Full length article

Automated measurements of whitecaps on the ocean surface from a buoy-mounted camera

CrossMark

M. Bakhoday-Paskyabi [a,*], J. Reuder [a], M. Flügge [a,b]

[a] Geophysical Institute, University of Bergen and Bjerknes Center for Climate Research, Bergen, Norway
[b] Christian Michelsen Research AS, Fantoftvegen 38, 5072 Bergen, Norway

HIGHLIGHTS

- Adaptive Thresholding Segmentation using histogram information for detection of whitecap areas.
- Iterative between class variance method to separate whitecap pixels from the darker water pixels.
- Estimates of whitecap fraction from threshold-based automated techniques.

ARTICLE INFO

ABSTRACT

We quantify the percentage of sea surface covered by whitecaps from images taken by a non-stationary camera mounted on a moored buoy using an Adaptive Thresholding Segmentation (ATS) method and an Iterative Between Class Variance (IBCV) approach. In the ATS algorithm, the optimal value for the threshold is determined as the last inflection point of the smoothed cumulative histogram of the scene. This makes the method more effective in finding the optimal value of the threshold and reduces the computational efforts compared to the conventional Automated Whitecap Extraction (AWE) technique. In the IBCV method, the optimum criterion for determining the value of the threshold corresponds to the measure of separability between the segmented water and whitecap pixels. In our experiments, the fraction of each image covered by the whitecap is determined using the aforementioned dynamical thresholding techniques for images taken under complex forcing and lighting conditions. Comparisons between different techniques suggest the effectiveness of the proposed methodolo-

---

* Corresponding author.
 E-mail address: mba098@uib.no (M. Bakhoday-Paskyabi).

1. Introduction

For wind speeds higher than 3 m s^{-1}, breaking of surface gravity waves ejects sea spray droplets into the marine atmospheric boundary layer and entrains air into the water column in a broad ranges of scales, from dissolving bubbles (with radii less than 100 μm) to the large buoyant and rising bubbles near the ocean surface (Thorpe, 1982; Deane and Stokes, 2002). The surface whitecaps, which constitute 2%–5% of the global coverage of the ocean surface (Blanchard, 1963) are the most important manifestation of such enhanced interactions between atmosphere and ocean during the sea surface breaking events (Melville, 1996; Scanlon and Ward, 2013). The ejection of droplets from breaking waves, the bursting of bubbles floating on the sea surface, and several air–sea exchange processes are closely linked to the fractional areas of the whitecap coverage at the sea surface. Within the marine atmospheric boundary layer during high wind events, turbulent mixing carries the ejected sea spray droplets up into the air where they are effectively transported over the sea surface and possibly collide with fixed and floating offshore structures (Jones and Andreas, 2012). Formation and evolution of surface whitecaps also influence many air–sea exchange processes, such as gas- and heat-transfer (Andreas, 1992), and wave energy dissipation into the water column (Hanson and Phillips, 1999). Therefore, quantifying the properties of whitecaps, in particular active whitecaps as strong manifestations of wave breaking crests, substantially contributes to the detailed understanding of several air–sea processes.

After the pioneering study of Monahan (1969) for the extraction of the areas of whitecaps from film, and subsequently video images, several investigations have then focused on the determination of the fraction of the sea surface covered by whitecaps in an attempt to link the whitecap fraction to other measurable physical quantities, such as wind speed (Monahan, 1971). Various empirical parameterisations of the whitecap fraction at the sea surface may be specified as a function of wind speed at 10 m height, U_{10}, i.e.

$$W_c(\%) = W_c^A(\%) + W_c^B(\%) = \alpha \times (U_{10} - \beta)^n, \tag{1}$$

where W_c is the total whitecap fraction in %, and α, β, and n are tuneable coefficients which can be determined by a curve-fitting technique (Stramska and Petelski, 2003). Several studies have confirmed that the coefficients α, β, and n in Eq. (1) depend on the location of the measurements, the specific experimental conditions, the source of wind speed data (e.g. sonic anemometer), the temperature difference between air and sea, the sea-state conditions, and the atmospheric stability in the lower layer of the marine atmospheric boundary layer (Monahan and O'Muircheartaigh, 1980, 2012; Salisbury et al., 2014; Paget et al., 2015). W_c^A and W_c^B indicate the active (highly reflective and bright surface elements) and residual (less bright and thinner vertically than the active whitecaps) whitecaps, respectively. While the active whitecaps are of importance for the parameterisations of dynamical processes at the sea surface, such as momentum transfer, dissipation, and spume drop production (Melville and Matusov, 2002), the residual whitecaps mainly influence heat exchange and some bubble-mediated processes such as gas transfer (Andreas et al., 2008) and jet droplet production (Woolf et al., 1987).

A key tool in camera-based whitecap measurement is the automatic identification of the whitecap covered areas from a sequence of images monitoring the sea surface. Callaghan et al. (2008) and Callaghan and White (2009) proposed an Automated Whitecap Extraction (AWE) approach to determine the total whitecap fraction value by the means of intensity thresholding. They showed that the uncertainty in determining the whitecap fraction is directly related to the number of images used in the processing and the sampling frequency. Particularly, they suggested a minimum measurement

Table 1
The used parameters to generate the idealistic histograms.

n	μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
3	-10	0.0	10	2.5	2.5	2.5
2	-10	0.0	–	2.5	2.5	–
1	-10	–	–	2.5	–	–

frequency between 0.25 and 0.33 Hz to achieve a statistical uncertainty of 5% for each value of W_c. Although the AWE technique is typically applied for calculation of the whitecap fraction, the assumptions behind this method are violated in the presence of strong uneven illumination and other atmospheric interferences. Therefore, we present two alternative fully automated techniques to identify the fraction of whitecaps from the background non-whitecap areas based on the intensity segmentation of each individual pixel. The optimal value of the threshold for each scene is obtained based on the statistical information of the greyscale histogram, and the separability distance between the two segmented classes, i.e. the water and the whitecap regions.

This paper is organised as follows. Section 2 presents the image processing background, and Section 3 describes the experiment and the measurement site. The results are presented and discussed in Section 4. The last section will provide a brief summary and outlook.

2. Digital image processing

The common digital image processing technique to calculate the whitecap coverage consists of several steps, e.g. Callaghan and White (2009):

1. The coloured images are converted into greyscale.
2. A threshold level is determined for each scene based on the sea surface brightness characteristics. This allows for the identification of whitecap regions as pixels with bright intensities above the determined optimal threshold.
3. The whitecap information are averaged over a certain number of images to determine statistically certain estimates for the values of whitecap coverage.

Due to the environmental complexity, occlusion effects, sun-light reflection, non-stationarity of whitecaps and change in the illumination at each scene, the principle assumption behind the thresholding methods that there should exist a distinct distribution difference between the water and the whitecap pixels may be violated. Therefore, it is essential to reduce the effects of such disturbances by applying appropriate pre-processing and quality control procedures before starting the feature extraction (Sugihara et al., 2007; Callaghan and White, 2009).

In order to validate the performance of methods, we discuss the application of the ATS and the IBCV algorithms to the ideal composite histograms generated by the sum of the Gaussian probability density functions as

$$h_s(x) = \sum_{i=1}^{n} \frac{a_i}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\frac{(x - \mu_i)^2}{\sigma_i^2} \right], \tag{2}$$

where $h_s(x)$ is the simulated histogram as a function of the bin value x, n denotes the number of modes, a_i (here $= 1$ for $i = 1 \cdots n$) are the respective proportion, and μ_i and σ_i represent the mean and standard deviation of the ith mode. Table 1 lists the parameters used for checking the performance of the proposed threshold-detection algorithms.

