Wave runup extraction on dissipative beaches
New video-based methods

van der Grinten, Meye J.; Christiaanse, Jakob C.; Reniers, Ad J.H.M.; Taal, Falco; Figlus, Jens; Antolínez, José A.A.

Wave runup extraction on dissipative beaches: New video-based methods

Meye J. van der Grinten [a], Jakob C. Christiaanse [a] [ID],*, Ad J.H.M. Reniers [a], Falco Taal [a], Jens Figlus [b], José A.A. Antolínez [a]

[a] Department of Hydraulic Engineering, Delft University of Technology, Stevinweg 1, Delft, 2628 CN, Netherlands
[b] Department of Ocean Engineering, Texas A&M University, 200 Seawolf Parkway, Galveston, 77554, TX, USA

ARTICLE INFO

ABSTRACT

Wave runup observations are important for coastal management providing data to validate predictive models of inundation frequencies and erosion rates, which are vital for assessing the vulnerability of coastal ecosystems and infrastructure. Automated algorithms to extract the instantaneous water line from video imagery struggle under dissipative conditions, where the presence of a seepage face and the lack of contrast between the sand and the swash impede proper extraction, requiring time-intensive data quality control or manual digitization. This study introduces two novel methods, based on color contrast (CC) and machine learning (ML). The CC method combines texture roughness — local entropy — with saturation. Images are first binarized using entropy values and then refined through noise reduction by binarization of the saturation channel. The ML method uses a convolutional neural network (CNN) informed by five channels: the grayscale intensity and its time gradient, the saturation channel, and the entropy and its time gradient. Both methods were validated against nine manually labeled, 80 min video time series. The CC method demonstrated strong agreement with manually digitized water lines (RMSE = 0.12 m, r = 0.94 for the vertical runup time series; RMSE = 0.08 m, r = 0.97 for the 2% runup exceedance ($R_{2\%}$); and RMSE = 3.88 s, r = 0.70 for the mean period ($T_{m-1,0}$)). The ML model compared well with the manually labeled time series (RMSE = 0.10 m, r = 0.96 for the vertical runup time series; RMSE = 0.09 m, r = 0.97 for $R_{2\%}$; and RMSE = 3.51 s, r = 0.79 for $T_{m-1,0}$). Furthermore, the computed $R_{2\%}$ values of both methods show a good agreement with the formula proposed by Stockdon et al. (2006) for extremely dissipative conditions, with RMSE-values lower than 0.13 m and correlations exceeding 0.70 for manual, CC, and ML estimates. While the CC method is deemed applicable for wave-by-wave analysis under similar dissipative conditions with a smooth seepage face and sufficient turbulent swash, the ML method still struggles with new, unseen data. However, it shows promise for a broader application and serves as a viable proof of concept. Together, these methods reduce the need for manual processing and enhance real-time coastal monitoring, contributing to more accurate predictive modeling of runup events and a better understanding of nearshore processes.

1. Introduction

Wave runup — the instantaneous position of the water's edge on the beach face — is a key factor in coastal flooding, especially during high tide and storm surges (Tomás et al., 2016; Gomes da Silva et al., 2016), and plays a critical role in sediment transport (Coco et al., 2014). Accurate runup observations are essential for deepening our understanding of nearshore processes (Gourlay, 1992; Raubenheimer et al., 1999; Bertin et al., 2018), and for coastal management, providing data to validate predictive models of inundation frequencies and erosion rates, which are vital for assessing the vulnerability of coastal ecosystems (Ware et al., 2019, 2021) and infrastructure (Najafi et al., 2021). Measuring runup, however, is challenging due to its highly

dynamic nature, particularly during storm events. Video imagery offers a solution (e.g., Bergsma et al., 2019), enabling remote monitoring of the coastline and thus emerging as the preferred method for runup observation (Mendes et al., 2022).

Numerous image preprocessing techniques have been shown to enhance nearshore features from timestack images—generated by stacking pixel arrays from sequential video frames over time, capturing temporal changes in wave runup. Manual selection is one of the oldest methods for extracting runup from video images and, despite its laborious nature, remains widely used (e.g., Holman et al., 1993; Huisman et al., 2011; Power et al., 2011; Atkinson et al., 2017; Yang et al., 2022). This continued reliance on manual methods underscores
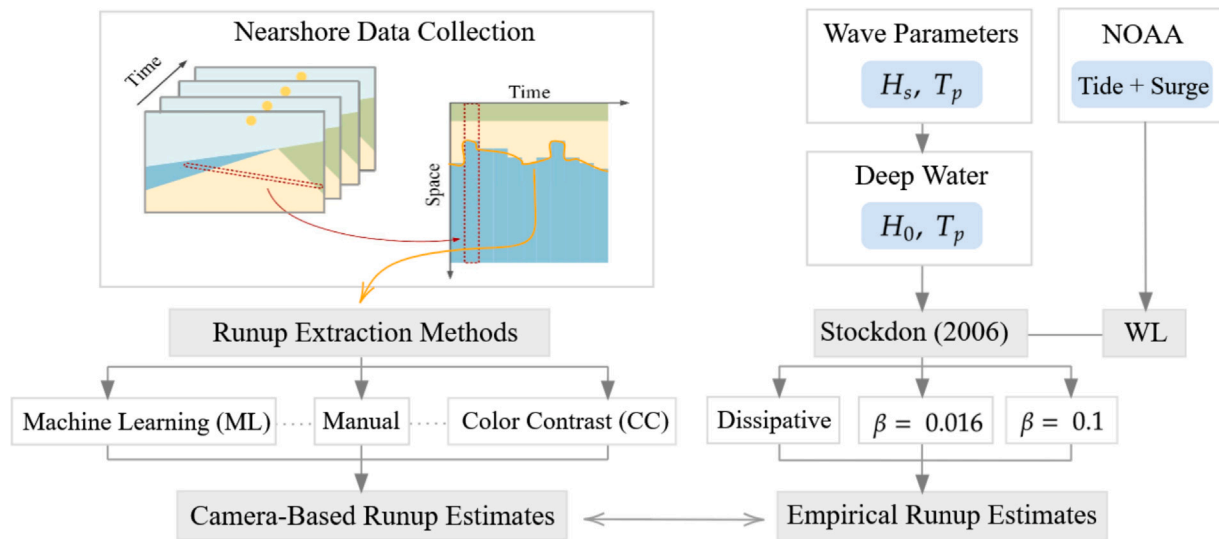

Fig. 1. Schematic overview of the development of runup analysis methods, with on the left the workflow of the camera-based runup estimates and on the right the empirical-based equivalent.

some limitations of automated approaches, despite recent advances in efficiency and accuracy.

Automated color contrast (CC) methods for runup extraction have been developed over the past few decades. Bailey and Shand (1994) proposed a grayscale image-based method that employed smoothing, edge detection, contrast normalization, and path-finding techniques. Though effective for relatively pristine data, the approach often smoothed smaller runup events and produced saw-tooth artifacts due to cut-off backwash. Color channel difference methods, such as in Turner et al. (2001), exploit color distinctions, computing red-blue channel differences to capitalize on the relative red dominance in sand and blue in water. Similarly, Simarro et al. (2015) introduced a pixel variance-based method, effective in many conditions but inconsistent 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 time-averaged images. This method translates the red-green-blue (RGB) color space to the hue-saturation-value (HSV) color space and segregates pixels based on the two-dimensional histograms of color and luminance. In another approach, Zhang and Zhang (2009) developed roughness metrics based on local entropy in the HSV space. However, its performance under dissipative conditions has yet to be fully evaluated.

Despite all these attempts, complications often arise from CC methods when applied on dissipative beaches. If the effluent line, marking the boundary where water seeps back into the sand, is decoupled from the waterline — which is common under dissipative conditions — a thin layer of water remains on the sand, complicating the distinction between water and sand. Furthermore, under dissipative conditions, waves lose most of their energy before reaching the shoreline, resulting in subtle gradients that are harder to distinguish from background noise, especially for the backwash (Huisman et al., 2011). Almar et al. (2017) successfully used a Radon transformation method, validating it on a dissipative beach with high-energy conditions. This method separates the uprush and backwash components of runup, but it depends on sufficient wave energy for effective edge detection.

Recent research leverages Convolutional Neural Networks (CNNs) for automated recognition of water, sand, and wrack pixels (Kang et al., 2024a), with potential applications in stochastic flooding predictions (Kang et al., 2024b). However, the lack of georeferencing currently limits precise measurement applications. 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. Although they achieved very accurate results, the use of LiDAR data on temporary field campaigns remains a limiting factor due to the high cost and technical challenges during deployment, as well as the high computational demand involved with postprocessing (Kuschnerus et al., 2024). The use of multi-channel input has not been implemented yet for regular video imagery data. This raises the question of whether image preprocessing techniques could serve as effective input channels for a CNN algorithm to extract wave runup from video imagery.

Runup extraction from video imagery is a challenging task, particularly under highly dissipative conditions where the instantaneous waterline is difficult to distinguish due to subtle gradients and low image contrast. Although advancements in video-based shoreline mapping have been made, as mentioned above, limitations are often observed due to sensitivity to environmental variability or a lack of robustness in dissipative settings. In this research, these gaps are addressed through the development and testing of automated frameworks tailored for such conditions. Video imagery collected from Galveston Island, Texas, along the Gulf of Mexico, was used to extract timestack images. Two novel methods were implemented: a CC method that leverages local entropy and the saturation channel from the HSV color space, and a multi-channel ML model incorporating preprocessed timestack images as input with manually labeled data for training. The methods were evaluated by comparing their runup metrics to manual delineations, and the measured runup statistics were compared with empirical estimates to provide insights into their accuracy and applicability under challenging dissipative conditions. Fig. 1 provides an overview of the research workflow.

