library or the *Matlab* toolbox from Bruder and Brodie (2020) are user-friendly options to obtain the camera intrinsics.

Once the intrinsics are known, the photogrammetric extrinsics can be determined. They consist of six unknowns relative to the camera orientation and position, defined by the three angles azimuth, tilt, swing, and the camera position in x, y, z. To solve the extrinsic parameters of the homogeneous photogrammetric equation one needs three distinct ground control points (GCPs) with known real-world and pixel coordinates. In practice, using more GCPs than the minimum required is often

beneficial to increase the accuracy of the result, and to have fallback options for erroneous measurements (Bruder and Brodie, 2020). The extrinsic parameters should be solved for every deployment individually, as small changes in the camera position and orientation greatly impact the results.

Once all intrinsic and extrinsic parameters are estimated a simple relation is obtained, linking each camera pixel with coordinates u, v to the real world in x, y, z. However, one should remember that the reliability of the result decreases further away from the GCPs. Hence, these should be placed around the area of interest.

2.3.2. Methods for runup extraction

Often used in runup extraction are Timex, Variance, and Timestack images (Andriolo, 2019). Time-exposure (Timex) images are acquired by averaging the RBG pixel intensity over a defined sampling period. Variance images correspond to the RBG pixel intensity standard deviation throughout sampling. A timestack image is created by extracting the same line of pixels from sequential frames of a video and stacking these lines side by side. The result is an image in time and space of a singular transect. These images are then further processed to extract nearshore properties such as the water line and the effluent line but also nearshore wave transformation domains (Andriolo, 2019), wave celerity (Almar et al., 2009), wave direction (Kuo et al., 2009), wave height (Colvin et al., 2020) or bathymetry (Aarninkhof and Ruessink, 2004; Uunk et al., 2010).

Manual selection is probably the oldest method for runup extraction from video images. A few examples of research in which runup was selected manually from timestack images are Holman et al. (1993); Huisman et al. (2011); Power et al. (2011); Atkinson et al. (2017); Yang et al. (2022). The redundancy of the manual method has not yet discouraged researchers from using it. Despite the persistence of manual selection, recent advancements have led to the development of several alternatives that enhance efficiency and accuracy in runup extraction.

Color Contrast methods

Color contrast (CC) methods have been abundantly developed over the past few decades to automate runup extraction. Bailey and Shand (1994) proposed a method based on processing techniques of grayscale images. The routine involved smoothing, edge detection, normalizing the contrast in time, blurring, and finding the least cost path. The method seemed to agree relatively well with the pristine data they presented. However, even for their almost noiseless timestack, the smaller runup events were often missed, the top of runup events was smoothed, and the backwash was cut off resulting in saw-tooth-shaped runup time series. Alternatively, methods based on Colour Channel Difference, CCD, have also been proposed. Turner et al. (2001) computed the difference between the red and blue color channels. The resulting image holds great information because sand pixels contain relatively more red than blue and vice versa for water pixels. Alternatively, Simarro et al. (2015) proposed a runup extraction method based on pixel variance. Their methods agreed with hand-picked results but did not work quite as well when runup occurred within the seepage face.

Pixel Intensity Clustering, PIC (Aarninkhof et al., 2005; Uunk et al., 2010), was implemented into the Argus coastal monitoring system to segregate water from sand on Timex images. This method translates the RGB color space to the HSV color space, separating color (Hue, Saturation) from the intensity (Value). Then, the two-dimensional histograms of color (Hue–Saturation) and luminance (Value–Value) are constructed. These typically show two peaks corresponding to the wet and the dry pixels. The histogram with the highest contrast is retained for pixel classification. For this, a line

passing through the saddle point of the histogram is drawn. Zhang and Zhang (2009) approached the problem from the HSV color space, from which they derived metrics for the image roughness. Then they used Otsu's method to classify pixels and extract the water line from timestack images.

Complications often arise from CC methods when applied on dissipative beaches because of the wet-dry boundary. If the effluent line is decoupled from the waterline, which can happen on dissipative beaches, a thin layer of water remains on the sand, hampering the distinction between water and sand. Furthermore, the relatively lower energy levels under dissipative conditions contribute to more subtle gradients, less easily distinguished from the background noise. This is especially true for the backwash (Huisman et al., 2011).

Radon Transform routine

Almar et al. (2017) showed promising results with a runup extraction method based on Radon transformation, which they validated on an energetic dissipative beach. This processed-based method separates the incident component (uprush) from the reflected component (backwash) on timestack images by using the Radon transform. Although the model returned a water line that agreed well with LiDAR results in the considered cases, it requires wave conditions to hold a certain level of energy. If not, the runup does not present foamy white characteristics and the Radon routine can't identify areas of high intensity to extract the water line.

Machine-learning algorithms

With recent technological advancements, runu[] extraction has leaped forward (Kingston, 2003; Rigos et al., 2014; Vousdoukas et al., 2011). In a more recent study Marti-Puig et al. (2024) extracted the waterline averaged in time from Timex images using a bidirectional Long Short-Term Memory (bi-LSTM) Network over cross-shore pixel arrays. Their results agreed well with the manually picked water line. Building on these technological strides, researchers are now extending these methods to address new challenges and improve the accuracy and efficiency of coastal monitoring.

Water, dry sand, wet sand, and wrack can easily be distinguished by the eye. Currently, CNN base architectures are being implemented to automate the detection process of these markers (Kang et al., 2024a), to develop Stochastic properties of coastal flooding (Kang et al., 2024b). For now, their research focused on developing pixel recognition but did not include georeferencing. Combining the two will facilitate the collection of precise runup measurements over extended periods. Collins et al. (2023) employed a CNN architecture, utilizing LiDAR data to extract the instantaneous waterline from timestack images, with four input channels: elevation, reflectance, and the variance of elevation across both space and time. They achieved impressive results, but the use of LiDAR data remains a limiting factor due to the high cost of the equipment and the computational demand. The use of multi-channel input has not been implemented yet for regular video imagery data.

2.3.3. Considerations on runup extraction from video images

Multiple CC-based runup extraction methods have been developed, highlighting the ongoing interest in this area of research. However, these methods often encounter difficulties on dissipative beaches, where seepage face disturbances complicate water line detection (Huisman et al., 2011). Despite these challenges, the growing number of publications on machine learning algorithms signals a promising shift towards more advanced techniques.

Traditionally, machine learning approaches for runup extraction have been limited to utilizing

RGB or grayscale images as input channels. In a significant advancement, Collins et al. (2023) demonstrated that machine-learning algorithms, when informed by multi-channel LiDAR data—including elevation, reflectance, and spatial-temporal elevation variance—can accurately delineate the water line. This success presents a notable step forward.

What if the same principles applied in multi-channel LiDAR data processing could be adapted for video imagery? Video imagery, which typically provides only RGB color channels is limited in the sense that it does not provide elevation or reflectance information. However, by preprocessing video images using insights gained from CC-based methods new channels could be generated. Combining these processed images to inform a multi-channel machine-learning algorithm, it may be possible to significantly enhance the accuracy and reliability of runup detection from video sources.
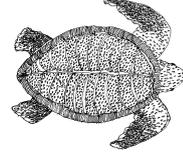
<div align="right">3</div>

Research Design and Methodology

This chapter outlines the methodology employed to analyze wave runup, from data collection with three different runup extraction methods to empirical estimations. The approach integrates traditional and advanced extraction techniques, including manual digitization, machine learning (ML), and color contrast (CC) methods to enhance the accuracy and reliability of runup measurements. Figure 3.1 presents a schematic overview of the runup analysis.
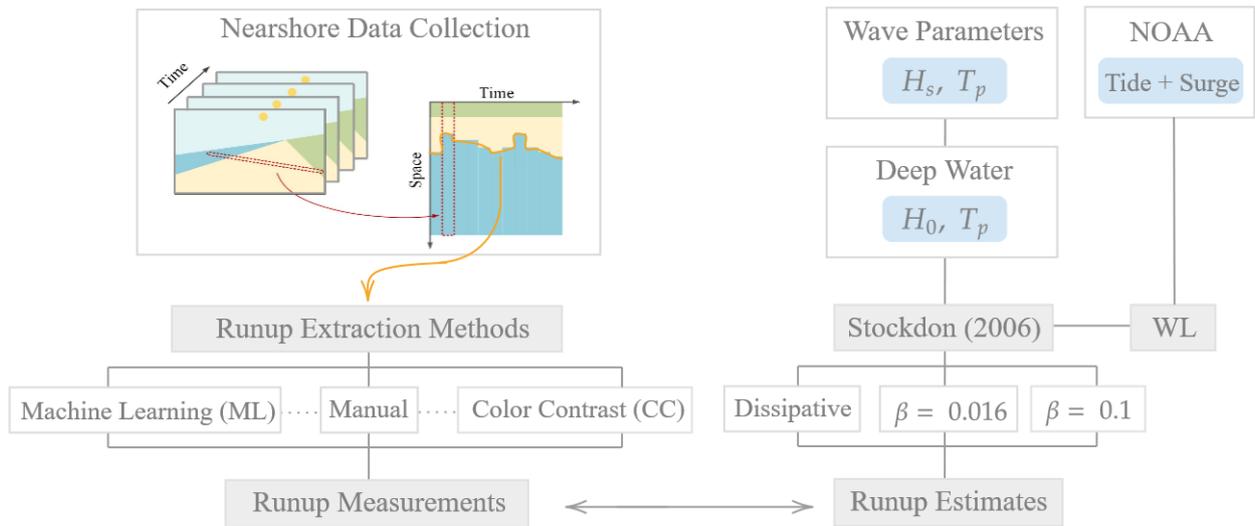

Figure 3.1: Schematic overview of the runup analysis. This diagram details the process from initial data collection through runup extraction methods from timestack images including Manual, Machine Learning (ML), and Color Contrast (CC), to the final runup measurements. It also depicts the workflow for empirical runup estimates using Stockdon et al. (2006) for dissipative conditions and with the general formula with beach slopes $\beta = 0.016$ and $\beta = 0.1$. As input for the empirical estimates, the measure offshore wave parameters (H_s and T_p) were computed back to deep water wave parameters (H_0 and T_p). The absolute runup height was obtained summation of the measured mean water level (tide + surge) to the Stockdon estimates (setup + swash).

The chapter is structured as follows: Section 3.1 characterizes the study site in Galveston, TX, USA. Section 3.2 details the collection and processing of video imagery into georectified timestack images and outlines the manual digitization process. Sections 3.3 and 3.4 describe the methodologies for runup extraction using color contrast (CC) and machine learning (ML), respectively. Section 3.5 discusses additional optimization steps to enhance both models. Finally, Section 3.6 examines

the runup characteristics and the application of the Stockdon et al. (2006) formulas for empirical estimation.

3.1. Site Characterization

3.1.1. Spatial synopsis

Galveston Island is located on the Gulf Coast of Texas, USA, along the Gulf of Mexico. Positioned approximately 50 miles southeast of Houston, it is a barrier island known for its sandy beaches. Figure 3.2 a cartographic representation delineating Galveston Island and its proximate geographical environs is presented for reference.
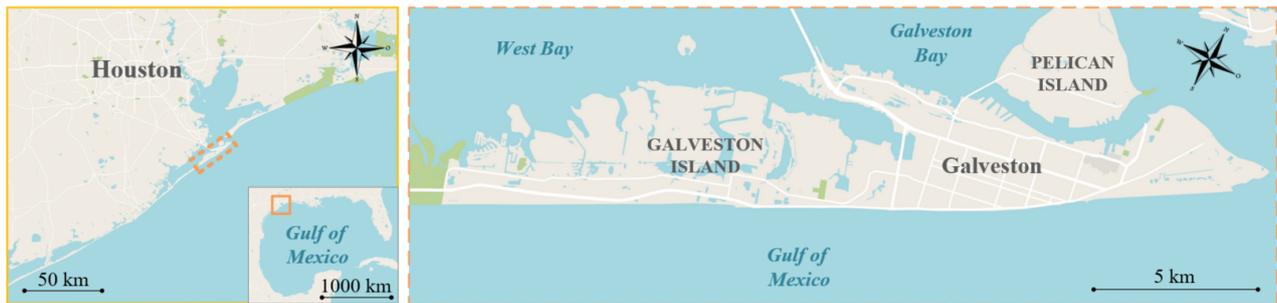

Figure 3.2: Cartographically delineates the Gulf of Mexico, the coastal region proximate to Houston, Texas, and the topography of Galveston Island.

On the southwest end of the Island, a natural tidal inlet delimits the island. Its position coincides with the location of the Brazos River channel belt 6000-4000 years ago. The inlet is extremely stable and existed for 2100 years (Dellapenna and Olszewski, 2023). At the northeast end, a long jetty delimits the Island and protects the entrance of the shipping channel into Galveston Bay.

3.1.2. Sea turtles in Galveston

Galveston Island provides an ideal study setting for exploring the hydrodynamics associated with sea turtle nesting. Five species of sea turtles are found in the Gulf of Mexico (Valverde and Holzwart, 2017), the Kemp's ridley (*Lepidochelys kempii*), loggerhead (*Caretta caretta*), green (*Chelonia mydas*), leatherback (*Dermochelys coriacea*), and hawksbill (*Eretmochelys imbricata*). Besides, nesting of sea turtles has been shown to increase along the northern Texas coastline (Valverde and Holzwart, 2017), including Galveston Island where the Kemp's ridley is regularly found to lay eggs. This conquest of the northern Texan coast by sea turtles may be interpreted as an extension of the Kemp's ridley nesting grounds or as the reinstatement of their original nesting habitat (Seney, 2008).

3.1.3. Hydrological Features

Surface currents in the Gulf of Mexico

Large-scale currents in the Gulf of Mexico play an important role in shaping the region's marine environment. Below, in Figure 3.3 the main surface currents averaged over the Spring-Summer and the Autumn-Winter periods are shown.

The loop current, a large, warm current that originates from the Caribbean Sea, turns around in a clockwise direction above Yucatan and leaves the Gulf through the Strait of Florida where it joins the Gulf Stream. Meandering Eddies detach from the loop current and travel westwards into the
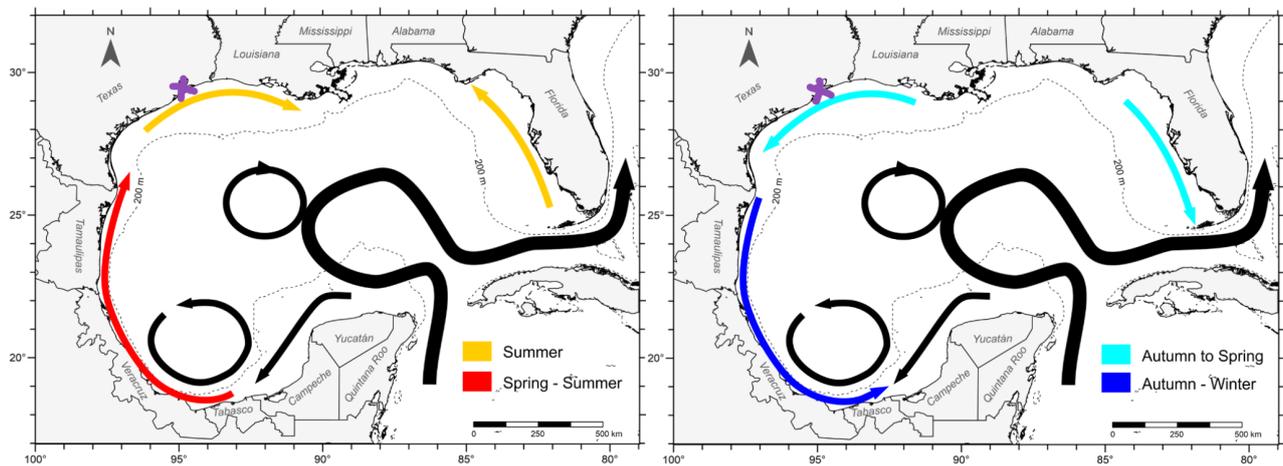
Figure 3.3: Schematic of the main surface circulation patterns in the Gulf of Mexico during the Spring-Summer and Autumn-Winter periods (Sanvicente-Añorve et al., 2018). Galveston is indicated with a purple cross.

Gulf of Mexico. Some of these Eddies are stable in time and are found back over seasonal averages.

The direction of the surface velocities shows a high correlation with the velocities at 30m depth (Sanvicente-Añorve et al., 2018). Hence, surface circulation patterns are thought to be representative of the direction of the shallow nearshore. In Figure 3.3 one can observe a seasonal shift in the longshore current of the Texan coast. During the Summer period, the averaged longshore current is directed towards the northeast while from Autumn to Spring, the averaged longshore current is directed towards the southwest.

Tidal properties

The tidal signal at Galveston Island is mixed semidiurnal. The Gulf of Mexico is characterized as a micro-tidal basin meaning that astronomical tide typically fluctuates within the range of 0.3 to 0.9 meters over a spring-neap cycle (NOAA, 2024b).

Despite being classified as a micro-tidal basin, tidal ranges within the Gulf of Mexico can vary significantly. The observed water level and the NOAA estimates regularly differ with a factor two (Lefevre et al., 2000), and the predictions have been found to jump from accurate to an error in this range in only a few hours (Huff et al., 2020). A drop in the atmospherical pressure over a significant area, wave-induced setup, and wind tides can have a great impact on local water levels (Huff et al., 2020) and are not always accounted for in predictions.

Wave climate

Historical data from 1993 to 2021 shows some seasonal variation in the wave direction at Galveston Island. The prevailing wave direction is southeast-south with a bias towards the southeast from November to February, and a bias towards the south from June to August (NOAA, 2024a). This coincides with the previously discussed seasonal variation in the direction of the current along the coast of Texas. Oriented towards the northeast during Summer and towards the southeast from Autumn to Spring.

Due to the protected Gulf environment wave energy remains relatively modest in Galveston during normal conditions. During normal conditions from 2007 to the present, 95% of the waves were smaller than 1.3m, 99% of the waves smaller than 2m, and 99.9% smaller than 3m (Surf-Forecast, 2024). However, Galveston is susceptible to the influence of tropical storms and hurricanes, especially

during the Atlantic hurricane season from June to November. During these events, offshore waves exceeding 10m have been measured (Nayak and Panchang, 2015).

3.1.4. Soil composition and geology

During the Pleistocene epoch, large ice sheets covered significant portions of the Earth's surface, resulting in a global sea level approximately 85 meters lower than present levels. As the ice sheets gradually melted, the sea level began to rise. This rise in sea level led to the formation of coastal ridges that eventually separated from the mainland, giving rise to new barrier islands. Among these newly formed barrier islands is Galveston Island, which protects salt marshes and an open bay against the sea.

Galveston Island began to accumulate 5500 years ago on top of the old Brazos Delta and continued growing seaward until 1200 years ago when the sediment supply must have run out. At that time, the coastal profile was rather steep but relative sea level rise and storms slowly eroded the island according to the model described by Bruun (1962); Swift (1968). The erosion process resulted in the redistribution of sediment, with a portion being washed over the island and filling its lagoon, while the remainder settled at the bottom of the Gulf of Mexico. This sediment redistribution was facilitated by multiple washover and storm surge channels, which enabled the exchange of water and sediment between the sea and the bay. As a result of this erosion and sediment redistribution, the coastal slope of Galveston Island became milder, leading to the formation of the dissipative beach observed today. The deltaic history of Galveston Island resulted in poorly graded beaches with a D_{50} ranging in between 0.07 to 0.1mm (US Amry Corps of Engineering, 2022).

3.1.5. Infrastructure and urbanization

In 1900, a category 4 hurricane struck Galveston Island, resulting in 10,000 casualties and destroying more than one-third of all the buildings. (National Geographic, 2024). This event led to the construction of a 15km long seawall starting at the jetty on the northeast side of the island to the southwest end of the city. The elevated position of the sea wall concerning the beach allows for a strategic position of the video system.

The seawall was intended to protect Galveston from storms. As an inevitable side effect, it caused the starvation of the longshore drift, exacerbating erosion for kilometers southwest of it. Paine et al. (2020) quantified erosion rates from 2000 to 2019. They reported retreating coastlines at about 2m/yr south of the seawall and advancing coastlines at similar rates along the seawall. The latter is mainly due to the consecutive suppletions intended to protect the seawall from storms and to maintain an aesthetically pleasing appearance. The last recorded sand deposit dates back to 2019 (US Amry Corps of Engineering, 2022) when sediment from the shipping channel composed of 38% fines was placed on the beaches of Galveston (Maglio et al., 2020).

In addition to the coastal erosion exacerbated by the seawall, Galveston Bay faces environmental challenges from pollution and habitat loss. During hurricanes, significant amounts of pollutants are released into Galveston Bay, impacting water quality (Du et al., 2020; Yang et al., 2021). Furthermore, salt marshes in the area are rapidly eroding due to various factors, including wave action, subsidence, eustatic sea-level rise, and insufficient sediment supply (Ravens et al., 2009). Salt marshes, that serve as vital sources of food for sea turtles.

3.2. Data collection

3.2.1. Overall setup

Monitoring the instantaneous water line was part of a data collection campaign motivated by the need to validate predictive models for the flooding of sea turtle nests (see Figure 3.4).
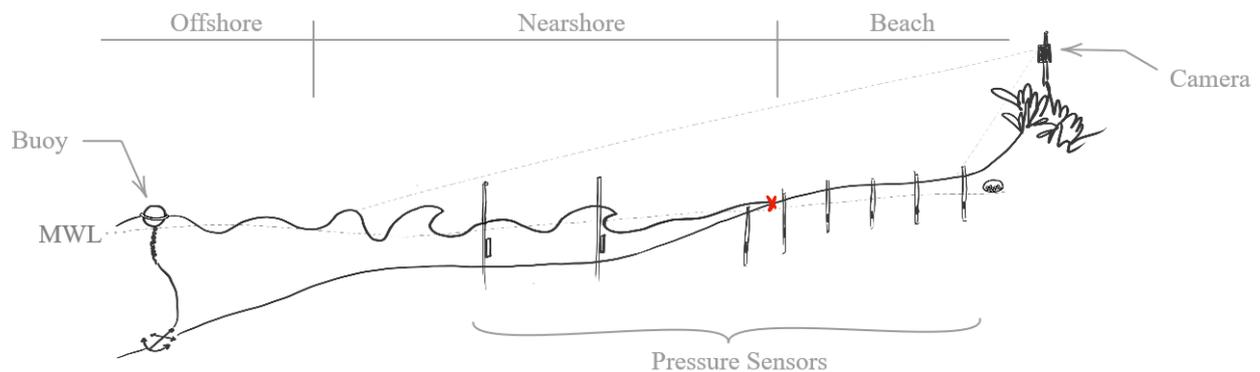

Figure 3.4: Setup for data collection campaign.

The campaign involved collecting offshore wave data by deploying a buoy around 1 km from the shoreline, tracking wave evolution in the nearshore area with submerged pressure sensors, extracting wave runup data using video imagery, tracking of the nearshore water surface elevation, and monitoring the groundwater table with pressure sensors installed 1 meter below the ground surface in wells.

3.2.2. Camera deployment for timestack generation

Camera choice and settings
The GoPro Hero 10 Black was available for this measurement campaign and was chosen as the preferred camera for its compact design, durable construction, and suitability for outdoor environments. Moreover, the GoPro's flat lens enhances the camera's ability to effectively shed raindrops, ensuring clear visibility during adverse weather conditions. The linear lens was adopted for filming to minimize distortion in captured images.