2.1. Extraction of whitecap fraction

The population of pixels with intensity k is defined as

$$J_k = \sum_{(u, v) \in R} [I(u, v) == k], \tag{3}$$

where $I(u, v)$ is the greyscale value of a pixel located at the point (u, v) in the image frame, and R is typically a local image region including the defined Region-of-Interest (ROI). The percentage of whitecap coverage is then calculated from

$$W_c(\%) = 100 \times \frac{\sum\limits_{k \geq T} J_k}{\sum\limits_{k \geq 0} J_k}, \tag{4}$$

where T denotes the value of threshold which is applied to divide the pixel space into the white objects and dark background. After determining the threshold, each pixel in the image is labelled as 1 if it corresponds to the object (whitecap) and 0 if it corresponds to the background (water). In the next section, we present different methods to determine the optimum threshold value.

2.2. Automated image segmentation

Although there are a wide range of sophisticated classification approaches, the pixel-based thresholding schemes are still common due to the simplicity of their algorithms (Sugihara et al., 2007; Callaghan and White, 2009). However, in images of the ocean surface in which scenes exhibit a strong change of illumination, the thresholding-based techniques are no longer appropriate for mapping of the pixel values due to particular surface events, i.e. the bursts of breaking crest and residual whitecaps. In such situations, the distinction between bright areas and the background water still remains a computational challenge for the vision-based techniques.

The major issue in the thresholding method is the accurate choice of the threshold for the separation between two groups of pixels, i.e. object (whitecap) and non-object pixels. A widespread approach is based on the determination of the local minimum of the intensity frequency histogram by assuming a bimodal histogram distribution, such as Otsu's method (Otsu, 1979). However, in a bimodal histogram where the valley between two peaks is wide and in the absence of strong illumination effect, the local minimum strategy will results in a range of choices for the value of the threshold instead of an optimal value. Furthermore, the histogram is not always bimodal and the optimal schemes consequently predict poor values for threshold in such cases.

2.2.1. Segmentation using AWE method

Callaghan and White (2009) developed the AWE algorithm to determine a suitable threshold intensity value to separate whitecaps from the background water pixels. This technique contains two major components: an image structure as the fraction of pixels with intensities greater than a given threshold; and an optimum threshold-selection system based on the first, the second, and the third-order derivatives of the image structure. According to our analysis for the images taken from a non-stationary camera, the image structure is controlled by the sunglint, sky reflection, and the strength of uneven illumination, see also Callaghan and White (2009). These features add brightness and texture to the vast majority of images used in our analysis and confine the identification of transition regions between whitecaps and background water due to the erratic nature of the image structure. The resulting disturbances in the image structure, especially when the threshold approaches the higher intensity values, cannot be sufficiently counteracted by the conventional smoothing procedures such as the iterative cropping methods or using contrast filtering techniques (Callaghan and White, 2009). We found the variations of $W_c(\%)$ as a function of the threshold value less sensitive to the atmospheric and platform motion-induced contaminations than the application of the image structure for our images. Hence, to avoid further quality control protocols to omit the noisy portions of the image structure, we replace hereafter the image structure in the AWE method by the percentage of the whitecap coverage (Eq. (4)) as a function of the threshold, see also Sugihara et al. (2007). The optimum threshold is then determined based on the first- or the second-order derivatives of the $W_c(\%)$ as a function of threshold. This modification leads to a very robust detection algorithm with minimum requirement to further image pre-processing operations.

In the AWE method used in this study, the percentage of whitecap coverage, $W_c(\%)$, is first determined over different normalised threshold values ranging between 0.01 and 1 (Callaghan and White,

2009). A decision graph for $W_c(\%)$, relative to the threshold values, is then constructed to extract the optimal threshold as a point at which the graph approaches a shallow tail (i.e. the last inflection point). Although the technique is effective, it misidentifies the whitecap area under certain conditions, such as strong pixel intensity and wave-induced camera motions, contamination from sunglint and sky reflection, raindrops, and uneven changes of illumination in the scene. These disturbances may lead to a substantial reduction in the frequency of images used for the extraction of $W_c(\%)$.

2.2.2. Segmentation using ATS

Motivated by Callaghan and White (2009), we propose an optimal threshold determination technique that utilises the information from a cumulative histogram of the greyscale images, see also Kleiss and Melville (2011) and Schwendeman and Thomson (2015). In this method, i.e. ATS, the threshold value is accomplished as the intensity value at which the curvature of the histogram graph is maximal or the change in the slope of the cumulative sum approaches asymptotically to zero. It is expected that using curvature information will result in a better estimation of the threshold compared to the histogram-based threshold detection techniques, particularly in the situations where the scene is multimodal or unimodal with weak well-separated peaks.

For the normalised (discrete) histogram h, the Cumulative Distribution Function (CDF) at point k is defined by

$$H(k) = \sum_{i=0}^{k} h(i). \tag{5}$$

Let H be a continuous function of variable x, the curvature of H is then given by

$$C(x) = H''(x) \left[1 + H'(x)^2 \right]^{-3/2}, \tag{6}$$

where H' and H'' are the first- and the second-order derivatives of H, respectively. If the curvature approaches to zero at point x, an inflection point at x is achieved at which the cumulative sum finds a change of slope and sign. In real maritime images without a clear gap between background and foreground textures, C is very noisy and oscillatory in the vast majority of cases at the tail area. This makes the detection of the correct inflection point very complicated. Therefore, to effectively determine the inflection point of H (zero of C at the tail area), the following inequality needs to be checked

$$\min\{x : |H''(x)| \geq \gamma\}, \tag{7}$$

where $|\cdot|$ denotes the absolute value function, and $\gamma > 0$ is selected as a constant small number (here $\gamma = 10^{-4}$) or is determined adaptively to avoid the effects of noise at the tail of $C(x)$ for each scene. It is obvious that the efficiency of this method is directly controlled by the quality of the CDF graph to avoid spurious threshold estimation. Consequently, we apply a four-point running mean to the raw histograms prior to constructing the threshold-detection algorithm.

In this study, images mostly exhibit a long tail with whitecaps forming a rather weak peak compared to the corresponding peak of the background water pixels. We can then determine the optimum value of T_{opt} using a triangle algorithm for $H''(x)$ or $H'(x)$ (Zack et al., 1977). Since $C(x)$, $H'(x)$, and $H''(x)$ share similar behaviour at the tail area, we apply the triangle algorithm to the graph of $H'(x)$, which seems less noisier than other two functions, as follows:

1. Fit a line between the peak of $H'(x)$ at bin T_{max} and the end of the longer tail of the $H'(x)$ graph at bin T_{min}.
2. Determine the perpendicular distance between the fitted line and every point in the graph of $H'(x)$ between T_{min} and T_{max}.
3. Set the point corresponding to the maximum distance as the optimum threshold, T_{opt}.

An example of a non-modal histogram segmentation using the ATS threshold detection algorithm is given in Fig. 1. Since the values of $H''(x)$ and $H'(x)$ approach asymptotically to zero, the ATS threshold voting-system adjusts the threshold through a few iterations from the triangle method so that the value of $|H''(T_{opt}) - H'(T_{opt})|$ for the optimum threshold T_{opt} is within a small offset, Fig. 1(d). It is worth to mention that the results in this study are less sensitive to this final adjustment.
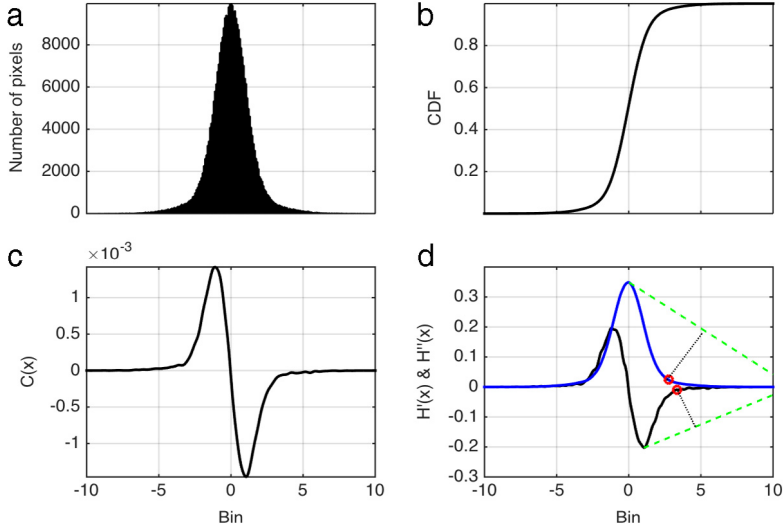
Fig. 1. Example of threshold detection of a unimodal histogram using ATS method. (a) Non-modal histogram as a function of bin; (b) the CDF graph; (c) the graph of curvature $C(x)$; and (d) graphs for the $H'(x)$, blue curve, and $H''(x)$, black curve, together with the triangle threshold detection phase of ATS. Green lines are the fitted lines and parameters used to generate the histogram in Fig. 1(a) are given in Table 1 for $n = 1$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2.3. Iterative Between-Class Variance (IBCV) segmentation