The paper is organized as follows: Section 2 describes the study area, details the data collection procedure, outlines the CC and ML methods, and describes the used runup metrics, while Sections 3 and 4 present and discuss the results. Finally, Section 5 summarizes the study's conclusions.

2. Methods

2.1. Site characterization

Galveston Island, located on the Gulf Coast of Texas, USA, is a barrier island with sandy beaches situated approximately 50 miles southeast of Houston (Fig. 2). The region experiences mixed semidiurnal tides, with astronomical tidal ranges typically between 0.3 and 0.9 m in the micro-tidal Gulf of Mexico (NOAA, 2024b). Local water levels
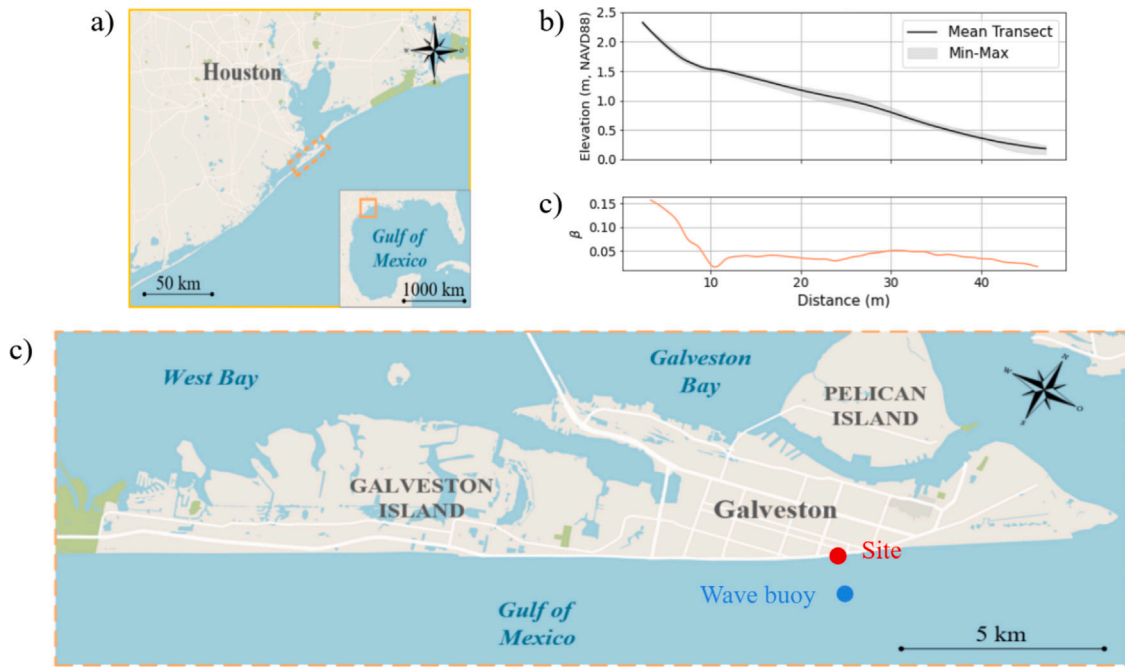
Fig. 2. In (a), the Gulf of Mexico and the coastal region near Houston, Texas; in (b), the mean elevation profile at the test location; in (c), the corresponding beach slopes; and in (d), a map of Galveston Island, with the campaign site indicated by a red dot and the Spotter wave buoy marked with a blue dot.

can vary significantly due to atmospheric pressure changes, wave-induced setup, and wind tides, often exceeding the NOAA predictions of the astronomical tide by a factor of two (Lefevre et al., 2000; Huff et al., 2020).

Seasonal variation influences wave and current direction, with southeast-south waves dominant from November to February and south-biased waves from June to August (NOAA, 2024a). During normal conditions, 95% of waves are under 1.3 m, and 99% are under 2 m (Surf-Forecast, 2024). However, Galveston is highly impacted by tropical storms and hurricanes during the June–November hurricane season, with offshore waves exceeding 10 m observed during these events (Nayak and Panchang, 2015).

The Galveston barrier island formed approximately 5500 years ago and evolved into its current dissipative beach profile due to sediment redistribution driven by relative sea-level rise and storm activity, as described by Bruun (1962) and Swift (1968). The coastline has been further shaped by both natural processes and anthropogenic influences, resulting in sandy dissipative beaches characterized by typical beach-face slopes of approximately 1:20 and dune-to-toe slopes of around 1:10 (Fig. 2c). Its poorly graded beaches, with a D_{50} ranging between 0.07–0.1 mm (US Amry Corps of Engineering, 2022), are characteristic of the region.

In 1900, a Category 4 hurricane devastated the island, causing 10,000 fatalities and destroying over one-third of all buildings (National Geographic, 2024). This tragedy prompted the construction of a 15 km seawall, extending from the northeast jetty to the southwestern edge of the city. The field campaign was conducted on the beach fronting the seawall, which protects against storm surges. While the seawall protects the city from storms, it disrupts longshore sediment transport, exacerbating erosion southwest of the structure at rates of 2 m/yr (Paine et al., 2020). Periodic beach nourishment along the seawall has mitigated this erosion, with the most recent effort in 2019 using sediment from the shipping channel and with more than 38% of fines (Maglio et al., 2020; US Amry Corps of Engineering, 2022).

2.2. Data collection

2.2.1. Camera deployment

Video imagery of the instantaneous water line was collected during the TURTLE field experiments on Galveston Island in November 2023.

Here, the part of the field campaign relevant to this study is summarized. A detailed description of the entire data collection campaign is provided in Christiaanse et al. (2025). The campaign included collecting offshore wave data with a *SOFAR Spotter* wave buoy (Sofar Ocean Technologies Inc., 2023) located about 1.3 km from the shoreline, extracting wave runup data using video imagery from a GoPro camera, and real-time kinematic (RTK) beach elevation profiles with a Leica GS08 GPS.

During the data collection period, the tidal range was approximately 0.8 m, with the local 10 min average water level (surge + tide) varying between 0.113 and 0.92 m NAVD88. Offshore wave heights (H_0) ranged from 0.82 to 1.76 m, with a mean of 1.11 m, while peak wave periods (T_p) varied between 7.3 and 8.52 s, averaging 8.1 s.

The GoPro was strategically positioned to capture a comprehensive view of the transect, spanning from the distant horizon to the base of the dune (Fig. 3). It was securely enclosed within a protective box and mounted to a pole for stability and height. Through the Quick app, the real-time stream of the camera's view was accessed, and the field of view was confirmed to meet all requirements.

Six ground control points (GCPs) were deployed within the camera's field of view. These were strategically positioned above the waterline, ensuring visibility. Their arrangement was designed to cover the area of interest and prevent collinearity as much as possible. The precise locations of the camera and the six GCPs were determined using the Leica GS08 GPS RTK device with an accuracy of 3 cm.

2.2.2. Camera settings

A GoPro Hero 10 Black camera (23MP CMOS camera sensor, 16 mm f2.8 lens; GoPro Inc., 2021) was chosen for the measurement campaign due to 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.

An important consideration in the camera selection process involved balancing temporal resolution and memory usage. GoPro cameras have a relatively high bit rate, which causes the memory to fill up quickly. The GoPro Hero 10 Black supports a maximum SD card storage of
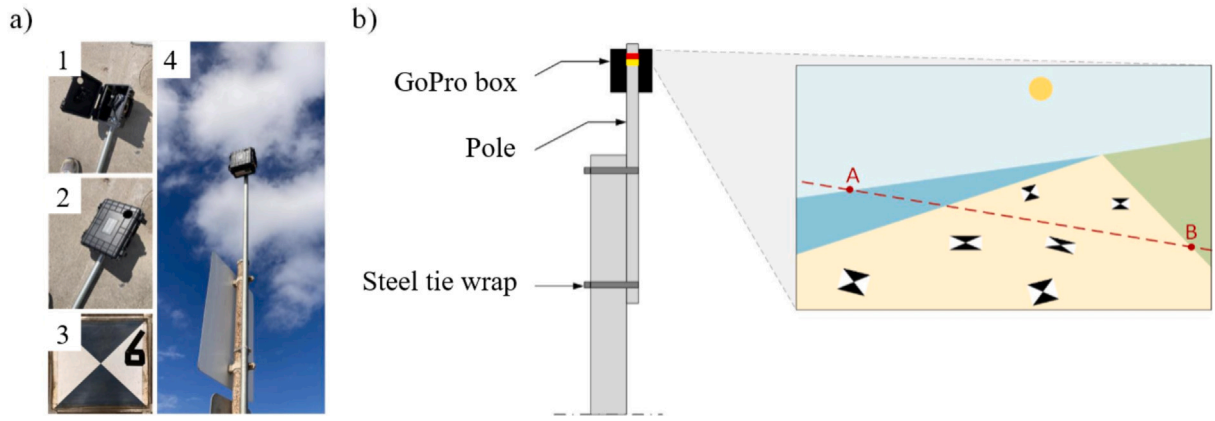
Fig. 3. Visualization of GoPro installation (Figure 5 in Christiaanse et al., 2025). In (a), pictures of the GoPro box (a.1,a.2), a GCP (a.3), and the installed system (a.4). In (b), a schematic of the GoPro installation with the transect spanning from the horizon (A) to the dune toe (B).

