

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Image-clustering analysis of the wave-structure interaction processes under breaking and non-breaking waves

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Formentin S.M., Gaeta M.G., De Vecchis R., Guerrero M., Zanuttigh B. (2021). Image-clustering analysis of the wave-structure interaction processes under breaking and non-breaking waves. PHYSICS OF FLUIDS, 33(10), 1-18 [10.1063/5.0065019].

Availability:

This version is available at: <https://hdl.handle.net/11585/853503> since: 2022-02-06

Published:

DOI: <http://doi.org/10.1063/5.0065019>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

1

Manuscript submitted to Physics of Fluids for review and possible publication

2

3

IMAGE-CLUSTERING ANALYSIS OF THE WAVE-STRUCTURE INTERACTION PROCESSES UNDER BREAKING AND NON-BREAKING WAVES

4

5

6

Sara Mizar Formentin¹, Maria Gabriella Gaeta¹, Roberto De Vecchis¹,
Massimo Guerrero¹ and Barbara Zanuttigh¹.

7

8

⁽¹⁾ Department of Civil, Chemical, Environmental and Materials Engineering, University of
Bologna, Viale del Risorgimento 2, Bologna 40136, Italy.

9

10

saramizar.formentin2@unibo.it

11

12

Abstract

13

This contribution presents the effectiveness and the potentialities of a consolidated technique – the video-cluster analysis – to the study of turbulent flow and breaking waves, in order to demonstrate its suitability as a low-cost, non-intrusive method to derive quantitative key parameters describing the wave-structure interaction processes at coastal defense structures. To this purpose, a new methodology, consisting of a series of pre- and post-processing techniques developed to optimize the automatic detection of clusters in video imagery, was designed to process the video-records of experiments of wave run-up and wave overtopping at sea-dikes subjected to irregular waves. The results of the cluster analysis were elaborated to reconstruct the instantaneous profiles of the free-surface elevations across the structure crest and derive simultaneous information on overtopping volumes, discharges, depths and velocities, and to get spatial-time maps of the concentration of the air entrapped in the liquid phase. The accuracy of the methodology is demonstrated by comparing the quantities derived from the cluster analysis to laboratory measurements performed with resistive gauges and acoustic Doppler profilers. The novelty of the work is either represented by the results of the application of the cluster-analysis and by the procedures of optimizations, whose ensemble may establish a best practice and represent a guideline for other applications.

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

Keywords: Videography; cluster analysis; wave-structure interaction; wave overtopping; air entrainment; crown walls

31

32

33 1. Introduction

34 The analysis of the interaction processes between waves and coastal structures is still a
35 challenging task, due to the highly turbulent flow conditions, the non-linear dynamics, the
36 occurrence of the wave breaking and the air entrainment (Aleixo et al., 2018; Raby et al., 2020;
37 Stringari et al., 2021). Time and spatial resolution of the measurement techniques poses
38 limitations to the type and accuracy of the observations such as the size and distribution of the
39 bubbles in the water flow (Na et al., 2015).

40 Different traditional measurements or combinations of measurements are available, including
41 impedance probes (Waniewski et al., 2001; Cox and Shin, 2003), conductivity probes (Chanson,
42 2002; Hoque and Aoki, 2005; Mori et al., 2007), optical fibre probes (Blenkinsopp and Chaplin,
43 2007; Lim et al., 2008), imaging and acoustics-based methods (Deane and Stokes, 2002; Gaeta
44 et al. 2020), imaging and optical fibre techniques (Rojas and Loewen, 2010; Na et al., 2015),
45 dual-tip resistance-type probe and Acoustic Doppler Velocimeters (Mori and Kakuno, 2008).
46 Since the 2010's, the widespread use of laser scanners prompted a number of successful
47 applications related to the monitoring of the wave motion (Blenkinsopp et al. 2012; Streicher et
48 al., 2013) and to the wave run-up and overtopping at coastal dikes (Hofland et al., 2015).

49 The development of Particle Image Velocimetry (PIV) in the 2000s provided a powerful
50 technology to analyze velocity and turbulence fields in breaking waves (a.o. Oh et al. 2008;
51 Drazen and Melville, 2009), with some limitations due to the unsteady, non-linear, high-speed
52 process. Techet and McDonald (2005) concluded that significant improvements to the air
53 seeding technique were needed to increase the flow resolution and resolve small scale vorticity.
54 Kimmoun and Branger (2007) observed that the main difficulties in the PIV application were to
55 automatically detect the free surface during wave breaking and to avoid laser light reflection by
56 the air bubbles toward the camera. Duz et al. (2020) analysed the kinematics of spilling and
57 plunging breakers, highlighting that the main measurement challenges are the reflections from
58 the air/water interface and from the air entrapped in the breaking region, while the main post-
59 processing challenges were due to the huge number of images and to the application of masking
60 algorithm. Recent works highlighted that the PIV measurements are still extremely "expensive"
61 (i.e. time consuming). Lim et al. (2015) measured the flow properties in the aerated crest of a
62 plunging breaker by using modified PIV and bubble image velocimetry (BIV). The authors spent
63 several months to complete the 20 repeated runs of the same tested condition to collect data
64 from all the available measurement devices, finding that the generated waves were highly
65 repeatable only before breaking.

66 Another peculiar technique to detect the wave breaking is represented by the use of infrared
67 imagery, based on the fact that the surface temperature of the ocean is generally a few tenths
68 of degrees Celsius less than the bulk water temperature immediately below (Robinson et al.,
69 1984). Breaking waves momentarily disrupt the cool skin layer and the surface temperature
70 becomes approximately equal to the bulk water temperature. The effectiveness of the infrared
71 imagery to the detection of the wave breaking has been investigated for deep water (Jessup et
72 al., 1997a), microbreaking waves (Jessup et al., 1997b) and, more recently, in the surf zone

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0065019

73 (Carini et al., 2015). Siddiqui et al. (2001) employed simultaneously the infrared imagery and the
74 measurements analysis to reconstruct the velocity fields of breaking waves, while Buscombe
75 and Carini (2019) applied deep convolutional neural networks to classify the wave breaking type
76 (e.g., non-breaking, spilling, plunging) from infrared imagery of the surf zone.

77 Summarizing, the direct measurement of the wave breaking characteristics requires the
78 introduction of numerous instruments during the same campaign, such as resistant gauges,
79 velocimeters, water volume trapping in tanks, pressure sensors, PIV – which may be
80 unaffordable for economic and practical reasons. Furthermore, especially at small scales of
81 laboratory tests, such a huge installation of instrumentation might not be recommendable due to
82 the potential disturbance to the investigated processes (Soares-Frazão et al., 2009).

83 A relatively economic and non-intrusive alternative to the traditional techniques is represented
84 by the use of video imagery, whose application in the coastal engineering is long-term lasting.
85 For several decades, the coastal image data collected by monitoring systems such as ARGUS
86 (Holman and Stanley, 2007) were elaborated to detect the wave characteristics (De Vries et al.,
87 2011; Bechle and Wu, 2011), and to measure the wave run-up (Almar et al., 2017) on beaches.
88 The diffusion of high-resolution digital camera allowed relatively simpler and cheaper
89 applications of the videography, from the detection of the wave breaking on beaches (Almar et
90 al., 2012) to the analysis of the wave viewing angle (Perugini et al., 2019). Vousdoukas et al.
91 (2014) combined the laser scanning technique to a video camera monitoring to model the free-
92 surface elevation and the wave-by-wave morphological changes in the swash and surf zones.
93 Advances on the stereo processing of sea surface waves (Benetazzo et al., 2015) highly reduced
94 the computational resources required to analyze a sequence of stereo images, allowing the
95 employment of fast stereo video imaging and 2D laser slope gauge study in laboratory to gain
96 information on the three-dimensional structure of the wind-wave field (Zavadsky et al., 2017).
97 Gaeta et al. (2020) combined monostatic ultrasound velocity profiler to low cost videography
98 which eventually provided the overtopping discharge at scaled dikes in laboratory. The results
99 of these hybrid technologies highlighted that the use of the video camera may add valuable
100 information to the laser scanner and improve the accuracy of the results.