This technique is similar to Otsu's method for the selection of the optimal threshold. According to Otsu's method, the normalised grey level histogram is regarded as a discrete probability distribution function $p(i) = n_i/M$, where n_i is the number of pixels at grey level i and M is the total number of pixels in the ROI. By dividing the histogram into 2 classes using the threshold level T, the BCV is defined as follows (Otsu, 1979):

$$BCV(T) = p_1(T)p_2(T)\left[m_2(T) - m_1(T)\right]^2,\tag{8}$$

where p_1 and p_2 are the probabilities of the respective classes (water and whitecap regions), and m_1 and m_2 denote the corresponding mean intensity values of the two classes. The optimal threshold can then be determined iteratively by maximising the BCV. It should be noted that the maximising (minimising) problem can be extended for a multi-threshold methodology (Kittler and Illingworth, 1986). For example, if we divide the histogram into three classes separated by two thresholds T_1 and T_2, the following Iterative BCV (IBCV) algorithm will approach to the optimal threshold value:

1. Initialise T_1 and T_2, e.g. the smallest and the largest possible values, respectively.
2. Divide an image into three regions using the values of T_1 and T_2: R_1 (with all pixel value $\leq T_1$), R_2 (with all pixels values greater than T_1 and less than T_2), and R_3 (with all pixel values $\geq T_2$).
3. Calculate the average intensity values m_1, m_2 and m_3 for each region.
4. Compute a new value of T_1 and T_2 as

$$T_1 \longleftarrow \frac{m_1 + m_2}{2}; \quad \text{and} \quad T_2 \longleftarrow \frac{m_2 + m_3}{2}.$$

5. Repeat steps above from 2 to 4 until the differences between two successive values of T_1 and T_2 become minimal.
6. Set $T = \alpha_1 T_1 + \alpha_2 T_2$

where α_1 and α_2 are weighted average coefficients (here set to 0.5). When the density distribution of the histogram shifts towards the water background end, the variations of grey-level tend to alter the modal shape of the histogram (i.e. unimodal shape) in a manner that the IBCV method fails to detect
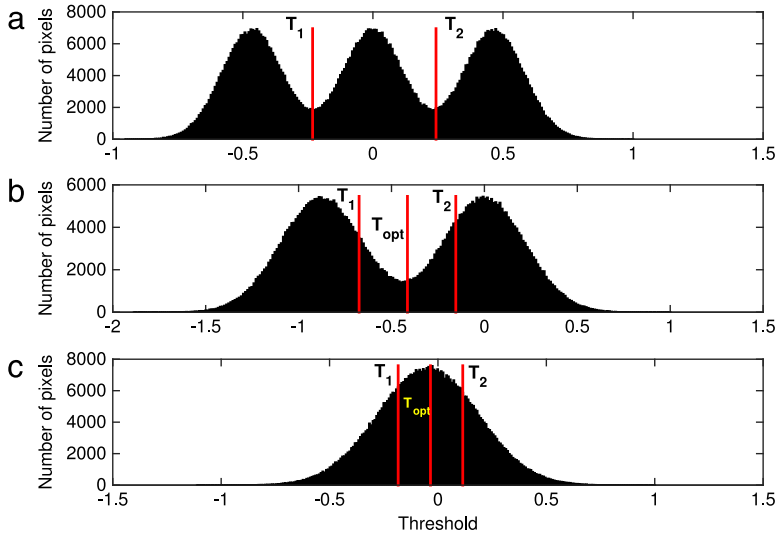
Fig. 2. Examples of threshold detection using the IBCV algorithm in the absence of illumination-correction phase for: (a) Trimodal histogram; (b) bimodal histogram; and (c) unimodal histogram using the classical IBCV thresholding search. Here, T_{opt} is the optimal threshold estimated from the classical IBCV method, i.e. from the step 6 of the algorithm. Parameters used to generate each histogram are given in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a good threshold. In order to increase the practicality of the IBCV algorithm, we shall incorporate an automatic search threshold phase after step 6. The IBCV algorithm is then proceeded by the following illumination-correction phase applied for the real images in this study:

1. Search from the histogram value at bin T (estimated from step 6) towards the end of the longer (foreground) tail of the histogram.
2. Estimate the value for the whitecap fraction at each search bin.
3. Set as the optimal threshold, T_{opt}, the bin at which the absolute error between two successive values of the estimated W_c (%) at the sequence of the whitecap fraction is within a small fixed offset.

Fig. 2(a) shows an ideal histogram with a trimodal distribution at which two adjacent distributions with the same variance overlap heavily in the middle. It is observed that the IBCV technique results in two equally good threshold values (vertical red lines) with less tendency to bias the true locations of the thresholds. Fig. 2(b) shows a bimodal histogram for which the classical IBCV method makes a decent estimate of the threshold value. Fig. 2(c) illustrates an application of the method when there is an unclear modality pattern in the shape of the composite histogram. While the weighted average of two determined thresholds provides the optimum threshold, the optimal value of the threshold is highly sensitive to the disturbances induced by the uneven change of illumination at each real maritime scene. In such cases, the method may effectively be converged to the optimum value by utilising the illumination-correction stage, see also Section 4.4.

3. Experimental approach

In November 2013, a field cruise was conducted with the research vessel R/V Håkon Mosby to Marstein Fyr, approximately 30 km southwest of Bergen, Norway (Fig. 3(a)). The objective of the cruise was the study of turbulent air–sea interaction processes in a coastal area. Several oceanographic instruments were deployed on both moving (submerged buoys) and fixed (bottom frame) sea going platforms. Moreover, a Furgo Oceanor Wavescan buoy was moored at the measurement site to record gravity wave statistics. In addition to the oceanographic measurements, a direct covariance flux (DCF)
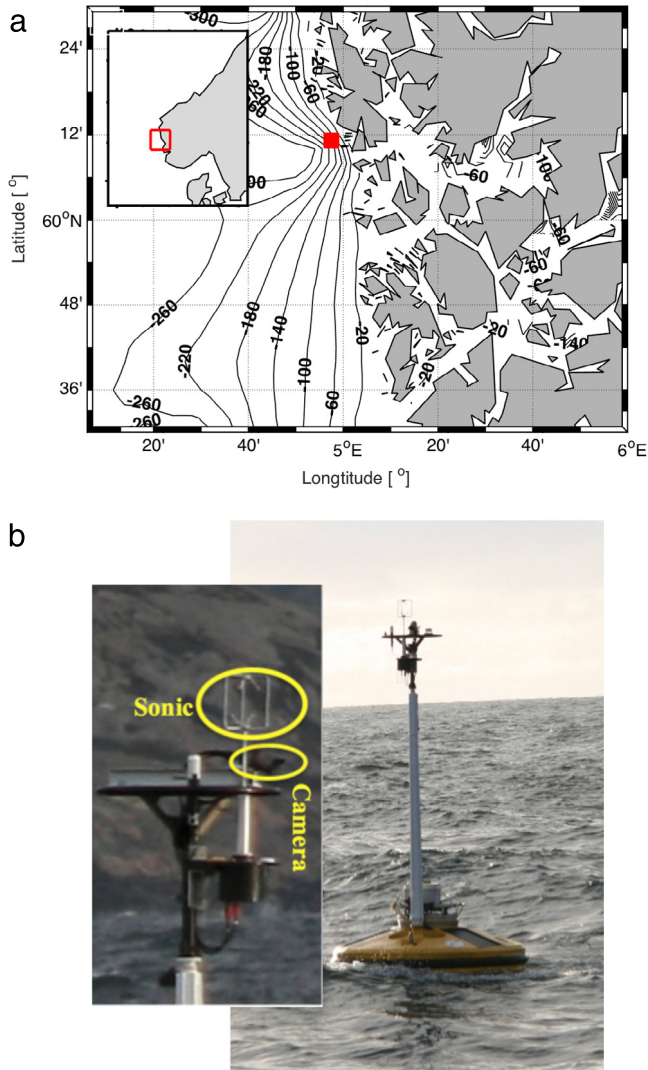
Fig. 3. (a) Map showing the deployment site (red square) at Marstein on the western coast of Norway. The inset shows the experiment site location in Norway; and (b) pictures of the Wavescan buoy and its mast with the installed sensors and the camera. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