128 GB. A frame rate of 2 frames per second (fps) was deemed sufficient to capture runup smoothly and was chosen through the time-lapse video option. The videos were recorded in 4 K to maximize the spatial resolution. With the above settings, one filming day (approximately 7 to 9 h) translated to a data storage requirement of about 30 GB. Acknowledging the limitation of the GoPro's internal battery life, lasting approximately 1 h and 30 min at 2 fps and 4 K resolution, an external battery was integrated into the setup, enabling daily monitoring without interruptions.

2.2.3. Camera calibration

Intrinsic camera calibration was performed individually for each camera and filming mode to tailor the calibration parameters to specific settings. The calibration procedure involved capturing a short video of a flat checkerboard pattern with known dimensions, observed under varying angles and distances. A selection of sharp frames, showing the complete checkerboard pattern, was extracted from the video, and the *OpenCV* for Python toolbox was 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 — were corrected using the *OpenCV* toolbox. This resulted in a camera-specific calibration matrix encompassing focal lengths and optical center location with a corresponding Root Mean Square (RMS) reprojection error, which is required to remain sub-pixel. The obtained calibration matrix was then applied to undistort all videos filmed with the same camera and identical setting combination.

Extrinsic camera calibration involves the georeferencing of two-dimensional undistorted images to real-world coordinates. The calibration process was 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 theoretically with three GCPs with known real-world and pixel coordinates, under the assumption that the camera and GCP positions are free of measurement noise and that the GCPs are not placed collinearly. However, in practice, additional GCPs are needed to improve the accuracy and reliability of the solution, as real-world conditions often introduce measurement errors and GCP placements that may approach collinearity. Six 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, resulting in an equation linking each pixel (u, v) to real-world coordinates (x, y, z).

2.2.4. Manual digitization

Once the photogrammetric equation was solved, an averaged xy cross-shore transect was computed over each deployment and converted into pixel coordinates. The corresponding RGB arrays from
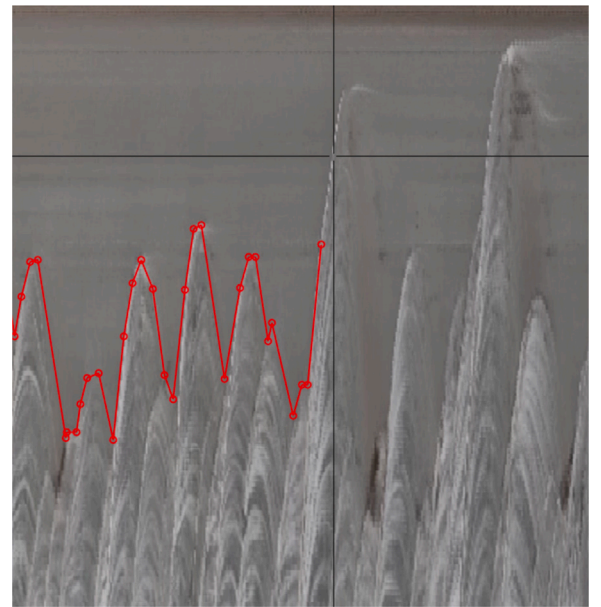

Fig. 4. Snapshot of the manual digitization process.

Table 1
Data compendium with days, video IDs, start and end times.

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

each frame were then extracted and stacked on top of one another to generate a timestack image, which provided a spatial and temporal representation of the transect.

The instantaneous position of the water line was manually digitized on nine timestack images of 80 min, except one that was 28 min and 25 s long. A moving window was used to move along the timestack in time and select the contours of individual swash extends produced by the waves (Fig. 4). The swash motion includes a decelerating uprush phase and an accelerating downrush phase (i.e., backwash). The
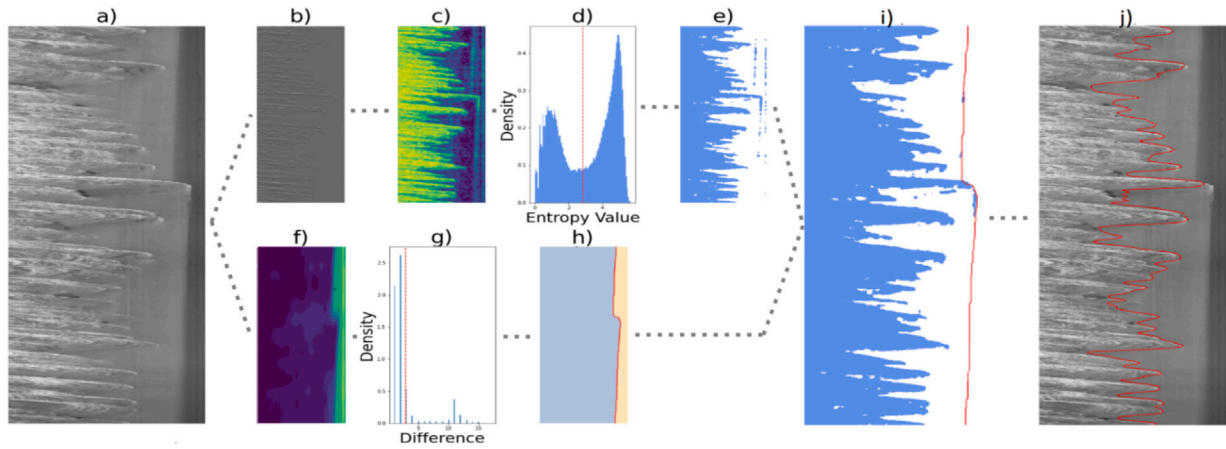
Fig. 5. Diagram of the runup extraction process utilizing local entropy and saturation channel operations. In (a) the grayscale image with pixel intensity I (b) dI/dt, (c) the local entropy, (d) the corresponding histogram with Otsu's threshold (Otsu, 1979), (e) the binarized image, (f) the saturation image, and in (g) the corresponding histogram with Otsu's threshold, (h) the binarized saturation image with the effluent line in red, (i) the binarized entropy image cleaned from noise above the effluent line, and (j) the grayscale image with the computed instantaneous water line in red.

wave uprush was easily detected by a sharp white/gray contrast. The backwash, however, was less obvious to the eye. The reason for this complication related to backwash detection originates from the dissipative character of the beach. When waves reach the shoreline, they have lost most of their energy already. As part of the swash motion, water infiltrates into the beach matrix during both the uprush and the backwash. The accelerating very shallow flow during backwash over an extremely gentle slope can result in an unclear delineation of the water and sand boundary. During manual digitization, the backwash was estimated by following the slope of the upper backwash where it was visible. In cases where the upper backwash was indistinct, an inverse slope to that of the uprush was applied to delineate the end of the wave before the next runup event.

While only one person was involved in the manual digitization of the runup, previous research by Buckley et al. (2024) showed that the RMSE of the $R_{2\%}$ was small, ranging from 0.04 m to 0.1 m relative to the amplitude of the swash. This suggests that, despite the use of a single digitizer, the approach produces reasonably accurate results. The manually digitized timestack images are summarized in Table 1, with video IDs corresponding to the published raw data by Christiaanse et al. (2025).

2.3. Runup extraction methods

2.3.1. CC method

The CC method combines the concepts of entropy and saturation into a straightforward model based on channel operations, enabling efficient runup extraction on dissipative beaches (Fig. 5). The process begins with binarizing the image based on local entropy values, followed by noise reduction achieved by removing all values above the effluent line, which corresponds to where groundwater intersects the beach surface, as determined through the binarization of the saturation channel of the image.

Local entropy was specifically employed to quantify the local complexity within a timestack image. The dynamics of the swash zone are characterized by turbulent movements with a large range of gradients in both time and space and a smoother appearance (high complexity), which can be distinguished from the less dynamic and more smooth sandy areas (low complexity). To quantify this complexity, the *entropy* function from the *skimage.filters.rank* module of the *Scikit-Image* library was applied to the time gradient of the grayscale image, dI/dt (Fig. 5b). This function computes the local entropy of an image, effectively highlighting areas of high complexity, such as those found in the swash zone.

A local neighborhood was defined with the *disk* function from the *skimage.morphology* module with a radius of 5 pixels. Then, the local normalized histogram was computed internally for each pixel according to:

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

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 was computed as:

$$E = -\sum_{i=0}^{n-1} p_i \log_2(p_i) \tag{2}$$

This resulted in an image where each pixel represents the complexity of the surrounding area, making it easy to distinguish between the swash and the beach (Fig. 5c). This was confirmed by the corresponding histogram, showing strong bimodality (Fig. 5d). After binarization by Otsu's method (Otsu, 1979), some noise persisted (Fig. 5e). Notably, certain areas above the effluent line had high entropy. These correspond to regions with a more textured sand surface (not smoothed by wave runup).

Under reflective conditions, the saturation or Red-Blue channel has been shown to segregate the swash from the sand quite well, as sand pixels contain more red than water pixels, while water pixels contain more blue (Turner et al., 2001; Aarninkhof and Ruessink, 2004; Uunk et al., 2010). However, on dissipative beaches, these methods do not delineate the swash as well, because the presence of a seepage face can disturb the signal (Huisman et al., 2011). This happens when the instantaneous water line is decoupled from the effluent line, resulting 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 saturation or the Red-Blue channel does not enhance the swash but delineates the effluent line instead. However, what initially appears to be a disadvantage can be turned into an advantage.