A crucial consideration in the camera selection process involved striking a balance between temporal resolution and memory usage. GoPro cameras have a relatively high bitrate which results from its compact but not-so-efficient storage algorithm which causes the memory to fill up quickly. The GoPro Hero 10 Black supports a maximum SD card storage of 128 GB. These aspects are vital for budgeting data storage resources throughout the monitoring period. A frame rate of 2 frames per second (fps) was deemed sufficient to capture runup smoothly. The videos were recorded in 4K to maximize the spatial resolution. Herewith, the cost of filming for one day, approximately 7 to 9 hours, translated to a data storage requirement of 30 GB. Acknowledging the limitation of the GoPro's internal battery life, lasting approximately 1 hour and 30 minutes at 2 fps and 4K resolution, an external battery was integrated into the setup. This addition served to extend the filming duration, enabling longer monitoring sessions without interruptions.

Camera installation
The beaches were monitored by deploying two GoPro cameras positioned to capture comprehensive views of the transects of interest. Figure 3.5 provides a visualization of the setup.
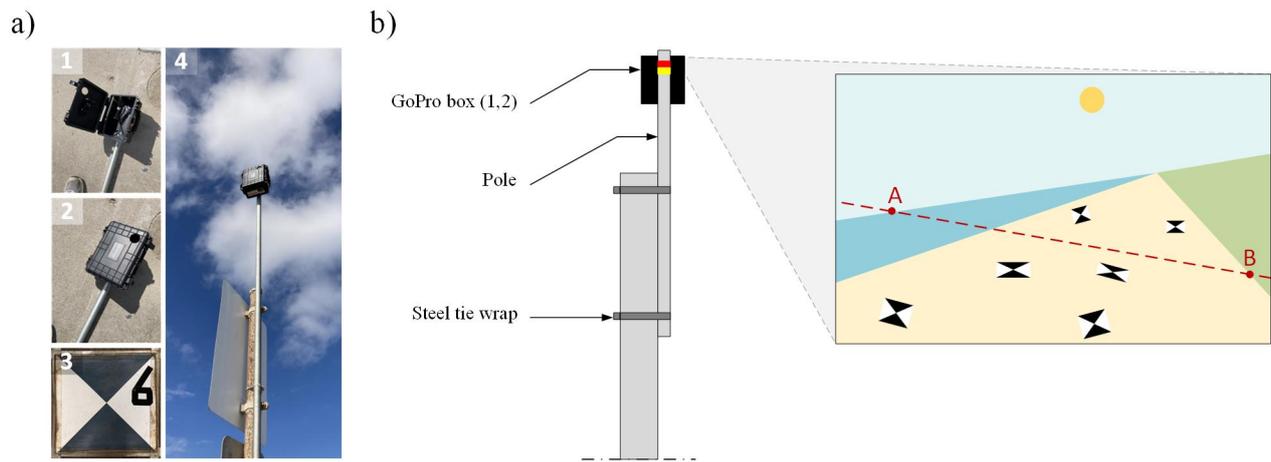
Figure 3.5: Deptiction of GoPro installation. In (a), pictures of the GoPro box (a.1,a.2), a GCP (a.3), and the installed system (a.4). In (b), a schematic of the GoPro installation with the field of view.

The camera is securely enclosed within a protective box, which is then mounted to a pole for stability and height. The camera is strategically positioned to capture a comprehensive view of the transect, spanning from the distant horizon to the base of the dune. Additionally, it includes six Ground Control Points (GCPs) within its field of vision. Through the Quick app, users can access a real-time stream of the camera's view and verify that the field of view satisfies the requirements.

The GCPs are strategically positioned above the waterline, ensuring they are visible to the camera. Their arrangement is designed to prevent collinearity as much as possible and to cover the area of interest. The precise locations of the camera and the six GCPs were determined using a Leica GS08 GPS RTK device with an accuracy of 3cm.

Camera calibration

Intrinsic camera calibration is performed individually for each camera and filming mode to tailor the calibration parameters to specific settings. The calibration procedure involves capturing a short video of a flat checkerboard pattern with known dimensions, observed under diverse angles and distances. A selection of sharp frames, showing the complete checkerboard pattern, are extracted from the video, and the *OpenCV* for Python toolbox is employed to recognize corner points in these frames. Radial and tangential distortions - the fisheye effect and non-parallel alignment of the lens and image plane - are corrected using the *OpenCV* toolbox. The outcome is a camera-specific calibration matrix encompassing focal lengths, optical center location, and the Root Mean Square (RMS) reprojection error. The RMS reprojection error serves as a criterion for the quality of calibration and is expected to be sub-pixel. The obtained calibration matrix is then applied to undistort all videos filmed with the same camera and with identical setting combinations.

Extrinsic camera calibration involves the georeferencing of two-dimensional undistorted images to real-world coordinates. The calibration process is accomplished through the solution of a photogrammetric algorithm with six unknowns, comprising three rotation angles of the camera and the real-world coordinates of the camera in XYZ dimensions. The equation can be solved with three ground control points (GCP) with known real-world and pixel coordinates. In practice, more GCPs are needed to improve the accuracy and reliability of the solution. Six ground control points (GCPs) were deployed each time the camera was used, allowing for adjustments in case of measurement errors. The *Matlab toolkit* from Bruder and Brodie (2020) was used to approximate the solution.

A timestack image is obtained by extracting a line of pixels from sequential frames from a video recording and putting the arrays one after another in chronological order. This way, one transect can be analyzed in space and time.

To obtain a timestack, the average transect is computed for each deployment and converted into pixel coordinates through the earlier solved photogrammetric equation, ensuring that the results are georectified. The corresponding RGB arrays from each frame are then extracted and stacked on top of one another to generate the timestack image, which provides a spatial and temporal representation of the transect. Figure 3.6 gives a visualization of the process.
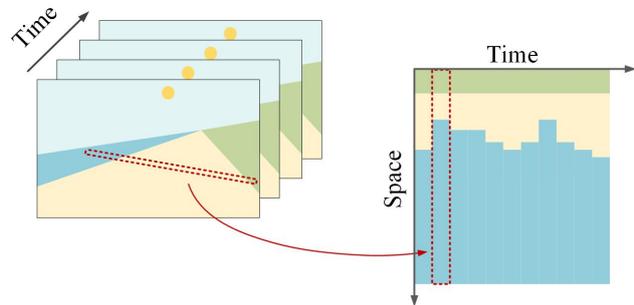

Figure 3.6: Schematic of timestack generation.

From timestack properties to xyz

An averaged xy cross-shore transect was computed over each deployment. The z-coordinates of the beach profiles were interpolated to the nearest points. Additionally, to achieve smooth interpolated profiles, a Savitzky-Golay filter was applied. The filter had a window size of 2001, indicating the number of data points used for smoothing, and a polynomial order of 3. The results are varying beach profiles over a fixed cross-shore transect in the horizontal plane.

The timestack images represent the mean cross-shore beach transect in time and space with fixed x- and y-coordinates. Once the water line is identified on this spatiotemporal image, it can be mapped back to the corresponding profile. On days where no profile was measured the closest GPS measurements in time were used.

3.2.3. Manual digitization

The instantaneous position of the water line water manually digitized on nine timestack images of 1 hour 30 minutes, except one that was 28 minutes 25 seconds long. A Matlab script was designed for this purpose. A moving window allows one to move along the timestack in time and select the contours of individual waves (Figure 3.7). The wave uprush was easily detected by a sharp white/gray contrast. The backwash however was less obvious to the eye but was determined as the darker region following each uprush. The reason for this complication related to backwash detection originates from the dissipative character of the beach. When waves attain the shoreline, they have lost most of their energy already, and, when they retract part of the wave infiltrates the beach matrix and part flows back slowly over an extremely gentle slope. Resulting in an unclear delineation of water and sand. During manual digitization, the backwash was precisely estimated in the darker region, and a natural decay function was consistently followed, maintaining approximately the same slope as the uprush. The timestack images that were manually digitized are summarized in Table 3.1.
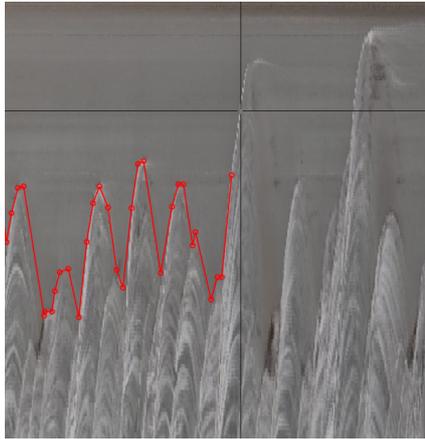
Figure 3.7: Snapshot of the manual digitization process.

Table 3.1: Data compendium with days, video IDs, start and end times, and whether or not the water line was manually labeled.

Day	Video ID	t_{Start}	t_{End}	Labeled
2023-11-13	GX010081	13:46:00	15:06:00	Yes
2023-11-13	GX020081	15:06:00	16:26:00	Yes
2023-11-13	GX030081	16:26:00	16:54:25	Yes
2023-11-14	GX010084	08:50:00	10:10:00	Yes
2023-11-14	GX020084	10:10:00	11:30:00	Yes
2023-11-14	GX030084	11:30:00	12:50:00	Yes
2023-11-14	GX040084	12:50:00	14:10:00	Yes
2023-11-14	GX050084	14:10:00	15:30:00	Yes
2023-11-14	GX060084	15:30:00	16:50:00	Yes

3.3. Development of a CC-based runup extraction method

3.3.1. Image processing techniques for nearshore features

The cleanest timestack image, taken from the video shot on November 14th (video ID: GX060084), was selected for the first steps of the optimization process. The underlying assumption is that processing techniques may not significantly enhance the swash signature in such a clean image and might yield poorer results on timestack images with additional disturbances, such as beach wrack, and variations in lightning. Conversely, successful feature extraction from a clean timestack image does not necessarily imply similar success on more noisy images. Various image processing techniques were applied to analyze nearshore features. Results from each processing technique were analyzed to determine their efficacy in enhancing and identifying nearshore features. Visual comparisons were drawn to assess which techniques provided the most significant insights into the dynamics and characteristics of the near-shore area and, were most promising to use as channel input for a CC-base method, and later, for a machine-learning algorithm.

Grayscale conversion and gradient calculation
Converting an RGB image to grayscale reduces the initial three channels to one intensity channel, corresponding to the third channel of the HSV color space. The used conversion from the *OpenCV* library reads as:

$$I = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \tag{3.1}$$

where I denotes the new grayscale channel derived from the original image, while R, G, and B correspond to the red, green, and blue channels, respectively. Intensity images distinguish lighter from darker features such as foam in the swash zone from the sand above it. Calculating the gradients in time, dI/dt, and space, dI/dx, reveals information about the dynamical processes within the images (Figure 3.8). These processing techniques emerged in the early dates of runup extraction, with among others Bailey and Shand (1994).

The gradient time and space enhance the swash signature in the timestack images. However, a considerable amount of noise persists in these images. Especially above the effluent line where the beach isn't smoothed out by the passage of waves. Although the swash is now easily distinguishable from the beach to the naked eye, the pixel composition of both areas isn't uniformly distributed,
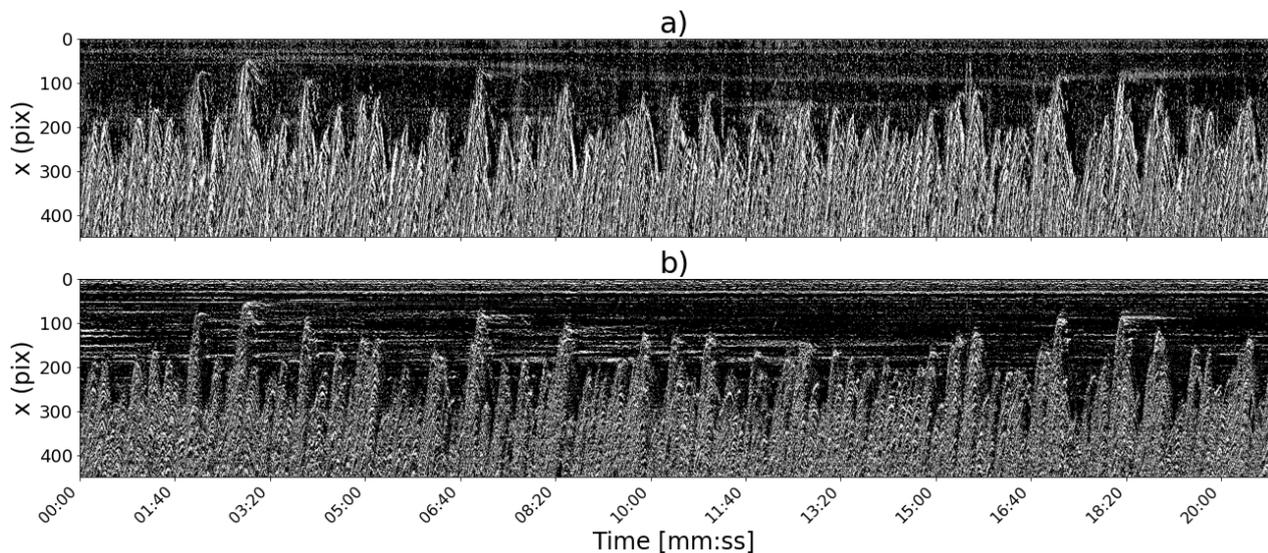
Figure 3.8: Panel (a) displays the temporal gradient of the grayscale image, denoted as dI/dt, while panel (b) shows the spatial gradient of the grayscale image, denoted as dI/dx. Both images were histogram-equalized for visualization purposes.

which hinders the binarization of the image. Even if the background seems dark compared to the light swash, both areas contain large amounts of dark and light pixels. Therefore, dI/dt or dI/dx of the grayscale images aren't suitable for extracting the water line from timestack images.

What catches the eye in Figure 3.8a and 3.8b is the difference in texture between the two areas. The swash appears rough and disordered, whereas the beach seems more organized and smooth, interrupted occasionally by a horizontal line. This observation is logical considering the highly turbulent nature of the swash zone and the smoothing effect of wave action on the beach.

Local entropy

A measure for roughness has already been implemented for runup extraction by Zhang and Zhang (2009). Generally, local entropy is used to assess texture roughness in images, helping to distinguish between different materials or conditions by capturing variations in pixel intensity. In this study, local entropy was employed as a measure of the complexity or disorder within an image.

The swash-zone dynamics are characterized by turbulent movements that create significant gradients in both time and space. These gradients are pronounced when compared to the more stable and less dynamic sandy areas. The *entropy* function from the *skimage.filters.rank* module of the *Scikit-Image* library was used here. This function computes the local entropy in an image, effectively highlighting areas of high complexity and disorder, such as those found in the swash zone.

First, a local neighborhood is defined. Therefore the *disk* function from the *skimage.morphology* module with a radius of 3 was implemented. Then, the local normalized histogram is computed internally for each pixel according to (Eq. 3.2):

$$p_i = \frac{h_i}{\sum_{j=0}^{n-1} h_j} \tag{3.2}$$

where p_i is the probability of intensity i in the local neighborhood, n is the number of distinct intensity levels, and h_i the pixel count per intensity entry i. Finally, the local entropy E of the local neighborhood is computed as (Eq. 3.3):
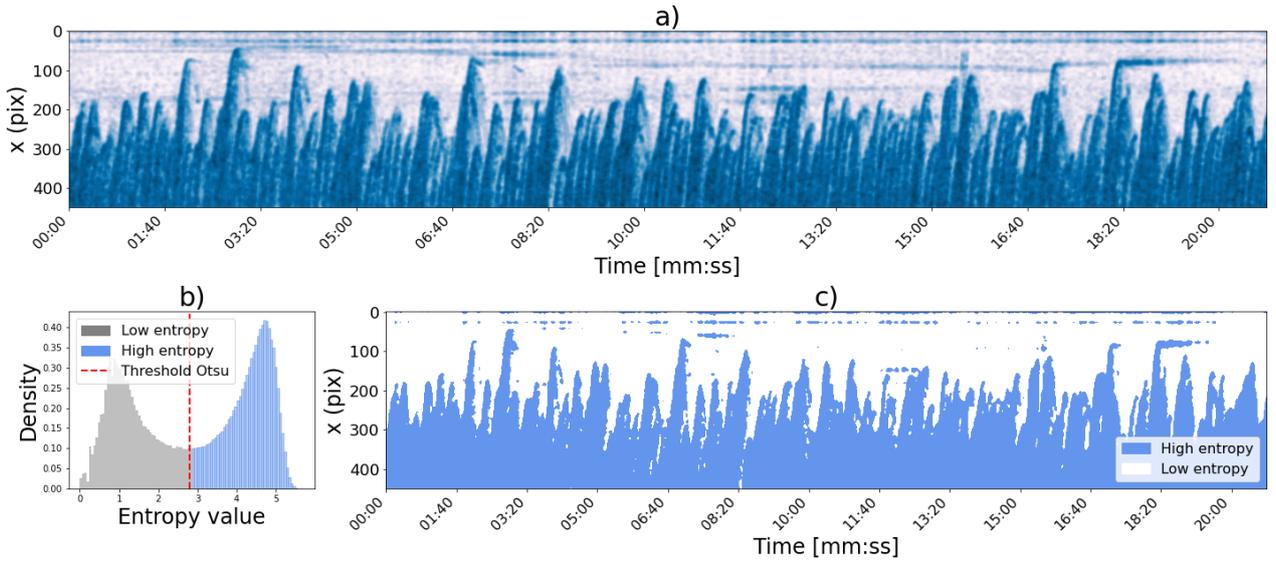
Figure 3.9: Swash zone delineation using the local entropy technique: (a) represents the dI/dt image after applying local entropy; (b) shows the histogram of local entropy values with the optimum threshold determined by Otsu's method (Otsu et al., 1975) marked; and (c) depicts the resulting binary classification, distinguishing between high and low entropy.
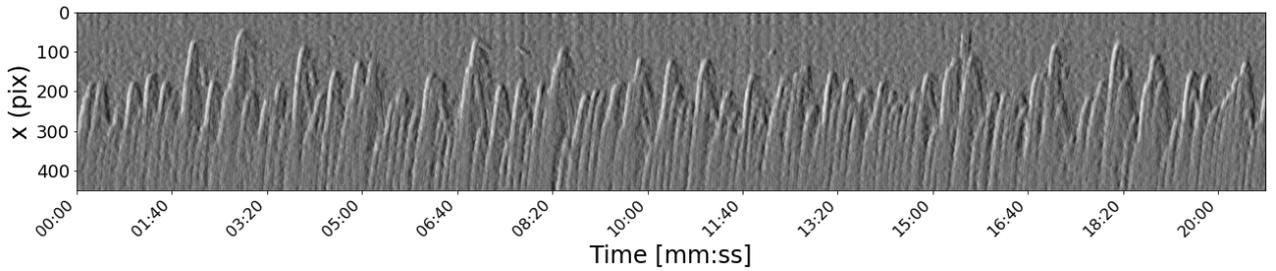

Figure 3.10: The gradient in time of the local entropy of dI/dt. Light values correspond to positive gradients while dark values correspond to negative gradients.

$$E = -\sum_{i=0}^{n-1} p_i \log_2(p_i) \tag{3.3}$$

The result is an image where each pixel describes the randomness of the surrounding area, allowing the inference that high entropy values correspond to regions of high dynamics. Hence, the turbulent swash motion. Figure 3.9 shows the binarization process for swash extraction from a timestack image using the local entropy of dI/dt, as the local entropy of dI/dx gives the almost identical image.

The swash and the beach of the entropy image seem to have a uniform pixel distribution (see Figure 3.9a). This is confirmed by the histogram in Figure 3.9b that shows a strong bimodality (see Figure 3.9c), meaning that the image can easily be binarised using Otsu's method (Otsu et al., 1975). After binarization, some noise persists. Especially, some areas above the effluent line were classified as high entropy.

Taking the gradient in time of the local entropy of dI/dt gives insightful results. This is not the case for the gradients of the entropy over space, which are not depicted here but tend to enhance noise. Figure 3.10 shows the dynamic changes in entropy over time, revealing the shape of individual waves passing by.

The uprush can be distinguished clearly by the light region on the image. The backwash, though

less striking, can be recognized by the darker regions that follow each uprush. This image is not suited for the binarization of water and sand but could be processed further using an edge detection algorithm to extract the waterline, or used to refine the machine-learning algorithm.

Saturation and Red-Blue channel operations

On reflective beaches with orange-yellowish sand, the Red-Blue channel operation can segregate the swash from the sand quite well, because sand pixels contain more red than the water pixels, in contrast to water pixels containing more blue. The saturation channel from the HSV color space also segregates water from sand effectively (Aarninkhof et al., 2005; Uunk et al., 2010). However, on dissipative beaches, the Red-Blue channel operation does not enhance the swash, because the seepage face disturbs the signal (Huisman et al., 2011). Decoupling of the water line from the effluent line results in a thin layer of water covering the beach above the swash and below the effluent line. Sand grains located in this area are then falsely classified as water. Hence, under dissipative conditions, the Red-Blue channel operation doesn't enhance the swash but delineates the effluent line instead, as does the saturation channel from the HSV color space. Figure 3.11 shows the implemented routine to extract the effluent line from the saturation channel.
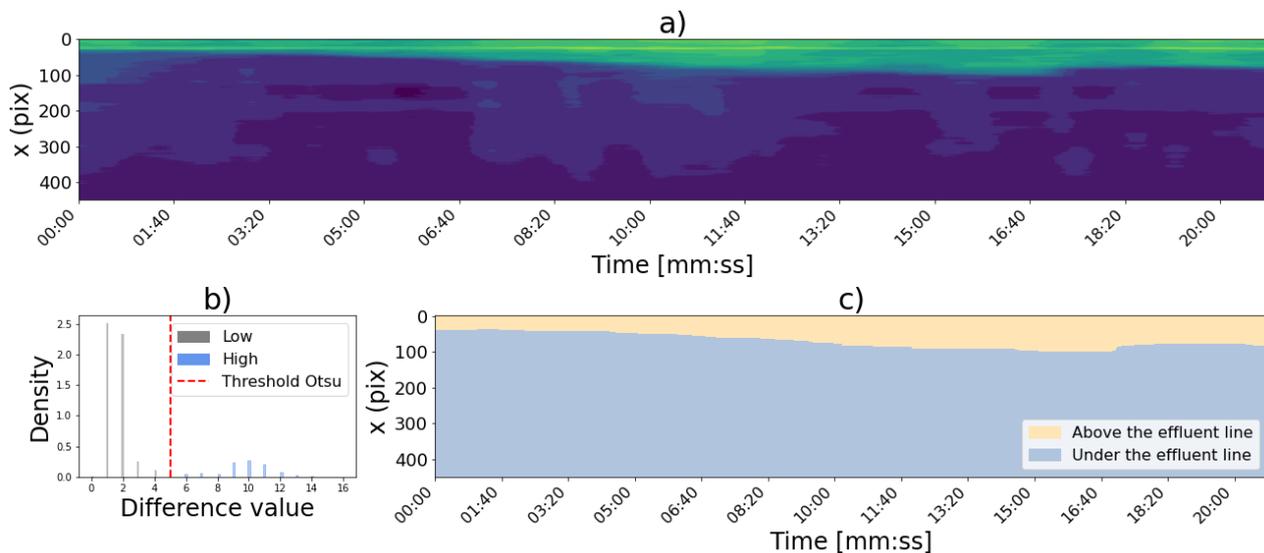

Figure 3.11: Effluent line detection via the Saturation channel: (a) the image after Gaussian-smoothing with a kernel of (201x1), (b) the histogram of the smoothed image, and (c) the resulting binary classification relative to the line according to Otsu's method (Otsu et al., 1975).