system was mounted at both the front bow of the research vessel and on top of the Wavescan buoy's meteorological mast (Fig. 3(b)). The DCF system, consisting of a sonic Gill R3 anemometer and an Inertial Measurement Unit (IMU), recorded atmospheric heat and momentum fluxes at approximately 8 m (vessel) and 3.5 m (buoy) above the sea surface. A Campbell scientific CC5MPX digital network camera was mounted next to the buoy's DCF system for visual inspection of the sea surface gravity waves and whitecapping events. The camera with the standard lens of 4–12 mm and a field of view from 27° to 80° was mounted with a downward tilting angle of approximately 10° relative to the horizon and programmed to collect snapshot images, 1 frame per second (fps), during daylight hours. Each full colour image with the size of $992 \times 1280 \times 3$ is cropped on all sides (i.e. $793 \times 1081 \times 3$) to reduce the atmospheric interferences depending on displacement from the vertical. The cropped
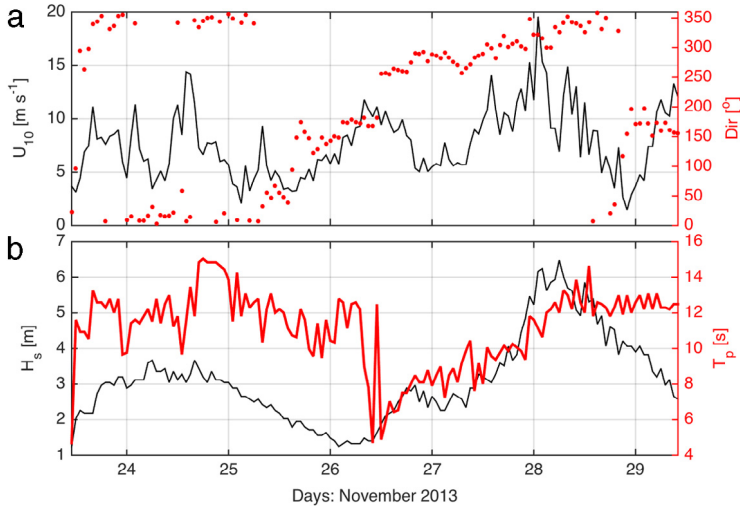
Fig. 4. Environmental forcing during the experiment. Time series of (a) wind speed and direction at 10 m height from the ship's meteorological mast; and (b) the significant wave height, H_s, and the wave peak period, T_p, recorded from the Furgo-Oceanor Wavescan buoy deployed nearby the ship station.

colour image is then converted to the greyscale and the colour depth is reduced to 8 bits resulting in a normal greyscale image with 256 levels.

The weather and sea conditions encountered during the cruise are summarised in Fig. 4. Wind speed and wave height varied considerably during the campaign, with maximum wind speeds reaching 20 m s^{-1} and a significant wave height up to 6 m during a storm, which passed over the deployment site on November 28. With the approach of the low pressure system, the wind direction turned from northerly winds at the beginning of the campaign towards southeasterly and finally westerly directions during the passage of the storm. Until November 26, and after the passage of the storm, the sea-state was characterised by swell waves (Fig. 4(b)) and varying cloud cover which provided acceptable conditions for the video camera to capture the wave field. During the storm, the tilt angles of the buoy were too large and the recorded images could not be further processed.

4. Results and comparisons

This section presents the results for the determination of the whitecap fraction using video images. Here, we use the images taken from the camera mounted on the meteorological mast of the moored buoy. While the algorithms have been rigorously tested for different sequences of images, we use a six-minute series of images in this section which is sufficient for a proof of concept demonstration.

4.1. Image structure

Fig. 5 shows an example of the smoothed image structure (using a four-point running average) of an image and its first- and second-order derivatives. The image structure shows the variation of the Percentage Increase in number of Pixels (PIP) versus threshold at level j as

$$\text{PIP}(j) = 100 \times \frac{\sum_{k=j}^{j+\Delta t} J_k}{\sum_{k \geq j} J_k}, \tag{9}$$

where J_k has been defined in Eq. (3), and $\Delta t = 0.01$ denotes the threshold increment (Callaghan and White, 2009). It is observed that the shape of the image structure relies on the grey-level variations
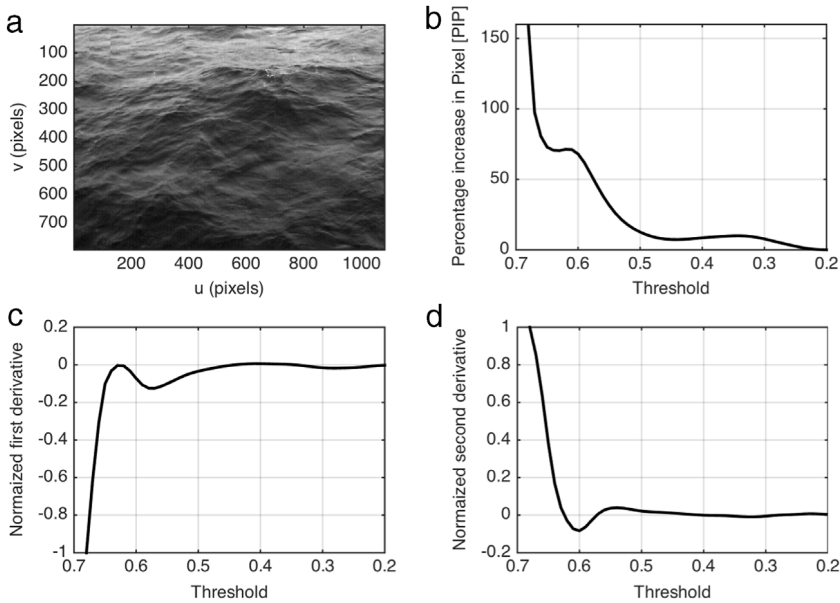
Fig. 5. The image structure of an image with strong non-uniform illumination pattern. (a) The original image; (b) the image structure estimated from Eq. (9); and (c and d) the first and the second-order derivatives of the image structure.