To handle the remaining noise in the entropy image, the saturation channel was obtained from the HSV color space with the *OpenCV* library. Subsequent application of Gaussian blur over time with a 201×1 kernel allowed for a clearer distinction between the regions above and below the effluent line (Fig. 5f). The strong bimodality of the corresponding histogram (Fig. 5g) further confirmed this distinction and allowed for binarization of the image using Otsu's method (1979). The results clearly distinguished between the wet and the dry region

(Fig. 5h) and were used to remove the noise above the effluent line from the entropy image (Fig. 5i). Finally, the summation of the water pixels resulted in the runup estimates of the entropy-saturation method, delineated in red in Fig. 5j.

All runup delineations with the CC method were initially performed on full 90 min timestack images, with binarization thresholds computed based on characteristics over the entire time frame. A sensitivity analysis was conducted on the computational window size to better capture local variations. Time windows ranging from 5 grid cells to 9000 grid cells were tested, and a window size of 500 grid cells (approximately 4 min) was selected for its optimal balance between capturing local characteristics and maintaining stable results.

Another approach to address local temporal variations involved column-wise normalization of the processed images. However, testing different configurations, including normalizing saturation, entropy, or both, revealed no clear improvement. Normalizing entropy or both channels simultaneously slightly decreased model performance, while normalizing only the saturation channel had no overall effect on the model performance. Although it improved handling of high-luminescence areas, it also reduced accuracy in regions where the seepage face was noisier. Given the lack of a consistent positive impact, column-wise normalization was not retained in the final method.

2.3.2. ML method

A second method was implemented to extract the runup position from timestack images using a relatively simple yet effective Convolutional Neural Network (CNN), prioritizing computational efficiency while ensuring sufficient spatial and temporal resolution to capture wave runup dynamics.

The CNN algorithm was implemented using the *keras.models* module from the *tensorflow* library, with the spatial–temporal input dimensionality set to small patches of 60×60 pixels across five channels: the grayscale image, I, the intensity over time, dI/dx, the saturation channel, S, the entropy image, E, and the entropy over time dE/dt, resulting in an input shape of $60 \times 60 \times 5$. This design balanced computational efficiency with adequate spatial and temporal resolution, effectively capturing the physical processes of wave runup. With a spatial resolution of 7 cm per pixel and a temporal resolution of 0.5 s, each patch covered 4.2 m and spanned over 30 s.

The network architecture is shown in Fig. 6. It is composed of a contracting path (on the left) and an expansive path (on the right). The contracting path adheres to the standard structure of a convolutional network, beginning with two 3×3 convolution layers with 32 and 64 filters, respectively. Each convolution is followed by a ReLU activation function and a 2×2 max pooling operation for downsampling. The expansive path features a 3×3 up-convolution with a stride of 2, which upsamples the feature map, followed by another 3×3 convolution with stride 2 that reduces the feature map size from 64 to 32. The final layer is a 1×1 convolution with a sigmoid activation to produce a binary segmentation of the image.

The model was focused 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 likely contained more inaccuracies than the broader classification of areas as either water or sand. Segregating water from sand presented a less error-prone task, as the extensive areas covered by each category provided more reliable and consistent data than the precise delineation of the water line, resulting in less class imbalance.

Initially, seven input channels were considered: the grayscale image, I, the intensity changes over time, dI/dt, the intensity changes over space, dI/dx, local entropy, E, the entropy changes over time, dE/dt, the original RGB image, and the saturation image, S.

During the first steps of the model development, k-fold cross-validation was applied to an 80 min timestack image for efficient computation (Fig. 7). The data was split into five equal parts (folds), with four used for training and one for testing. This process was
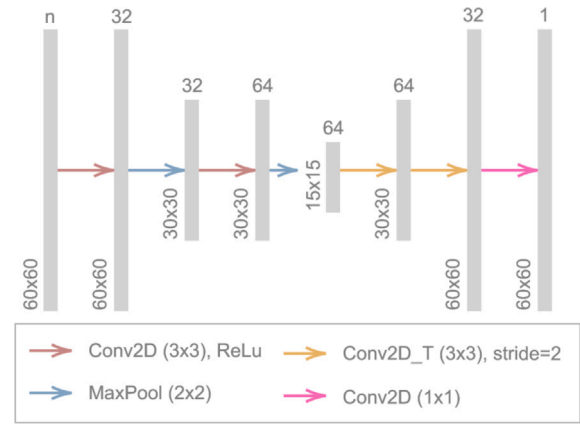

Fig. 6. Network architecture of ML model.

repeated for all folds, and the average performance guided model selection, assessed by accuracy and precision. Accuracy was defined as the percentage of correctly predicted water and sand pixels relative to the manually labeled timestack and is expressed as:

$$\text{Accuracy} = \frac{\text{Number of correctly predicted pixels}}{\text{Total number of pixels}} \times 100\% \qquad (3)$$

and the standard deviation of the accuracy in between folds:

$$\text{Precision} = \sigma_{acc} = \sqrt{\frac{\sum_{i=1}^{k}(x_i - \mu_{acc})^2}{k-1}} \qquad (4)$$

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

First, the optimal dimensions for Gaussian smoothing kernels on the grayscale image were assessed to improve noise reduction while preserving relevant features. The smoothing process involved applying weighted averages to pixels based on a Gaussian function, with closer pixels having more influence. Using *OpenCV*'s *GaussianBlur* function, two kernel types were evaluated: a horizontal kernel (n ×1) and a square kernel (n ×n), with different values of n tested via k-fold cross-validation. The best-performing kernels according to accuracy and precision were retained: a 3-pixel horizontal kernel (time smoothing) and a 3×3 square kernel (time and space smoothing).

Second, for every considered input channel the horizontal and squared smoothing kernels were evaluated to the no-smoothing baseline. A pairwise comparison evaluated the effect of smoothing on model accuracy and variance of every single-channel ML model, using a t-test for mean accuracy and an f-test for variance. A significance threshold of $\alpha = 0.05$ was applied. If the square kernel did not reduce accuracy or increase variance, it was preferred because it maximized smoothing while preserving key features. If the square kernel led to accuracy loss or increased variance, the horizontal kernel was chosen as an alternative. If neither smoothing method provided a performance advantage, the no-smoothing baseline was used. The grayscale (I) and RGB images showed no improvement with smoothing, so the no-smoothing baseline was retained. For the entropy (E), dE/dt, dI/dx, and saturation (S) images, the square kernel maintained accuracy and precision, making it the preferred choice. However, dI/dt experienced a precision drop with the square kernel but performed well with a horizontal time-axis kernel, avoiding excessive loss of details.

Third, all 127 combinations of the seven input channels were evaluated. Beforehand, the input channels were normalized to ensure that all inputs contributed equally to the learning process, preventing features with larger scales from dominating. On a laptop equipped with an Intel Core i7 processor and 16 GB of RAM, this took approximately 10 h to complete. Finally, the channel combination with the highest accuracy was retained: a model informed by I, dI/dx, S, E, and dE/dt.
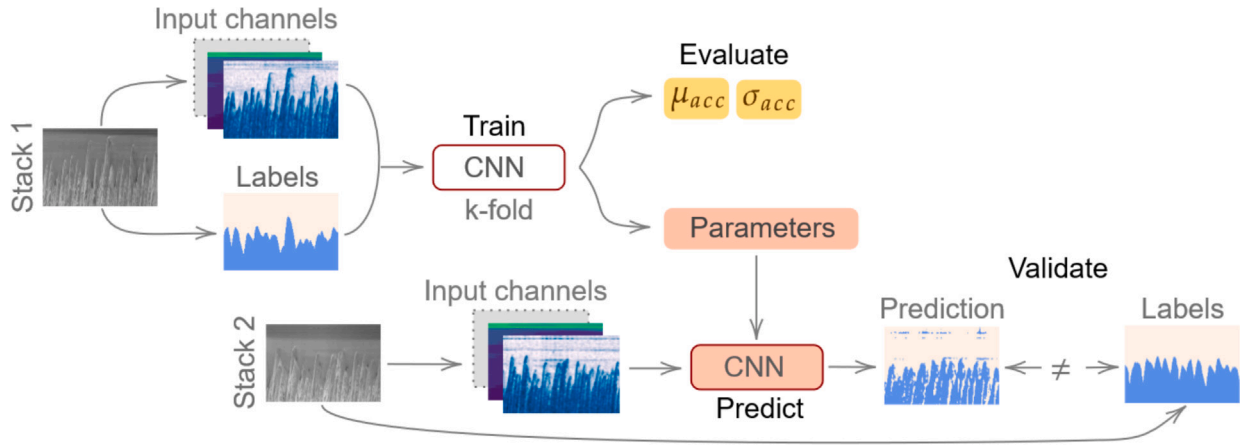
Fig. 7. Optimization and evaluation scheme of ML model for runup extraction.

Fourth, the training set was expanded to include a broader range of conditions. During the first optimization steps, the ML algorithm was developed by k-fold cross-validation on a single timestack (video ID: GX060084), to accelerate the optimization process. This model (ML_1) was validated on a new timestack (video ID: GX050084) and the results were compared to a model trained on all remaining timestack images (ML_8). To assess the performance μ_{acc} and σ_{acc} metrics of ML_8 a 9-fold cross-validation was performed using all labeled timestacks. Additionally, the added value of a multi-channel informed model was assessed by comparing it to a single-channel model informed solely by a grayscale image and trained on eight timestacks ($ML_{8;I}$).