101 Imagery-based analyses often require a significant amount of calibration to yield reliable results
102 and less generically applicable results (Den Bieman et al., 2020). The wide diffusion of machine-
103 learning techniques, instead, led in the last years to numerous and innovative applications of the
104 image processing (e.g., Li et al., 2021). Among others, Stringari et al. (2019, 2021) made use of
105 machine-learning techniques based on the pixel intensity peaks to detect the wave breaking;
106 Buscombe et al. (2020) applied deep learning techniques to estimate wave heights and periods
107 from imagery of waves in the surf zone; Den Bieman et al. (2020) used segmentation algorithms
108 to obtain quantitative measurements of the free surface elevation and of the wave run-up from
109 the video-analysis of physical model tests. These latest studies demonstrate the reliability of the
110 automatic processing of videography, highlighting its effectiveness and advantages with respect
111 to more traditional techniques.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

112 Following the stream of the machine-learning algorithms, this paper investigates the
113 effectiveness and the potentialities of the automatic video-cluster analysis as low-cost and non-
114 intrusive method to model the wave-structure interaction processes in alternative to more
115 traditional or expensive techniques. In particular, the aim of the present contribution is to
116 demonstrate that such technique can be easily and successfully applied to study the wave run-
117 up and wave overtopping phenomena, under wave breaking, high-turbulence and high-aerated
118 flow conditions, providing accurate quantitative estimations of the main overtopping parameters
119 (discharge, volumes, velocities) and, for the first time, quali-quantitative estimations of the air
120 entrainment concentration during the several phases of the wave and flow propagation. To the
121 authors' purpose, the video-records of laboratory tests of wave overtopping at sea-dikes
122 (Formentin et al., 2019a, 2019b, see Section 2) were elaborated with a frame-by-frame cluster
123 analysis algorithm (Section 3) to reconstruct the free-surface elevation signals across the
124 structure profile (Section 4) and to get spatial-time mapping of the areas of the flow subjected to
125 air entrainment (Section 5). Specific pre- and post-processing filtering techniques were also
126 implemented to improve the automatic recognition of the image features (Section 3), outlining a
127 new set of procedures that may represent new state-of-the-art guidelines for the optimization of
128 the automatic cluster analysis. The main results, the potentialities and the limits of the image
129 processing procedure are summarized (Section 6).

130 **2. Description of the tests**

131 **2.1. Tested configurations**

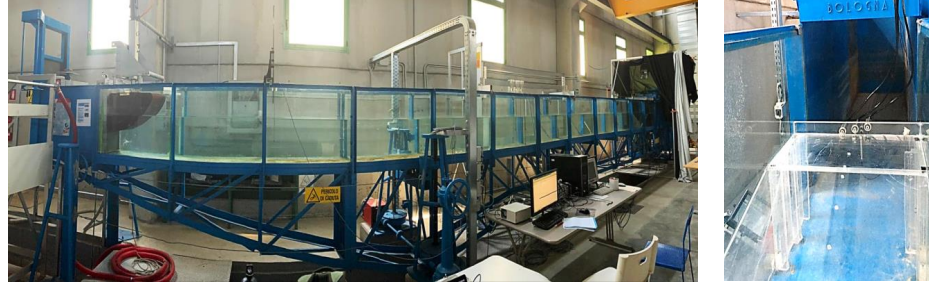
132 The videography-modelling technique presented in this contribution was applied to a total of 184
133 small scale experiments of irregular wave attacks against smooth dike-type structures conducted
134 in the wave flume of the Hydraulic Laboratory of the University of Bologna (Unibo). The
135 experiments consisted of 56 tests of wave run-up and wave overtopping at dikes characterized
136 by a plain, trapezoidal cross-section and 128 further tests of wave overtopping and wave loads
137 against the same dikes upgraded with crown walls, included in correspondence of the on-shore
138 edges.

139 A wide-angle picture of the wave flume and a front-view picture of the structure with the crown
140 wall installed at the end of the flume are shown in Figure 1.

141

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

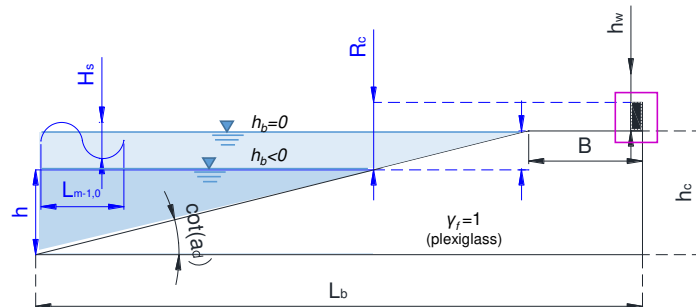
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0065019



142
143 Figure 1. Wide-angle view of the wave flume (left) and front-view of the structure with crown wall
144 installed at the end of the flume (right).

145

146 Figure 2 shows the scheme of the cross-section of the tested configurations, providing reference
147 to the symbols and parameters which will be adopted hereinafter. Following the EurOtop (2018)
148 classification of the structure types, and in line with published articles on the Unibo experiments
149 (e.g. Formentin and Zanuttigh, 2019), the crest width of the dikes is hereinafter considered and
150 schematized as a berm (width B , emergence to the swl $h_b < 0$).



151
152 Figure 2. Schematic layout of the tested configurations (not in scale) with reference to the
153 symbols and parameters.

154

155 All the tested cross-sections presented the same structure height from the bottom of the channel
156 to the berm level (h_c in Figure 1) equal to 0.35 m and the same distance between the berm off-
157 shore edge and the wave maker ($L_b - B$) equal to 10.75 m. The range of variations of all the other
158 structural parameters were made varying: dike slope $\cot(\alpha_d) = 2$ and 4; berm width $B = 0.15$ and
159 0.30 m; berm relative emergence to the wave height, $h_b/H_{m0} = 0, -0.5$ and -1 (this latter case, only
160 for tests without walls); crown wall height $h_w = 0, 0.04$ and 0.05 m.

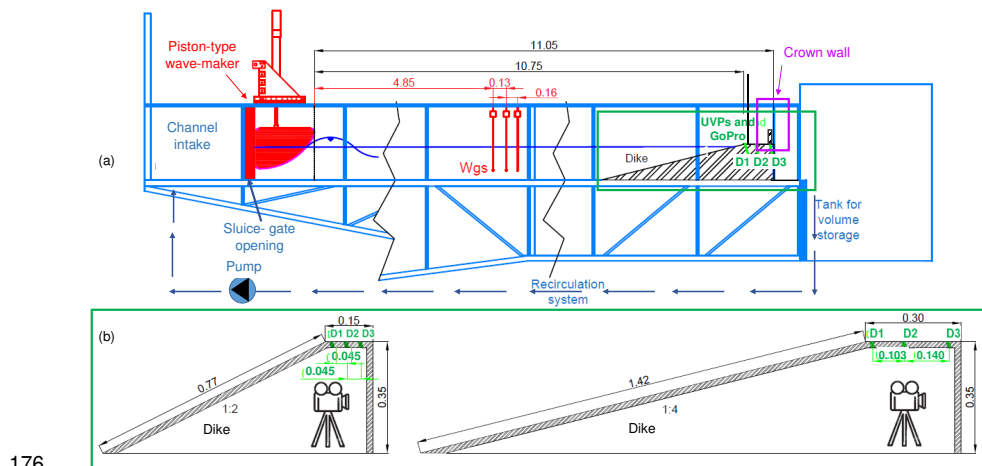
161 The plain dike configurations were subjected to 6 target wave conditions realized by varying the
162 significant wave height, $H_{m0} = 0.04, 0.05$ and 0.06 m, and the spectral wave periods $T_{m-1,0}$ to get
163 the 2 target wave steepnesses $s_{m-1,0} = H_{m0}/L_{m-1,0} = 0.03$ and 0.04 , where $L_{m-1,0}$ is the wave length

164 calculated from $T_{m-1,0}$. The configurations including also crown walls were instead subjected to 4
 165 wave conditions ($H_{m0}=0.05$ and 0.06 m and $s_{m-1,0}=0.03$ and 0.04). The wave attacks were all
 166 characterized by Jonswap spectra, with a peak enhancement factor of $\gamma=3.3$ and consisted of
 167 approximately 500-600 waves (according to the target peak wave period, T_p).