which substantially alter the illumination pattern of the image, Fig. 5. Therefore, prior to applying the detection algorithm, it is necessary to compensate the effects of the non-uniform illumination and sunglint or discard those contaminated images from the image dataset (Callaghan and White, 2009). However, a typical feature for almost all images used in this study is the strong uneven illumination of the background water with intensity values comparable to the intensity values of whitecaps, Fig. 5(a). This effect makes it very difficult to efficiently distinguish between the whitecap areas and water pixels using the PIP and its derivatives, Fig. 5(b), (c), and (d). Our analysis based on the cropped images suggests that the replacement of the image structure by $W_c(\%)$ in the AWE method may result in a more robust procedure for images taken from a non-stationary camera in the presence of contaminations induced by the sky reflection and uneven illumination. Hereafter, we use the AWE method based on $W_c(\%)$ instead of employing the image structure, and postpone further investigations based on the image structure using this dataset elsewhere.
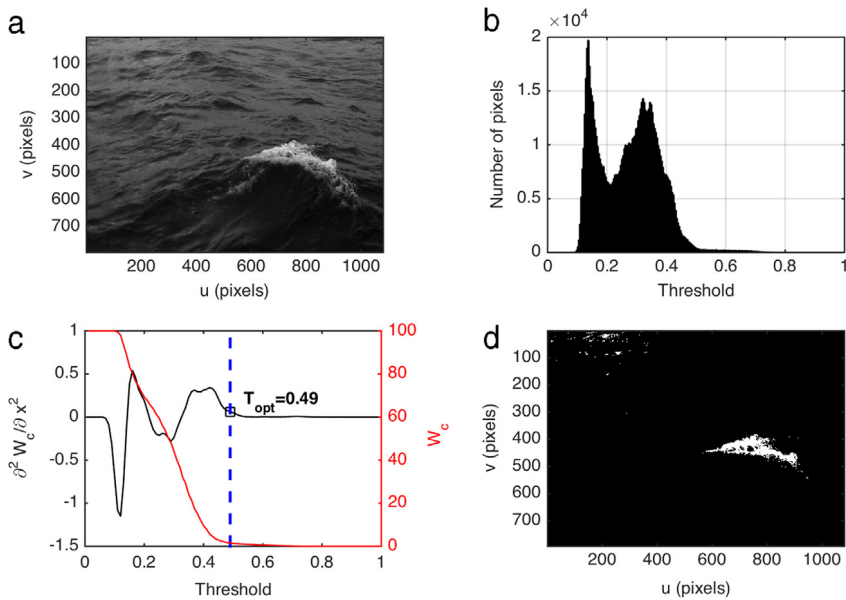
4.2. The AWE method

Fig. 6 illustrates a snapshot greyscale image including visible active whitecap events with a relatively long lifetime. This plunging plume can be captured in a digital image-processing algorithm by defining a global threshold to separate the background pixels (non-whitecapping) from the whitecapping pixels. In order to check the sensitivity of whitecap coverage to the threshold magnitude, three different threshold values in the range of [0, 1] are applied to the original greyscale image. The area of foam and the whitecap coverage are estimated by Eq. (4). In Fig. 6(b), the brightness of the binary image is relatively high due to the sunlight, and the non-whitecapping areas are identified as part of the whitecapping region with a coverage of approximately 80% for $T = 0.3$. By using a threshold of $T = 0.53$ (Fig. 6(c)), the contributions due to the sunglint substantially decline in the upper part of the image. Changing the value of threshold to 0.6 (Fig. 6(d)) results in further discrimination between whitecapping and non-whitecapping regions.

Fig. 6 highlights the importance of choosing an appropriate threshold level to distinguish between different illumination conditions. In Fig. 7, we show details of the optimal threshold-selection system based on the AWE approach.
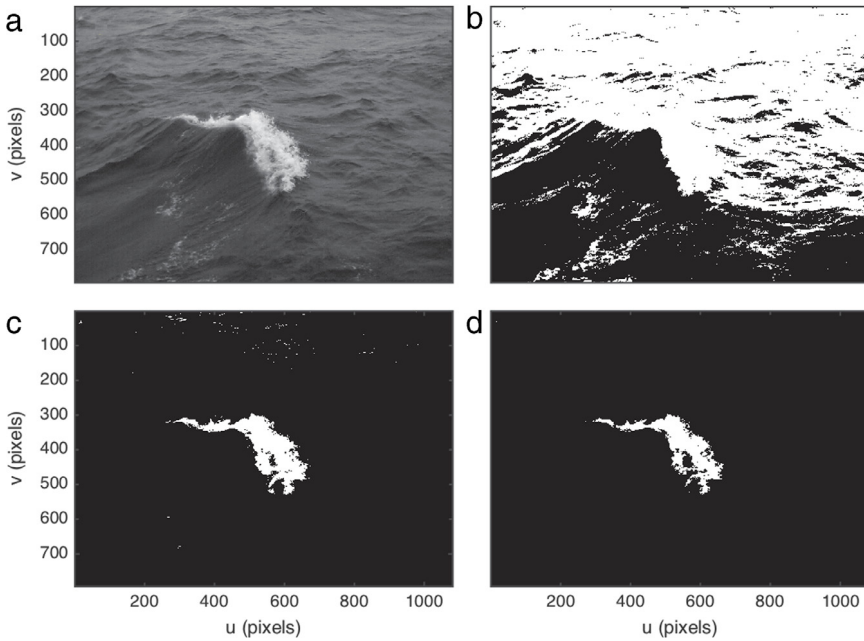
Fig. 6. Picture of whitecaps variations as a function of threshold. (a) Original greyscale image; (b) the binary image with $T = 0.3$ and $W_c = 80\%$; (c) the binary image with $T = 0.53$ and $W_c = 19\%$; and (d) the binary image with $T = 0.6$ and $W_c = 15\%$.

Fig. 7. (a) Original grey scale image; (b) image histogram; (c) variation of W_c based on the AWE method and its smoothed second-order derivative as function of the threshold, $\partial^2 W_c / \partial x^2$, where x denotes the dummy variable related to the histogram bin; and (d) the binary image with optimum threshold of 0.53. The threshold increments set to 0.01. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
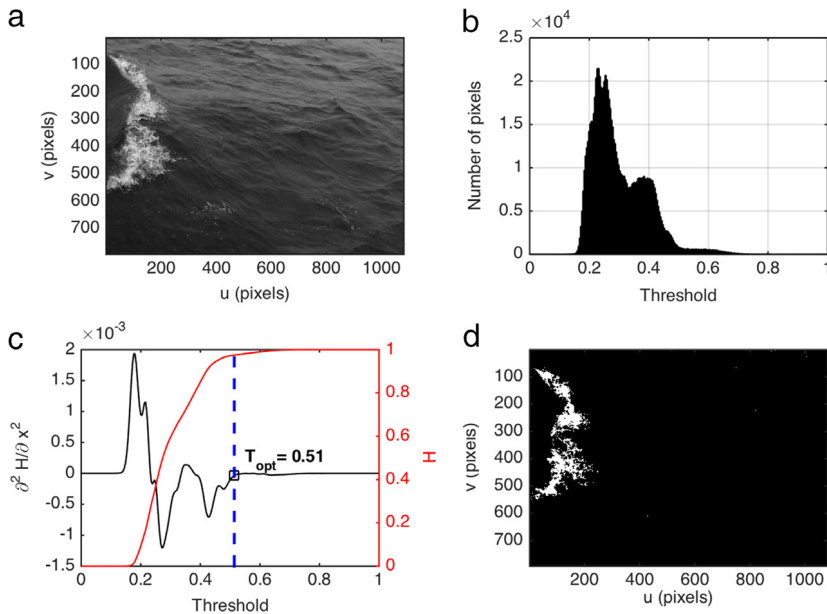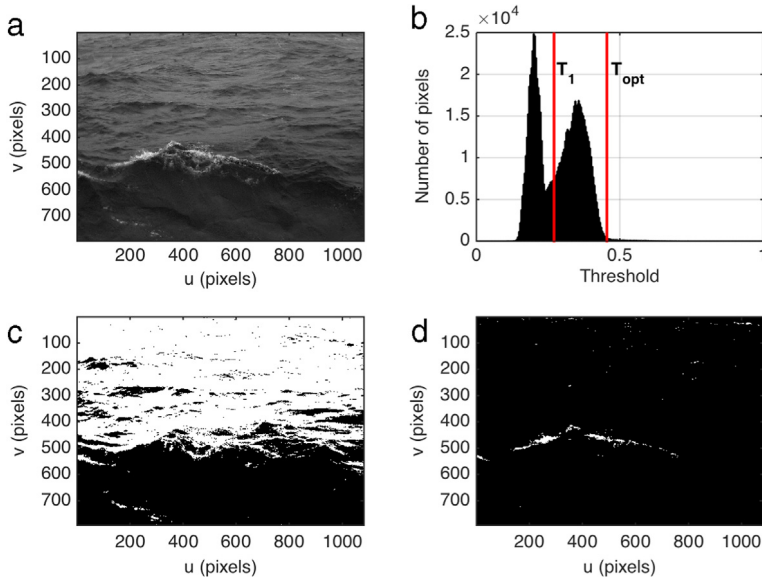
Fig. 8. (a) Original greyscale image; (b) image histogram; (c) variation of cumulative histogram, H, and its smoothed second-order derivative as a function of grey level, $\partial^2 H/\partial x^2$, where x denotes the dummy variable related to the histogram bin; and (d) the binary image with the optimal value of threshold, normalised in the range of [0, 1], calculated from the ATS algorithm.