2.4. Evaluation of runup measurements

2.4.1. Comparison with manual results

Once the instantaneous shoreline was delineated from the timestack images, the pixel coordinates were converted to real-world coordinates with the solved photogrammetric equation (Section 2.2.3). An elevation time series $R(t)$ was obtained from which runup statistics ($R_{2\%}$ and $R_{50\%}$) were computed.

The mean runup period $T_{m-1,0}$ of the runup delineation methods is assessed through spectral decomposition of the runup time series. The spectra are computed over a duration D of 30 min with blocks of length D/16 which satisfies a trade-off between spectral resolution and reliability. Then, the mean period $T_{m-1,0}$ is computed as m_{-1}/m_0. With m_0 and m_{-1}, the zeroth and the -1st spectral moments. The mean runup period $T_{m-1,0}$ is often used in wave runup as it gives relatively more weight to the lower frequencies, which dominate the runup processes, especially on gently sloping beaches. The runup metrics obtained using both the CC and ML methods were evaluated against the manual results by calculating the root-mean-square error (RMSE) and the correlation coefficient (r).

2.4.2. Empirical estimates

The ultimate goal is to predict wave runup, so the measured runup values were compared to the most common empirical estimates of $R_{2\%}$: the general expression from Stockdon et al. (2006):

$$R_{2\%} = 1.1 \left(\left(0.35\beta_f (H_0 L_0)^{1/2}\right)^2 + \left[\frac{H_0 L_0 \left(0.563\beta_f^2 + 0.004\right)}{2} \right]^{1/2} \right)^{1/2} \quad (5)$$

and the reduced formula for extremely dissipative conditions (applicable when the surf similarity parameter $\xi < 0.3$):

$$R_{2\%} = 0.043(H_0 L_0)^{1/2} \quad (6)$$

where β_f is the foreshore slope or beach slope, H_0 is the deep-water significant wave height, and L_0 is the deep-water wave length.

Table 2
Comparison of multi- and single-channel ML algorithms.

	ML model	# Training	$r(-)$	$\mu_{acc}(\%)$	$\sigma_{acc}(\%)$
ML_1	Multi-channel	1	0.70	96.09	0.42
ML_8	Multi-channel	8	0.77	93.03	1.40
$ML_{8,I}$	Single-channel	8	0.66	90.92	2.23

The deep-water wave conditions, H_0 and T_p, were derived from the wave data recorded by the Spotter buoy, approximately 1.3 km offshore from the site at a depth of ± 8 m during low tide. The significant wave height at that location was translated back to the deep-water significant wave height H_0 iteratively through the dispersion relation. Refraction was disregarded under the assumption of alongshore uniformity and normal incident waves. Energy losses due to white capping or wave breaking were also neglected.

Mean water level data (tide + wind surge) were obtained from a nearby NOAA station at the North Jetty of the Galveston Bay entrance (Station ID: 8771341, Location: 29.357°N, 94.725°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). Three empirical runup estimates were evaluated. Two were based on the general Stockdon formula, using beach slopes of 0.016, the smallest measured slope at the dune toe (Fig. 2b) representing the most dissipative conditions, and 0.1, a standard value for more reflective conditions, to assess the effect of the input slope on the results. Additionally, the reduced formula for extremely dissipative conditions was applied, which does not require a beach slope input.

3. Results

3.1. ML method optimization

When considering the ML method itself, it appears that the multi-channel ML approach significantly outperformed the single-channel model according to the correlation coefficient, μ_{acc}, and σ_{acc} metrics (see Table 2).

Expanding the training dataset in multi-channel models improved model performance, as shown by an increase in the correlation coefficient with the manual results from 0.70 to 0.77. However, μ_{acc} decreased, and σ_{acc} increased when the training set was 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
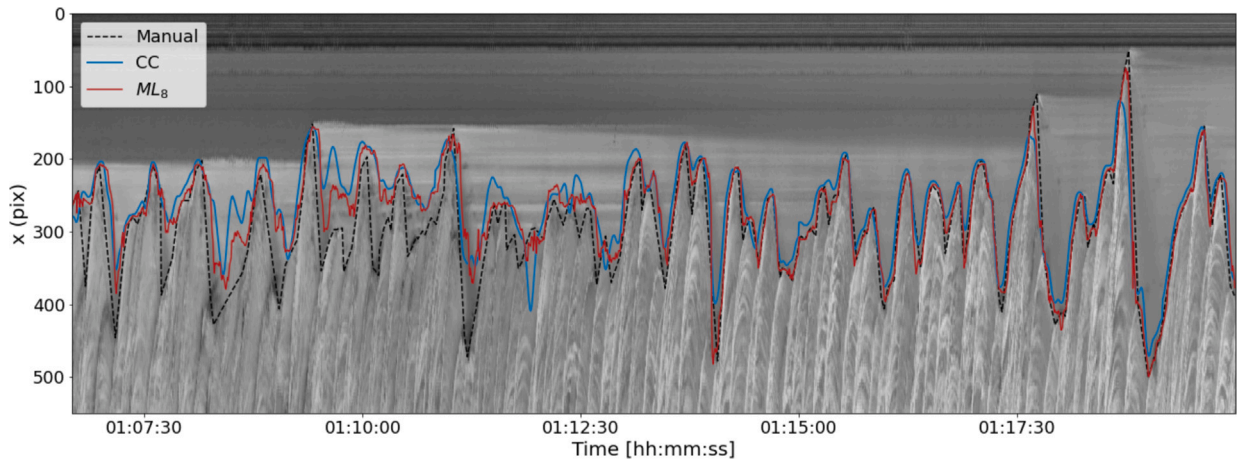
Fig. 8. Section timestack from November 14th (video ID: GX050084), with the manually digitized water line (back dashed), and the CC (blue solid) and ML results (red solid).

Table 3
Correlation and RMSE of $R(t)$, $R_{2\%}$, $R_{50\%}$, and $T_{m-1,0}$ from the ML and CC methods compared to manually digitized for all measurements on November 13th and 14th.

	$R(t)$		$R_{2\%}$		$R_{50\%}$		$T_{m-1,0}$	
	r (–)	RMSE (m)	r (–)	RMSE (m)	r (–)	RMSE (m)	r (–)	RMSE (s)
ML	0.96	0.10	0.97	0.09	0.99	0.05	0.79	3.51
CC	0.94	0.12	0.97	0.08	0.99	0.05	0.70	3.88

consistent performance. However, when zooming in on the individual folds of the multi-channel model trained on the extended set, the accuracy was lower when the validation timestacks were from November 13th, with fewer timestacks. Specifically, the average accuracy was 91.21% with a standard deviation of 0.78%. In contrast, November 14th had more timestacks, an average accuracy of 93.94%, and a standard deviation of 0.41%. Possible causes for this discrepancy might be that the validation timestacks from November 13th had higher data variability, or that the model did not have enough similar samples from that day compared to November 14th. This suggests that the model was better tested for under- and over-fitting when cross-validated on a larger and more diverse set of timestacks, capturing a wider range of conditions, hence improving generalization.

3.2. Comparison with manual delineation

A section of the timestack from November 14th was selected to illustrate the results of both the CC and the ML method, alongside the manually digitized runup (Fig. 8). Both methods demonstrate very strong correlations with the manual results. For the full runup time series $R(t)$, the correlation coefficients are 0.96 for the ML method and 0.94 for the CC method, indicating an excellent linear relationship between the predicted and manual values. The corresponding RMSE values are approximately 10 cm and 12 cm for the ML and CC methods, respectively, representing an error of around 16% relative to the maximum runup amplitudes, which is considered reasonable (Table 3).

However, the performance of both methods fluctuates over time. In the left section of the image, where the swash appears less distinctly white and horizontal seepage lines are visible, both methods demonstrate difficulties 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.

The estimated runup percentiles $R_{2\%}$ and $R_{50\%}$ show good agreement with the manual method, which is confirmed by high correlation

coefficients of 0.97 and 0.99, respectively, indicating a solid linear relationship between the ML and CC runup statistics and manual values. 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 3). This low error metric further validates the accuracy of the ML and CC methods in estimating runup heights, enhancing their reliability for practical applications in coastal management. This is especially true for $R_{2\%}$, which is often used in extreme value analysis.

From Fig. 9a, it is clear that over the considered period, most energy is contained in the infragravity band of the spectra. On November 14th from 12:30, some more energy is found in the sea-swell part of the spectra located above the infragravity frequency band. In Fig. 9b and c, 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 method. 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). Overall, the methods tend to slightly underestimate the energy contained around the peak, except for the ML method on November 13th around 17:00, 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 are not significant. The methods do not present a significant difference in the sea-swell part of the spectrum. They all show reduced oscillation values in the sea-swell domain, which is typical as this component represents higher frequency waves that generally contain less energy in the nearshore due to breaking and dissipation.

From the plots, it appears that the Very Low Frequency (VLF) 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 VLF components fall out of the resolution range.