168 The structures were all realized in plexiglass (giving a roughness factor $\gamma_r=1$) and the walls (when
 169 presents) were deeply clamped in the berm to avoid overturning.

170 2.2. Laboratory equipment

171 The wave flume at Unibo, largely described in Formentin and Zanuttigh (2019a) and Gaeta et al.
 172 (2020), is 10 m long, 0.5 m wide and 1.0 m deep. It is equipped by a piston-type wave-maker
 173 with a special cuneiform shape, which generates the waves by its vertical movements. The
 174 scheme of the wave flume, including the installed structures and instrumentation for the present
 175 tests is shown in Figure 3.



176 Figure 3. (a) Scheme of the wave flume, structures and instruments installed; (b) particular of
 178 the dikes ($\cot(\alpha)=2$ and 4), position of the UVPs (D1, D2, D3 in the 2 berms, $B=0.15$ and 0.30
 179 m) and of the camera Measures in m.

180

181 With reference to Figures 1, 2 and 3, the instrumentation adopted in the laboratory consisted of:

- 182 • a channel intake – located upstream the wave-maker and separated from the channel by a
 183 vertical sluice gate with a small opening at bottom and connected to the recirculation flow –
 184 with turbulence dissipation in a still water volume; the system pump-recirculation flow allows
 185 to keep the water level difference in the channel within a ± 4 mm range for each test.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 186 • a water tank downstream the wave flume and connected to the recirculation conduit, to
187 collect the overtopping volumes and discharge, with a precision in the measure of q of
188 approximately $1 \cdot 10^{-5} \text{ m}^3/\text{s}$.
189 • 6 resistive wave gauges (wgs) so installed: the first one in the settling chamber to measure
190 and regulate the recirculation flow; the second one at approximately half wave length
191 ($0.5 \cdot L_{m-1,0}$) from the wave-maker to control the generated free-surface elevation; the
192 following 3 at $\approx 1.5 \cdot L_{m-1,0}$ to separate the incident and reflected waves (the exact position is
193 shown in Figure 3a with red color); the last one in correspondence of the mid-section of the
194 berm width to measure the overtopping layer thickness. The sampling frequency of 100 Hz
195 was selected for all the gauges.
196 • 3 Ultrasonic Doppler Velocimeters (UVPs) installed along the structure berm to record the
197 time series of the vertical profiles of the horizontal flow velocities u and track the free surface
198 elevation h . The positions of the three UVPs, shown in Figure 2b in green color and
199 referenced as D1, D2 and D3, were selected to reconstruct the statistics of h and u in
200 proximity of berm off-shore edge (D1), in the berm mid-section (D2) and close to its in-shore
201 edge (D3).
202 • a 30 Hz full-HD camera (resolution 1080x1920 pixels) employed to film the wave run-up, the
203 flow over the crest, the overtopping process and the wave impacts at the walls (when
204 present); the camera was placed at the exactly perpendicular position to the side window of
205 the wave flume where the dike was located, see Fig. 3b; a “darkroom” was set up around
206 the channel in correspondence of the camera – visible in Figure 1 to the right – to optimize
207 the light condition for filming.
208

209 2.3. Laboratory measurements

210 The following “direct” measurements were performed during the experimental tests:

- 211 • the time series of the free-surface elevations (h) at the wgs installed in the channel;
212 • the average specific wave overtopping discharges (q_{tank} , $\text{m}^3/(\text{ms})$) calculated from the
213 volumes stored in the tank;
214 • the time series of the overtopping flow characteristics (depths h and velocities u) over the
215 berm from the records of the 3 UVPs;
216 • the records of the experiments filmed with the full-HD camera.

217 From the elaboration of the “direct” measurements, the following further results are available:

- 218 • significant incident wave heights (H_{m0}) in front of the structures and bulk wave reflection
219 coefficients (K_r);
220 • the instantaneous overtopping wave volumes for unit width (V , m^3/m) from the integration of
221 the h -signals in the time-domain recorded at D1, D2 and D3 from the reflecting water-air
222 interface of the projected acoustic beams;
223 • the average specific overtopping discharges, q_{Vol} ($\text{m}^3/(\text{m s})$) at D1, D2 and D3 obtained as
224 sum of the corresponding instantaneous V -values to the duration of the test (480 s);

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 225 • the wave celerities (c), the celerities of propagation of the water front above the berm width,
 226 from the coupling of the UVPs records of the h -values at D1 and D3; for each wave
 227 overtopping the berm, the instantaneous wave celerity is calculated as the ratio between the
 228 known-distance between D1 and D3 ($diswg=0.090$ and 0.243 m for $B=0.15$ and 0.30 m,
 229 respectively, see Figure 3b) and the corresponding time lag between detected h wave peaks
 230 at D1 and D3;
 231 • the specific overtopping discharges, q_{cel} ($m^3/(m\ s)$) calculated from the integration of the
 232 wave celerities c with the water depths h ; for shallow water conditions, under the hypothesis
 233 of kinematic wave propagation, flow velocities and wave celerities can be assimilated.
 234

235 The ranges of the main parameters resulting from the experiments are reported in Table 1,
 236 where:

- 237 • $\xi_{m-1,0}$ is the Iribarren-Battjes breaker parameter resulting from the values of the spectral wave
 238 height (H_{m0}) and spectral wave length ($L_{m-1,0}$) measured at the toe of the structure;
 239 • R_c/H_{m0} is the structure relative freeboard (considering the crown wall height) as measured
 240 in the lab, i.e. accounting for the mean swl in the wave flume during the test; the negative
 241 values of $R_c/H_{m0} < 0$ (tests on dikes without walls only) are obtained when the mean swl
 242 increased during the test, giving a slightly submerged berm level;
 243 • the symbol "*" following the q_{tank} -values (tests with walls) indicates that the value reported in
 244 the Table is the minimum non-zero value, whereas some tests actually gave no overtopping
 245 ($q_{tank}=0$);

246 Table 1. Ranges of the main parameters resulting from the experiments. The symbol "*" indicates
 247 that the minimum non-zero value is reported.

	$\xi_{m-1,0}$	R_c/H_{m0}	q_{tank} [m^3/sm]
# tests without wall ($h_w=0$)	[1.19; 3.72]	[-0.19; 1.12]	[1.90e-4; 5.00e-3]
# tests with wall	[1.23; 4.03]	[0.66; 2.51]	[1.09e-5*; 7.00e-4]

248
 249 Further reference to the laboratory measurements and results can be found in: i) Formentin and
 250 Zanuttigh (2018) and Formentin and Zanuttigh (2019a) for the characterization of the data of q ;
 251 ii) Formentin et al. (2019a) and Gaeta et al. (2020) for, respectively, the presentation of the
 252 overtopping flow processes across the berm (data of h , u and V) and a detailed description of
 253 the elaboration technique of the data from the UVPs; iii) Formentin and Zanuttigh (2019b) as
 254 regards the wave coupling technique, the wave celerities and related quantities (e.g., q_{cel}).