First, the saturation channel is obtained from the HSV color space with the *OpenCV* library. Subsequent application of Gaussian blur over time with a 201×1 kernel allows for a clearer distinction between the regions above and below the effluent line, depicted in Figure 3.11a. The strong bimodality of the corresponding histogram in Figure 3.11b further confirms this distinction and allows for binarization of the image thanks to Otsu's method (Otsu et al., 1975). The results clearly distinguish between the region above the effluent line and the area below (Figure 3.11c).

3.3.2. A CC-based entropy-saturation method

Combining entropy and saturation results in a simple method for runup extraction for dissipative beaches, where the image is first binarized according to local entropy values, and noise reduction is
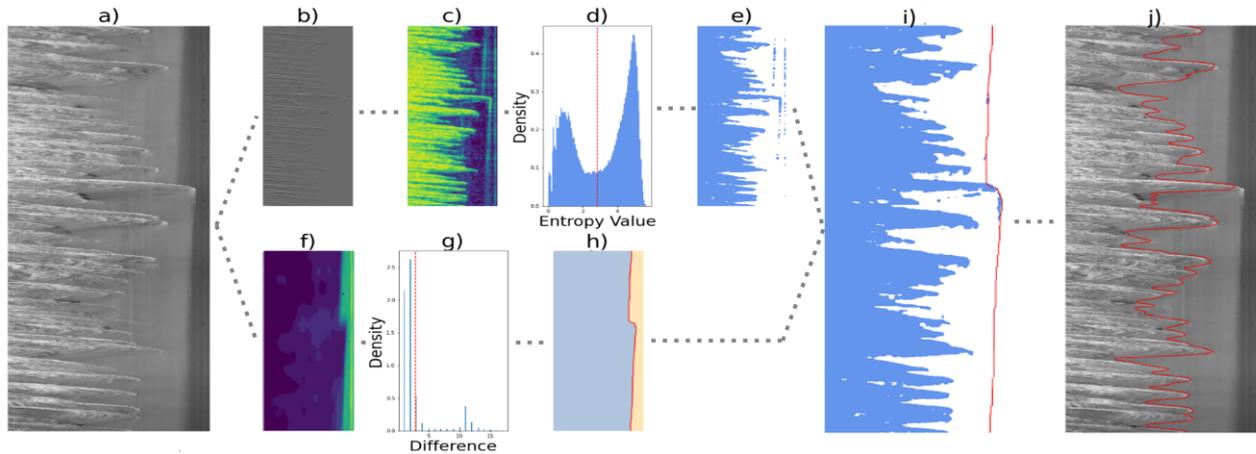
Figure 3.12: Diagram of runup extraction process utilizing local entropy and differential red-blue channel operation. In (a) the grayscale image with pixel intensity I (b) dI/dt of the grayscale image with pixel intensity I, (c) the local entropy image computed with a disk of 5, (d) the corresponding histogram with Otsu's threshold (Otsu et al., 1975), (e) the binarized image, (f) the saturation image from the HSV color space after gaussian blurring in time with (g) the corresponding histogram with Otsu's threshold, (h) the binarized image with the effluent line in red, (i) the results binarized entropy image with all high entropy values above the effluent line deleted and (j) the grayscale image with the computed water line in red.

achieved by deleting all values above the effluent line computed with the Red-Blue channel operation. Figure 3.12 gives a visualization of the method.

For this particular timestack image, the uprush, backwash, and local maxima are captured by the method. However, the highest peak is shortened a little due to the noise reduction process. Due to Gaussian blurring over the time axis, the limit between wet and dry sand is smeared out (see Figure 3.12f), and when deleting the noise above the effluent line the top of the larger runup events are cut off as well.

3.4. Development of a ML runup extraction method

3.4.1. Optimization and model evaluation framework

Optimization scheme

The resulting image has each pixel describing the randomness of the surrounding area, allowing the inference that high entropy values correspond to regions of high dynamics. The question now is whether these images can serve as input channels to a machine learning (ML) algorithm for runup extraction, similar to what Collins et al. (2023) did with LiDAR data. The considered images are the grayscale image, I, the intensity changes over time, dI/dt, the intensity changes over space, dI/dx, local entropy, the entropy changes over time, dE/dt, the original RGB image, and the saturation image, S. Figure 3.13 displays a schematic of the optimization and evaluation scheme.

The model is trained to recognize sand from water because labeling only the water line would introduce significant errors. This is due to the difficulty in precisely selecting the boundary between sand and water. Additionally, focusing solely on the water line would lead to a large class imbalance, which could negatively impact the model's performance. The model is first evaluated on its ability to accurately predict classes through the mean and standard deviation of the accuracy across folds. This encompasses the search for an adequate smoothing function for each input channel, as well as the determination of the optimal combination of channels for the model. Later, the parameters from
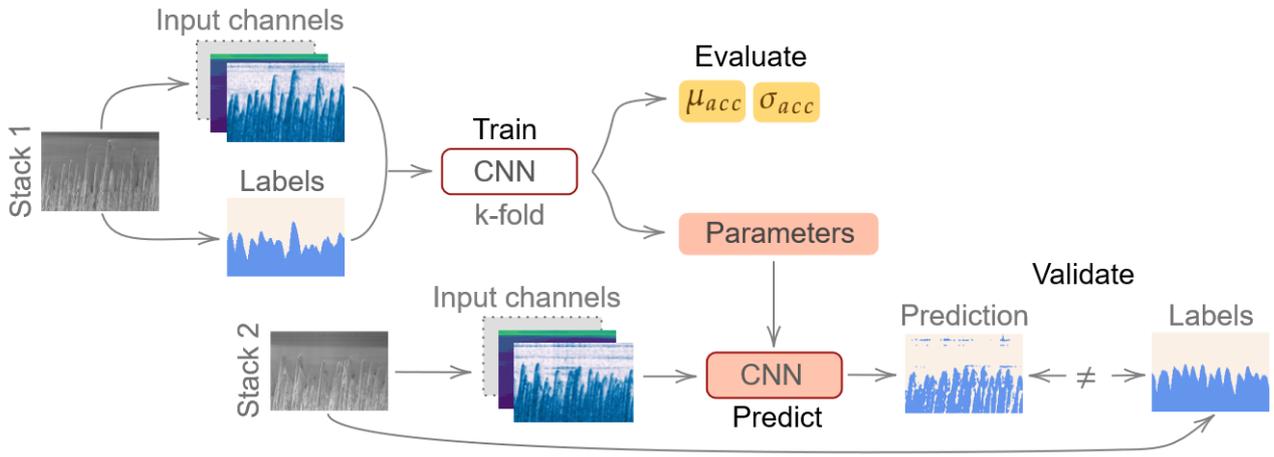
Figure 3.13: Optimization and evaluation scheme of ML model for runup extraction.

the optimum model are used to predict water and sand on new timestack images. The resulting water line is then compared to the hand-picked one, and the performance is assessed by local correlation. A more detailed explanation is provided in Section 3.5.1.

CNN architecture

A machine-learning algorithm designed to take n input channels according to a classic Convolutional Neural Network (CNN). A *Sequential* architecture was implemented using the *tensorflow.keras.models* python library.

The input dimensionality of the model was set to small patches, specifically 60×60 pixels, and 'n' input channels. Resulting in a $60{\times}60xn$ input. This decision balances computational efficiency with sufficient spatial and temporal resolution, ensuring that the physical processes at play are effectively captured. The spatial resolution of the input channel's height corresponds to 4.2 meters in real-world coordinates, while the width represents 30 seconds.

A rudimentary CNN was implemented. Convolutional layers are a standard for image processing as they capture features efficiently thanks to their shared weights but local receptive field. Setting up 32 and 64 successive convolutional layers seemed to result in sufficient learning capability for water and sand detection during input channel testing. A standard kernel 3x3 was chosen to capture local features without losing too much spatial resolution. To avoid overfitting and reduce memory load, each convolutional layer was followed by a MaxPooling layer downsampling the feature map in height and width by a factor of 2 by taking the maximal value over a window of 2x2.

Consequently, the local receptive field widens, allowing the network to build a more global view of the earlier computed features. After, the feature maps are transposed back to recover the dimension lost in earlier layers so the output is of the same dimension as the original image. This process involves upscaling the image by inserting zeros between pixels and then applying a convolutional kernel to fill in the blanks. Then, finally to achieve binary image segmentation the final Conv2D with a 1x1 kernel and sigmoid activation, typical for binary image segmentation.

It was determined that the model should focus on distinguishing between water and sand, rather than directly identifying the water line. This decision was based on the observation that the manually selected water line will likely contain more inaccuracies than the broader classification of areas as either water or sand. Segregating water from sand presents a less error-prone task, as the extensive

areas covered by each category provide more reliable and consistent data than the precise delineation of the water line.

k-fold cross-validation

The performance and generalizability of our image processing models were assessed through k-fold cross-validation. This involves the partitioning of the data into k equally sized subsamples. The cross-validation process is repeated k times, where every subsample is retained once as testing data while the remaining k-1 subsamples are used for training. This process allows for a statistical description of the training procedure. Here, the k resulting accuracies were averaged and the corresponding variances were computed and used as a metric for model performance. During the initial optimization stages, k-fold cross-validations were conducted within a single timestack lasting 1 hour and 30 minutes, using k = 5. This approach aimed to strike a balance between statistical reliability and computational costs. Subsequently, the training dataset was expanded to include 9 timestacks. K-fold cross-validation was then performed across these 9 timestack images to further enhance robustness and validate the model's performance across a broader temporal scope.

Evaluation metrics

During the first part of the optimization process, model performance was assessed through the model's accuracy and precision. Where the accuracy was defined as the percentage of correctly predicted pixels, such that:

$$\text{Accuracy} = \frac{\text{Number of correctly predicted pixels}}{\text{Total number of pixels}} \times 100\% \tag{3.4}$$

and the standard deviation of the accuracy in between folds:

$$\text{Precision} = \sigma_{acc} = \sqrt{\frac{\sum_{i=1}^{k}(x_i - \mu_{acc})^2}{k-1}} \tag{3.5}$$

where μ_{acc} is the mean accuracy in between folds i, and k the total number of folds.

3.4.2. Input-channel assessment

In the first optimization steps, the ML algorithm is trained on a single timestack of 1 hour and 30 minutes (video ID: GX060084) and validated on a new timestack from the same day (video ID: GX050084). The k-folds are thus done by splitting up one single timestack image.

Sensitivity to smoothing-kernel dimensions

The optimum smoothing kernel dimensions were assessed through the grayscale image only, based on the assumption that all relevant features shared similar spatial and temporal dimensions throughout input channels. Different Gaussian smoothing kernels were applied. This smoothing method is commonly used to reduce noise in images. It involves applying a weighted average to the pixels within a specified neighborhood around each pixel in the image. The weights are determined by the Gaussian function, which ensures that pixels closer to the center of the neighborhood have a higher influence on the smoothed value. The *GaussianBlur* function from the *OpenCV* library was used

A horizontal kernel in time of size $n \times 1$, and a square kernel of dimension $n \times n$ were evaluated for different values of n using a k-fold cross-validation. The best n was then chosen based on the

model's accuracy and precision as described just above. The resulting best horizontal and square kernels were then applied to all input channels and assessed as described in the following section.

Input-channel sensitivity to smoothing techniques

To evaluate the efficacy of different image-smoothing techniques on the model's accuracy for different input channels, a pairwise comparison between smoothing methods and a baseline to which no smoothing was applied. To this end, two variants of the Gaussian blurring method were examined; one employing a horizontal kernel, smoothing across time only, with a length of 3 pixels, and the other using a square kernel, smoothing across both time and space, with dimensions of 3x3 pixels. These resulted from earlier sensitivity analysis of the model's accuracy and variance due to kernel size. For each input channel, the mean accuracy and the variances of the two methods were individually compared to the no-smoothing baselines, according to the following hypothesis:

- h_0: *The mean accuracy and the standard deviation of smoothing method X in between folds equals the no-smoothing method.*
- h_1: *The mean accuracy or the standard deviation of smoothing method X in between folds is lower/higher than the no-smoothing method.*

The mean and the variance were assessed through a t-test and a f-test respectively. The one-sided t-test evaluates whether the mean of method X in between folds is significantly lower than the no-smoothing baseline. The f-test assesses whether the standard deviation of the accuracy of method X is larger than that of the no-smoothing baseline. The threshold for statistical significance was established at $\alpha = 0.05$ for both tests. A p-value above that threshold implies no significant difference between the accuracy of the smoothing method and the no-smoothing baseline. A p-value below that threshold is a substantial detriment to the model's accuracy due to the blurring method. To choose the optimum smoothing preprocessing technique for each input channel, preference was given to the squared kernel if it did not result in accuracy loss. The underlying assumption is that maximal smoothing of the original image provided it does not obscure relevant shapes, will facilitate easier feature extraction. This approach aims to reduce noise and irrelevant details, enhancing the algorithm's ability to discern and process key features more effectively. If the square kernel resulted in a significant decline in accuracy, the horizontal kernel was envisaged. When the horizontal kernel did not result in a substantial decline in accuracy, it was chosen for the given input channel, otherwise, the no smoothing method was chosen.

Assessing input-channel combinations

Out of the seven input channels, all 127 combinations were evaluated. Channel normalization is crucial when combining different types of preprocessed images. Standardizing the scale of the features across input channels ensures that all inputs contribute equally to the learning process, preventing features with larger scales from dominating the learning. The model was trained using k-fold cross-validation for each of the 127 combinations. On a laptop equipped with an Intel Core i7 processor and 16 GB of RAM, this took approximately 10 hours to complete. After, a cluster of best combinations was selected based on accuracy and standard deviation. This resulted in 24 combinations that gave similar performances. Finally, the channel combination with the highest accuracy was retained.

3.5. Further optimization of CC-based and ML method

3.5.1. Assesment of input statistics for model performance

Model performance according to the hand-picked water line

Until now, the model was optimized according to the mean and standard deviation of the accuracy across folds. These metrics give information about the number of correctly predicted pixels across categories. However, they do not disclose much about the shape of the runup estimate computed from these pixel estimates. Therefore, once the best combination of input channels was selected the ML model was validated on a new timestack image. The water line was computed by column-wise summation of the sand pixel, binarized as 1, and compared to the manual and the CC methods. No additional smoothing was applied to the obtained water line to evaluate the standalone performance of the ML method. For both methods, the Pearson correlation r coefficient between the hand-picked water line and the model outcome was computed over a moving window of 100 grid cells in time, as (Eq. 3.6):

$$r = \frac{\sum_{i=1}^{100}(x_i - \bar{x})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^{100}(x_i - \bar{x})^2 \sum_{i=1}^{100}(m_i - \bar{m})^2}} \tag{3.6}$$

with x_i the pixel runup values of method X, $\bar{x}$ the mean runup value of methods X, m_i the pixel runup values of the manually selected water line, and $\bar{m}$ the mean runup value of the manual results. The edges of the runup estimates were padded to ensure that the resulting correlation coefficients time series had the same temporal dimension as the runup time series (see right side of Figure 3.14). The correlation coefficients with the entire hand-picked water line were computed to assess the models' overall performance.

Correlation of model performance and input statistics

To identify the eventual causes of low model performances, the Pearson correlation coefficient between the full correlation times series and the different input statistics were computed for each model and each input channel (see Figure 3.14). Therefore, the input channels were reduced to the time dimension by taking the mean, standard deviation, maxima, and minima over each column. Next, the correlation coefficient of each statistical measure was calculated against the correlation coefficient between the model and the hand-picked data. The result is one correlation coefficient per input statistic.
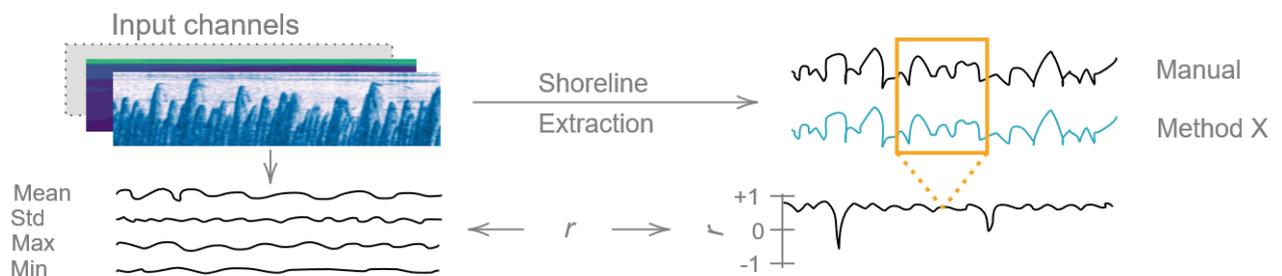

Figure 3.14: Schematic of the evaluation of model performance input statistics. First, the water line is extracted and compared to the manual results. This results in a correlation time series that is then correlated to the statistics of the different input channels.

The correlation coefficient served as a guideline for new optimization measures that could be under-

taken to improve the CC and ML methods. Some of these measures were assessed and gave birth to a more robust model, others were left aside for discussion.

3.5.2. Optimization of CC-based method

Window size reduction
All runup extractions with the CC method were first estimated on full timestack images of 1 hour 33 min. Meaning, that the thresholds were computed according to characteristics over this entire time frame. To try and refine the thresholds to local characteristics a sensitivity analysis of the computational window size was performed. A range of windows from 9000 grid cells down to 5 grid cells in time were tested and evaluated on their correlation with the manual results.

Column-wise normalization
Another attempt to accommodate local temporal variations was a column-wise normalization of the processed images. Three cases were compared to the non-normalized baseline: one where the saturation image was normalized, one where the entropy image was normalized, and one where both the saturation and entropy images were normalized.

3.5.3. Further optimization of ML method

Column-wise normalization
Similar to the CC method, column-wise normalization was implemented for the ML method to accommodate local temporal variations. This time three cases were also compared to the non-normalized baseline, namely, one where the saturation image was normalized, one where the entropy image was normalized, and one where all input channels were normalized. The input channels begin the grayscale image (I), dI/dx, saturation (S), entropy (E), and dE/dt.

Training set expansion
Previously, the ML algorithm was trained on a single timestack of 1 hour and 30 minutes (video ID: GX060084) and validated on a new timestack of the same duration (video ID: GX050084). In total, nine timestacks were digitized. To assess whether expanding the training set improved overall performance, 9-fold cross-validation was performed using the labeled timestacks. The results from the fold where GX050084 was the test timestack (M_1) were compared to the results of the model trained on a single timestack (M_8). The efficacy of this multi-channel expanded model was assessed by comparing it to a single-channel model. Informed only by a grayscale image, I, and trained on eight timestacks as well. Finally, the predictions of the k-fold cross-validation were stitched together and converted to real-world coordinates to assess the ML model performance in terms of runup statistics.

3.6. Assessment through runup characteristics

3.6.1. Analysis of runup statistics

The pixel coordinates were converted to real-world coordinates using the photogrammetric equation (see Section 3.2.2). From this transformation, an elevation time series was obtained from which runup statistics were computed.
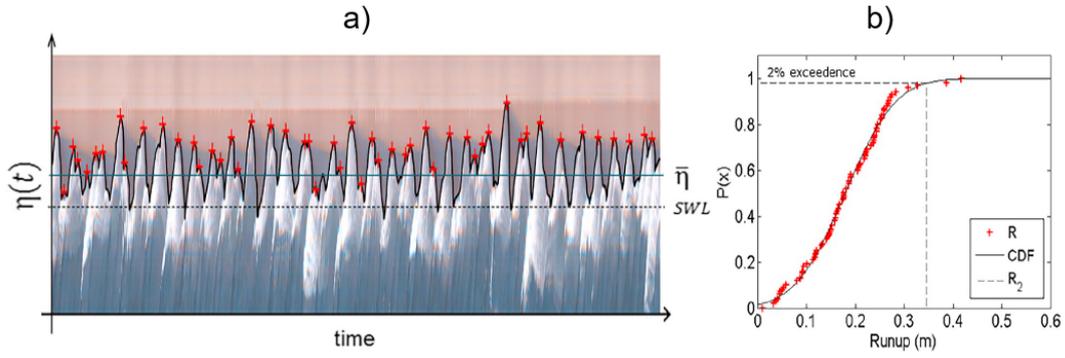
Figure 3.15: In (a), the runup time series with the still water level (SWL) and the mean water level ($\bar{\eta}$), and in (b) the resulting CDF with runup height exceeded by 2% of the events ($R_{2\%}$) from Gomes da Silva et al. (2020).

The runup height was calculated by finding the local maxima (see Figure 3.15a). From the CDF of the runup heights (Figure 3.15b) the runup height exceeded by 2% ($R_{2\%}$) and 50% ($R_{50\%}$) are then derived by taking the 98th and 50th percentile respectively. The runup statistics were evaluated using a 30-minute moving window across the runup time series. Subsequently, a rolling average over 1 hour was applied, and comparisons between methods were made against the manually digitized water line. This comparison was quantified using the correlation coefficient and the root mean square error (RMSE).

3.6.2. Spectral decomposition

The spectra of the runup time series are computed over a duration D of 30 minutes with blocks of length D/16 which satisfies a trade-off between spectral resolution and reliability. The 90%-confidence interval resulting from the variance between blocks was also computed. The frequency ranges of 0.004 to 0.04 Hz and 0.04 to 1 Hz were used to characterize the infragravity waves and the incoming swash, respectively Bertin et al. (2018). To assess the timing aspect of the runup extraction methods the mean period $T_{m-1,0}$ was computed as (Eq. 3.7):

$$T_{m-10} = \frac{m_{-1}}{m_0} \tag{3.7}$$

where m_0 and m_{-1} are the zeroth and the -1th spectral moments. The mean runup period $T_{m-1,0}$ is often used in wave runup as is gives relatively more weight to the lower frequencies, which are more relevant for the runup process on gently sloping beaches for predicting the swash oscillations. The n-th spectral moment is computed as (Eq. 3.8):

$$m_n = \int_0^\infty f^n E(f)\, df \tag{3.8}$$

where f is the frequency, and $E(f)$ is the spectral density function. At first, the spectrum on November 14th from '15:00' to '15:30' was computed and the ML and CC methods were evaluated according to the manually digitized water line. Then for a discontinuous runup time series from November 13th '13:46' to '16:54' and from November 14th '13:46' to '10:10', the S_{ig} and S_{ig} were compared and the correlation and RMSE to the manual results were assessed.