The model's performance is satisfactory when considering the mean wave period ($T_{m-1,0}$). The corresponding mean wave frequency ($f_{m-1,0}$) derived from both the ML and CC methods closely follows that obtained from the manual results, as shown in Fig. 9b and c. The ML model shows a correlation coefficient of 0.79, while the CC method has a slightly lower correlation coefficient of 0.70. Also, the differences remain tolerable as the magnitude of the mean wave period is not too far off. The RMSE for $T_{m-1,0}$ remains within 10% of the observed values
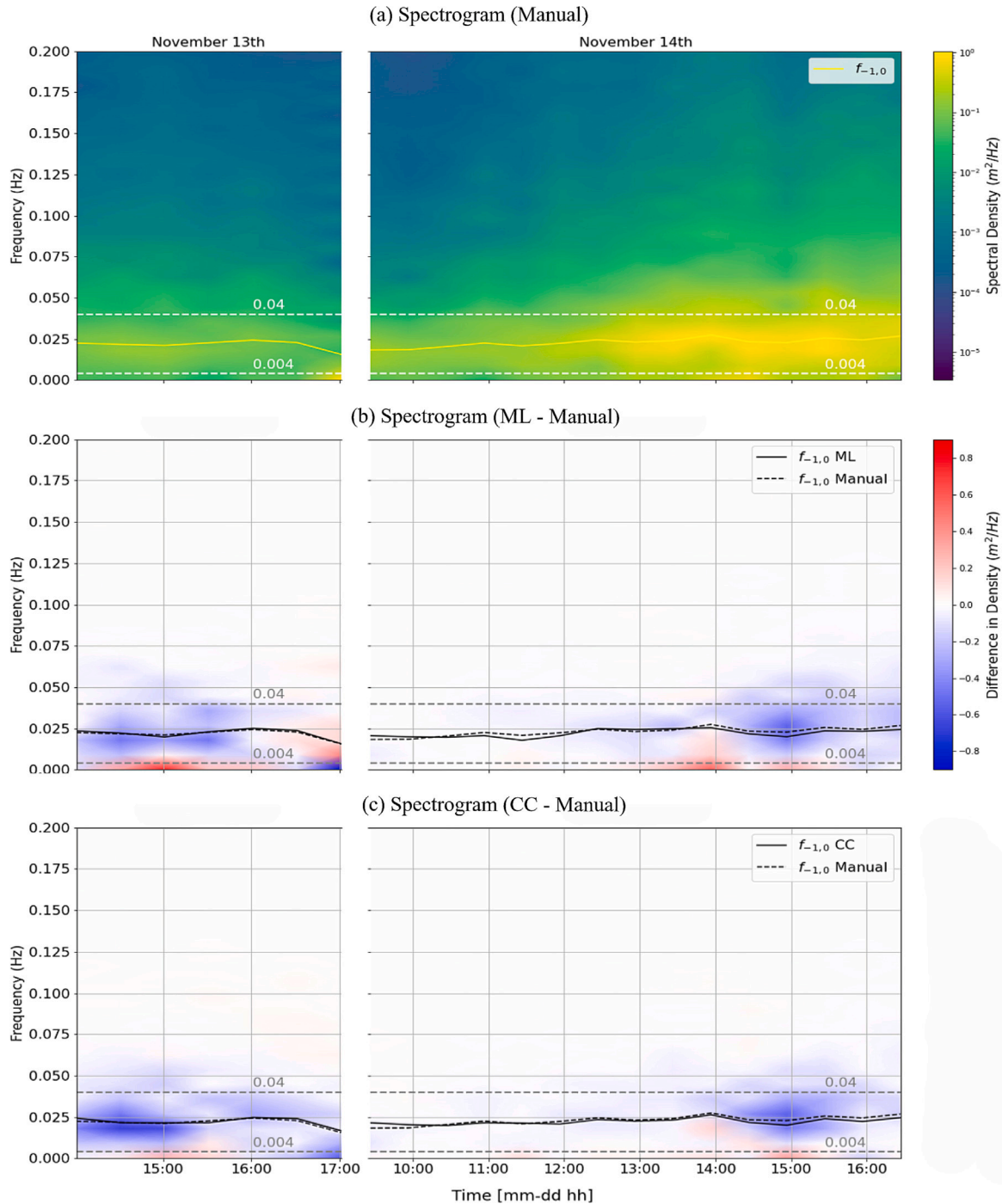
Fig. 9. Spectrogram of the extracted runup: manually digitized results (a), and the spectral difference between ML and Manual (b), and CC and Manual (c). The horizontal dashed lines indicate the limits of the infragravity band, while the solid lines represent the peak frequency ($f_{-1,0}$). From 2023-11-13 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.

($T_{m-1,0} \sim 40$ s), at 3.51 s for the ML model and 3.88 s for the CC method (Table 3). 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.

3.3. Runup measurements and empirical estimates

Since the ultimate objective is to predict wave runup, the runup observations from the manual, ML, and CC methods were evaluated

against empirical estimates of 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. The results are displayed in Fig. 10.

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 (tide + surge) from the NOAA station. However, there are notable
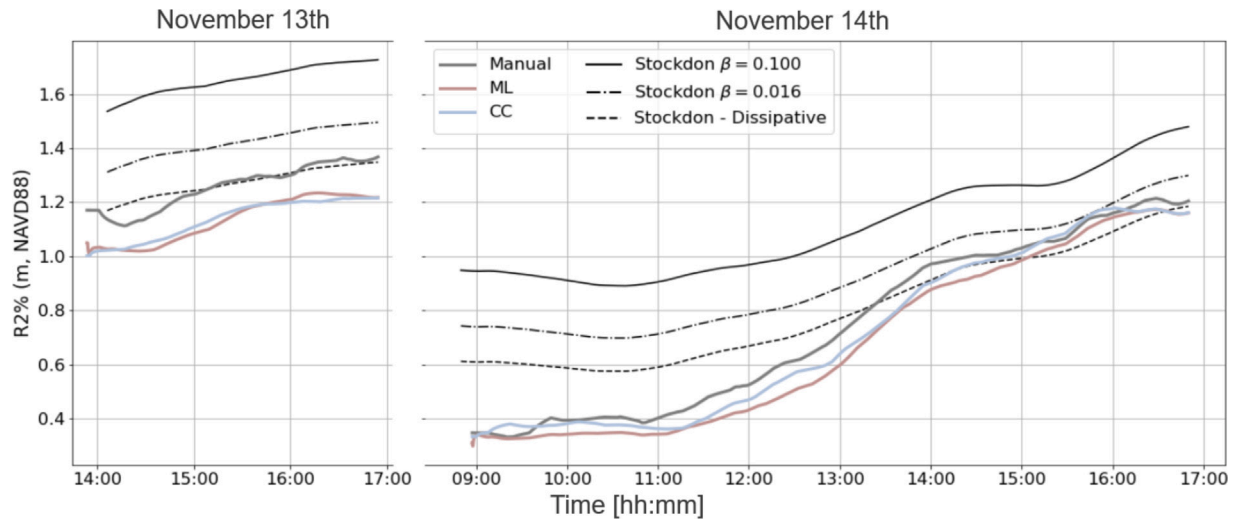
Fig. 10. Runup observations and estimates of $R_{2\%}$ on November 13th and November 14th, 2023. Observations include manual, ML, and CC methods. Estimates are based on the general Stockdon formula (with $\beta = 0.1$ and $\beta = 0.016$), and the Stockdon formula for extremely dissipative conditions. 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'.

differences in the $R_{2\%}$ values predicted by the different 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 methods, respectively.

The dissipative Stockdon formula 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 the observations during mid- and high tide more closely. This trend is also evident 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. This aligns 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.

4. Discussion

The results of this study highlight the potential of machine learning (ML) and color contrast (CC) methods to accurately delineate runup contours in timestack images derived from video imagery under the challenging conditions of dissipative environments. Both methods showed a strong correlation with manually digitized waterline data, as evidenced by the full runup time series $R(t)$, runup statistics $R_{2\%}$ and $R_{50\%}$, and mean runup period $T_{m-1,0}$. However, the findings also reveal critical challenges and opportunities for enhancing the accuracy and reliability of these methodologies.

Entropy, combined with saturation to reduce noise above the effluent line, proves effective for runup extraction. Previous studies demonstrate that spatial pattern analysis aids shoreline detection in satellite images (Fuse and Ohkura, 2018), and topographical LiDAR data identifies the swash zone through surface roughness variations (Wang et al., 2023). As a roughness metric, local entropy shows strong potential

for isolating the turbulent swash zone from sand, highlighting again the efficacy of roughness-based approaches for swash detection on dissipative beaches (Zhang and Zhang, 2009).

However, the CC 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 enhanced the model performance and is an essential metric. Column normalization of the saturation channel improved the CC method's performance in certain sections of the timestack images but diminished it in others. Further research is needed to assess the robustness of this approach.

As for the ML model, the primary challenge is related to the training set size and composition. Initial attempts to train the ML model on a limited dataset revealed that a single timestack lacked the variability needed to capture complex coastal conditions, such as high luminescence or irregular seepage faces. The implementation of k-fold cross-validation across nine labeled timestacks significantly improved correlation with the manual results. However, this increased data also introduced an imbalance, with a disproportionate number of timestacks from November 14th compared to November 13th, which may have led to overfitting, which led to a decrease in mean accuracy as the training set was enlarged. Future work should prioritize balanced training datasets, potentially through data augmentation techniques, to ensure robust model performance across diverse environmental conditions.

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. These errors remained infrequent and did not significantly impact broader runup statistics over extended periods, such as trends across tidal cycles or seasonal variations. For wave-by-wave analysis, future research should focus on extending the training to more extensive and diverse 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, called data augmentation, could also help in raising a flag when strange patterns are encountered.