255

256 **2.4. Characteristics of the selected tests**

257 The accuracy of the image processing methodology is assessed in Section 4 by showing the
258 results of the cluster analysis applied to the reconstruction of the free-surface elevation for 8
259 tests selected among the whole database.

260 The selected tests, whose characteristics are reported in Table 2, cover the variety of the
261 geometrical configurations tested in the laboratory (the 2 slopes, $\cot(\alpha_d)=2$ and 4; the 2 berm
262 relative submergences $h_b/H_{m0}=0$ and 0.5 and the 2 berm widths, $B=0.15$ and 0.30 m). The 2
263 slopes were selected to specifically investigate the results of the procedure under both breaking
264 ($\cot(\alpha_d)=4$) and non-breaking ($\cot(\alpha_d)=2$) wave conditions, as determining completely different
265 run-up and overtopping flow processes.

266 The wave conditions are the same for all the 8 tests (target wave height $H_s=0.05$ m and target
267 wave steepness $s_{m-1,0}=0.03$) and all the configurations refer to structures without walls. For these
268 tests, the main overtopping flow characteristics were by elaborating the free-surface elevation
269 signals recorded at virtual gauges conveniently set up above the structure berm (Sub-section
270 4.1). The “virtual” overtopping flow characteristics are then compared to the corresponding
271 quantities measured in the lab from “traditional” techniques (Sub-sections 4.2 to 4.4).

272

273 Table 2. Summary of the 8 experiments elaborated with the video-cluster analysis.

Test ID	Test ID	$\cot(\alpha_d)$ [-]	B [m]	h_b/H_{m0}
R00H05s3B30c2	TD1	2	0.30	0
R00H05s3B15c2	TD2	2	0.15	0
R05H05s3B30c2	TD3	2	0.30	0.5
R05H05s3B15c2	TD4	2	0.15	0.5
R00H05s3B30c4	TD5	4	0.30	0
R00H05s3B15c4	TD6	4	0.15	0
R05H05s3B30c4	TD7	4	0.30	0.5
R05H05s3B15c4	TD8	4	0.15	0.5

274

275 **3. Description of the methodology**

276 This Section describes the image-processing methodology developed to reconstruct a series of
277 parameters and data related to the wave overtopping and the wave impact processes from the
278 analysis of the video-records of the experiments conducted at Unibo. The methodology is based
279 on the cluster analysis of the frames and on the automatic extraction of “features” from the
280 images and includes a number of pre- and post-processing techniques specifically designed to
281 improve the automatic detection phase. The image clustering technique, which was originally
282 used by Gaeta et al. (2020) to reconstruct the water depth at dike crest, has been upgraded to
283 meet the following objectives i) tracking the free-surface of the water along the slope and above
284 the berm; ii) reconstructing the shape and the height of the up-rushing jet along the wall during

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

285 the impact events; iii) estimating the amount of air bubbles entrapped in the overtopping flow
286 and beneath the water tongue during the impacts.

287 The open-source toolbox “PRTools” (Pattern recognition Tools,
288 <http://prtools.tudelft.nl/?from=www.website80.com>; Duin and Pekalska, 2015) developed by TU
289 Delft was adopted for the cluster analysis and the features extraction from the videos.

290 All the implemented algorithms and the image filtering techniques are tuned to fit the specific
291 site conditions, the camera characteristics (resolution, frequency of acquisition, etc.) and the
292 environmental features (light, blur, channel shape and size, etc.). However, the ensemble of the
293 methodology might represent a “best practice” exportable to other applications related to the
294 image processing of video-records of experiments on coastal structures.

295 The main steps composing the methodology are illustrated in the following Sub-sections.

296 **3.1. Calibration**

297 The first step prior to any image processing consists in the camera calibration. This step is
298 necessary to evaluate and correct the image distortion due to the camera lens and to extract the
299 information required to convert the data minable from the images from the camera intrinsic
300 coordinate system (pixels) to world points (m). The camera calibration requires a specific video
301 to be performed after each power on of the camera and before starting filming the experiments.
302 The resolution and the size of the calibration video must be the same set up to record the
303 experiments.

304 The calibration of the camera was performed following the approach indicated by Bouget (2015)
305 (see Figure 4): a planar checkerboard was moved in front of the camera in correspondence of
306 the structure in the wave flume. A few images presenting the checkerboard in different positions
307 with respect to the structure were taken. The intrinsic coordinates (pixels) of the corners of the
308 squares are automatically detected (red dots in Figure 4c) and the corresponding world
309 coordinates are generated (Figure 4b) by knowing the metric extrinsic size of the checkerboards
310 (30 mm). Finally, the camera calibration parameters are calculated: the rotation translation matrix
311 and the translation vectors to be used to convert all the data from the videos of the experiments
312 from intrinsic to world coordinates. The camera calibration process involves the evaluation of the
313 skewness and of the tangential distortion and returns the estimation of the mean projection error
314 per image (Figure 4b) due to lens distortion, optical errors and 2D–3D projections. In the example
315 of Figure 4b such error ranges approximately between 1 and 2.5 pixel per image, i.e. between
316 0.3 and 0.75 mm. Generally, it was verified that the projection error was always less than 1 mm
317 per image.

318 To evaluate the accuracy of the calibration process before carrying out and filming the
319 experiments, it is suggested to select a sample frame for validation. By comparing the main
320 geometrical distances (e.g., width of the berm or slope angle of the dike) of the picture to the
321 corresponding real dimensions, the relative errors characterizing the conversion process can be

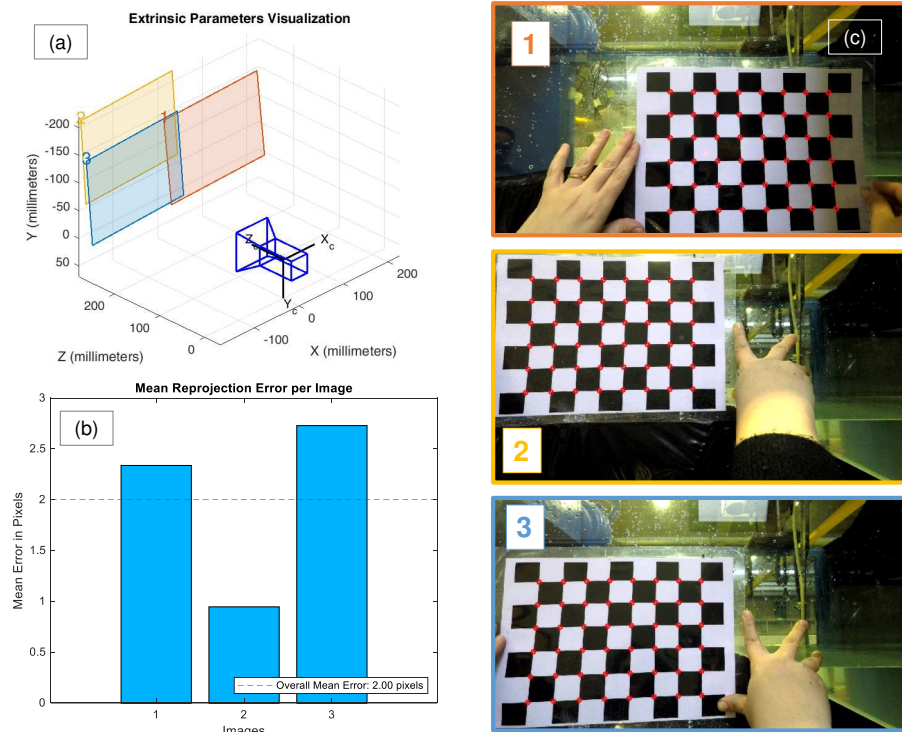
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

322 computed. In the present application, relative errors between 0.6 and 2% were found for the
 323 lengths and between 0.5 and 2.5% for the slope angles.