The histogram of the greyscale image illustrates a clear separation between bright and dark pixels. The higher peak at $T = 0.21$ represents the pixel corresponding to the darker colour with the larger intensity variation, and the information of the histogram at the valley between two peaks may be captured by the peak-selection algorithm as the optimum value of the threshold (Fig. 7(c), and Section 2.1). Fig. 7(c) illustrates the relationship between the threshold level and the fraction of whitecap coverage for the greyscale image, with a threshold increment of 0.01. The whitecapping coverage curve (red) remains flat until a threshold of approximately 0.15, before declining rapidly in the mid-section. Beyond the threshold value of $T = 0.5$, the curve asymptotically approaches towards zero. Finally, the optimum value of the normalised threshold, T_{opt}, is chosen as the last inflection point of W_c which is calculated using the smoothed second-order derivative of $W_c (\%)$ (black line). While in this example, the AWE method perfectly determines the separation between whitecaps and water pixels, the histogram-based methods generally fail to successfully distinguish between the whitecapping and the non-whitecapping segments due to the valley caused by the illumination-induced contamination (Fig. 7(b)).

4.3. The ATS scheme

An example of using the ATS scheme for the extraction of bright pixels is shown in Fig. 8. This adaptive technique attempts to pick up the largest value for the threshold by assuring the appropriate detection of whitecap areas. The image greyscale histogram exhibits a valley region between two clear peaks as a result of contaminations from the uneven illumination. In such situations, there are several pixels in the whitecap patch which share the similar intensity distribution with the background water, and there is no distinct difference between the whitecap and water areas (see also Fig. 9(c)). Therefore, applying the Otsu's method will result in a non-optimal value for the threshold (Fig. 8(b)). An outline of the threshold selection procedure in the ATS method is shown in Fig. 8(c). After smoothing a greyscale image (using four-point running mean) to remove the noise and very high frequency content, the cumulative histogram, H, and its second-order derivative, $\partial^2 H/\partial x^2$, where x is the histogram bin

Fig. 9. An example of applying the IBCV threshold algorithm for the images with strong inhomogeneous background intensity distribution. (a) Original greyscale image; (b) image histogram and the threshold values from the standard IBCV algorithm (T_1) and the method including the illumination-correction phase, T_{opt}; (c) the binary image obtained by the standard IBCV algorithm; and (d) the binary image obtained by the illumination-correction version of the IBCV algorithm.

variable, are calculated from the smoothed image. The strong gradient of both whitecap and water regions can then be separated using a proper selection of the threshold value as a point in which the change in the $\partial^2 H / \partial x^2$ slope approaches zero. While the optimal point can be chosen using small value of γ in Eq. (7), the optimum value for the threshold is objectively selected using the fully automated triangle method presented in Section 2.2.2.

4.4. The IBCV method

We perform an analysis as shown in Fig. 9 to demonstrate the ability of the IBCV algorithm for an image with a strong background illumination pattern. It is observed that in the presence of strong inhomogeneous background conditions, the standard IBCV algorithm exhibits poor performance to capture the optimal value of the threshold. This may be attributed to the modulation of the image intensity histogram by the sudden and uneven changes of the image illumination, and nonuniform reflected sky conditions (Fig. 9(b) and (c)). While there are several approaches to handle the uneven illumination effects, we have found the IBCV method less vulnerable to the uneven illumination changes for each scene (Fig. 9(d)). It should be noted that we use the IBCV method together with the illumination-correction phase to obtain the optimal value of the threshold for all of maritime images in this study.

4.5. Comparisons and performance analysis

A comparison of the different approaches used in the whitecap detection process is shown in Fig. 10. Here we show the results from the AWE (Fig. 10(b)), the ATS (Fig. 10(c)), and the IBCV (Fig. 10(d)) techniques. All techniques determine suitable thresholds to separate between whitecaps and unbroken background water under uneven light conditions. While most of the whitecap pixels have been detected, none of the techniques could identify several sparse residual whitecaps in the scene. This is because these pixels have intensity values between the estimated threshold value and the intensity of the dark water pixels which may not be easily recognised by using the conventional
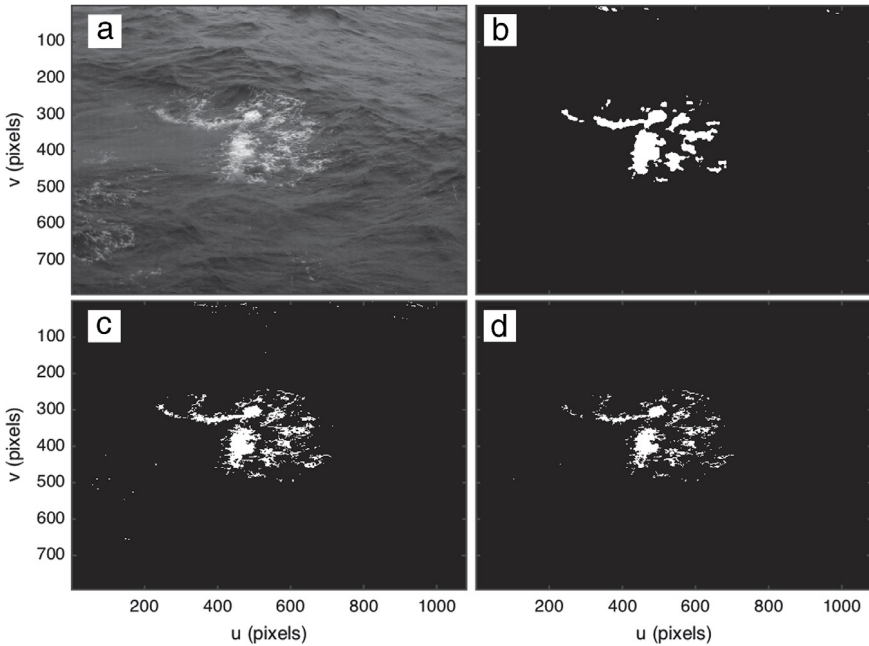
Fig. 10. The comparison between different W_c extraction techniques. (a) Original greyscale image; (b) the binary image estimated by the AWE method; (c) the binary image obtained from AST approach; and (d) the binary image estimated by the IBCV method.

thresholding-based techniques. Application of clustering algorithms, such as the k-mean algorithm, prior to the threshold-based segmentation may mitigate this drawback.

From Fig. 11, it can be seen that there is high positive correlation between the optimum thresholds estimated from two automated techniques and one from the ground truth analysis, which is determined manually using the thresholding-based technique. Generally, there might be some situations in which the ATS and the AWE schemes may result in erroneous values for the threshold when getting shifted towards the lower values of intensity. This can happen when the scene contains low noise levels and the cumulative spectrum does not exhibit any gradient response. However, such situations were not observed in the images used in this study.