Another critical aspect of model optimization is the establishment of reliable ground truth data. Although manually digitized water lines

serve as a useful baseline, they are inherently subjective and may not accurately represent the true water line, especially in dissipative beach environments where distinguishing backwash can be challenging. Integrating alternative measurement techniques, such as LiDAR or resistance wire data, could provide more accurate ground truth references, facilitating further improvements in the ML model.

Collins et al. (2023) developed a CNN model incorporating multi-channel LiDAR data, including elevation. While video imagery lacks direct height maps, solving the photogrammetric equation enables projection from pixel coordinates to 3D coordinates if accurate beach profile measurements are available. This elevation map can then serve as an additional channel, aiding in distinguishing sand from water along the cross-shore. If profile data are unavailable, a linear function could approximate the effect.

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).

Another possible improvement could be the addition of other covariates such as temperature, wind speed, offshore wave characteristics, atmospheric pressure, or humidity. While these variables may not be available at the same timescale as the runup events, they could be obtained from local measurement stations and incorporated as one-dimensional metrics in the learning process to improve the model's accuracy and robustness. This process, known as statistical downscaling, has been successfully implemented for medium-term shoreline prediction (Antolínez et al., 2019).

In this study, the best channel combination was chosen based on the highest accuracy. It might be beneficial to conduct a feature importance analysis to identify and prioritize the most significant channels (Verdinelli and Wasserman, 2023), so they can be weighted dynamically during the training (Turali et al., 2024). Also, the channels were combined linearly. Exploring non-linear combinations could enhance the model's ability to capture complex interactions in runup phenomena, potentially leading to more accurate predictions. Non-linear methods can better represent underlying dynamics, though they also increase model complexity and the risk of overfitting. A larger training set would be needed to offset these risks, and the associated computational costs should be considered.

The CNN architecture in this study was intentionally kept shallow, and with 60×60 input patches to accommodate computational constraints, limiting the model's ability to capture larger spatial patterns crucial for runup continuity. Using larger patches could improve generalization on unseen data, reduce edge effects, and benefit from advanced padding techniques to better manage border-related challenges (Hamwood et al., 2018). Combining larger patches with a deeper architecture could help uncover hidden patterns, enabling the model to more effectively identify complex spatial and temporal relationships. Future research could refine the CNN approach by experimenting with different pooling methods, activation functions, and optimizers to enhance feature extraction and convergence rates. For example, incorporating a mixed loss function that combines the initial classification of water and sand with regression of runup values could create a more targeted training strategy. Additionally, adding a third classification for the seepage face could improve predictive accuracy by helping the model better differentiate between water, sand, and seepage faces. Recent work by Kang et al. (2024a) classified coastal video images into three categories: 'water', 'sky', and 'background' (or beach), with the 'sky' category, included to reduce misclassified water pixels.

5. Conclusions

Accurate runup observations are essential for deepening our understanding of nearshore processes and supporting coastal management,

offering critical data to validate predictive models of inundation frequencies and erosion rates. The goal of this study was to develop a method for extraction of wave runup from video imagery under the challenging swash conditions of dissipative beaches. To this end, two methods were developed, based on color contrast (CC) and machine learning (ML). The CC method, using local entropy and saturation, showed strong agreement with manual results, achieving an RMSE of 0.12 m and r-value of 0.94 for the full time-series (R_t), 0.08 m RMSE and r-value of 0.97 for $R_{2\%}$, and 3.88 s RMSE and r-value of 0.70 for $T_{m-1,0}$, indicating its applicability to similar beaches along the northern Gulf of Mexico. However, it is important to recognize that when the seepage face is not smooth, high entropy values can lead to overestimated water lines and high-frequency noise, while insufficient turbulence in the swash zone may cause an underestimation of the water line.

This study also demonstrated the promising potential of image pre-processing techniques in enhancing a CNN for wave runup extraction from video timestack images. The developed ML method, informed by five input channels (I, dI/dx, S, E, and dE/dt) showed strong agreement with manual results at the test site, achieving an RMSE of 0.10 m and an r-value of 0.96 for R_t, 0.09 m RMSE and r-value of 0.96 for $R_{2\%}$, and an RMSE of 3.51 s with an r-value of 0.79 for $T_{m-1,0}$. When compared to a single-channel input model, the added value of multiple input channels was evident, as it led to an increased correlation with the manual results (from 0.66 to 0.77). These results establish the model as a viable proof of concept for runup delineation under the given local conditions, despite occasional inaccuracies when applied to new, unseen datasets. To improve the model robustness, future research should focus on expanding the training dataset, exploring data augmentation techniques, and utilizing more reliable ground truth data such as LiDAR. Further optimization of the model could involve experimenting with different input channels, pooling methods, activation functions, and optimizers, as well as increasing patch size and network depth to capture larger spatial patterns. Additionally, integrating a mixed loss function that combines regression and classification tasks and adding a third classification for the seepage face could enhance predictive accuracy. Finally, incorporating external covariates like environmental variables or conducting feature importance analysis could further improve the model's generalization and performance.

Additionally, the observed $R_{2\%}$ values show strong agreement with the formula proposed by Stockdon et al. (2006) for dissipative conditions, with RMSE values lower than 0.13 m and r-values exceeding 0.70 for manual, CC, and ML estimates. These results further support the reliability of our methods, which enable wave-by-wave runup extraction even in highly dissipative environments. This improved detection of short-term fluctuations enables a better understanding of swash zone dynamics, erosion, and local flooding events and contributes to more precise predictive modeling. By providing more accurate runup measurements, our approach strengthens coastal hazard assessments and supports better informed management strategies.

Based on the current results, the CC method is recommended as a ready-to-use approach, providing reliable and accurate results under dissipative beach conditions with a smooth seepage face and enough turbulence in the swash zone. In contrast, the ML method shows significant potential for broader applications, particularly as it is further refined and trained on larger datasets, which will enhance its robustness and reliability, extending its applicability beyond dissipative conditions.

CRediT authorship contribution statement

Meye J. van der Grinten: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jakob C. Christiaanse:** Writing – review & editing, Supervision, Resources,

Data availability

The GoPro footage is available at: https://data.4tu.nl/collections/93256801-ed01-4627-9e49-8607967a0853. Code is available through this GitHub repo: https://github.com/MeyevanderGrinten/coastal-run-up.

References

Aarninkhof, S.G., Ruessink, B.G., 2004. Video observations and model predictions of depth-induced wave dissipation. IEEE Trans. Geosci. Remote Sens. 42, 2612–2622. http://dx.doi.org/10.1109/TGRS.2004.835349.

Aarninkhof, S., Ruessink, B., Roelvink, J., 2005. Nearshore subtidal bathymetry from time-exposure video images. J. Geophys. Res.: Ocean. 110 (C6), http://dx.doi.org/10.1029/2004JC002791.

Almar, R., Blenkinsopp, C., Almeida, L.P., Cienfuegos, R., Catalán, P.A., 2017. Wave runup video motion detection using the radon transform. Coast. Eng. 130, 46–51. http://dx.doi.org/10.1016/j.coastaleng.2017.09.015.

Antolínez, J.A., Méndez, F.J., Anderson, D., Ruggiero, P., Kaminsky, G.M., 2019. Predicting climate-driven coastlines with a simple and efficient multiscale model. J. Geophys. Res.: Earth Surf. 124 (6), 1596–1624. http://dx.doi.org/10.1029/2018JF004790.

Atkinson, A.L., Power, H.E., Moura, T., Hammond, T., Callaghan, D.P., Baldock, T.E., 2017. Assessment of runup predictions by empirical models on non-truncated beaches on the south-east Australian coast. Coast. Eng. 119, 15–31. http://dx.doi.org/10.1016/j.coastaleng.2016.10.001.

Bailey, D.G., 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. Vol. 2, pp. 1–2.

Bergsma, E.W., Conley, D.C., Davidson, M.A., O'Hare, T.J., Almar, R., 2019. Storm event to seasonal evolution of nearshore bathymetry derived from shore-based video imagery. Remote. Sens. 11 (5), 519. http://dx.doi.org/10.3390/rs11050519.

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-Sci. Rev. 177, 774–799. http://dx.doi.org/10.1016/j.earscirev.2018.01.002.

Bruder, B.L., Brodie, K.L., 2020. CIRN quantitative coastal imaging toolbox. SoftwareX 12, 100582. http://dx.doi.org/10.1016/j.softx.2020.100582.

Bruun, P., 1962. Sea-level rise as a cause of shore erosion. J. Waterw. Harb. Div. 88 (1), 117–130. http://dx.doi.org/10.1061/JWHEAU.0000252.

Buckley, M.L., Buscombe, D., Birchler, J.J., Palmsten, M.L., Swanson, E., Brown, J.A., Itzkin, M., Storlazzi, C.D., Harrison, S.R., 2024. Wave runup and total water level observations from time series imagery at several sites with varying nearshore morphologies. Coast. Eng. 193, 104600. http://dx.doi.org/10.1016/j.coastaleng.2024.104600.

Christiaanse, J.C., Antolínez, J.A., van der Grinten, M.J., Taal, F., Figlus, J., Dellapenna, T.M., Ritt, B., Marshall, C.D., Tereszkiewicz, P.A., Cohn, N., et al., 2025. Measurements of groundwater, hydrodynamics, and sand characteristics at a dissipative sea turtle nesting beach. Sci. Data 12 (1), 123. http://dx.doi.org/10.1038/s41597-025-04455-5.