3.6.3. Comparison with empirical estimates

Two formulas were used to predict the wave runup height exceeded by 2% of the runup events. The first is the general expression from Stockdon et al. (2006) used earlier by Ware et al. (2019, 2021) to predict the flooding of sea turtle nests. Which reads as (Eq. 3.9):

$$R_2 = 1.1 \left(\left(0.35 \beta_f (H_0 L_0)^{1/2} \right)^2 + \left[\frac{H_0 L_0 \left(0.563 \beta_f^2 + 0.004 \right)}{2} \right]^{1/2} \right)^{1/2} \tag{3.9}$$

where β_f is the foreshore slope or beach slope, H_0 the deep-water significant wave height, and L_0 the deep-water wave length. The second is the formula from Stockdon et al. (2006) for extremely dissipative conditions, that is applicable for $\xi < 0.3$, and read as (Eq. 3.10):

$$R_2 = 0.043 (H_0 L_0)^{1/2} \tag{3.10}$$

for ξ, the Iribarren number or surf similarity parameter that reads as (Eq. 3.11):

$$\xi = \frac{\tan \alpha}{\sqrt{H/L_0}} \tag{3.11}$$

for α the seaward beach slope. It is worth noticing that the formula for extremely dissipative conditions does not require a beach slope to predict $R_{2\%}$.

The offshore wave conditions needed to drive the runup formulas were retrieved from a buoy deployed at approximately 1km from the site at 7.5 m depth. The significant wave height at that location was translated back to the deep-water significant wave height H_0. Refraction was disregarded under the assumption of alongshore uniformity and energy losses due to white capping or wave breaking were also neglected. First, the wave period at 7.5 m depth is computed by iterating over the dispersion relation (Eq. 3.12):

$$L_{n+1} = \frac{g T_p^2}{2\pi} \tanh \left(\frac{2\pi d}{L_n} \right) \tag{3.12}$$

with the initial guess L_1:

$$L_1 = \frac{g T_p^2}{2\pi} \tanh \left(\frac{2\pi d}{L_0} \right) \tag{3.13}$$

Then, the group velocities at a depth of 7.5 m, c_g, and in deep water, c_{g0}, are computed. From this the shoaling coefficient K_s can finally be computed as (Eq. 3.14):

$$K_s = \sqrt{c_{g0}/c_g} \tag{3.14}$$

And, the deep-water significant wave height is derived (Eq. 3.15):

$$H_0 = H_s / K_s \tag{3.15}$$

Mean water level data (tide + wind surge) were obtained from a nearby station at San Luis Pass (NOAA Station ID: 8771972, Location: 29.080°N, 95.130°W). The total water level was computed as the sum of the observed mean water level and the wave runup predicted by either the general

expression or the formula for extremely dissipative conditions (Stockdon et al., 2006). Table 3.2 provides an overview of all implementations for predicting $R_{2\%}$, and Figure 3.16, displays the mean profile and corresponding beach slopes β.

Figure 3.16: In (a) the mean interpolated profile elevation and variation band during deployment. In (b) the corresponding profile slope.
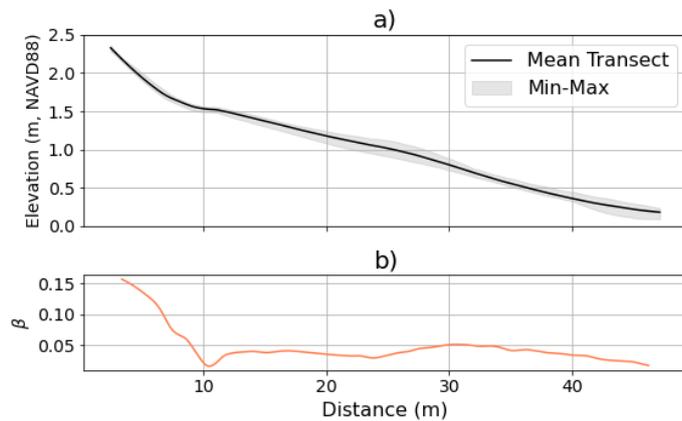
Table 3.2: Different Stockdon implementations for $R_{2\%}$ predictions with Stockdon et al. (2006).

Stockdon formula	β (-)
General expression	0.100
General expression	0.016
Extremely dissipative	-

In Ware et al. (2019), the nest slope –corresponding to the slope at the dune toe– was found to give best estimates for $R_{2\%}$, which corresponds to the flattest section over the entire beach profile and a slope of 0.016 (see Figure 3.16b at 10m distance). To asses how different beach slopes affect the results of the general expression a slope of 0.1 was also evaluated. Finally, the formula for extremely dissipative conditions was also applied, which does not require a beach slope as input.

4
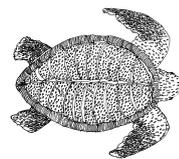
Results

This chapter presents the findings and analyses derived from the development and evaluation of multi-channel informed machine learning (ML) and color contrast (CC) methods for runup extraction. In Section 4.1, the ML method is developed. Section 4.2 focuses on the optimization processes for both the ML and CC methods. Finally, in Section 4.3, the resultant runup time series are rigorously compared with manually digitized values and empirical estimates, providing a comprehensive validation of the proposed approaches.

4.1. Development of a multi-channel informed ML model

Here, the efficacy of channel inputs to inform a machine-learning model is evaluated. The considered channels are the grayscale image, I, the intensity changes over time, dI/dt, the intensity changes over space, dI/dx, local entropy, the entropy changes over time, dE/dt, the original RGB image, and the saturation image. In Section 4.1.1, a smoothing kernel will be selected for each channel individually to reduce noise without losing relevant details. In Section 4.1.2, the predictive performance of each input will be assessed. In Section 4.1.3 the optimal combination of channels for the multi-channel informed machine-learning algorithm will be selected.

4.1.1. Sensitivity to smoothing-kernel

Finding the appropriate smoothing kernel for each channel relied on the assumption that all relevant features shared similar spatial and temporal dimensions. Various kernel sizes, denoted as n, were applied to find the horizontal and squared kernels that best suited the model's accuracy and precision of the grayscale channel (Figure 4.1a). After, the resulting best kernels were applied to all input channels individually through k-fold analysis (Figure 4.1b).

Figure 4.1a shows how the horizontal kernel size of the Gaussian smoothing filter impacts the model's accuracy for a one-dimensional smoothing kernel in time only and for a two-dimensional kernel in time and space. In both scenarios, a slight increase in accuracy and a decrease in standard deviation are observed as the kernel size increases to 3. Beyond this point, accuracy decreases, and the standard deviation increases. This suggests that a smoothing kernel with a temporal dimension of 3 effectively enhances the model's stability and performance without losing relevant details. This value was chosen to assess the impact of pre-smoothing on different input channels.

The boxplot in Figure 4.1b displays the results of the k-fold cross-validation for all input channels with no preprocessing techniques, and pre-smoothed in time with a 3×1 kernel or in time and space
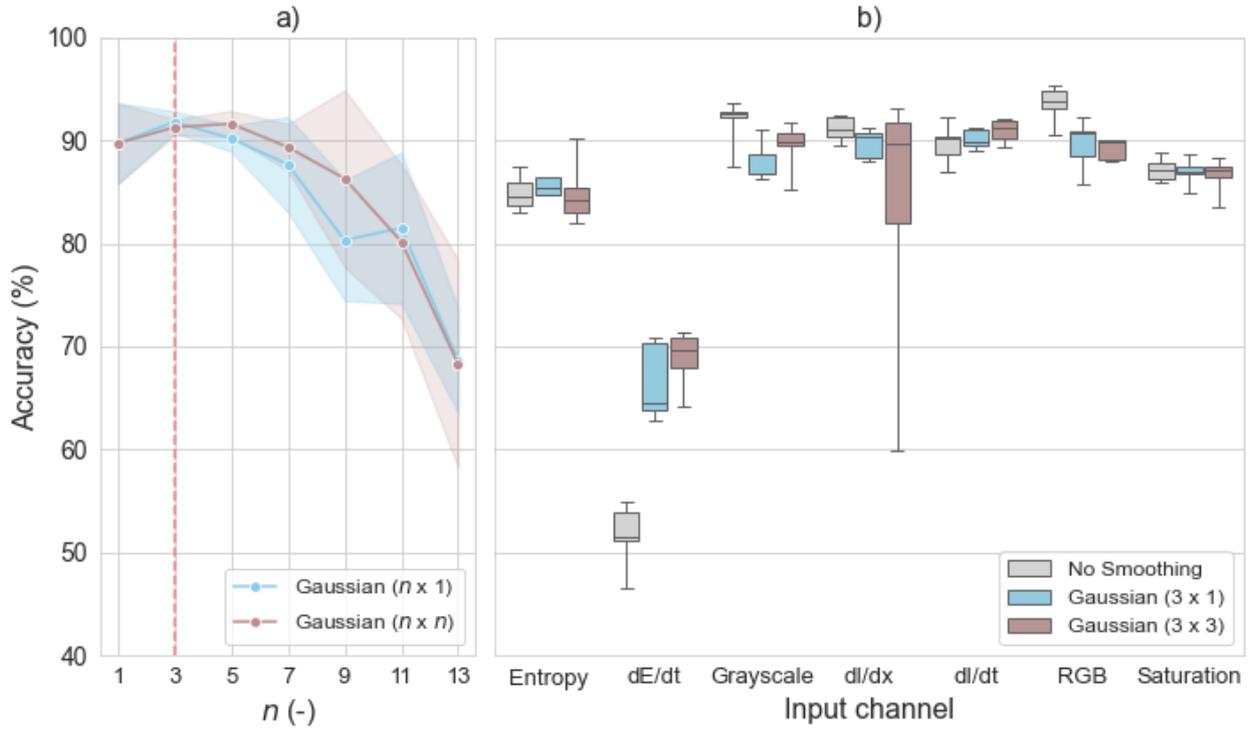
Figure 4.1: In (a), the graph illustrates the mean accuracy as a function of the horizontal kernel size for a pre-smoothed grayscale image using two different Gaussian kernels: a one-dimensional kernel applied in time ($n \times 1$) and a squared kernel that smooths both time and space ($n \times n$). In (b), the boxplot illustrates the k-fold cross-validation outcomes for all input channels under three distinct preprocessing scenarios: a baseline without any smoothing, a Gaussian smoothing with a kernel size of 3x1, and a squared kernel of size 3x3.

with a 3×3 kernel. For the Entropy, Grayscale, dI/dt, RGB, and Saturation input channels the smoothing methods seem to have little impact on the accuracy and the spreading of the results. For the gradient of entropy over time (dE/dt), accuracy increased substantially for both smoothing methods, while the spreading showed only minor differences. For the gradient of the grayscale image over space (dI/dx), the mean accuracy and spreading remained roughly equal for both methods compared to the no-smoothing baseline. However, prepossessing with a 3×3 smoothing kernel resulted in large variations in accuracy. The appropriate smoothing kernel for each input channel was determined by evaluating whether the mean accuracy or the spreading of the accuracy did not significantly decrease/increase upon applying one of the two smoothing filters. The results are summarized in Table 4.1.

The 'Kernel Choice' column indicates the final smoothing kernel choice for each input channel based on the results of the T-test and the F-test. The grayscale image and the RGB image did not seem to benefit from any smoothing method. Hence, the no smoothing baseline was maintained for these input channels. The entropy image, the dE/dt image, the dI/dt, and the saturation image did not suffer in accuracy or precision due to either smoothing method. For these input channels, the square kernel was thus retained. The input channel dI/dt experienced a notable decrease in precision with the square kernel. However, it performed satisfactorily when applying a horizontal kernel along the time axis. Smoothing in the spatial dimension led to excessive detail loss; hence, the horizontal kernel was the preferred choice for dI/dt.

Table 4.1: Paired T-test and F-test results for different smoothing kernels across all input channels, compared to a no smoothed baseline. The kernel choice column indicates the selected kernel based on the p-values of both tests, with a preference for the square kernel when both options meet the criteria.

Input Channel	Kernel Choice	Horizontal kernel (3x1)		Square kernel (3x3)	
		T-test p-value	F-test p-value	T-test p-value	F-test p-value
Entropy	square	0.7618	0.8931	0.5195	0.1414
dE/dt	square	0.9988	0.3709	0.9994	0.5652
Grayscale	no smoothing	0.0171	0.6304	0.0060	0.4837
dI/dx	horizontal	0.1003	0.3531	0.1450	0.0001
dI/dt	square	0.7145	0.9099	0.9556	0.8497
RGB	no smoothing	0.0414	0.2698	0.0024	0.8408
Saturation	square	0.3411	0.3775	0.1873	0.2079

4.1.2. Assessment of single-channel estimates

Each input channel was evaluated individually on its ability to predict water and sand pixels. The smoothing options from Table 4.1, aimed at improving pixel prediction accuracy by reducing noise, were applied to each channel individually. A snapshot of the predicted timestack for all input layers, along with the corresponding confusion matrices that display prediction accuracy and error rates, is presented in Figure 4.2

It appears that the Grayscale (I), dI/dt, dI/dx, and Entropy channels detect individual waves with a discernible degree of precision. However, the Grayscale and the RGB input channels introduce noise between the effluent and the water line. Other channels fail to capture these rapid variations of the water line over time and only provide a rough segregation between water, at the bottom of the time stacks, and sand, at the top. The Grayscale, the dI/dx, and the RGB input channel present grid-like features related to the applied patches of 60x60. This suggests model overfitting due to patch-related discontinuities during training.

Table 4.2: Accuracies of input channels

Input Channel	Accuracy (%)
Grayscale	85
dI/dt	90
dI/dx	84
Saturation	81
RGB	86
Entropy	87
dE/dt	80

Remarkable is that the other predictions do not present these grid-like structures. Grayscale, Saturation, and RGB effectively recognize all sand pixels above the effluent line. This outcome was expected for the Saturation input channel, as it has been previously demonstrated how effectively this input channel segregates wet from dry regions. For the Grayscale and the RGB input channels, this is more surprising. The dE/dt input channel introduces a lot of noise in the predictions. This was expected as the dE/dt image only delineates the water line but does not give any information about water and sand elsewhere over the timestack image.

Based on Table 4.2 the best single-input channel would be dI/dt as it has the highest accuracy. Additionally, when comparing the confusion matrices in Figure 4.2 this is also the input channel that balances best the false positives and the false negatives. Meaning that there is no overprediction of water or sand pixels. While the various channels individually exhibit distinct strengths in feature
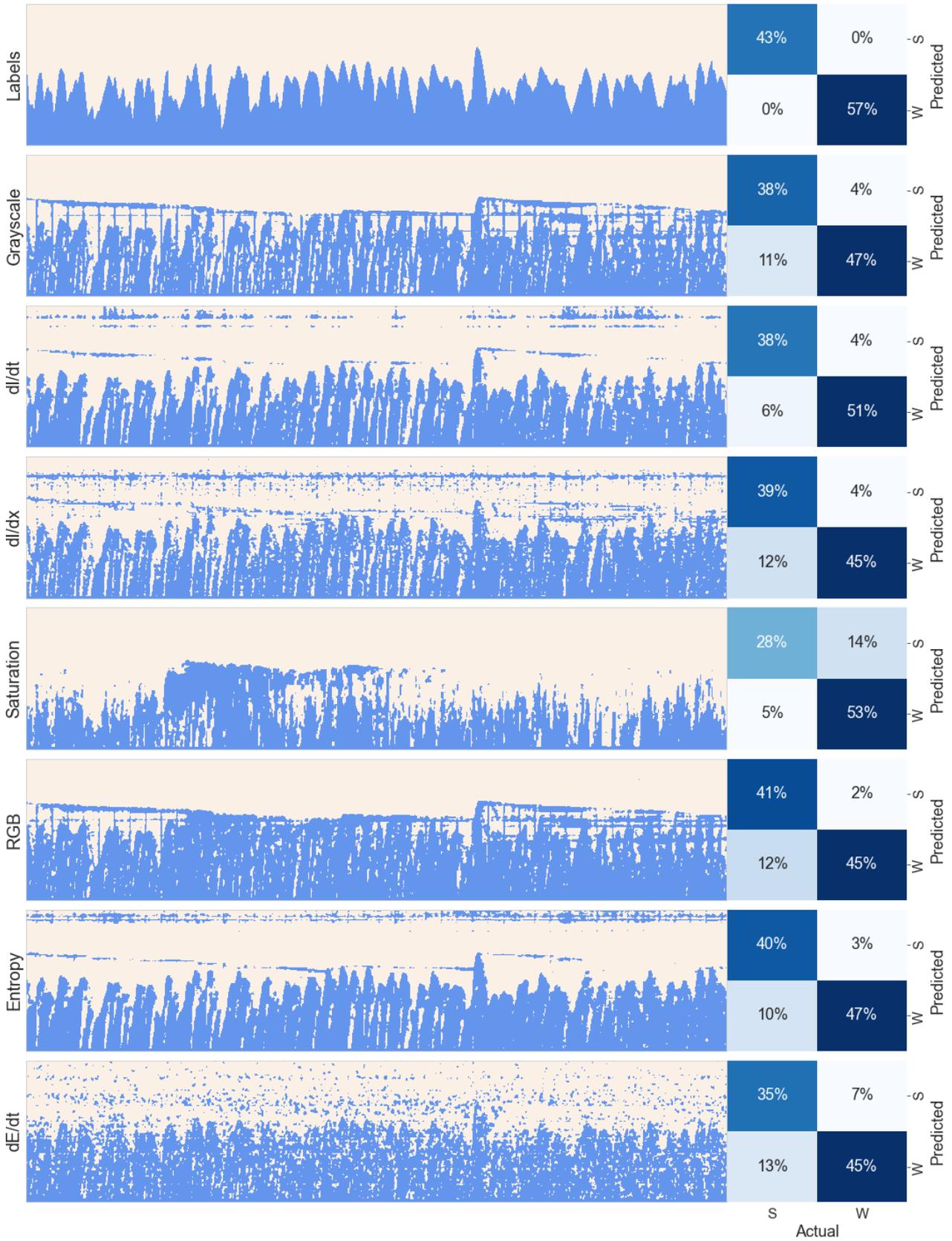
Figure 4.2: Labels and predictions for all single inputs, and corresponding confusion matrices with predicted and actual sand (S) and water (W) pixels.

representation within the timestacks, their unique attributes may complement each other. Integrating multiple input channels could enhance the model's precision in classifying water and sand pixels.

4.1.3. Analysis of multi-channel inputs

A total of 127 channel combinations were evaluated through k-fold cross-validation. Figure 4.3a presents the resulting mean accuracy and standard deviation. Subsequently, the 24 best combinations were selected based on arbitrary thresholds of $\mu_{acc} > 95.5\%$ for mean accuracy and $\sigma_{acc} < 0.5\%$ for standard deviation. These correspond to the blue dots in Figure 4.3a. The frequency of occurrence of every input channel in these 24 best combinations was then computed, and displayed in Figure 4.3b.
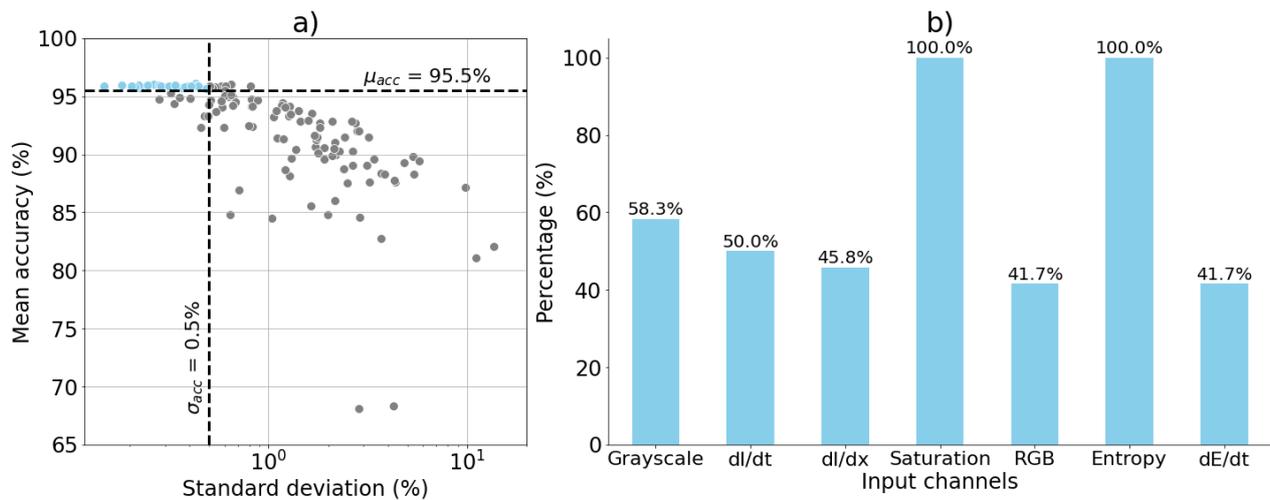

Figure 4.3: In (a), the results of a k-fold cross-validation analysis across all possible combinations of input channels, illustrate the standard deviation of the results as a function of the mean accuracy. Arbitrary thresholds were set at a $\mu_{acc} = 95.5\%$ and $\sigma_{acc} = 0.5\%$ to select the best combinations of input channels. In (b), the frequency of occurrence of input channels in these best combinations.

In Figure 4.3a, the results of the k-fold cross-validations show large disparities between the different combinations. The accuracy varies from 68% up to 96% and the standard deviation from 0.2% up to 10%. A cluster of channel combinations that showed an accuracy higher than 95.5% and a standard deviation smaller than 0.5% are depicted in blue in Figure 4.3a. Inspection of the frequency of occurrence of the different input channels over these best combinations showed that both the saturation and the entropy input channels were present in every one of the best combinations. This is surprising for the Saturation channel, which was performing poorly on its own, as demonstrated in Figure 4.2. These channels seem to enhance nearshore features and give complementary information to the model.

Also, among the 24 best combinations, 17% included six inputs, 29% five inputs, 33% four inputs, 17% three inputs, and only 4% (one combination) utilized just two channels: Saturation and Entropy. Which, once more advocates the efficacy of those channels to enhance nearshore features advantageous for runup extraction under the presented conditions. From the 24 best combinations the one with the highest accuracy was retained for comparison with manual selection and the CC-based Saturation-Entropy method. This best model combined the grayscale image with dI/dx, saturation, entropy, and dE/dt as input channels.

4.2. Optimization of ML and CC model

4.2.1. Model performance and correlation to input statistics

Until now the ML results were assessed only through accuracy: the percentage of correctly predicted pixels. To further refine the model, the local performance in time was evaluated by comparing the ML-predicted water line with the hand-picked results and CC estimates. The time correlations of the two methods with the hand-picked water line are displayed in Figure 4.4.
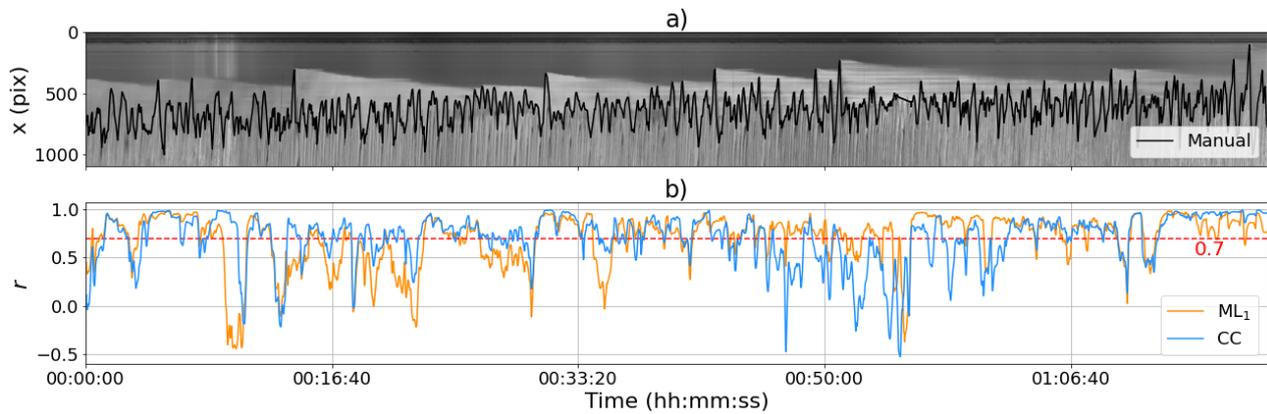

Figure 4.4: In (a), the timestack image along with the hand-picked water line, and in (b), the corresponding Pierson correlation coefficient of the ML and CC methods in a moving window of 100 grid cells.