324 Once calibrated the camera, it was possible to run and film the experiments. For each set of
 325 tests performed after each power on of the camera, the same calibration parameters, conversion
 326 factors and projection errors were assumed.

327



328 Figure 4. (a) Extrinsic camera visualization (world coordinates) of the 3 images used for the
 329 calibration; (b) mean camera projection error per image (pixels); (c) images (1-3) used for the
 330 camera calibration and checkerboard points (red dots) automatically detected.

332

333 3.2. Pre-clustering optimization techniques

334 The pre-clustering techniques were designed to improve the accuracy of the automatic detection
 335 of the different objects (water, air entrainment, structure, etc.) as much as possible. The following
 336 techniques were applied to each frame of each video and in the following order.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 337 i) Image subtraction: the first frame – representing the structure in the channel and the initial
338 swl conditions before running the waves and named “frame 0” hereinafter – was subtracted
339 from each frame of the record. Subtracting an image from another image means subtracting,
340 pixel by pixel, the digital numeric value of the second image from the first one. This process
341 allows levelling uneven sections of an image, such as shadows, reflexes or blurs, and
342 removing all the fixed elements present in all the frames (for example, the shape of the
343 structure or the border of the walls of the channel). After the subtraction, only the elements
344 differing from the “frame 0” are kept, i.e. the water and its evolution.
- 345 ii) Conversion into grayscale, to reduce the size of the image avoiding redundant information
346 stored in the color-scale picture. The use of grayscale images is also recommended to ease
347 the pattern recognition step and for the application of the functions related to the image
348 adjustment (i.e. noise add, light and contrast adjustment, application of morphological filters,
349 etc.).
- 350 iii) Noise add: when large portions of the images result too uniform for the pattern recognition,
351 the training of the cluster model fails. The application of a Gaussian white noise is useful to
352 solve this shortcoming. In the present case, a Gaussian noise with a mean of 0 and a
353 variance from 0.001 to 0.003 was applied. The variance values should be tuned for each set
354 of experiments filmed after the same camera power-on through a trial-and-error process
355 commanded by the cluster training phase. The more uniform the images, the higher the
356 variance of the noise to be included.
- 357 iv) Enhancement of the contrast of the grayscale image using a contrast-limited adaptive
358 histogram equalization (uniform histogram with a contrast enhancement limit = 0.03).

359 Figure 5a shows an example frame as taken from the camera, while Figure 5b shows the same
360 picture as resulting from the application of the pre-processing techniques.

361

362 3.3. Image clustering

363 The clustering is the task of grouping different “objects” (or patterns) of a picture into clusters so
364 that the objects in the same group are more “similar” to each other than to those in other groups
365 (clusters). An ideal clustering would recognize, for example, all the pixels composing the water
366 as actually belonging to the same cluster. An automatic clustering can detect patterns
367 (similarities among objects) in the images that are “invisible” to the human eye and extract
368 quantitative information otherwise hardly accessible (e.g., estimate the amount of air entrapped
369 in the water phase).

370 Numerous clustering algorithms are available and several categorizations can be done upon
371 different aspects (Duda et al., 2012). In the present study, the image clustering was performed
372 with the toolbox PRTTools by selecting the Expectation–Maximization (EM) algorithm (Dempster
373 et al., 1977): it is an unsupervised, probabilistic, iterative method to partition data into clusters
374 based on maximizing the likelihood to find the statistical parameters of underlying sub-

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

375 populations in the dataset. Once defined the total number of clusters, the EM algorithm classifies
376 the objects into temporary clusters (E step), which are re-assigned iteratively after the evaluation
377 of the "similarity" (or distance) of the objects within the same cluster (M step). The iterations may
378 stop after a certain number of attempts (early stopping) or when the subdivision of the objects
379 into clusters is stable, i.e. when an object is definitely assigned to a cluster. The EM algorithm
380 performs a partitioning clustering and, requiring the *a priori* definition of the number of clusters,
381 relies on the analyst's knowledge to classify the clusters in a meaningful way. According to the
382 syntax of PRTools, the labeling "crisp" was selected to apply a boosted version (generalized *k*-
383 *means* algorithm) of the traditional EM algorithm.

384 To perform a cluster analysis of a video record, it is necessary to train a cluster model map, i.e.
385 to create an example of clustering of the objects composing a frame to be used as reference
386 (model map). The frame selected for the training must be representative of all the frames of the
387 record and, in particular, must include all the clusters of objects that may appear in a frame
388 (water, solid structure, air entrapment, etc.). The number of clusters to be identified in each frame
389 depends on the image quality (e.g., light exposure, flow patterns over the crest, drops on the
390 walls of the channel), on the expected turbulence level and – in general – on the objective of the
391 analyses. The summary of the features selected for the training of the cluster models is reported
392 in Table 3.

393 The analysis of the experiments on dikes without walls were focused on the tracking of the free
394 surface only, disregarding the analyses of the areas of the flow affected by air entrainment. For
395 these tests, the frames for the training were therefore selected to reproduce an instant of wave
396 overtopping characterized by a pseudo-steady overflow process and no (or modest) turbulence
397 and no bubbles. A number of 2 clusters was setup for the training of the cluster model map,
398 which basically corresponds to the objects "water" and "non-water". The edges of such clusters
399 represent the interfaces water-structure and water-air, and this latter one can be directly taken
400 as free-surface. An example of a frame selected for training is shown in Figure 5a (original
401 picture). The cluster model map of this frame returned with the training is shown in Figure 5c:
402 here, the 2 clusters are clearly distinguished and the free-surface interface is sharp.

403 On the contrary, the video-analysis of the experiments on dikes with walls was also aimed at
404 tracking the profile of the water jets along the walls and at detecting the areas of the flow more
405 frequently subjected to turbulence and air entrainment. To this purpose, the training frames were
406 selected among images representing a fully-broken wave propagating above the berm,
407 characterized by a high air entrainment rate. An example of training frame is reported in Figure
408 6a. This example shows that the bi-phase flow is characterized by different shades of color and
409 brightness. To "capture" all the different shades and to correctly cluster all the area of bi-phase
410 flow as such, it was necessary to define "redundant" clusters of objects. Specifically, 5 clusters
411 were set up for the experiments on walls to ideally group the image objects into the classes
412 "water", "structure" and "air entrainment", where the latter class "air entrainment" was expected
413 to appear into up to 3 clusters. The association cluster-objects varied for each experiment,
414 because each experiment requires the training of a specific cluster model map, and the training
415 is a random, iterative process that vary at each run. Figure 6b shows the trained cluster model

416 map of Figure 6a: in this map, the areas of the flow that may be concerned by air entrainment
 417 correspond to the clusters 3 and 5 (light blue and yellow, respectively). For other experiments,
 418 3 clusters were necessary to correctly detect the areas with air.

419 Independently of its level of complexity, the association cluster-objects relies on the human
 420 supervision in the end. All the cluster model maps must be interpreted by the human intelligence
 421 after the training to associate one object to one (or more) clusters.

422

423 Table 3. Summary of the features adopted for the cluster training.

Dataset	Training algorithm	Number of clusters	Analyses performed (features)
# tests without wall	EM-crisp	2	free-surface tracking
# tests with wall	EM-crisp	5	free-surface tracking; bi-phase flow area estimation; wave-impact reconstruction (analysis in progress)

424

425 3.4. Post-clustering optimization techniques

426 After the image-clustering, the frames are transformed into cluster maps similar to the examples
 427 of Figures 5c or 6b, where each pixel is associated to a label. Post-clustering techniques were
 428 designed to convert the information embedded in the cluster maps of the frames into data for
 429 further elaboration.