In order to provide more insights into the accuracy of the linear regression and correlation coefficients (ρ), we conduct statistical testing of the significance such as the goodness-of-fit of the linear model (e.g. $y = a_0 x + b_0$) by calculating the error bound, and the significance of correlation coefficient by calculating the p-value and t-test. By assuming that both the ground truth and the estimated thresholds are Gaussian distributed, the t-value (i.e. Student's t-distribution) for the threshold values is determined using the following t-statistic as $t_{calc} = \rho \sqrt{(N-2)/(1-\rho^2)}$, where N is the number of images used for the inferential analysis (Emery and Thomson, 2001). The correlation coefficient is then significant if t_{calc} is greater than the critical t-value, t_{crit}, (for the significance level of 0.05 and $N - 2$ degrees of freedom). The p-value for all threshold-detection methods is approximately zero, suggesting that the correlation coefficient is significant. Accordingly, the t-test confirms further the statistical significance of all methods for the significance level of 0.05, i.e. $t_{calc} \gg t_{crit}$. Table 2 includes estimates for the standard deviation of the error in predicting the observed threshold at 5% level of significance. While all methods suggest nearly identical information on the accuracy of the threshold-detection procedure, the ATS method gives overall better improvement in predicting of the true threshold compared to the other techniques.

The superiority of the ATS and IBCV techniques relative to the AWE method used for this study is that they do not need any prior information about the values of the threshold and $W_c (\%)$ to
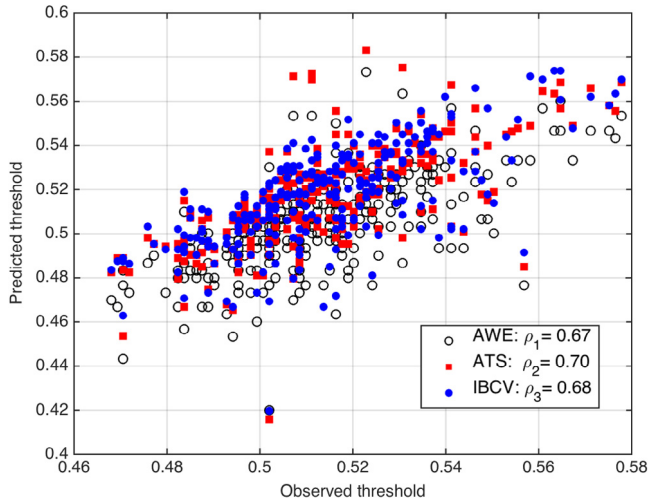
Fig. 11. The scatter plot between the normalised predicted thresholds, from the AWE method, the ATS scheme, the IBCV method, and the manually segmented ground truth data for 261 images. ρ_1, ρ_2, and ρ_3 indicate the values for the correlation between the observed and the predicted normalised thresholds.

Table 2

The results of statistical testing based on linear regression and correlation coefficient. a_0 and b_0 denote the slope and intercept of the linear model, respectively, ρ_i is the correlation coefficient for the ith method, the confidence interval error for the significant level of 0.05 is presented by Error (95%), the significance of the correlation analysis is given by p-value, and t_{calc} and t_{crit} are the t-test statistics.

Method	a_0	b_0	Error (95%)	p-value	ρ_i	t_{calc}	t_{crit}
AWE	0.72	0.14	0.0331	0.0	0.67	14.49	1.65
ATS	0.77	0.12	0.0323	0.0	0.70	15.83	1.65
IBCV	0.75	0.13	0.0330	0.0	0.68	15.09	1.65

generate the optimal threshold. Therefore, these methods are able to predict the optimum value of the threshold with less computational complexity than the AWE method, while preserving the same level of accuracy. Fig. 12 illustrates the computational performance of three presented threshold-detection methods implemented on a MAC OS X 296 10.9.5 with Intel Core i7 3 GHz CPU, and 16 GB 1600 DDR3 RAM. According to Eqs. (4) and (9), we conclude that the conventional AWE method (i.e. Callaghan and White, 2009) and the one used in the present study have approximately identical computational complexity. Consequently, using the ATS and the IBCV methods cause considerably less CPU usage than the time spent by the conventional AWE method to detect the optimum value of the threshold for each image.

To further assess the accuracy and performance of the automated segmentation techniques, we calculate the evaluation metrics for each image by comparing its segmented binary image with the manually segmented ground truth, which is, of course, subject to human error and the inter-observer variability. The information of a Confusion Matrix (CM) for each scene is used to estimate the number of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) pixels (Fig. 13(b), (c), and (d)). The calculated metrics are then used to estimate the precision and recall given by:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}; \qquad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{10}$$

where TP = CM(1, 1), FP = CM(2, 1), FN = CM(1, 2), and TN = CM(2, 2). Fig. 13 represents an example of the CM information estimated from the manually segmented image and the corresponding binary image obtained by applying the AWE, ATS, and IBCV techniques, respectively. The results show that the optimal threshold from the AWE method results in a high value for the true positive
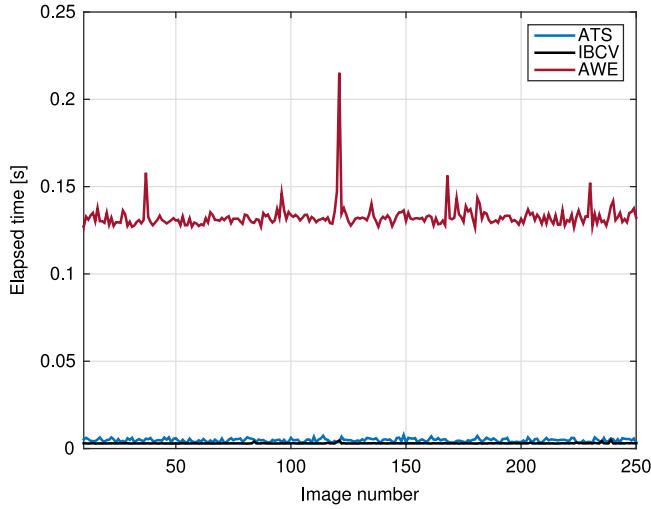
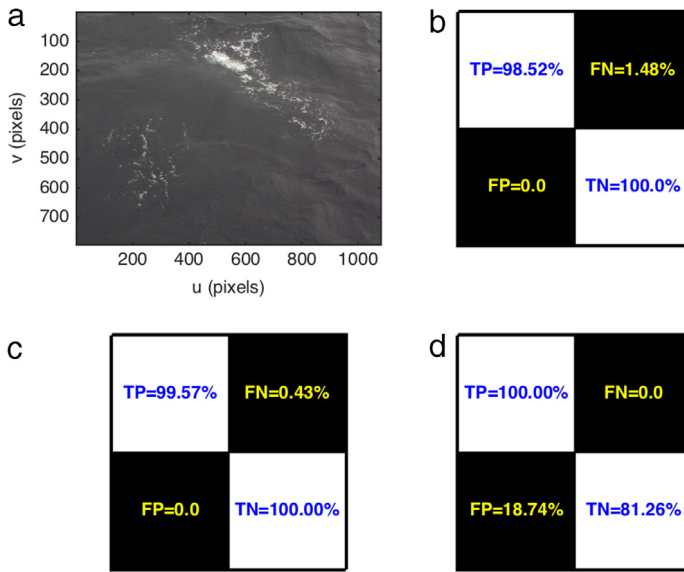**Fig. 12.** Elapsed time of three threshold-detection methods for a sequence of images.