Over the entire timestack, the ML method achieves a correlation of 0.7 with the handpicked data, whereas the CC method achieves a correlation of 0.68. The correlation coefficient in time seems negatively skewed, with many values close to one, and outliers that reach negative correlations. The percentage of correlation coefficients above 0.7 is 67% for the ML method and 62% for the CC method. Hence, most data points align strongly with the manual results for both methods. However, large sections remain where they fail to capture the water line accurately. The results provide a baseline for further comparison, encouraging further analysis of why certain segments show discrepancies and which features may be responsible.

To assess how the input channels affected the correlation between the models and the hand-picked results (Figure 4.4b), the input channels were reduced to the time dimension by taking the mean, standard deviation, maxima, and minima over each column. Next, the correlation coefficient of each statistical measure was calculated against the correlation coefficient between the model and the hand-picked data.

Performance ML method and correlation to input statistics
The correlation coefficients between the input channel statistics and the machine-learning model performance are summarized in Table 4.3. The column minima of the saturation channel, $S_{\min,j}$, with a correlation of -0.255 suggests that the minimum value in the saturation channel might negatively impact the match between machine-learning and hand-picked results, possibly indicating that lower saturation levels might be problematic for the model. The column average of the saturation channel, $S_{\mean,j}$, with a negative correlation of -0.213 shows that higher average saturation over space is associated with lower model performance. Also, $dE/dt_{\min,j}$ with a positive correlation of 0.156 shows that a higher minimal value might enhance the predictive performance of the machine-learning model. On the other hand, $\sigma_{dE/dt,j}$ with a correlation of -0.191 indicates that the minimal variations

in entropy over time have a moderately positive relationship with better-matching results, which might suggest that stable regions of low variability in texture are easier for the model to predict accurately. The extremes of the grayscale channel, $I_{\min,j}$ and $I_{\max,j}$, also show positive correlations, 0.142 and 0.138 respectively, suggesting that extreme grayscale values, whether high or low, correlate positively with better matches between the two methods. However, $\sigma_{I,j}$ shows a negative correlation, implying that high variability in pixel intensity could coincide with poor model results. Furthermore, dI/dt_{mean}, $\sigma_{E,j}$, $E_{\min,j}$, and $dE/dt_{mean,j}$ show very weak correlations close to zero. This indicates that these metrics might not strongly or directly influence the correlation between machine learning and hand-picked results in your current setup.

Table 4.3: Correlation coefficients organized by input feature types and statistical measures of the input channels.

Statistic	Correlation Coefficient				
	Grayscale (I)	dI/dt	Saturation (S)	Entropy (E)	dE/dt
Mean	0.082	0.008	-0.213	0.110	-0.008
Std	-0.176	-0.054	-0.027	-0.009	-0.191
Min	0.142	0.129	-0.255	-0.007	0.156
Max	0.138	-0.022	-0.110	-0.026	-0.108

The negative correlations with saturation-related metrics (S_{mean}, S_{min}) could imply that the machine learning model may not handle well the variations in saturation. Reviewing how the model processes saturation might uncover opportunities for model improvement. A positive correlation with $dE/dt_{min,j}$ but negative with $\sigma_{dE/dt,j}$ suggests that while the model handles stable textures well, it struggles with high variability in texture over time. This might point to incorporating more robust features in the model that can handle textural changes better.

Performance CC method and correlation to input statistics

The correlation coefficients between the input channel statistics and the CC method performance are summarized in Table 4.4. The column-averaged entropy, $E_{\mathrm{mean},j}$, with a low negative correlation coefficient of -0.183 suggests that high randomness and complexity tend to decrease the effectiveness of the CC-based method. On the other hand, $\sigma_{E,j}$ with a moderate positive correlation indicates that high variability of entropy in space im-

Table 4.4: Correlation coefficients for different statistics and input features.

Statistic	Correlation Coefficient	
	Saturation (S)	Entropy (E)
Mean	0.130	-0.183
Std	-0.090	0.300
Min	0.098	-0.191
Max	-0.138	0.014

proves contrast detection, most likely because of bimodality in the entropy spectra. The column-minima of the entropy channel, $E_{\min,j}$ has a coefficient of -0.191. This moderate negative correlation suggests less complexity in the least complex regions could enhance the method's effectiveness. Regarding the correlation with the saturation statistics, it might be that high column-average $S_{\mathrm{mean},j}$

positively influence the performance and that extreme values $S_{\mathrm{max},j}$ negatively impact the model's result.

4.2.2. Towards a more robust ML model

Column-wise normalization

From the correlation of model performance with input channel statistics, it was concluded that the most potential lies in preprocessing the entropy or the saturation channel. Several processing techniques were mentioned that could eventually enhance model performance. Table 4.5 resumes the results when applying column normalization to the entropy channel, the saturation channel, and all channels simultaneously.

Column-wise normalization of the entropy channel and all channels simultaneously decreases the model performance without any gain along the timestack. When only the saturation channel is normalized, overall performance remains constant. Although the model appears to handle areas of high luminescence more effectively, this adjustment concurrently leads to decreased performance across other sections of the timestack, notably in regions exhibiting a noisy

Table 4.5: Correlation coefficients of column-wise normalized input channels.

Normalized columns	r $(-)$
None	0.70
Entropy	0.66
Saturation	0.70
All channels	0.64

seepage face. The example of column-wise normalization shows how complex optimization through channel preprocessing can be. A measure can solve one issue while creating new ones. As the column-normalization of input channels did not seem to have any conclusive positive impact on the model's performance, the measures were not retained.
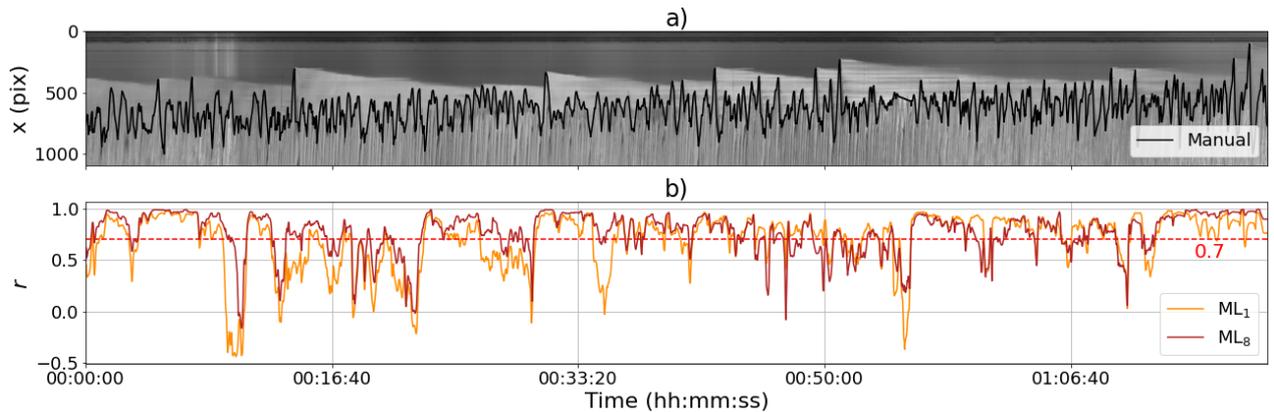
Figure 4.5: In (a), the timestack image along with the hand-picked water line, and in (b), the corresponding Pierson correlation coefficient of the ML method when trained on either one (ML_1) and eight (ML_8) timestacks of 1 hour 33 minutes over a moving window of 100 grid cells.

Training set expansion

Earlier, the water line was extracted by training the ML algorithm with small data sections, namely, a timestack of 1 hour and 33 minutes. The results were promising, but expanding the dataset seems unavoidable to improve the model's robustness and accuracy, ensuring it generalizes well to diverse conditions. The model was trained on all timestack images from Table 3.1 except one left for

validation (video ID: GX050084).

Overall, the correlation curve shifts upwards when increasing the training set to nine timestacks (Figure 4.5). The correlation coefficient goes from 0.7 when training the machine-learning model on a single timestack image to 0.77 when training on eight timestack images, meaning that training set expansion improves the model's performance.

Comparison to single-channel input

Until now, efforts have been focused on optimizing a multi-channel ML model. The training set was expanded, resulting in substantial model improvement. However, to evaluate the added value of a multi-channel informed model, it is compared against a baseline single-channel informed model. Therefore, an additional model was trained on the expanded set, using only the grayscale image as input. The results are summarized in Table 4.6.

Table 4.6: Comparison of multi- and single-channel input for the ML algorithm with corresponding correlation coefficients with the manual results and the mean accuracy (μ_{acc}) and standard deviation of the accuracy (σ_{acc}).

ML model	# Training	r $(-)$	μ_{acc} (%)	σ_{acc} (%)
Multi-channel	1	0.70	96.09	0.42
Multi-channel	8	0.77	93.03	1.40
Single-channel	8	0.66	90.92	2.23

When comparing the machine-learning results using channels I, dI/dx, S, E, and dE/dt with those using only I, it becomes evident that the multi-channel approach significantly enhances the performance. The correlation coefficient over the entire timestack with the hand-picked water line is 0.77 for the multi-channel model and 0.66 for the single-channel. Overall, only 42% of the single-channel model results have a correlation coefficient above 0.7 with the hand-picked results. For the multi-channel informed model, this was 72%.

When considering the results of the multi-channel models, it appears that the accuracy decreases when the training set is expanded to all labeled data. Initially, when k-fold cross-validation was performed within a single timestack, the model achieved an accuracy of 96.09% with a standard deviation of 0.42%. However, with the expanded training set, the mean accuracy dropped to 93.03%, and the standard deviation increased significantly, indicating greater variability and less consistent performance. However, when zooming in on the individual folds of the multi-channel model trained on the extended set, it appears that the accuracy is lower when the validation timestacks are from November 13th, 2023, which had fewer timestacks. Specifically, the accuracies for this day were 91.97%, 91.52%, and 90.13%, with an average of 91.21% and a standard deviation of 0.78%. In contrast, the accuracies for November 14th, 2023, which had more timestacks, ranged from 93.33% to 94.56%, with an average of 93.94% and a standard deviation of 0.41%. This suggests that the model performs better when validated on a larger and more diverse set of timestacks, capturing a wider range of conditions and improving generalization. Possible causes for this discrepancy might be that the validation timestacks from, November 13th, 2023, have higher data variability, or that the model does not have enough similar samples from that day compared to November 14th, 2023.

4.2.3. Towards a more robust entropy-saturation model

Window size reduction

The entropy-saturation model shows the highest correlation with the manual water line for a window size of 500 ($r = 0.72$), corresponding to approximately 4 minutes of measurements. The results of the CC method exhibit rapid degradation for window sizes smaller than 100, whereas the model performance gradually declines for windows larger than 500. The window size of 500 was retained in later computations.

Column-wise normalization

Normalization of the entropy and saturation columns slightly enhances the method. However, the increase in correlation with the hand-picked results remains modest. While beneficial, normalization alone may not sufficiently elevate model performance. Consequently, a more thorough investigation into additional or alternative enhancements is warranted to optimize the correlation outcomes.
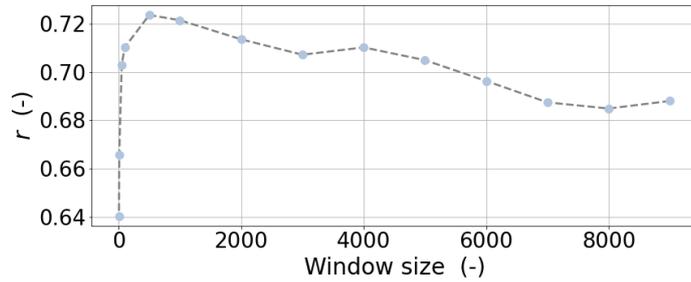

Figure 4.6: Correlation coefficient of CC method with hand-picked water line, shown as a function of temporal window size.

Figure 4.7: Correlation of CC methods with different normalized columns.

Normalized columns	r $(-)$
None	0.72
Entropy	0.72
Saturation	0.74
Entropy & Saturation	0.74

4.3. Comparison with manual and empirical estimates

4.3.1. Instantaneous runup

Figure 4.8 shows a time series plot comparing runup excursions measured by manual digitization and computed by ML and CC methods. The runup estimates of the ML and CC methods appear to closely follow the manually digitized water line, suggesting good agreement in detecting the position of the runup.
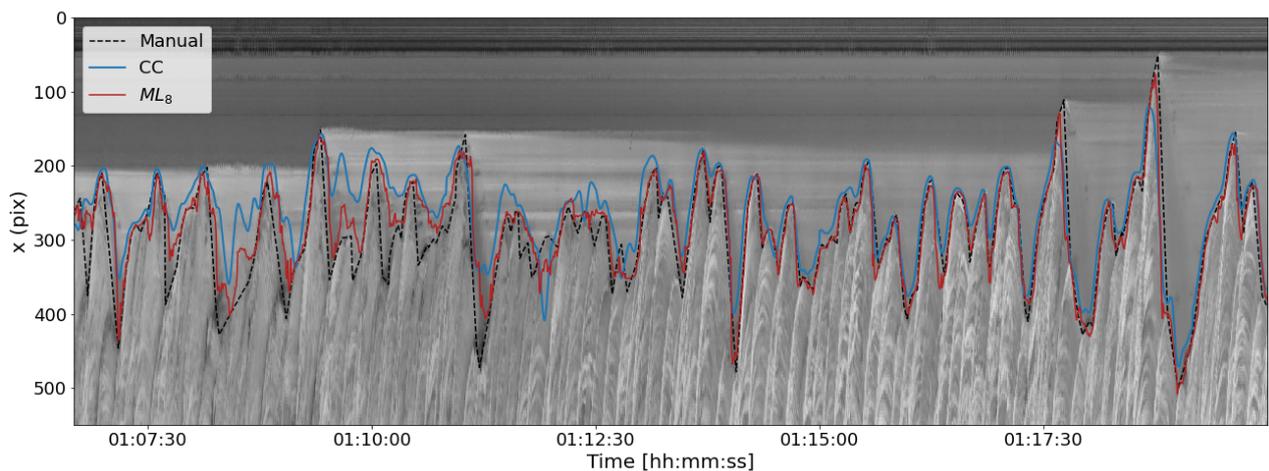

Figure 4.8: Section from the timestack from November 14th (video ID: GX050084), 2023, with the manually digitized water line, and the results of the CC and the ML methods.

There are minor discrepancies between the methods at certain points. In the left section of the image, where the swash appears less distinctly white and horizontal seepage lines are visible, both models demonstrate challenges in accurately estimating the runup. They tend to overestimate the shoreline position compared to the manually digitized results. This discrepancy suggests that the models may struggle with areas where the swash zone has lower contrast and with non-wave features in the seepage face such as horizontal lines.

Table 4.7: Swash characteristics from Manual, ML and CC: runup exceeded by 2% of the waves ($R_{2\%}$) and corresponding RMSE, and correlation coefficient (r) of ML and CC Methods with the manually digitized water line from '2023-11-14 14:10:00' to '2023-11-14 15:30:00'.

	$R_{2\%}$ (m)	$RMSE$ (m)	r $(-)$
Manual	0.84	–	–
ML	1.01	0.17	0.75
CC	0.95	0.17	0.72

From Table 4.7 it seems that the $R_{2\%}$ are consistent across methods with correlation coefficient of 0.75 and 0.70 for the ML and the CC methods respectively. However, the RMSE value stands at 0.17 meters for both methods throughout this period, which is considerable relative to the amplitude of the wave runup ($\sim 1m$). This level of error may still impact the reliability of runup estimations in practical applications. Table 4.8, summarizes the comparison between the ML and the CC methods over the full runup time series on November 13th and November 14th.

Table 4.8: RMSE, and correlation coefficient (r) of ML and CC Methods with the manually digitized runup time series. From '2023-11-13 13:46:00' to '2023-11-13 16:54:24.5' and from '2023-11-13 13:46:00' to '2023-11-14 10:10:00'.

	r $(-)$	$RMSE$ (m)
ML	0.96	0.10
CC	0.94	0.12

Both methods demonstrate very strong correlations with the actual data. With correlation coefficients of 0.96 and 0.94 for the ML and CC method respectively, which indicates an excellent linear relationship between the predicted and actual values of the runup time series. The RMSE values provide a measure of the average magnitude of the errors between the predicted values by each method and the actual values of the runup time series. The ML and the CC method show predictions deviating from the actual values by about 10 and 12 cm on average, which corresponds to an error of approximately 16% with respect to the maximal runup amplitudes.

4.3.2. Runup statistics

In general, the estimated runup quantiles show good agreement with the manual method. In Figure 4.9, for the two methods, runup exceeded by 2% of the waves ($R_{2\%}$), and runup exceeded by 50% of the waves ($R_{50\%}$) show generally good agreement with the manually digitized water line.

This is confirmed by high correlation coefficients of 0.97 and 0.99, respectively, which indicates a solid linear relationship between the ML and CC runup statistics and manual values, suggesting that both models are highly effective in predicting the spatial extent of wave uprush on the beach. Additionally, the RMSE values for runup heights are relatively low (0.05-0.09 m), indicating minor

deviations from the manually digitized water line (Table 4.9). This low error metric further validates the accuracy of the ML and CC models in estimating runup heights, enhancing their reliability for practical applications in coastal management. Especially in extreme-value applications.
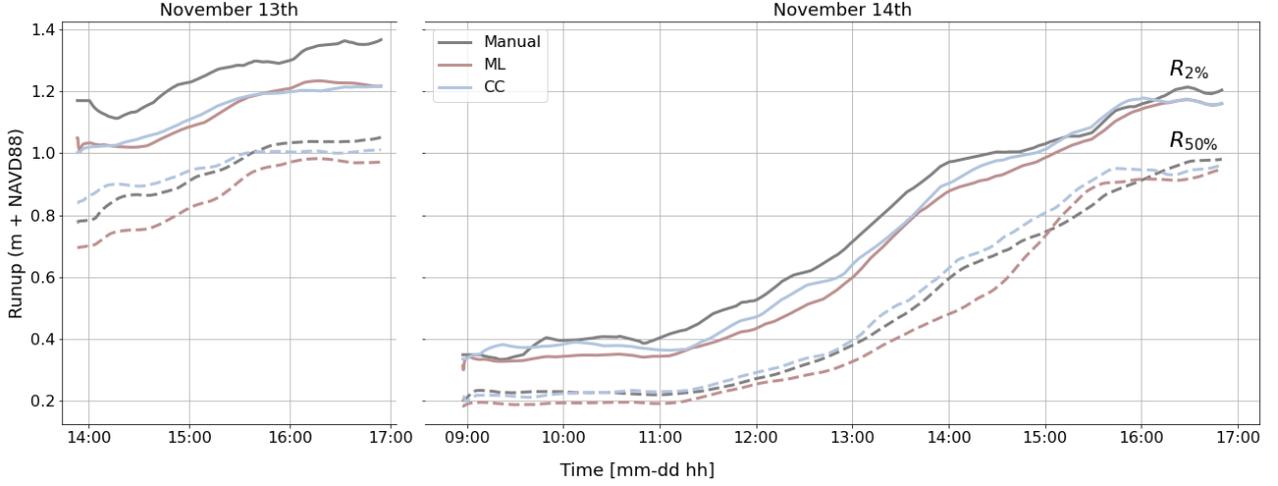

Figure 4.9: Runup statistics ($R_{2\%}$ and $R_{50\%}$) on November 13th and November 14th, 2023, including Manual, ML, and CC methods. The time ranges are from '2023-11-13 13:46:00' to '2023-11-13 16:54:24.5' and from '2023-11-13 13:46:00' to '2023-11-14 10:10:00'.

On November 13th, the manual method yielded a higher $R_{2\%}$ compared to the ML and CC methods. Overall, the difference was reasonable, approximately 0.15 m, corresponding to a horizontal excursion of 2 m. On this day, the uprush was distinguishable on the timestack images. However, the backwash was not, and it appears that both the ML and CC methods struggled with larger runup events as the effect is especially pronounced in the $R_{2\%}$ metric. Additionally, it raises inquiries regarding the reliability of the manual findings.

On November 14th, the $R_{2\%}$ values are slightly overestimated by the ML and CC methods. This overestimation is likely due to the method by which $R_{2\%}$ is extracted from the runup time series. The ML and CC methods introduce higher frequency components into the runup time series, causing single large runup events to exhibit multiple local maxima. Consequently, this biases the $R_{2\%}$ estimates to-

Table 4.9: Correlation and RMSE of ($R_{2\%}$) and ($R_{50\%}$) from ML and CC methods with manually digitized swash characteristics from November 13th and November 14th.

	$R_{2\%}$		$R_{50\%}$	
	r (-)	rmse (m)	r (-)	rmse (m)
ML	0.97	0.09	0.99	0.05
CC	0.97	0.08	0.99	0.05

wards higher values. On November 14th from '13:00' to '15:00' the ML method slightly underestimates the $R_{50\%}$. During this period, the timestack images exhibit extremely low contrast between the water and the sand. The absence of white foam in the swash zone likely leads to an increased number of incorrectly classified sand pixels. This effect is most pronounced during smaller runup events, which are characterized by less energy and thus less turbulent motion, thereby influencing the $R_{50\%}$. On November 14th around '15:00' both methods overestimate $R_{50\%}$ slightly. This corresponds to the timestack described in Section 4.3.1, where disturbances in the seepage face result in high entropy values, leading to more pixels being classified as water.

4.3.3. Spectral decomposition of runup

The runup spectra were computed on November 14th from '15:00' to '15:30' for all three methods, ass well as the mean long-wave oscillations, S_{ig}, and the mean short-wave oscillations, S_{inc}. The results are displayed in figure 4.10.