430 3.4.1. Free surface tracking

431 For the free-surface tracking and the reconstruction of the shape of the water jet during the wave
 432 impacts at the walls, the following post-clustering techniques were applied.

433 i) Segmentation of the areas of the image automatically labelled as “water”: the segmentation
 434 returns a binary image containing 1 where the function finds edges and 0 elsewhere. The
 435 selected image detection method is based on the “Canny” algorithm (Canny, 1986). Figure
 436 5d shows in red the edges of the “water” resulting from the segmentation of one frame
 437 relative to a test on dikes without walls. In Figure 5d, the contour of the free-surface is
 438 correctly detected, though several, unwanted areas of the frame (drops on the walls of the
 439 flume or other disturbing elements) have been wrongly labelled as “water”, as marked by the
 440 red contour.

441 ii) To reduce the number of gaps in the edge of the free-surface, the filling of the “holes”
 442 procedure was performed.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 443 iii) To remove the disturbing elements, an ad hoc filter based on the area of the “connected
444 components” of the image was designed: this filter removes all the connected components
445 that have fewer than 200 pixels, where 200 is approximately the largest realistic area of the
446 drops on the walls. The preliminary operation of holes filling (ii) ensures that all (or most) of
447 the adjacent pixels composing the edge of the free-surface are connected and that the
448 “water” is not removed by the filter.
- 449 iv) To remove specifically-shaped disturbing elements in the images “survived” to the extraction
450 of the connected components, the “morphological filtering” was applied. In a morphological
451 operation, each pixel of the image is adjusted based on the value of other pixels in its
452 neighborhood. By choosing the size and shape of the neighborhood, it is possible to
453 construct a morphological operation that is sensitive to specific shapes in the input image.
454 The following 2 morphological operations were built: 1) elimination of all the parts of the
455 figure containing a vertical element of at least 4-pixel length (image erosion), to remove the
456 edges of the walls of the channel or of the wires; 2) dilation of all the parts of the figures
457 containing a horizontal element at least 5 pixel-long (image dilation) to enhance and connect
458 the edges of the berm. The result of the subsequent holes filling, extraction of the connected-
459 components and morphological filtering of the edges of the “water” is shown in Figure 5e. In
460 this Figure, it is possible to detect all the edges removed from the original output of the
461 image clustering and segmentation (Figure 5d) and the marked profile of the structure berm.
- 462 v) Automatic procedure for the free-surface tracking. An if-else, customized procedure was
463 developed to reconstruct the effective profile of the free-surface, abscissa after abscissa,
464 from the dike slope to the wall height (when present), disregarding the edges of all the
465 remaining objects improperly classified as water and which were not eliminated by the
466 procedures ii) to iv). Such procedure is based on the definition of an initial time condition
467 and follows a sort of “forward scheme” to identify the nearby points of the free-surface, based
468 on the assumption of the continuity of the free-surface profile itself. The initial condition is
469 the actual profile of the free-surface of the first frame, which is prompted by the direct, human
470 identification of the coordinates through a graphical user interface. The initial condition is
471 used for comparison for the following frame: for each abscissa, the ordinate of the edges of
472 the “water” is kept as free-surface of the current frame if it is “sufficiently” close (± 10 pixels)
473 to the corresponding ordinate of the initial condition. The free-surface is updated for the
474 current frame and is used for comparison for the following one, and so on. When no value
475 of the edge of the “water” of the current frame is sufficiently close to the value of the free-
476 surface at the same abscissa of the previous frame, a comparison among adjacent pixels of
477 the current frame is performed (forward scheme). Starting from the left-most abscissa (wave
478 inlet), the ordinate of the nearby pixel of the “water” edges is kept if sufficiently close (± 10
479 pixels) to the ordinate of the previous abscissa, otherwise it is discarded. If no points are
480 accepted for a certain abscissa, a value of “NaN” is assigned for that position. The “NaNs”
481 correspond to holes in the profile of the free-surface. The value of ± 10 pixels was set after
482 a trial-and-error procedure and based on the time sampling frequency of the video (30 Hz)
483 and the spatial resolution of the frame (1080x1920 pixels).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

484 vi) Filtering. Median-filters and adjacent-pixel-filters were designed and applied to the records
485 of the free-surface as resulting from step v) to fill the holes (NaNs) and smooth the free-
486 surface profile, removing spikes and discontinuities. The time series of the filtered free-
487 surface elevation values, for the whole duration of the video, is the final output of the post-
488 clustering techniques applied to each test. Figures 5f and 6c display the filtered free-surfaces
489 (blue crosses) for the corresponding frames. The comparison between Figure 5d and Figure
490 5f shows the effects of the application of the steps i)-vi) to the outputs of the cluster analysis.

491 3.4.2. Bi-phase flow area detection

492 The detection of the areas of the flow subjected to entrainment of air bubbles was performed for
493 the experiments on walls only, through the application of the following post-clustering
494 techniques.

495 i) Segmentation of the areas of the image labelled as "air entrainment". With reference to the
496 example of Figure 6b, the areas belonging to the clusters 3 and 5 were labelled as "air
497 entrainment" upon eye examination of the cluster map.

498 ii) Extraction of the connected components to clear up the areas of "air entrainment" from all
499 the connected components consisting of less than 20 pixels of area extent (small drops
500 escaped from the image subtraction).

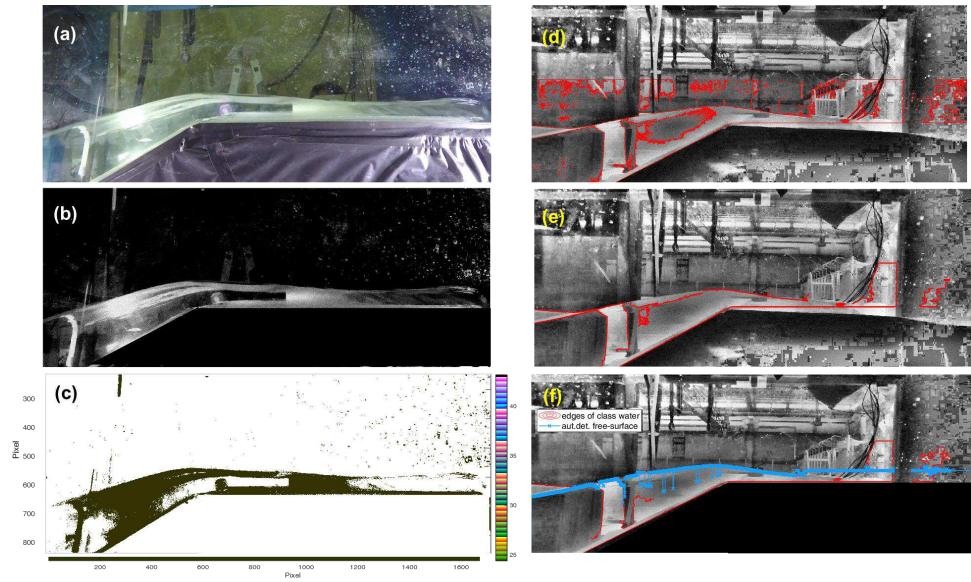
501 iii) Morphological filtering of all the mono-dimensional, 10-pixel-length features, to remove the
502 elements composing the contour of the structure which were wrongly assigned to the "air
503 entrainment" clusters because their intensity was similar to the one characterizing the bi-
504 phase flow.

505 iv) If-else, loop-procedure for the actual estimation of the areas of the flow potentially concerned
506 by air entrainment. Such procedure processed, abscissa after abscissa, all the pixels
507 labelled as "air entrainment" survived to the filtering procedures. Each pixel was accepted
508 as a potential air-element if its ordinate was lower than the ordinate of the water free-surface
509 for the same abscissa. Therefore, the detection of the bi-phase flow areas is necessarily
510 posterior to the free-surface tracking.