Coco, G., Senechal, N., Rejas, A., Bryan, K.R., Capo, S., Parisot, J., Brown, J.A., MacMahan, J.H., 2014. Beach response to a sequence of extreme storms. Geomorphology 204, 493–501. http://dx.doi.org/10.1016/j.geomorph.2013.08.028.

Collins, A.M., O'Dea, A., Brodie, K.L., Bak, A.S., Hesser, T.J., Spore, N.J., Farthing, M.W., 2023. Automated extraction of a depth-defined wave runup time series from lidar data using deep learning. IEEE Trans. Geosci. Remote Sens. 61, 1–13. http://dx.doi.org/10.1109/TGRS.2023.3244488.

Fuse, T., Ohkura, T., 2018. Development of shoreline extraction method based on spatial pattern analysis of satellite SAR images. Remote. Sens. 10 (9), 1361. http://dx.doi.org/10.3390/rs10091361.

GoPro Inc., 2021. HERO10 Black Product Manual. Technical Report.

Gourlay, M., 1992. Wave set-up, wave run-up and beach water table: Interaction between surf zone hydraulics and groundwater hydraulics. Coast. Eng. 17 (1), 93–144. http://dx.doi.org/10.1016/0378-3839(92)90015-M.

Hamwood, J., Alonso-Caneiro, D., Read, S.A., Vincent, S.J., Collins, M.J., 2018. Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers. Biomed. Opt. Express 9 (7), 3049–3066. http://dx.doi.org/10.1364/BOE.9.003049.

Hashem, S., 1997. Optimal linear combinations of neural networks. Neural Netw. 10 (4), 599–614. http://dx.doi.org/10.1016/S0893-6080(96)00098-6.

Holman, R.A., Sallenger, A.H., Lippmann, T.C., Haines, J.W., 1993. The application of video image processing to the study of nearshore processes. Oceanography 6 (3), 78–85. http://dx.doi.org/10.5670/oceanog.1993.02.

Huff, T., Feagin, R., Figlus, J., 2020. Enhanced tide model: improving tidal predictions with integration of wind data. Shore & Beach 88 (2), 40–45. http://dx.doi.org/10.34237/1008824.

Huisman, C.E., Bryan, K.R., Coco, G., Ruessink, B., 2011. The use of video imagery to analyse groundwater and shoreline dynamics on a dissipative beach. Cont. Shelf Res. 31 (16), 1728–1738. http://dx.doi.org/10.1016/j.csr.2011.07.013.

Kang, B., Feagin, R.A., Huff, T., Durán Vinent, O., 2024a. Stochastic properties of coastal flooding events–Part 1: convolutional-neural-network-based semantic segmentation for water detection. Earth Surf. Dyn. 12 (1), 1–10. http://dx.doi.org/10.5194/esurf-12-1-2024.

Kang, B., Feagin, R.A., Huff, T., Durán Vinent, O., 2024b. Stochastic properties of coastal flooding events–Part 2: Probabilistic analysis. Earth Surf. Dyn. 12 (1), 105–115. http://dx.doi.org/10.5194/esurf-12-105-2024.

Kuschnerus, M., de Vries, S., Antolínez, J.A., Vos, S., Lindenbergh, R., 2024. Identifying topographic changes at the beach using multiple years of permanent laser scanning. Coast. Eng. 193, 104594. http://dx.doi.org/10.1016/j.coastaleng.2024.104594.

Lefevre, F., Le Provost, C., 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. J. Geophys. Res.: Ocean. 105 (C4), 8707–8725. http://dx.doi.org/10.1029/1999JC900281.

Maglio, C., Das, H., Fenner, F., 2020. Modal grain size evolution as it relates to the dredging and placement process-galveston island, texas. Coast. Eng. Proc. (36v), http://dx.doi.org/10.9753/icce.v36v.papers.7, 7–7.

Mendes, D., Andriolo, U., Neves, M.G., 2022. Advances in wave run-up measurement techniques. In: Advances on Testing and Experimentation in Civil Engineering: Geotechnics, Transportation, Hydraulics and Natural Resources. Springer, pp. 283–297. http://dx.doi.org/10.1007/978-3-031-05875-2_12.

Najafi, M.R., Zhang, Y., Martyn, N., 2021. A flood risk assessment framework for interdependent infrastructure systems in coastal environments. Sustain. Cities Soc. 64, 102516. http://dx.doi.org/10.1016/j.scs.2020.102516.

National Geographic, 2024. Hurricanes in History: Galveston Hurricane. National Geographic website, URL: https://www.nationalgeographic.org/forces-nature/hurr-history.html.

Nayak, S., Panchang, V., 2015. A note on short-term wave height statistics. Aquat. Procedia 4, 274–280. http://dx.doi.org/10.1016/j.aqpro.2015.02.037.

NOAA, 2024a. National data buoy center: Galveston pleasure pier. https://www.ndbc.noaa.gov/station_page.php?station=42035. (Accessed 23 January 2024).

NOAA, 2024b. NOAA tides and currents galveston pleasure pier. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm.

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics 9 (1), 62–66. http://dx.doi.org/10.1109/TSMC.1979.4310076.

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

Power, H.E., Holman, R.A., Baldock, T.E., 2011. Swash zone boundary conditions derived from optical remote sensing of swash zone flow patterns. http://dx.doi.org/10.1029/2010JC006724.

Raubenheimer, B., Guza, R., Elgar, S., 1999. Tidal water table fluctuations in a sandy ocean beach. Water Resour. Res. 35 (8), 2313–2320. http://dx.doi.org/10.1029/1999WR900105.

Gomes da Silva, P., Dalinghaus, C., González, M., Gutiérrez, O., Espejo, A., Abascal, A.J., Klein, A.H., 2016. Estimating flooding level through the Brazilian coast using reanalysis data. J. Coast. Res. (75), 1092–1096. http://dx.doi.org/10.2112/SI75-219.1.

Simarro, G., Bryan, K.R., Guedes, R.M., Sancho, A., Guillen, J., Coco, G., 2015. On the use of variance images for runup and shoreline detection. Coast. Eng. 99, 136–147. http://dx.doi.org/10.1016/j.coastaleng.2015.03.002.

Sofar Ocean Technologies Inc., 2023. SOFAR SPOTTER User Guide. Technical Report.

Stockdon, H.F., Holman, R.A., Howd, P.A., Sallenger, Jr., A.H., 2006. Empirical parameterization of setup, swash, and runup. Coast. Eng. 53 (7), 573–588. http://dx.doi.org/10.1016/j.coastaleng.2005.12.005.

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. J. Geol. 76 (4), 444–456. http://dx.doi.org/10.1086/627342.

Tomás, A., Méndez, F., Medina, R., Jaime, F., Higuera, P., Lara, J., Ortiz, M., Álvarez de Eulate, M., 2016. A methodology to estimate wave-induced coastal flooding hazard maps in Spain. J. Flood Risk Manag. 9 (3), 289–305. http://dx.doi.org/10.1111/jfr3.12198.

Turali, M.Y., Lorasdagi, M.E., Koc, A.T., Kozat, S.S., 2024. AFS-BM: Enhancing model performance through adaptive feature selection with binary masking. arXiv preprint arXiv:2401.11250.

Turner, I.L., Leyden, V.M., Symonds, G., Mcgrath, J., Jackson, A., Jancar, T., Aarninkhof, S., Elshoff, I., 2001. Predicted and observed coastline changes at the Gold Coast artificial reef. In: Coastal Engineering 2000. pp. 1836–1847. http://dx.doi.org/10.1061/40549(276)143.

US Amry Corps of Engineering, 2022. Galveston 204 CAP – Beach nourishment. URL: https://www.swg.usace.army.mil/Portals/26/HHGalv204AppxA5may22%20-%20DRAFT_1.pdf, Engineering Appendix.

Uunk, L., Wijnberg, K., Morelissen, R., 2010. Automated mapping of the intertidal beach bathymetry from video images. Coast. Eng. 57 (4), 461–469. http://dx.doi.org/10.1016/j.coastaleng.2009.12.002.

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

Wang, J., Wang, L., Feng, S., Peng, B., Huang, L., Fatholahi, S.N., Tang, L., Li, J., 2023. An overview of shoreline mapping by using airborne LiDAR. Remote. Sens. 15 (1), 253. http://dx.doi.org/10.3390/rs15010253.

Ware, M., Ceriani, S.A., Long, J.W., Fuentes, M.M., 2021. Exposure of loggerhead sea turtle nests to waves in the Florida Panhandle. Remote. Sens. 13 (14), 2654. http://dx.doi.org/10.3390/rs13142654.

Ware, M., Long, J.W., Fuentes, M.M., 2019. Using wave runup modeling to inform coastal species management: An example application for sea turtle nest relocation. Ocean. Coast. Manag. 173, 17–25. http://dx.doi.org/10.1016/j.ocecoaman.2019.02.011.

Yang, H., Zhang, C., Shi, J., Cao, X., Li, Y., 2022. Laboratory observations of swash signatures using video imagery. J. Mar. Sci. Eng. 10 (12), 1833. http://dx.doi.org/10.3390/jmse10121833.

Zhang, S., Zhang, C., 2009. Image analysis for wave swash using color feature extraction. In: 2009 2nd International Congress on Image and Signal Processing. IEEE, pp. 1–4. http://dx.doi.org/10.1109/CISP.2009.5304530.