The spectra of the Manual, ML, and CC methods closely follow each other across the frequencies, indicating that each method captures similar dynamics, and although some deviations are observed, they are negligible when considered within the 90%-confidence interval (Figure 4.10a).
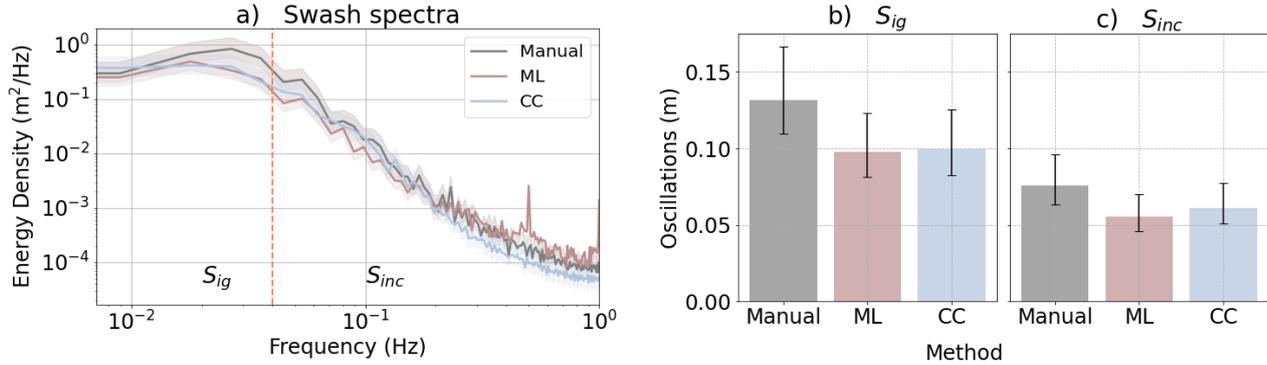

Figure 4.10: In (a), the wash energy density spectra for manual, ML, and CC methods with associated 90%-confidence interval. In b) the mean long-wave oscillations, S_{ig}, and in c) the mean short-wave oscillations, S_{inc}. All from '2023-11-14 15:00:00' to '2023-11-14 15:30:00'.

The infragravity part of the spectra (left of the dashed line at 0.05 Hz) shows less variability between the methods, suggesting that all approaches consistently capture the lower frequency motions. For all three methods, the spectral decay is approximately -2. It corresponds to nearshore characteristics, usually less steep than a full-grown offshore spectrum due to the energy losses associated with wave breaking and wave-wave interaction, transporting energy away from the peak.

In terms of energy contained in the spectra, the manually digitized water line shows the highest oscillation values among the three methods, suggesting it might be capturing a larger amplitude wave runup. This may be due to the tendency to manually digitize the waves retreating further after each backrush, and because manual digitization follows the peaks more closely than the ML and CC methods, which tend to round them off. However, the difference across methods remains within the 90%-confidence interval of the energy density spectra, meaning that the deviations aren't significant. The methods do not present a significant difference along the short-wave components. They all show reduced oscillation values in the S_{inc} category, which is typical as this component represents higher frequency waves that generally contain less energy in the nearshore due to breaking and dissipation.

From Figure 4.11a, it is clear that most energy is contained in the infragravity band of the spectra. On November 14th from 12:30, some more energy is found in the incoming wave part of the spectra. In Figure 4.11b and Figure 4.11c, the manual spectra and the differences between the ML and CC spectra compared to the manual spectra are shown (ML - manual and CC - manual). Positive values (red) correspond to overestimation by the method, while negative values (blue) correspond to underestimation by the methods. For both methods, the difference in energy density is largest around the peak frequency, located in the infragravity band (0.004 Hz - 0.04 Hz). This is confirmed by the RMSE of the S_{IG} which is higher than the RMSE of S_{inc} (0.02 m - 0.01 m for both methods). Overall, the methods tend to slightly underestimate the energy contained around the peak, except
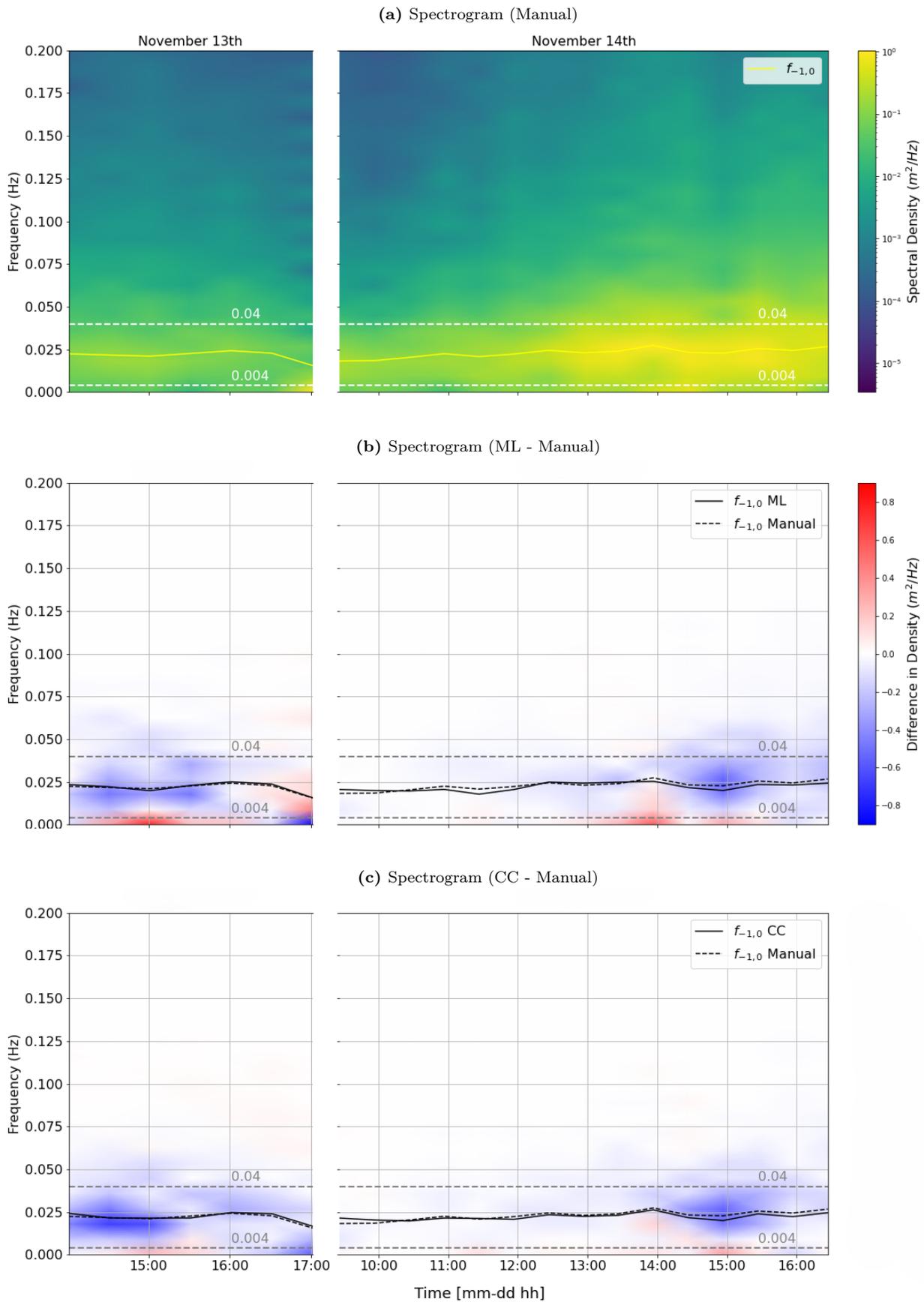
Figure 4.11: Spectrogram of the extracted runup: manually digitized results (a), and the spectral difference between ML and Manual (b), and CC and Manual (c). From 2023-11-13 13:46:00 to 2023-11-13 16:54:24, and from 2023-11-13 13:46:00 to 2023-11-14 10:10:00.

for the ML method on November 13th around 17:00.

From the plots, it appears that the sub-infragravity band is slightly underestimated by both methods. However, this is an artifact of the interpolation method of the plotting options, as the frequency resolution (δf) is 0.009 Hz, meaning that sub-infragravity components fall out of the resolution range.

The models' performance is satisfactory when considering the mean wave period ($T_{m-1,0}$). The ML model shows a correlation coefficient of 0.79, while the CC method has a slightly lower correlation coefficient of 0.70. Although these values are slightly lower than those for the runup extremes ($R_{2\%}$ and $R_{50\%}$), they are still very high. Hence, despite the models tending to underestimate the infragravity part of the spectrum, the differences remain tolerable as the mean frequency ($f_{m-1,0}$) is not too far off. The RMSE for $T_{m-1,0}$ remains within 10% of the observed values ($T_{m-1,0} \sim 40$ seconds), at 3.51 seconds for the ML model and 3.88 seconds for the CC method. This indicates that the models accurately predict the timing aspects of the runup, which is crucial for dynamic assessments, such as predicting the frequency of inundation of sea turtle nests.

Table 4.10: Comparison of S_{ig}, S_{inc} and $T_{-1,0}$ for ML and CC methods.

	$\boldsymbol{S_{ig}}$		$\boldsymbol{S_{inc}}$		$\boldsymbol{T_{m-1,0}}$	
	r (-)	rmse (m)	r (-)	rmse (m)	r (-)	rmse (s)
ML	0.90	0.02	0.88	0.01	0.79	3.51
CC	0.90	0.02	0.92	0.01	0.70	3.88

4.3.4. Empirical estimates for runup

The runup observations from the Manual, the ML, and the CC methods were compared to empirical estimates according to Stockdon et al. (2006). The runup predictions were made according to the general formula with two different beach slopes ($\beta = 0.1$, and $\beta = 0.016$), and the formula for extremely dissipative conditions (Figure 4.12).

All methods and predictions exhibit an increasing trend in $R_{2\%}$ throughout the day on November 13th and November 14th. Each observation method demonstrates a high correlation (greater than 0.7) with the empirical estimates, which is expected given that a significant portion of the estimates is derived from the observed mean water levels from the NOAA station. However, there are notable differences in the $R_{2\%}$ values predicted by the various methods and Stockdon implementations.

The Stockdon formula with $\beta = 0.1$ consistently predicts the highest $R_{2\%}$ values, which is confirmed by the higher RMSE values for all three methods: 0.38 for Manual, 0.51 for ML, and 0.50 for CC. For a beach slope of $\beta = 0.016$, the predicted $R_{2\%}$ values are lower than those for $\beta = 0.1$. Correspondingly, the RMSE values are smaller, at 0.16, 0.28, and 0.27 for Manual, ML, and CC, respectively.

The Dissipative Stockdon method appears to align best with the observations. It shows the lowest predictions and maintains a high correlation across all observation methods, with r values of 0.71 for Manual, 0.83 for ML, and 0.83 for CC. This runup formula also minimizes the RMSE, achieving values of 0.08 for Manual, 0.13 for ML, and 0.13 for CC.

Notably, all runup estimates show the least alignment with runup observations during low tide, while they follow more closely the observations during mid- and high tide. This trend is also evident
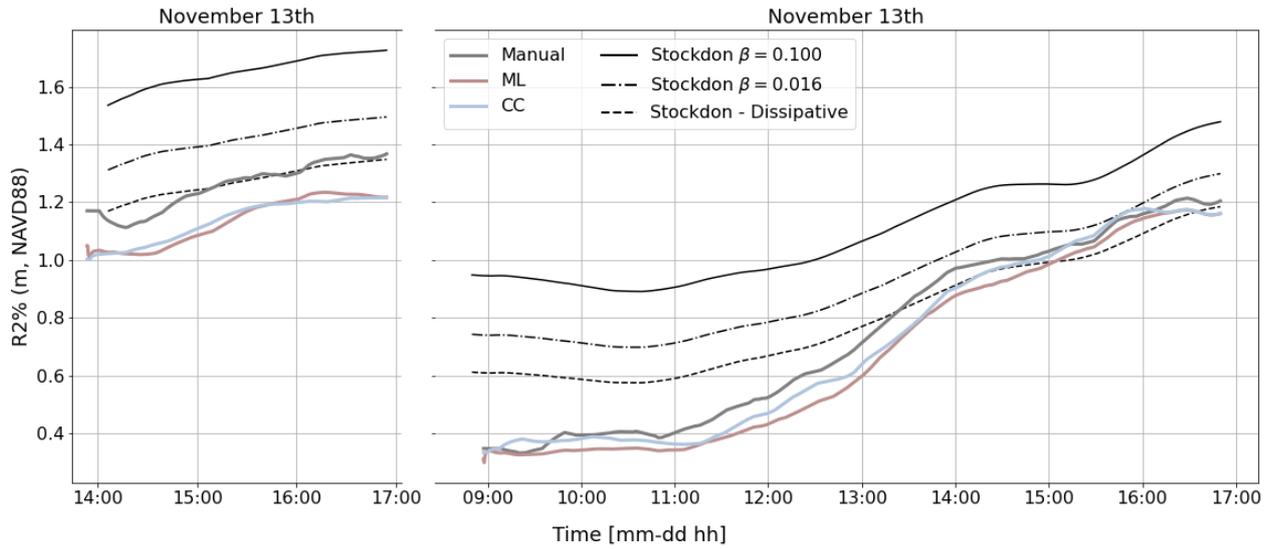
Figure 4.12: Runup observations and estimates of $R_{2\%}$ on November 13th and November 14th, 2023. Observations include Manual, ML, and CC methods. Estimates are based on the general Stockdon formula (with $\beta = 0.1$ and $\beta = 0.016$), and the Stockdon formula for extremely dissipative conditions. The time ranges are from '2023-11-13 13:46:00' to '2023-11-13 16:54:24.5' and from '2023-11-13 13:46:00' to '2023-11-14 10:10:00'.

in the runup estimates derived from the Stockdon formula for extremely dissipative conditions. During high tide, the estimates align almost perfectly with the observations, whereas during low tide, discrepancies can reach up to 20 cm.

Table 4.11: Comparison of Manual, ML, and CC methods with different Stockdon implementations.

	Stockdon					
	Dissipative		$\beta = 0.016$		$\beta = 0.1$	
	r (-)	rmse (m)	r (-)	rmse (m)	r (-)	rmse (m)
Manual	0.71	0.08	0.72	0.16	0.74	0.38
ML	0.83	0.13	0.83	0.28	0.83	0.51
CC	0.83	0.13	0.82	0.27	0.82	0.50
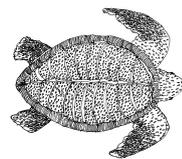
5

Discussion

A CC and ML method has been successfully implemented, demonstrating generally good agreement with manually digitized water lines. However, this discussion aims to address lingering questions. It will explore the relevance and applicability of these ML and CC methods (Section 5.1), along with their potential real-world applications and future research directions, including empirical estimates for sea turtle nest flooding. Key considerations will encompass the impact of training set size and composition, challenges and opportunities in preprocessing techniques, refining ground truth for model assessment, and investigating additional input channels and their combinations to further enhance the ML algorithm (Section 5.2).

5.1. Relevance of multi-channel input for water line detection

5.1.1. Insights from traditional runup extraction methods

Traditional CC methods for runup extraction showed limited performance under dissipative conditions due to the presence of a seepage face and low energy waves, which show little contrast with the sand, especially along the backwash of waves (Huisman et al., 2011). The Radon Transform (RT) method developed by Almar et al. (2017), showed good agreement with LiDAR data on an energetic dissipative beach, and was able to delineate the water line where the CC methods failed to. However, in most timestacks on this particular beach, the Radon method was not able to capture runup effectively and demonstrated high sensitivity to the input minimum and maximum runup values.

Earlier work has shown that spatial pattern analysis, can effectively be implemented for shoreline detection in satellite images (Fuse and Ohkura, 2018) and that on topographical LiDAR data, the swash zone can be identified thanks to variations in surface roughness (Wang et al., 2023). The local entropy, as a measure of roughness, has demonstrated significant potential in isolating the swash zone due to its turbulent nature compared to the sand above it. This underscores the efficacy of roughness measures in swash detection on dissipative beaches, as observed by Zhang and Zhang (2009). Combined with saturation, to reduce noise above the effluent line, entropy can be an efficient method for runup extraction.

The estimates on the full runup time series from November 13th and 14th aligned well with the hand-picked water line, with a correlation coefficient (r) of 0.94 and root-mean-square-error (RMSE) of 0.12m. The runup extreme values showed even better agreement with a r-value of 0.97 and an RMSE of 0.08m for the $R_{2\%}$, and a r-value of 0.99 and RMSE of 0.05 m for the $R_{50\%}$. Regarding temporal accuracy, the CC predictions were satisfactory, with a r value of 0.70 and an RMSE of 3.88

seconds.

However, the model still has limitations. It requires a smooth seepage face to perform well; otherwise, high entropy values in the seepage face lead to an overestimated water line and the introduction of high frequencies. Additionally, the model needs a certain level of turbulence in the swash zone. Without it, the swash is characterized by patches of low entropy, leading to an underestimated water line. This issue is particularly pronounced in the backwash of singular runup events.

Window size reduction and column normalization of the saturation channel enhanced the CC method's performance over the considered data. However, the correlation increment with the manually digitized water line remained small, and further research is needed to prove the robustness of these measures, and whether they generalize well on other dissipative beaches.

5.1.2. Enhanced runup detection with multi-channel approaches

The multi-channel informed ML model showed superior results to the single-channel model in terms of accuracy and standard deviation. The idea, derived from Collins et al. (2023) with their multi-informed CNN algorithm with LiDAR data showed promising results, and aligned well with the hand-picked data ($r = 0.77$), while the single-channel informed model returned poorer performance ($r = 0.66$). This improvement could be attributed to the increased informational content captured by these channels, which likely enhances the model's ability to distinguish between sand and water in various environmental conditions.

In total 7 input channels and 127 different combinations were assessed. Out of these, the 24 best combinations all included the entropy and saturation channels, and one combination consisted solely of these two channels. This highlights the critical importance of these channels for a multi-channel informed ML model for runup extraction and suggests they warrant further investigation. Our findings corroborate the work of Zhang and Zhang (2009) who highlighted the the role of texture and color for runup extraction.

The model with the highest accuracy out of the 127 combinations was retained for further optimization, while other models had similar accuracies and lower standard deviations. Furthermore, the best choice was based on k-fold cross-validation performed on one timestack of 1 hour 30 minutes. It might as well be that other models would have performed better on different training sets. Therefore, the training set should be extended during this phase for a more generalized approach, and ensure a more robust selection of combinations. Unfortunately, the limited amount of labeled data and the computational tools at hand did not allow for this during this research, as evaluating the 127 combinations took 10 hours for a CPU implemented CNN with *tensorflow.keras.model.Sequential* on an Intel Core i7 processor with 16 GB of RAM when training on a single timestack image.

The ML learning model was cross-validated on nine timestacks and the estimates of instantaneous runup were concatenated. The resulting runup time series from November 13th and 14th aligned well with the hand-picked water line, with a correlation coefficient of 0.96 and RMSE of 0.10m. The runup extreme values showed even better agreement with a correlation of 0.97 and a RMSE of 0.09 m for the $R_{2\%}$, and a r-value of 0.99 and RMSE of 0.05 for the $R_{50\%}$. Similarly to the CC method, when it comes to the temporality of the ML prediction, the model performed quite well concerning the manually observed water line, with a correlation of 0.79 and a RMSE of 3.51 seconds for the significant wave period ($T_{m-1,0}$).

One limitation of this study is the relatively small dataset. While it is sufficient to demonstrate proof-of-concept, it may not capture the full variability of coastal environments needed for a robust model for wave-by-wave analysis. When the estimations were extrapolated to new time series, the ML model occasionally failed to accurately capture the water line due to specific characteristics of the new timestack images. However, these errors were infrequent and should not pose a problem for long-term runup analysis. For wave-by-wave analysis, future research should focus on extending the training to a more extensive and diverse set of coastal regions and environmental conditions to improve the model's robustness. Alternatively, to overcome the incorrect runup extraction due to new unknown artifacts, it might be beneficial to add synthetic artifacts during training on the timestack images. This process is called data augmentation and could also help in raising a flag when strange patterns are encountered.

5.1.3. Real world applications and future research direction

Both the ML and the CC method demonstrated good agreement with the manually digitized runup time series, as well as with the extreme runup heights $R_{2\%}$, and $R_{50\%}$. The CC method is anticipated to be applicable across similar beaches along the northern Gulf of Mexico for extreme value analysis and wave-by-wave analysis; however, it is recommended to visually verify the runup estimates from the CC method to ensure accuracy. Despite that the ML model sometimes failed to capture the water line when extrapolated outside the k-fold training sets, due to artifacts never encountered before, it might be used for extreme value analysis under similar environmental conditions. However, visual data control is a prerequisite, and there is a risk of losing large portions of the time series that are not accurately estimated by the model, and filling these gaps manually will lead to redundant and time-consuming work. Alternatively, one should accept that a portion of the extracted water line may be somewhat inaccurate.

These models contribute to the optimization of runup data acquisition and build forth on a trend directing towards a more stochastic approach for species management, whether probabilistic (Ware and Fuentes, 2018; Ware et al., 2019) or numerical (Dédina, 2023). Reliable data is key in validating these predictions. The study by Dédina (2023) focused on predicting the likelihood of nest inundation but was limited by data scarcity, which impeded the direct validation of the runup estimates produced.

However useful extreme value analysis might be in coastal species management, the timing aspect of the estimates is also of importance, especially in groundwater processes. Unfortunately, this study did not encompass the analysis of groundwater table fluctuations. However, the literature has shown how intricate these processes can be. At first, the beach matrix was considered as a low pass filter (Raubenheimer et al., 1999). This means that in general low-frequency fluctuations such as the tide damp out slower in the onshore direction, and that the corresponding time scales are most important in the determination of the groundwater table. However, a myriad of other processes alter this simple statement. For instance, Nielsen and Hanslow (1991); Cartwright et al. (2004) showed that the presence of a capillary fringe allows groundwater waves to propagate faster and decay slower and found that this phenomenon is even more significant for high-frequency waves. Hence, it is crucial to accurately represent the periodicity of runup events, to ensure the accurate modeling of these subtle groundwater processes. Understanding what frequencies are translated into the beach matrix will help determine which time scales are relevant and will support precise management strategies.

Also in the context of shoreline erosion, the periodicity of runup is of importance and an accurate spectral description is necessary, as waves stir up the sediment that is then transported by current (Bosboom and Stive, 2021). Higher frequency waves result in a constant agitation of the swash which prevents the settlement of sediment and makes it available for transport. A better spectral representation of the swash zone will enable more accurate estimates of coastal erosion. For now, both the ML and the CC methods introduce high frequencies that aren't extracted manually. Post-processing of the runup time series could help overcome this issue. For instance, smoothing techniques could be applied that preserve the amplitudes of larger events, such as polynomial filters or median filters.

To enable a better understanding of the hydrodynamics it is necessary to further refine auto-mated runup extraction methods. While it may be unrealistic to develop a universal CC method adaptable to all beach types and environmental conditions, one significant challenge identified in this study is the fragmented information regarding the applicability of various models. It would be extremely beneficial to have a comprehensive tool that compiles all relevant preprocessing techniques and extraction methods developed to date. Such a tool would accept a timestack image as input, apply all available methods, and then provide a visual representation of each method's performance. Currently, researchers must independently navigate the same extensive investigation process, often grappling with various programming languages. This redundancy likely leads to suboptimal solutions, which is unfortunate given the wealth of existing research that remains underutilized amidst the continual influx of new studies. However, it must be recognized that attempts have been made to generalize the applicability of methods by developing user-friendly toolboxes (Vousdoukas et al., 2011; Nuyts et al., 2023).