511 The example of Figure 6c shows in red the contour of the areas of the flow characterized by air
512 entrainment as returned after the application of the steps i)-iv). The eye analysis of this picture
513 confirms that these areas do actually correspond to the areas of the flow affected by air
514 entrainment. The eye-examination of tens of frames randomly selected from the video records
515 of the tests under different light and wave attack conditions has confirmed that the procedure
516 detects the areas of bi-phase flow with the same accuracy shown in Figure 6c.

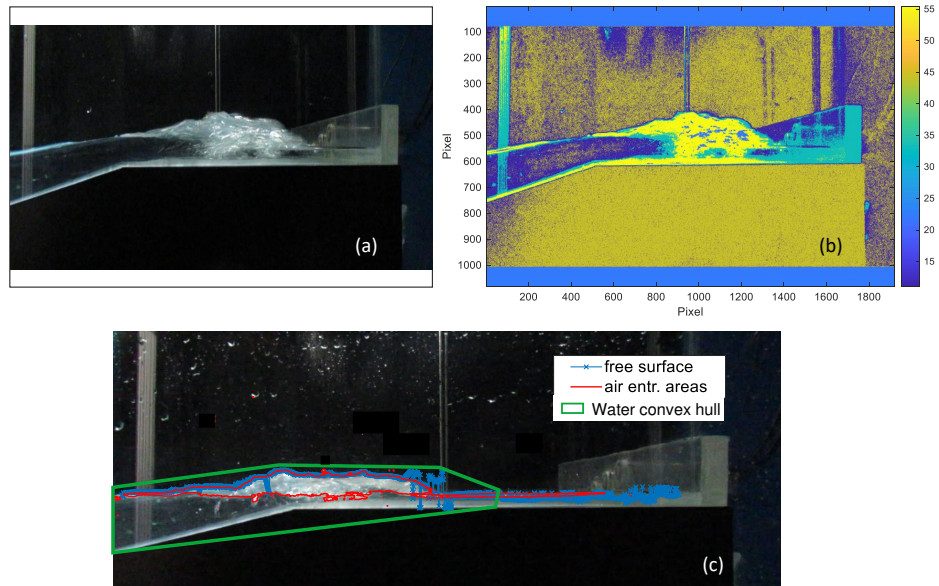
517 The main application of the bi-phase flow detection analysis was the estimation of the percentage
518 of the water flow affected by air entrainment. Such percentage was calculated for each frame of
519 each experiment, by dividing the extent of the areas of air entrainment to the total area of the
520 flow, estimated as the extent of the convex hull of the object "water" (green line in Figure 6c).
521 The details of the air entrainment analysis are illustrated in Section 5.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



522
523 Figure 5. a-b-c) Example frame selected for training the cluster model for a test without wall: a) Image as taken from the camera; b) image as
524 resulting after the application of the pre-processing techniques; c) clustered-model map of the image. d-e-f) Example frame processed with the
525 cluster analysis: d) edges of the cluster "water" resulting from the image segmentation (red); e) edges of the cluster "water" (red) after the
526 application of the post-clustering techniques; e) profile of the free-surface elevation as result of the whole image processing procedure (blue).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



527
528 Figure 6. a-b) Example frame selected for training the cluster model for a test with wall: a) image in real-scale colours and light conditions as
529 taken from the camera; b) clustered-model map of the image. c) Example frame processed with the cluster analysis showing: the profile of the
530 filtered free-surface elevation (blue); the contour of the object "air" (red) after the application of the post-clustering techniques; convex hull of
531 the object "water" (green).

532 4. Results and validation of the methodology

533 4.1. Image elaboration and creation of the virtual gauges

534 To elaborate the results of the videography for each of the tests reported in Table 2, it is firstly
535 necessary to convert the values of the free-surface elevations resulting from the video analysis
536 from pixels to “real world” coordinates. The conversion is carried out based on the parameters
537 of calibration of the camera (see Sub-section 3.1).

538 2 “virtual gauges” were defined to extract the time series of the overtopping flow depths (h) in
539 correspondence of the off-shore (OE) and in-shore (IE) edges of the berm and, specifically in
540 correspondence of the position of the UVPs, D1 and D3 (see Figure 3b). These time series were
541 elaborated to derive the statistics of practical interest (mean, standard deviation, upper-2%
542 percentiles, $h_{2\%}$, maximum envelopes), to reconstruct the individual overtopping volumes (V), to
543 calculate the wave celerities (c) and to estimate the average and instantaneous wave
544 overtopping discharges (q). The comparison among these quantities and the corresponding
545 measurements from the lab is used to verify the accuracy of the results of the videography
546 analysis and assess their uncertainty (Sub-sections 4.1.1 to 4.1.3). Figure 7a illustrates, for an
547 example frame, the position of the virtual gauges and Figure 7b provides 2 exemplary time series
548 of the h -values reconstructed at the 2 virtual gauges from the video-cluster analysis.

549 The quantities c , V and q were calculated from the time series of the h -values by applying the
550 procedure of threshold-down-crossing and wave-coupling developed by Formentin and Zanuttigh
551 (2019b). This procedure was applied to both the virtual signals and the records of the UVPs. The
552 procedure consists in the identification of the individual overtopping events based on the
553 definition of thresholds (shown as circles in Figure 7b) and in the coupling of the events at the 2
554 gauges based on the comparison of the shape of the h -signals and on the time lag occurring to
555 an event to propagate from the first to the second gauge, knowing the distance between the 2
556 gauges (*diswg*) themselves. An overtopping event is identified in the time domain by 2
557 consecutive crossings of a threshold (namely, t_i and t_{i+1}). By integrating the h -signal between t_i
558 and t_{i+1} , it is possible to calculate the volume (V_i) of each i -th overtopping event:

$$559 \quad V_i = \int_{t_i}^{t_{i+1}} h(t) dt, \quad [\text{m}^3/\text{m}] \quad (1)$$

560 The sum of the individual volumes for the whole duration of the experiment, gives the total
561 volume of wave overtopping, V_{tot} :

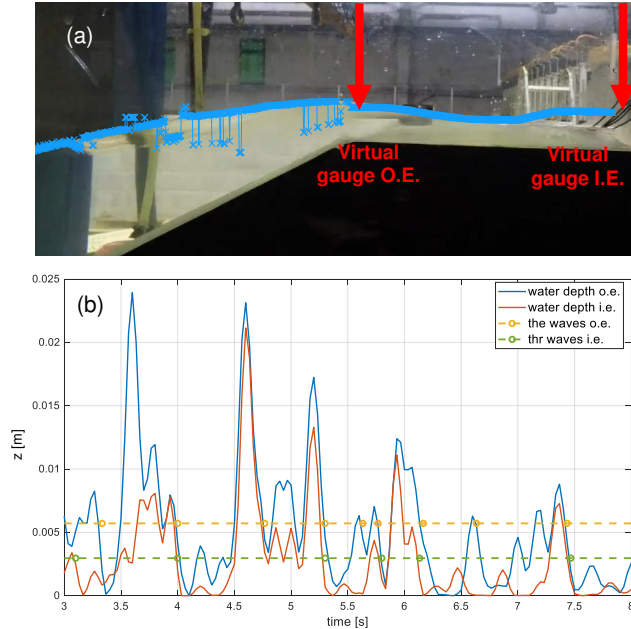
$$562 \quad V_{tot} = \sum_{i=1}^{Now} V_i, \quad [\text{m}^3/\text{m}] \quad (2)$$

563 where *Now* is the number of overtopping waves (i.e. of the overtopping volumes). By dividing
564 V_{tot} for the duration of the experiment (480 s), it is possible to estimate the “volumetric” average
565 specific wave overtopping discharge, q_{vol} :

$$566 \quad q_{vol} = \frac{V_{tot}}{t_{tot}}, \quad [\text{m}^3/(\text{s}\cdot\text{m})] \quad (3)$$

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



567

568 Figure 7. a) Definition of the “virtual gauges” for the extraction of the overtopping flow depths, h ,
 569 from the videography. The blue crosses represent the free-surface as resulted for this example
 570 frame after the application of the filtering techniques. b) Example time series of the overtopping
 571 flow depths (h) at the 2 virtual gauges (blue, OE edge; orange IE edge). The yellow and green
 572 dashed lines are the threshold for the wave detection at the o.e. and i.e., respectively.