Fig. 13. The comparison between the estimated confusion matrix for different segmentation techniques compared with the manually (threshold-based) segmented ground truth: (a) Original greyscale image; (b) the AWE method; (c) the ATS method; and (d) the IBCV scheme. The performance metrics are defined as: T-Positive (TP): the whitecap pixel (foreground) is correctly identified as whitecap pixel; F-Positive (FP): the water pixel (background) is incorrectly diagnosed as the whitecap pixel; T-Negative (TN): the water pixel is correctly diagnosed as the water pixel; and F-Negative (FN): the whitecap pixel is incorrectly identified as the water pixel. The diagonal elements in the CM matrix correspond to the correct classifications and the off-diagonal elements denote the incorrect classifications.

rate with the false negative rate of approximately 1.48% (Fig. 13(b)). The value of the FN rate is completely reduced by using the ATS and IBCV methods which preserve the higher values of TP rates than that of from the AWE method for this sample. Generally, the ATS classifier with 99% true positive rate and 0.43% false negative rate achieves better performance than the other automated segmentation techniques (see also Fig. 11). Furthermore, the implementation of the performance

test for three segmentation schemes over 261 images shows that all automated techniques achieve a high level of performance with an approximate value of 98% for both precision and recall, with respect to the manually segmented ground truth. It should be noted that although promising, the performance analysis is constrained to the choice of the ground truth datasets, the general limitations of the thresholding methods, and the uncertainty of using a non-stationary camera in the evaluation-scheme. Investigation of such uncertainty resources is beyond the scope of the present study and will be addressed in more details elsewhere, i.e. (Bakhoday-Paskyabi et al., submitted for publication).

5. Summary

In this study, we analysed a sequence of images acquired from a non-stationary camera mounted on a moored Wavescan buoy to measure the coverage of whitecaps at the sea surface. We proposed two threshold-based algorithms to discriminate between the whitecap and non-whitecap populations at each maritime scene with no need to know the camera's intrinsic and extrinsic parameters. The first method is an Adaptive Thresholding Segmentation (ATS) technique which uses the cumulative histogram information to extract the value of the optimal threshold as the last inflection point of the smoothed cumulative sum. From the implementation of this method, we observed that this technique not only makes it possible to efficiently divide each image into whitecap and non-whitecap classes, but also demonstrated that it is computationally very fast and robust relative to the conventional Automated Whitecap Extraction (AWE). Furthermore, we used another automated multi-thresholds algorithm to determine the optimal threshold by maximising the between-class variance (BCV) of the whitecap and the water pixels. An iterative strategy used in this algorithm allowed an efficient and robust bi-level segmentation of images with less computational demand and higher accuracy than the conventional Otsu's search method. A correction phase was further calculated in the standard Iterative BCV (IBCV) method to avoid the influence of image distortion by the uneven illumination. Comparisons between different techniques showed that the ATS method gives an overall better performance than the IBCV and the AWE methods.

Although the accuracy of the algorithms were precisely checked by using the confusion matrix entries, the analysis needs yet to be proven based on more accurate ground truth images less affected by the human errors and limitations of the conventional segmentation techniques. Since the investigation on the accurate choice of an empirical relation between the wind and $W_c(\%)$ and the best approach to the elimination of the contaminations induced by the non-stationary camera are beyond the scope of the present work, we will address them elsewhere.

References

Andreas, E.L., 1992. Sea spray and the turbulent air-sea heat fluxe. J. Geophys. Res. 97, 11429–11441.
Andreas, E.L., Persson, P.P.G., Hare, J.E., 2008. A bulk turbulent air-sea flux algorithm for high-wind, spray conditions. J. Phys. Oceanogr. 38, 1581–1596. http://dx.doi.org/10.1175/2007JPO3813.1.
Bakhoday-Paskyabi, M., Flügge, M., Schwendeman, M., Reuder, J., 2016. Visual attitude estimation from a shipboard moving camera: Applications in turbulent flux measurements. J. Atmos. Ocean. Technol. submitted for publication.
Blanchard, D., 1963. The electrification of the atmosphere by particles from bubbles in the sea. J. Prog. Oceanogr. 1, 71–202.
Callaghan, A.H., de Leeuw, G., O'Dowd, C.D., 2008. Relationship of oceanic whitecap coverage to wind speed and wind history. Geophys. Res. Lett. 35, http://dx.doi.org/10.1029/2008GL036165.
Callaghan, A.H., White, S., 2009. Automated processing of sea surface images for the determination of whitecap coverage. J. Atmos. Ocean. Technol. 26, 383–394.

Deane, G.B., Stokes, M.D., 2002. Scale dependence of bubble creation mechanisms in breaking waves. Nature 418, 839–844. http://dx.doi.org/10.1038/nature00967.

Emery, W.J., Thomson, R.E., 2001. Data Analysis Methods in Physical Oceanography, second ed. Elsevier Science, Netherlands, p. 638.

Hanson, J.L., Phillips, O.M., 1999. Wind sea growth and dissipation in the open ocean. J. Phys. Oceanogr. 29, 1633–1648.

Jones, K.F., Andreas, E.L., 2012. Sea spray concentrations and the icing of fixed offshore structures. Q. J. R. Meteorol. Soc. 138, 131–144.

Kittler, J., Illingworth, J., 1986. Minimum error thresholding. Pattern Recognit. 19 (1), 41–47.

Kleiss, J.M., Melville, W.K., 2011. The analysis of sea surface imagery for whitecap kinematics. J. Atmos. Ocean. Technol. 28, 219–243.

Melville, W.K., 1996. The role of surface-wave breaking in air-sea interaction annual review of fluid mechanics. Annu. Rev. Fluid Mech. 28, 279–321. http://dx.doi.org/10.1146/annurev.fl.28.010196.001431.

Melville, W.K., Matusov, P., 2002. Distribution of breaking waves at the ocean surface. Nature 417, 58–63. http://dx.doi.org/10.1038/417058a.

Monahan, E.C., 1969. Fresh water whitecaps. J. Atmospheric Sci. 26, 1026–1029.

Monahan, E.C., 1971. Oceanic whitecaps. J. Phys. Oceanogr. 1, 139–144.

Monahan, E., O'Muircheartaigh, I.G., 1980. Optimal power-law description of oceanic whitecap coverage dependence on wind speed. J. Phys. Oceanogr. 10, 2094–2099.

Monahan, E., O'Muircheartaigh, I.G., 2012. Oceanic whitecaps and the 10m-elevation wind speed: Toward improved power-law descriptions for use in climate modeling. In: Presented at 18th Air–Sea Interaction Conference. Amer. Meteorol. Soc., Boston, Mass, 19–13. Jul.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Mach. Learn. Appl. 17–21.

Paget, A.C., Bourassa, M.A., Anguelova, M.D., 2015. Comparing in situ and satellite-based parameterizations of oceanic whitecaps. J. Geophys. Res. Oceans 120, 2826–2843. http://dx.doi.org/10.1002/ 2014JC010328.

Salisbury, D.J., Anguelova, M.D., Brooks, I.M., 2014. Optimal power-law description of oceanic whitecap coverage dependence on wind speed. Geophys. Res. Lett. 41, 1616–1623. http://dx.doi.org/10.1002/2014GL059246.

Scanlon, B., Ward, B., 2013. Oceanic wave breaking coverage separation techniques for active and maturing whitecaps. Methods Oceanogr. http://dx.doi.org/10.1016/j.mio.2014.03.001.

Schwendeman, M., Thomson, J., 2015. Observations of whitecap coverage and the relation to wind stress, wave slope, and turbulent dissipation. Geophys. Res. Oceans 12, http://dx.doi.org/10.1002/2015JC011196.

Stramska, M., Petelski, T., 2003. Observations of oceanic whitecaps in the north polar waters of the Atlantic. J. Geophys. Res. 108 (C3), http://dx.doi.org/10.1029/2002JC001321.

Sugihara, Y., Tsumori, H., Ohga, T., Yoshioka, H., Serizawa, S., 2007. Variation of whitecap coverage with wave-field conditions. J. Mar. Syst. 66, 47–60.

Thorpe, S.A., 1982. On the clouds of bubbles formed by breaking wind-waves in deep water, and their role in air-sea gas transfer. J. Philos. Trans. R. Soc. Lond. Ser. A 304, 155–216.

Woolf, D.K., Bowyer, E.K., Monahan, E.C., 1987. Discriminating between the film-drops and jet-drops produced by a simulated whitecap. J. Geophys. Res. 92, 5142–5150.

Zack, G.W., Rogers, W.E., Latt, S.A., 1977. Automatic measurement of sister chromatid exchange frequency. J. Histochem. Cytochem. 25 (7), 741–753.