While knowledge about hydrodynamics is a prerequisite, a better understanding of the resilience of specific species to environmental influences is also required. Presently, bringing it back to sea turtle eggs, our knowledge of species-specific tolerance to hypoxic events is limited, and, factors such as frequency, duration, and timing relative to embryonic development remain understudied (Caut et al., 2010; Foley et al., 2006; Pike et al., 2015; Ware and Fuentes, 2018). A multidisciplinary approach is essential to investigate both the stressors affecting coastal systems and their resilience, whether these relate to turtle nests or other species.

5.1.4. Empirical estimates for flooding of sea turtle nests

In prior studies, Ware et al. (2019, 2021) utilized the empirical formula of Stockdon et al. (2006) to estimate $R_{2\%}$ and assess the necessity for relocating nests. Their evaluation involved comparing runup estimates against GPS measurements of the high water line, leading to the conclusion that the optimal beach slope β for applying the Stockdon formula was at the dune toe, the gentlest slope across the profile.

In this research, the Stockdon formula was applied to two distinct beach slopes: one corresponding to the dune toe ($\beta = 0.016$) and the other to the dune slope itself ($\beta = 0.1$). Additionally, the formula for extremely dissipative conditions, as suggested by Stockdon et al. (2006), was examined. The most favorable outcomes were observed with the formula adapted for extremely dissipative beaches, applied to data collected on November 13th and 14th. These settings yielded the lowest RMSE values across all three runup observation techniques, with the manual method recording 0.08 m, the machine learning (ML) approach 0.13 m, and the cross-correlation (CC) technique 0.13 m. These

results corroborate with Ware et al. (2019), however the formula for extremely dissipative conditions outperformed the general formula with the β as the dune toe slope.

Comparing the two beach slope estimates, the gentler slope at the dune toe ($\beta = 0.016$) consistently outperformed the steeper dune slope ($\beta = 0.1$). Lower β values significantly reduced the RMSE, enhancing the model's performance. Specifically, for $\beta = 0.016$, the RMSE was $0.16\,m$ for Manual, $0.28\,m$ for ML, and $0.27\,m$ for CC; whereas, for $\beta = 0.1$, the values escalated to $0.38\,m$, $0.51\,m$, and $0.50\,m$ respectively.

Furthermore, all three estimates exhibited a strong linear correlation with the measured runups (exceeding 0.7). It is important to note that a significant component of the runup estimates was derived from the recorded mean water level (tide + surge) at the NOAA station, which induced notable similarities in the shape of the three estimates for $R_{2\%}$.

Despite the high correlation between the $R_{2\%}$ estimates and actual runup measurements, the alignment was not uniform throughout the tidal cycle. During low tide, the estimations tended to overestimate $R_{2\%}$ more than during mid- to high tide. This pattern was also observed with the formula for dissipative conditions, where the measured values closely matched the estimates during high tide but overestimated $R_{2\%}$ by up to $0.2\,m$ during low tide. These findings align with observations by Stockdon et al. (2006), who noted a significantly lower correlation between estimates and observed values during low tide ($\rho = 0.29$) compared to mid- to high tide ($\rho = 0.52$), attributable to the shallower depth of the nearshore platform causing waves to break further offshore.

5.2. Prospect for multi-channel ML model

5.2.1. Importance of training set size and composition

Initially, the idea was to train the ML model on a small data selection of 1 hour 30 minutes on a particular day and then compute estimates of the remaining timestacks of that same day. However, one timestack image did not seem to provide enough variability. The model performed poorly under certain conditions not present in the training set, such as high luminescence or a non-smooth seepage face. To overcome these limitations a k-fold cross-validation was performed on all 9 labeled timestacks from November 13th and 14th. The results were promising but prompted some additional considerations.

On the one hand, the prediction of the considered timestack from November 14th gave a higher correlation with the hand-picked water line when the training set was extended to 8 timestacks. Namely, 0.77 instead of 0.70. This means that additional data helped the model generalize to data from the same day.

On the other hand, the mean accuracy across folds decreased from 96.09% to 93.03%, when the training set was increased. Suggesting that the model's performance became less consistent with expanded training. This outcome can be attributed to the high quality and similarity of the original training timestack to the validation data, which resulted in high initial accuracy. However, the imbalance within the extended training set might as well have caused the model to overfit to the timestacks of November 14th. In total, 9 labeled timestacks were used of which 6 were shot on November 14th, and only 3 on November 13th. The over-represented class had a mean accuracy of 93.94% with a standard deviation of 0.41%, while the under-represented class had an accuracy of 91.21% with a standard deviation of 0.78%. The imbalance, with more data from November 14th, likely affected the model's learning, making it less effective when validating the November 13th

timestacks. Therefore in future work, the model should be trained with balanced data or either naturally or by data augmentation, and regularization techniques could be implemented to address this issue and avoid overfitting. Another explanation would be the lesser quality of the timestack images from November 13th, where the swash is characterized by an undefined backwash.

Further training set expansion is recommended to increase the model's ability to overcome all environmental conditions and anomalies it could encounter. When predicting new unlabeled timestacks, areas that follow the water line accurately are observed, while others have estimates that are off by a few meters in the horizontal plane. Often, visual inspection reveals anomalies in the timestack that likely caused these inaccurate predictions, such as the presence of a truck or a person that remained in the section for a while and was not present in the training sets.

5.2.2. Challenges and opportunities related to preprocessing techniques

Analysis of the correlation of the model's performance to the hand-picked water line with the input channels statistics revealed that most potential lay in improving the model's handling of features with high variability in saturation and entropy, as these channel statistics correlate strongest with low model performances. Additional preprocessing steps that could help balance these aspects within our input data could be evaluated.

In this research, the timestacks were normalized over the full images. Column-wise normalization or clipping could reduce the variability of the timestack images and reduce the impact of sudden changes in time of for instance high luminescence. Normalization techniques are commonly used to ensure that all the data is on a similar scale and that gradients do not become too large, which can impede learning (Chng, 2022). Other preprocessing steps would be additional smoothing, transformation, standardizing, or combining channels by taking the mean of two channels.

The example of column-wise normalization of the entropy and the saturation channel showed how complex these measures operate, and, that a measure that seemingly improves the model's performance on one section of a timestack concurrently degrades other areas. A comprehensive sensitivity analysis that evaluates the impact on all input channels and preprocessing measures may provide more clarity.

5.2.3. Refining ground truth for better model assessment

The model was trained towards a manually digitized water line, and, after optimization areas of low correlation remained. It is important to recognize that hand-picked data shouldn't be treated as absolute ground truth. Instead, it functioned as a baseline for optimization in the absence of a more accurate water line, and, comparison with other data sources would be beneficial to improve the ML model further.

Manual digitization is straightforward and remains widely used (Simarro et al., 2015; Atkinson et al., 2017; Yang et al., 2022; Collins et al., 2023). However, on dissipative beaches such as the study case, distinguishing the backwash can be challenging. This can be mediated by assessing the averaging of the manually digitized waterline over multiple specialists. However, the labor involved with manual digitization is tremendous, and unrealistic if one wishes to extend the training set further. Preferably, another measurement method such as LiDAR measurements or resistance wires should be installed and used as ground truth.

5.2.4. Additional input channels and combinations

In this study, only 7 input channels were evaluated on their joint ability to segregate water from sand in timestack images. These channels were the grayscale image (I), dI/dt, dI/dx, saturation, RGB, entropy (E), and dE/dt. It took 10 hours to complete the k-fold cross-validation for all 127 combinations with an Intel Core i7 processor and 16 GB of RAM. Adding more input channels will result in an exponential increase in the computational time and should be performed on a higher-performance machine. Despite the increased computational demands, evaluating extra input channels is essential for further developing a multi-channel informed ML model for runup extraction.

In research by Collins et al. (2023) a CNN model was informed by multiple channels from LiDAR data. One of which was the elevation. Video imagery does not return height maps directly. However, by solving the photogrammetric equation, it is possible to project the two-dimensional pixel coordinates back to three-dimensional world coordinates -if the beach profile is measured accurately. The resulting elevation map can be included as an extra channel, which would naturally assign weight to the distribution of sand and water in the cross-shore direction. Alternatively, if regular profile measurements are unavailable, a linear function could be implemented to approximate this effect. The function should be implemented carefully to ensure a balanced delineation between water and sand, meaning that the water line should be positioned approximately at the center of the timestack. Otherwise, this additional channel might adversely impact the results by misrepresenting the boundary, leading to inaccurate predictions of water or sand regions.

Another possible improvement could be the addition of other covariates such as temperature, wind speed, offshore wave characteristics, atmospheric pressure, or humidity. These variables are easily obtained from local measurement stations and could be included as one-dimensional metrics in the learning process. This process, known as statistical downscaling, has been successfully implemented for medium-term shoreline prediction (Antolínez et al., 2019). By incorporating these additional covariates, the model could achieve better accuracy and robustness in predicting runup.

The Hue channel from the HSV color model was not evaluated here because the resulting image resembled the Saturation channel for the considered timestacks. The limited computational power led to choosing the Saturation over the Hue. However, studies (Aarninkhof et al., 2005) have used this channel in the context of shoreline and wet/dry boundary detection. It might be worth investigating whether or not it holds any added value compared to the saturation channel.

Also, combining channels into new ones by averaging might help balance certain extreme features by reducing the variance of the newly generated input channel. This method is called ensemble averaging and 'has been frequently compared to the individual estimators, and in many cases improves the resultant model accuracy' (Hashem, 1997).

Here the best channel combination for runup extraction was chosen based on the highest accuracy from a sensitivity analysis. It might be beneficial to conduct a feature importance analysis to identify and prioritize the most significant channels (Verdinelli and Wasserman, 2023). Alternatively, channels could be weighted dynamically during training (Turali et al., 2024).

During this research, the different channels were combined linearly. However, exploring non-linear combinations might improve the model's ability to predict runup. Non-linear combinations can capture more complex interactions and relationships between variables, which linear methods might miss. This could lead to more accurate and robust predictions by allowing the model to better understand and represent the underlying dynamics of runup phenomena. However, non-linearity in-

creases the model's complexity and herewith the chances of overfitting. A non-linear implementation is more likely to be beneficial if paired with and larger training set or measures to avoid overfitting. Also, the additional computational costs associated with a non-linear approach should be considered.

5.2.5. Further develop the machine-learning algorithm

Limitations, related to the computational time led to a stripped-down version of a CNN architecture, and hyperparameter tuning was left out of the scope of this research. The model was trained on small patches of 60×60 which internally limits the ML model's ability to capture the entire scales. It is advised to reduce the number of patches by using, for instance, square patches matching the spatial dimension. This way the ML model would be given a broader context, which is crucial in runup extraction where spatial and temporal continuity of larger patterns play an important role. Models trained on larger patches tend to generalize better on unseen data as they learn from a more comprehensive view of the timestack (Hamwood et al., 2018). Also, reducing the number of patches concurrently reduces the edge effects, observed in the predictions based on some of the input channels. To overcome the edge effects, diverse padding methods could also be evaluated, to better handle the border of the patches.

Adaptations in the CNN architecture could further enhance the model. For instance, augmenting the number of layers and filters could increase its ability to capture more complex patterns and the amount of information that the network can process. The size of filters influences the field of view of each convolutional operation. And it might be that a more suited kernel size exists that could help the model recognize specific patterns for the scales at play in the timestack images. Our research used Max pooling to reduce the dimension of the input data and widen the field of view of the neurons. Other pooling methods such as average pooling, might result in better results. Especially because extreme values in the entropy and the saturation channel correlated with low model performance. Averaging might help balance this issue. Additional model improvement could involve varying the activation functions (e.g., ReLU, sigmoid, tanh), or the optimizers (e.g., Adam, SGD, RMSprop) to speed up the convergence and improve stability.

The end model returns the most wrongly predicted water pixels in the seepage face. This corroborates with most CC-based methods that fail to extract the water line on dissipative beaches due to the decoupling of the water line from the effluent line (Huisman et al., 2011). To address this issue, the addition of a third class to the model is proposed, which distinguishes the seepage face from the dry sand. It might help to segregate the characteristics of the seepage face from the sand and the water as it is very different from both classes to the eye. In recent literature (Kang et al., 2024a) analyzed coastal video images and classified them according to the following three classes: 'water', 'sky', and 'background' (or beach). Here, the 'sky category was added to reduce the amount of [...] wrongly predicted water pixels'.
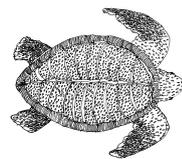
6

Conclusion
& Recommendations

The goal of this study was to extract the instantaneous runup from video imagery under the challenging conditions of dissipative beaches, to enable more precise coastal management strategies. A more traditional color contrast (CC) method based on local entropy and saturation, demonstrated good performance and is considered applicable under similar conditions. The results from the developed machine learning (ML) algorithm validate it as a viable proof of concept and challenge us to enhance its adaptability and accuracy across varied environmental conditions. When applied to other datasets, the CNN method occasionally failed to accurately capture the water line due to specific characteristics of the new timestack images. Despite these limitations, the CNN method can be effectively implemented for long-term runup analysis under similar conditions. Both methods demonstrate potential in reducing the time required to extract the instantaneous runup from video imagery and contribute to the enhancement of real-time monitoring and predictive modeling of coastal processes.

6.1. Conclusion

- Image preprocessing techniques have proven effective in creating new input channels for a convolutional neural network (CNN), enhancing its ability to accurately extract wave runup from video imagery. A multi-channel informed machine learning (ML) algorithm was developed to utilize the grayscale image (I), its spatial derivative (dI/dx), saturation (S), entropy (E), and the temporal derivative of entropy (dE/dt) as input. This multi-channel ML model significantly outperformed the single-channel approach using only grayscale as input, achieving a correlation coefficient of 0.77 compared to 0.66.

- The ML model demonstrated strong agreement during cross-validation with manually labeled time series and corresponding runup statistics and mean period ($T_{m-1,0}$). The model achieved a root-mean-square-error (RMSE) of 0.10 m and a correlation coefficient (r) of 0.96 for the runup time series, an RMSE of 0.09 m, and an r-value of 0.97 for $R_{2\%}$, and an RMSE of 3.51 s with an r-value of 0.79 for $T_{m-1,0}$. However, when applied to other datasets, the method occasionally failed to accurately capture the waterline due to specific characteristics of the new timestack images. Despite these small limitations, the ML method could be implemented for long-term runup analysis.

- The CC method demonstrated strong agreement with the manual runup time series, achieving an RMSE of 0.12 m and an r-value of 0.94. For the $R_{2\%}$ metric, the method showed an RMSE of 0.08 m and an r-value of 0.97. Additionally, for $T_{m-1,0}$, the method achieved an RMSE of 3.88 s and an r-value of 0.70. These results suggest that the CC method is likely to be applicable to similar beaches along the northern Gulf of Mexico.

- Among the three empirical estimates for $R_{2\%}$ (Stockdon et al., 2006), the formula for extremely dissipative conditions best agreed with the measured runup (Manual, ML, and CC), with RMSE-values lower than 0.13 m and r-values exceeding 0.7. For the general formula, the dune toe slope provided the best results—similarly concluded by Ware et al. (2021)—although the alignment was not as accurate as that achieved with the dissipative formula.

6.2. Recommendations

- In general, the field of runup extraction would benefit from a holistic literature review of all the methods developed to date. Including their availability (whether open source or proprietary), the specific environmental conditions under which they are applicable, and the prerequisite knowledge required for proper implementation, such as familiarity with the relevant programming languages or the availability of a guided user interface (GUI).

- Based on the obtained results, it is recommended to use the formula for extremely dissipative conditions from Stockdon et al. (2006) under similar environmental conditions as those presented here. However, it is important to note that the $R_{2\%}$ estimates will likely be overestimated during low tide.

- Future research should refine the automation of runup extraction routines to produce even more reliable data for model validation, focusing on the timing aspect of the runup time series. For the ML model, it is suggested to further extend the training set to enhance robustness, investigate other possible input channels and additional preprocessing techniques, use more reliable ground truth data such as LiDAR, and further develop the ML algorithm. This development could include increasing the patch size and the network depth, changing the pooling method, or adding the seepage face as a third prediction class.

- Aside from enhancing runup extraction and understanding hydrodynamic processes, further development of knowledge on species' resilience to withstand partial flooding events is necessary, focusing on aspects such as frequency, duration, and timing relative to embryonic development.
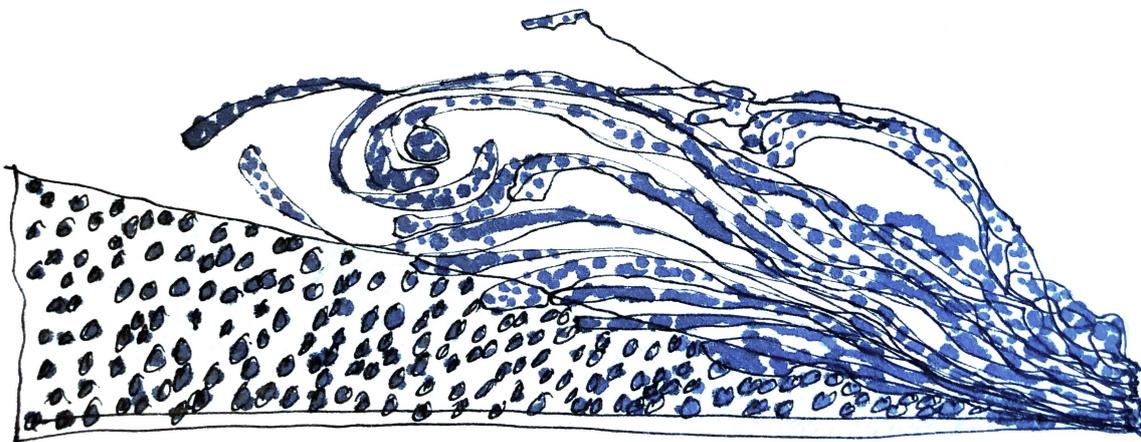
Bibliography

Aarninkhof, S., Ruessink, B., and Roelvink, J. (2005). Nearshore subtidal bathymetry from time-exposure video images. *Journal of Geophysical Research: Oceans*, 110(C6).

Aarninkhof, S. G. and Ruessink, B. G. (2004). Video observations and model predictions of depth-induced wave dissipation. *IEEE Transactions on Geoscience and Remote Sensing*, 42:2612–2622.

Ahles, N. and Milton, S. L. (2016). Mid-incubation relocation and embryonic survival in loggerhead sea turtle eggs. *The journal of wildlife management*, 80(3):430–437.

Almar, R., Blenkinsopp, C., Almeida, L. P., Cienfuegos, R., and Catalán, P. A. (2017). Wave runup video motion detection using the radon transform. *Coastal Engineering*, 130:46–51.

Almar, R., Bonneton, P., Senechal, N., and Roelvink, D. (2009). Wave celerity from video imaging: A new method. In *Coastal Engineering 2008: (In 5 Volumes)*, pages 661–673. World Scientific.

Andriolo, U. (2019). Nearshore wave transformation domains from video imagery. *Journal of Marine Science and Engineering*, 7.

Antolínez, J. A., Méndez, F. J., Anderson, D., Ruggiero, P., and Kaminsky, G. M. (2019). Predicting climate-driven coastlines with a simple and efficient multiscale model. *Journal of Geophysical Research: Earth Surface*, 124(6):1596–1624.

Arienzo, M. (2023). Progress on the impact of persistent pollutants on marine turtles: A review. *Journal of Marine Science and Engineering*, 11(2):266.

Atherton, R. J., Baird, A. J., and Wiggs, G. F. S. (2001). Inter-tidal dynamics of surface moisture content on a meso-tidal beach. *Source: Journal of Coastal Research*, 17:482–489.

Atkinson, A. L., Power, H. E., Moura, T., Hammond, T., Callaghan, D. P., and Baldock, T. E. (2017). Assessment of runup predictions by empirical models on non-truncated beaches on the south-east australian coast. *Coastal Engineering*, 119:15–31.

Bailey, D. G. and Shand, R. D. (1994). Determining wave run-up using automated video analysis. In *Proceedings of the Second New Zealand Conference on Image and Vision Computing*, volume 2, pages 1–2.

Baird, A. J. and Low, R. G. (2022). The water table: Its conceptual basis, its measurement and its usefulness as a hydrological variable. *Hydrological Processes*, 36.

Baird, A. J., Mason, T., and Horn, D. P. (1998). Validation of a boussinesq model of beach ground water behaviour. *Marine Geology*, 148(1-2):55–69.

Bakhtyar, R., Brovelli, A., Barry, D. A., and Li, L. (2011). Wave-induced water table fluctuations, sediment transport and beach profile change: Modeling and comparison with large-scale laboratory experiments. *Coastal Engineering*, 58:103–118.

Benson, C. H., Gunter, J. A., Boutwell, G. P., Trautwein, S. J., and Berzanskis, P. H. (1997). Comparison of four methods to assess hydraulic conductivity. *Journal of Geotechnical and Geoenvironmental Engineering*, 123(10):929–937.

Bertin, X., de Bakker, A., Van Dongeren, A., Coco, G., André, G., Ardhuin, F., Bonneton, P., Bouchette, F., Castelle, B., Crawford, W. C., et al. (2018). Infragravity waves: From driving mechanisms to impacts. *Earth-Science Reviews*, 177:774–799.

Bosboom, J. and Stive, M. J. (2021). Coastal dynamics.

Bruder, B. L. and Brodie, K. L. (2020). Cirn quantitative coastal imaging toolbox. *SoftwareX*, 12:100582.

Bruun, P. (1962). Sea-level rise as a cause of shore erosion. *Journal of the Waterways and Harbors division*, 88(1):117–130.

Caillouet Jr, C. W., Raborn, S. W., Shaver, D. J., Putman, N. F., Gallaway, B. J., and Mansfield, K. L. (2018). Did declining carrying capacity for the kemp's ridley sea turtle population within the gulf of mexico contribute to the nesting setback in 2010- 2017? *Chelonian Conservation and Biology*, 17(1):123–133.

Cartwright, N., Baldock, T., Nielsen, P., Jeng, D.-S., and Tao, L. (2005). Swash-aquifer interaction in sandy beaches. In *Coasts and Ports 2005: Coastal Living-Living Coast; Australasian Conference; Proceedings*, pages 351–356. Institution of Engineers, Australia Barton, ACT.

Cartwright, N., Nielsen, P., and Li, L. (2004). Experimental observations of watertable waves in an unconfined aquifer with a sloping boundary. *Advances in Water Resources*, 27:991–1004.

Caut, S., Guirlet, E., and Girondot, M. (2010). Effect of tidal overwash on the embryonic development of leatherback turtles in french guiana. *Marine Environmental Research*, 69(4):254–261.

Cheng, I.-J., Wang, H.-Y., Hsieh, W.-Y., and Chan, Y.-T. (2019). Twenty-three years of sea turtle stranding/bycatch research in taiwan. *Zoological Studies*, 58.

Chng, Z. (2022). Using normalization layers to improve deep learning models. *MachineLearningMastery. com*.

Christiaanse, J. C., Antolínez, J. A., Luijendijk, A. P., Athanasiou, P., Duarte, C. M., and Aarninkhof, S. (2024). Distribution of global sea turtle nesting explained from regional-scale coastal characteristics.

Collins, A. M., O'Dea, A., Brodie, K. L., Bak, A. S., Hesser, T. J., Spore, N. J., and Farthing, M. W. (2023). Automated extraction of a depth-defined wave runup time series from lidar data using deep learning. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–13.

Colvin, J., Lazarus, S., and Splitt, M. (2020). Extracting nearshore wave properties from video: A new method for coastal estuaries. *Estuarine, Coastal and Shelf Science*, 246:107053.

da Silva, P. G., Medina, R., González, M., and Garnier, R. (2018). Infragravity swash parameterization on beaches: The role of the profile shape and the morphodynamic beach state. *Coastal Engineering*, 136:41–55.

Dalleau, M., Ciccione, S., Mortimer, J. A., Garnier, J., Benhamou, S., and Bourjea, J. (2012). Nesting phenology of marine turtles: insights from a regional comparative analysis on green turtle (chelonia mydas).

Darquea, J. J., Ortiz-Alvarez, C., Córdova-Zavaleta, F., Medina, R., Bielli, A., Alfaro-Shigueto, J., and Mangel, J. C. (2020). Trialing net illumination as a bycatch mitigation measure for sea turtles in a small-scale gillnet fishery in ecuador. *Latin american journal of aquatic research*, 48(3):446–455.

Dédina, D. (2023). Sea turtle nest inundation in ras baridi: Improving flood risk modeling in data-limited coastal regions.

Dellapenna, T. and Olszewski, T. (2023). Field trip notes: Barrier island depositional environments of the upper texas coast.

Dewald, J. R. and Pike, D. A. (2014). Geographical variation in hurricane impacts among sea turtle populations. *Journal of Biogeography*, 41(2):307–316.

Didier, D., Bernatchez, P., Marie, G., and Boucher-Brossard, G. (2016). Wave runup estimations on platform-beaches for coastal flood hazard assessment. *Natural Hazards*, 83:1443–1467.

Didier, D., Caulet, C., Bandet, M., Bernatchez, P., Dumont, D., Augereau, E., Floc'h, F., and Delacourt, C. (2020). Wave runup parameterization for sandy, gravel and platform beaches in a fetch-limited, large estuarine system. *Continental Shelf Research*, 192:104024.

Dodet, G., Leckler, F., Sous, D., Ardhuin, F., Filipot, J. F., and Suanez, S. (2018). Wave runup over steep rocky cliffs. *Journal of Geophysical Research: Oceans*, 123(10):7185–7205.

Du, J., Park, K., Yu, X., Zhang, Y. J., and Ye, F. (2020). Massive pollutants released to galveston bay during hurricane harvey: Understanding their retention and pathway using lagrangian numerical simulations. *Science of The Total Environment*, 704:135364.

Fish, M. R., Côté, I. M., Gill, J. A., Jones, A. P., Renshoff, S., and Watkinson, A. R. (2005). Predicting the impact of sea-level rise on caribbean sea turtle nesting habitat. *Conservation biology*, 19(2):482–491.

Foley, A. M., Peck, S. A., and Harman, G. R. (2006). Effects of sand characteristics and inundation on the hatching success of loggerhead sea turtle (caretta caretta) clutches on low-relief mangrove islands in southwest florida. *Chelonian Conservation and Biology*, 5(1):32–41.

Fuentes, M., Limpus, C., and Hamann, M. (2011). Vulnerability of sea turtle nesting grounds to climate change. *Global Change Biology*, 17(1):140–153.

Fuentes, M. M., Meletis, Z. A., Wildermann, N. E., and Ware, M. (2021). Conservation interventions to reduce vessel strikes on sea turtles: A case study in florida. *Marine Policy*, 128:104471.

Fuse, T. and Ohkura, T. (2018). Development of shoreline extraction method based on spatial pattern analysis of satellite sar images. *Remote Sensing*, 10(9):1361.

Geng, X., Heiss, J. W., Michael, H. A., and Boufadel, M. C. (2017). Subsurface flow and moisture dynamics in response to swash motions: Effects of beach hydraulic conductivity and capillarity. *Water Resources Research*, 53:10317–10335.

Geng, X. and Michael, H. A. (2021). Along-shore movement of groundwater and its effects on seawater-groundwater interactions in heterogeneous coastal aquifers. *Water Resources Research*, 57(12):e2021WR031056.

Glen, F., Broderick, A., Godley, B., and Hays, G. (2003). Incubation environment affects phenotype of naturally incubated green turtle hatchlings. *Journal of the Marine Biological Association of the United Kingdom*, 83(5):1183–1186.

Gomes da Silva, P., Coco, G., Garnier, R., and Klein, A. H. (2020). On the prediction of runup, setup and swash on beaches. *Earth-Science Reviews*, 204.

Grand, J. and Beissinger, S. R. (1997). When relocation of loggerhead sea turtle (caretta caretta) nests becomes a useful strategy. *Journal of Herpetology*, 31(3):428–434.

Guedes, R. M., Bryan, K. R., and Coco, G. (2012). Observations of alongshore variability of swash motions on an intermediate beach. *Continental Shelf Research*, 48:61–74.

Guedes, R. M. C., Bryan, K. R., Coco, G., and Holman, R. A. (2011). The effects of tides on swash statistics on an intermediate beach. *Journal of Geophysical Research: Oceans*, 116(C4).

Guest, T. B. and Hay, A. E. (2017). Vertical structure of pore pressure under surface gravity waves on a steep, megatidal, mixed sand-gravel-cobble beach. *Journal of Geophysical Research: Oceans*, 122:153–170.

Guza, R. T. and Thornton, E. B. (1982). Swash oscillations on a natural beach. *Journal of Geophysical Research: Oceans*, 87(C1):483–491.

Hamwood, J., Alonso-Caneiro, D., Read, S. A., Vincent, S. J., and Collins, M. J. (2018). Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of oct retinal layers. *Biomedical optics express*, 9(7):3049–3066.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.

Hasan, G. J. and Takewaka, S. (2014). Foreshore applications of x-band radar. In *Remote Sensing and Modeling: Advances in Coastal and Marine Resources*, pages 161–191. Springer.

Hashem, S. (1997). Optimal linear combinations of neural networks. *Neural networks*, 10(4):599–614.

Hays, G. C. (2017). Ocean currents and marine life. *Current Biology*, 27(11):R470–R473.

Hegge, B. J. and Masselink, G. (1991). Groundwater-table responses to wave run-up: An experimental study from western australia. *Journal of Coastal Research*, 7(3):623–634.

Holland, K., Raubenheimer, B., Guza, R., and Holman, R. A. (1995). Runup kinematics on a natural beach. *Journal of Geophysical Research: Oceans*, 100(C3):4985–4993.

Holman, R. A., Sallenger, A. H., Lippmann, T. C., and Haines, J. W. (1993). The application of video image processing to the study of nearshore processes. *Oceanography*, 6(3):78–85.

Horn, D. P. (2002). Beach groundwater dynamics. *Geomorphology*, 48(1):121–146. 29th Binghamton Geomorphology Symposium: Coastal Geomorphology.

Horn, D. P. (2006). Measurements and modelling of beach groundwater flow in the swash-zone: a review.

Huff, T., Feagin, R., and Figlus, J. (2020). Enhanced tide model: Improving tidal predictions with integration of wind data. *Shore and beach*, pages 40–45.

Huisman, C. E., Bryan, K. R., Coco, G., and Ruessink, B. (2011). The use of video imagery to analyse groundwater and shoreline dynamics on a dissipative beach. *Continental shelf research*, 31(16):1728–1738.

International Union for Conservation of Nature (2021). Red list assessments: IUCN-SSC marine turtle specialist group. https://www.iucn-mtsg.org/statuses. Online; accessed 19 December 2023.

Jensen, M. P., Eguchi, T., FitzSimmons, N. N., McCarthy, M. A., Fuentes, M. M., Hamann, M., Limpus, C. J., Bell, I. P., and Read, M. A. (2022). Integrating climate change and management scenarios in population models to guide the conservation of marine turtles. *Bulletin of Marine Science*, 98(2):131–154.

Kang, B., Feagin, R. A., Huff, T., and Durán Vinent, O. (2024a). Stochastic properties of coastal flooding events–part 1: convolutional-neural-network-based semantic segmentation for water detection. *Earth Surface Dynamics*, 12(1):1–10.

Kang, B., Feagin, R. A., Huff, T., and Durán Vinent, O. (2024b). Stochastic properties of coastal flooding events–part 2: Probabilistic analysis. *Earth Surface Dynamics*, 12(1):105–115.

Kingston, K. S. (2003). *Applications of complex adaptive systems approaches to coastal systems*. PhD thesis, University of Plymouth.

Kranenborg, J. W., Pauli, T., Jacobsen, N. G., van der Werf, J. J., Antonio, S. D., Campmans, G. H., Reniers, A. J., and Hulscher, S. J. (2023). Measurements and modeling of pore-pressure gradients in the swash zone under large-scale laboratory bichromatic waves. *Journal of Geophysical Research: Oceans*, 128.

Krumbein, W. and Monk, G. (1943). Permeability as a function of the size parameters of unconsolidated sand. *Transactions of the AIME*, 151(01):153–163.

Kuo, C., Hwung, H.-H., and Chien, C. (2009). Using time-stack overlooking images to separate incident and reflected waves in wave flume. *Wave Motion*, 46(3):189–199.

Lefevre, F., Le Provost, C., and Lyard, F. (2000). How can we improve a global ocean tide model at a regional scale? a test on the yellow sea and the east china sea. *Journal of Geophysical Research: Oceans*, 105(C4):8707–8725.

Limpus, C. J., Baker, V., and Miller, J. D. (1979). Movement induced mortality of loggerhead eggs. *Herpetologica*, pages 335–338.

Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., and Aarninkhof, S. (2018). The state of the world's beaches. *Scientific reports*, 8(1):1–11.

Lutcavage, M. E. (2017). Human impacts on sea turtle survival. In *The Biology of Sea Turtles, Volume I*, pages 387–409. CRC press.

Lyddon, C., Bird, C., Brown, J., and Plater, A. (2021). Application of x-band radars for deriving intertidal bathymetries and characterising coastal behaviours.

Maglio, C., Das, H., and Fenner, F. (2020). Modal grain size evolution as it relates to the dredging and placement process-galveston island, texas. *Coastal Engineering Proceedings*, (36v):7–7.

Marti-Puig, P., Serra-Serra, M., Ribas, F., Simarro, G., and Caballeria, M. (2024). Automatic shoreline detection by processing planview timex images using bi-lstm networks. *Expert Systems with Applications*, 240:122566.

Matsuzawa, Y., Sato, K., Sakamoto, W., and Bjorndal, K. (2002). Seasonal fluctuations in sand temperature: effects on the incubation period and mortality of loggerhead sea turtle (caretta caretta) pre-emergent hatchlings in minabe, japan. *Marine Biology*, 140:639–646.

Maulany, R., Booth, D., and Baxter, G. (2012). The effect of incubation temperature on hatchling quality in the olive ridley turtle, lepidochelys olivacea, from alas purwo national park, east java, indonesia: implications for hatchery management. *Marine Biology*, 159:2651–2661.

Mendes, D., Andriolo, U., and Neves, M. G. (2022). Advances in Wave Run-Up Measurement Techniques. In *Advances on Testing and Experimentation in Civil Engineering*, pages 283–297. Springer, Cham, Switzerland.

Mrosovsky, N. (2006). Distorting gene pools by conservation: assessing the case of doomed turtle eggs. *Environmental Management*, 38(4):523–531.

National Geographic (2024). Hurricanes in history: Galveston hurricane. National Geographic website.

Nayak, S. and Panchang, V. (2015). A note on short-term wave height statistics. *Aquatic Procedia*, 4:274–280.

Nielsen, P. and Hanslow, D. J. (1991). Wave runup distributions on natural beaches. *Journal of Coastal Research*, 7(4):1139–1152.

NOAA (2024a). National Data Buoy Center galveston pleasure pier. https://www.ndbc.noaa.gov/station_page.php?station=42035.

NOAA (2024b). NOAA Tides and Currents galveston pleasure pier. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm.

Nooren, H. and Claridge, G. (2002). Guidelines for turtle hatchery management. *Turtle Foundation, Hauptstr*, 1.

Nuyts, S., Almar, R., Morichon, D., Dealbera, S., Abalia, A., Muñoz, J. M., Abessolo, G. O., and Regard, V. (2023). Coastcams: A matlab toolbox making accessible estimations of nearshore processes, mean water levels, and morphology from timestack images. *Environmental Modelling Software*, 168:105800.

Otsu, N. et al. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27.

Paine, J. G., Caudle, T. L., Costard, L., Elliott, B. A., Woodruff Jr, C. M., Woodruff Jr, C., and Tinker, S. W. (2020). Texas statemap program summary, fy19 (2019–2020).

Parmenter, C. (1980). Incubation of the eggs of the green sea turtle, chelonia mydas, in torres strait, australia: the effect of movement on hatchability. *Wildlife Research*, 7(3):487–491.

Patrício, A. R., Hawkes, L. A., Monsinjon, J. R., Godley, B. J., and Fuentes, M. M. (2021). Climate change and marine turtles: recent advances and future directions. *Endangered Species Research*, 44:363–395.

Pheasey, H., Glen, G., Allison, N. L., Fonseca, L. G., Chacón, D., Restrepo, J., and Valverde, R. A. (2021). Quantifying illegal extraction of sea turtles in costa rica. *Frontiers in Conservation Science*, 2:705556.

Pike, D. A., Roznik, E. A., and Bell, I. (2015). Nest inundation from sea-level rise threatens sea turtle population viability. *Royal Society Open Science*, 2(7):150127.

Poate, T. G., McCall, R. T., and Masselink, G. (2016). A new parameterisation for runup on gravel beaches. *Coastal Engineering*, 117:176–190.

Power, H. E., Gharabaghi, B., Bonakdari, H., Robertson, B., Atkinson, A. L., and Baldock, T. E. (2019). Prediction of wave runup on beaches using gene-expression programming and empirical relationships. *Coastal Engineering*, 144:47–61.

Power, H. E., Holman, R. A., and Baldock, T. E. (2011). Swash zone boundary conditions derived from optical remote sensing of swash zone flow patterns.

Raubenheimer, B., Guza, R., and Elgar, S. (1999). Tidal water table fluctuations in a sandy ocean beach. *Water Resources Research*, 35(8):2313–2320.

Ravens, T. M., Thomas, R. C., Roberts, K. A., and Santschi, P. H. (2009). Causes of salt marsh erosion in galveston bay, texas. *Journal of Coastal Research*, 25(2):265–272.

Rigos, A., Andreadis, O. P., Andreas, M., Vousdoukas, M. I., Tsekouras, G. E., and Velegrakis, A. (2014). Shoreline extraction from coastal images using chebyshev polynomials and rbf neural

networks. In *Artificial Intelligence Applications and Innovations: 10th IFIP WG 12.5 International Conference, AIAI 2014, Rhodes, Greece, September 19-21, 2014. Proceedings 10*, pages 593–603. Springer.

Rivas, M., Spínola, M., Arrieta, H., and Faife-Cabrera, M. (2018). Effect of extreme climatic events resulting in prolonged precipitation on the reproductive output of sea turtles. *Animal Conservation*, 21(5):387–395.

Ruessink, B., Kleinhans, M., and Van den Beukel, P. (1998). Observations of swash under highly dissipative conditions. *Journal of Geophysical Research: Oceans*, 103(C2):3111–3118.

Ruggiero, P., Holman, R. A., and Beach, R. (2004). Wave run-up on a high-energy dissipative beach. *Journal of Geophysical Research: Oceans*, 109(C6).

Sanvicente-Añorve, L., Zavala-Hidalgo, J., Allende-Arandía, E., and Hermoso-Salazar, M. (2018). Larval dispersal in three coral reef decapod species: Influence of larval duration on the metapopulation structure. *PloS one*, 13(3):e0193457.

Segura, L. N. and Cajade, R. (2010). The effects of sand temperature on pre-emergent green sea turtle hatchlings.

Senechal, N., Coco, G., Bryan, K. R., and Holman, R. A. (2011). Wave runup during extreme storm conditions. *Journal of Geophysical Research: Oceans*, 116(C7).

Seney, E. E. (2008). *Population dynamics and movements of the Kemp's ridley sea turtle, Lepidochelys kempii, in the northwestern Gulf of Mexico.* Texas A&M University.

Senko, J. F., Burgher, K. M., del Mar Mancha-Cisneros, M., Godley, B. J., Kinan-Kelly, I., Fox, T., Humber, F., Koch, V., Smith, A. T., and Wallace, B. P. (2022). Global patterns of illegal marine turtle exploitation. *Global change biology*, 28(22):6509–6523.

Simarro, G., Bryan, K. R., Guedes, R. M., Sancho, A., Guillen, J., and Coco, G. (2015). On the use of variance images for runup and shoreline detection. *Coastal Engineering*, 99:136–147.

Sinaei, M., Zare, R., Talebi Matin, M., and Ghasemzadeh, J. (2021). Marine debris and trace metal (cu, cd, pb, and zn) pollution in the stranded green sea turtles (chelonia mydas). *Archives of Environmental Contamination and Toxicology*, 80:634–644.

Sönmez, B., Karaman, S., and Turkozan, O. (2021). Effect of predicted sea level rise scenarios on green turtle (chelonia mydas) nesting. *Journal of Experimental Marine Biology and Ecology*, 541:151572.

Sous, D., Petitjean, L., Bouchette, F., Rey, V., Meulé, S., Sabatier, F., and Martins, K. (2016). Field evidence of swash groundwater circulation in the microtidal rousty beach, france. *Advances in Water Resources*, 97:144–155.

Stark, N., Mewis, P., Reeve, B., Florence, M., Piller, J., and Simon, J. (2022). Vertical pore pressure variations and geotechnical sediment properties at a sandy beach. *Coastal Engineering*, 172.

Stockdon, H. F., Holman, R. A., Howd, P. A., and Sallenger Jr, A. H. (2006). Empirical parameterization of setup, swash, and runup. *Coastal engineering*, 53(7):573–588.

Surf-Forecast (2024). Surf-Forecast galveston surf stats. https://www.surf-forecast.com/breaks/Galveston/surf-stats.

Swift, D. J. (1968). Coastal erosion and transgressive stratigraphy. *The Journal of Geology*, 76(4):444–456.

Tanner, C. E., Marco, A., Martins, S., Abella-Perez, E., and Hawkes, L. A. (2019). Highly feminised sex-ratio estimations for the world's third-largest nesting aggregation of loggerhead sea turtles. *Marine Ecology Progress Series*, 621:209–219.

Tomillo, P. S., Oro, D., Paladino, F. V., Piedra, R., Sieg, A. E., and Spotila, J. R. (2014). High beach temperatures increased female-biased primary sex ratios but reduced output of female hatchlings in the leatherback turtle. *Biological Conservation*, 176:71–79.

Turali, M. Y., Lorasdagi, M. E., Koc, A. T., and Kozat, S. S. (2024). Afs-bm: Enhancing model performance through adaptive feature selection with binary masking. *arXiv preprint arXiv:2401.11250*.

Turner, I. L., Coates, B. P., and Acworth, R. I. (1997). Tides, waves and the super-elevation of groundwater at the coast. *Source: Journal of Coastal Research*, 13:46–60.

Turner, I. L., Leyden, V. M., Symonds, G., Mcgrath, J., Jackson, A., Jancar, T., Aarninkhof, S., and Elshoff, I. (2001). Predicted and observed coastline changes at the gold coast artificial reef. In *Coastal Engineering 2000*, pages 1836–1847.

Turner, I. L. and Nielsen, P. (1997). Rapid water table fluctuations within the beach face: Implications for swash zone sediment mobility? *Coastal Engineering*, 32(1):45–59.

Turner, I. L., Russell, P. E., and Butt, T. (2008). Measurement of wave-by-wave bed-levels in the swash zone. *Coastal Engineering*, 55(12):1237–1242.

US Amry Corps of Engineering (2022). Galveston 204 cap – beach nourishment. Engineering Appendix.

Uunk, L., Wijnberg, K., and Morelissen, R. (2010). Automated mapping of the intertidal beach bathymetry from video images. *Coastal Engineering*, 57(4):461–469.

Valverde, R. A. and Holzwart, K. R. (2017). Sea turtles of the gulf of mexico. *Habitats and Biota of the Gulf of Mexico: Before the Deepwater Horizon Oil Spill: Volume 2: Fish Resources, Fisheries, Sea Turtles, Avian Resources, Marine Mammals, Diseases and Mortalities*, pages 1189–1351.

Veelenturf, C. A., Sinclair, E. M., Paladino, F. V., and Honarvar, S. (2020). Predicting the impacts of sea level rise in sea turtle nesting habitat on bioko island, equatorial guinea. *Plos One*, 15(7):e0222251.

Verdinelli, I. and Wasserman, L. (2023). Feature importance: A closer look at shapley values and loco. *arXiv preprint arXiv:2303.05981*.

Vousdoukas, M. I., Ferreira, P. M., Almeida, L. P., Dodet, G., Psaros, F., Andriolo, U., Taborda, R., Silva, A. N., Ruano, A., and Ferreira, Ó. M. (2011). Performance of intertidal topography video monitoring of a meso-tidal reflective beach in south portugal. *Ocean Dynamics*, 61:1521–1540.

Wang, J., Wang, L., Feng, S., Peng, B., Huang, L., Fatholahi, S. N., Tang, L., and Li, J. (2023). An overview of shoreline mapping by using airborne lidar. *Remote Sensing*, 15(1):253.

Ware, M., Ceriani, S. A., Long, J. W., and Fuentes, M. M. (2021). Exposure of loggerhead sea turtle nests to waves in the florida panhandle. *Remote Sensing*, 13(14):2654.

Ware, M. and Fuentes, M. M. (2018). Potential for relocation to alter the incubation environment and productivity of sea turtle nests in the northern gulf of mexico. *Chelonian Conservation and Biology*, 17(2):252–262.

Ware, M., Long, J. W., and Fuentes, M. M. (2019). Using wave runup modeling to inform coastal species management: An example application for sea turtle nest relocation. *Ocean and Coastal Management*, 173:17–25.

Whitmore, C. P. and Dutton, P. H. (1985). Infertility, embryonic mortality and nest-site selection in leatherback and green sea turtles in suriname. *Biological conservation*, 34(3):251–272.

Wood, D. W. and Bjorndal, K. A. (2000). Relation of temperature, moisture, salinity, and slope to nest site selection in loggerhead sea turtles. *Copeia*, 2000(1):119–119.

Yang, H., Zhang, C., Shi, J., Cao, X., and Li, Y. (2022). Laboratory observations of swash signatures using video imagery. *Journal of Marine Science and Engineering*, 10(12):1833.

Yang, S.-H., Chen, C.-H., and Chu, K.-H. (2021). Fecal indicators, pathogens, antibiotic resistance genes, and ecotoxicity in galveston bay after hurricane harvey. *Journal of Hazardous Materials*, 411:124953.

Yntema, C. and Mrosovsky, N. (1982). Critical periods and pivotal temperatures for sexual differentiation in loggerhead sea turtles. *Canadian Journal of Zoology*, 60(5):1012–1016.

Zhang, S. and Zhang, C. (2009). Image analysis for wave swash using color feature extraction. In *2009 2nd International Congress on Image and Signal Processing*, pages 1–4. IEEE.