573

574 The quantities V_i , V_{tot} and q_{vol} can be calculated for each test and for both the virtual gauges.
 575 The comparison between the corresponding quantities at the 2 virtual gauges is used as
 576 evidence of the mass conservation. The same quantities are also compared to the corresponding
 577 results from the elaboration of the UVPs signals (see Sub-sections 4.3 and 4.4).

578 The wave coupling performed by the procedure of Formentin and Zanuttigh (2019b) associates
 579 the instants of threshold-down-crossing of each overtopping event detected at the 2 virtual
 580 gauges (and at the 2 UVPs D1 and D3) namely t_{i1} for the virtual gauge at the OE and t_{i2} for the
 581 virtual gauge at the IE. Given $diswg$, it is possible to calculate the celerity of propagation of each
 582 overtopping event:

$$583 \quad c_j = \frac{diswg}{t_{i2} - t_{i1}}, \quad [\text{m/s}] \quad (4)$$

584 For each experiment, the mean, the maximum and the upper 2%-values of c can be derived from
 585 the statistical analysis of the c_i -values. The wave celerities can be used, to some extent, as
 586 estimators of the overtopping flow velocities u (Formentin and Zanuttigh, 2019b).

20

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

587 Finally, the celerities can be used to estimate the values of the wave overtopping discharges of
588 each overtopping event ($q_{cel,i}$) and the average overtopping discharges (q_{cel}):

$$589 \quad q_{cel,i} = C_i \cdot \frac{\int_{t_i}^{t_{i+1}} h(t) dt}{t_{i+1} - t_i}, \quad [\text{m}^3/(\text{s}\cdot\text{m})] \quad (5)$$

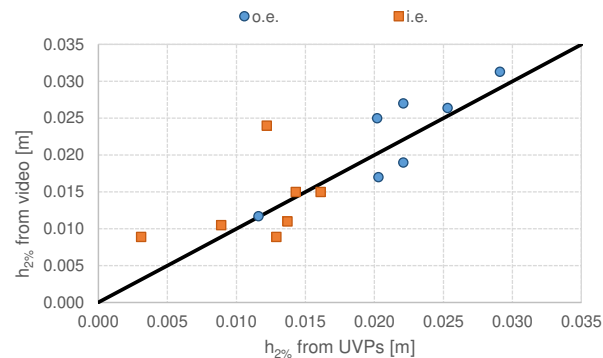
$$590 \quad q_{cel} = \frac{\sum_{i=1}^{N_{ov}} q_{cel,i}}{t_{tot}}, \quad [\text{m}^3/(\text{s}\cdot\text{m})] \quad (6)$$

591 where the second term on the right-side of Eq. (5) is the time-average of the h -values of the i -th
592 overtopping event. The average q_{cel} can be compared to q_{Vol} (both the quantities from the
593 videography and from the UVPs) and to q_{tank} (see Sub-section 4.2 and Figure 8).

594

595 4.2. Overtopping flow characteristics

596 Figure 8 illustrates the comparison among the values of $h_{2\%}$ at the OE and IE as resulting from
597 the measurements of the UVPs in the lab (abscissas) and from the videography (ordinates) for
598 the tests TD1-TD8 of Table 2. The average relative errors among the videography and the UVP
599 values of $h_{2\%}$ are +4.1% and +36% for the OE and the IE respectively. In both cases, the positive
600 sign indicates that the videography tends to give higher estimations of the overtopping flow
601 depths than the UVPs. The modest error at the OE actually indicates a very good agreement
602 among the results from the 2 techniques, while the higher uncertainty at the IE can be explained
603 with the single outlier of Figure 8 relative to the test TD2 (relative error +90%). It was verified
604 that the cause of the outlier is related to the particularly bad lighting conditions affecting the
605 quality of the video records, especially in correspondence of the outlet of the wave flume (i.e.
606 closer to the IE).



607

608 Figure 8. Values of $h_{2\%}$ at the OE (blue circles) and IE (orange squares) measured in the lab with
609 the UVPs (abscissa) and reconstructed from the cluster analysis of the video of the experiments
610 (ordinate). Tests TD1-TD8 of Table 2.

611 **4.3. Overtopping discharge**

612 Figure 9 provides the comparison among the values of q measured in the water tank in the lab
 613 (q_{tank} , abscissa) and the corresponding estimations derived from the videography for the tests
 614 TD1-TD8. In this Figure, the triangles and the diamonds represent the values of q calculated
 615 from, respectively, the overtopping volumes (q_{vol}) and the wave celerities (q_{cel}). The agreement
 616 among the videography and the water tank values is quantified by the relative errors +19% and
 617 -25% in case of q_{vol} and q_{cel} , respectively. Whereas the absolute value of the errors is similar,
 618 the signs reveal that q_{vol} and q_{cel} tend to respectively overestimate and underestimate the
 619 overtopping discharges.

620 Similarly to the case of $h_{2\%}$ of Figure 8, the average overestimation observed for q_{vol} is induced
 621 by the single outlier relative to the test TD2, clearly visible in Figure 9 (relative error +78%). On
 622 the contrary, the underestimation trend associated to q_{cel} is observed for almost all the tests, and
 623 in particular when $q_{tank} > 3 \cdot 10^{-3}$. These underestimations are probably due to the higher level of
 624 complexity of the procedure necessary to extract the q_{cel} -values with respect to q_{vol} , involving the
 625 coupling of the overtopping events between the 2 virtual gauges. It is likely that some of the
 626 waves propagating faster between the OE and the IE are not caught by the video recording at
 627 30 Hz. Indeed, Figure 7b shows that the time-lag between the records of h at the OE and at the
 628 IE is generally very small and, in some cases, the overtopping events identified at the OE are
 629 nearly coincident to the events identified at the IE (compare, for example, the peaks of the events
 630 recorded at the 2 virtual gauges around 4.6 and 5.2 s). The upper limit in the wave coupling is
 631 imposed by the Nyquist frequency, which is one-half the recording frequency, i.e. 15 Hz. Since
 632 2 frames are necessary to perform the wave coupling, the upper-limit is reduced of a further one-
 633 half. Eventually, the maximum wave celerity that can be caught by the video record can be
 634 evaluated as follows:

635
$$c_{max} = \frac{\text{acquisition frequency}}{4} \cdot diswg. \quad (7)$$

636 For the Unibo experiments, $c_{max} = 7.5 \cdot diswg = 0.675$ m/s in case of $B=0.15$ m ($diswg=0.09$ m, see
 637 Fig. 3b) and 1.82 m/s in case of $B=0.30$ m ($diswg=0.243$ m). As shown in Formentin et al. (2019),
 638 the maximum wave celerities values calculated from the coupling of the signals at the UVPs were
 639 around 1.2 m/s. Therefore, the limit of c_{max} mainly affects the tests with $B=0.15$ m.

640 The limit by Eq. (7) affects in turns, the values of the overtopping discharges q_{cel} , which depends
 641 on the c -values. Nonetheless, Figure 9 shows that the q_{cel} -values follow the same trend of q_{tank} ,
 642 revealing that the elaboration of the video records works in principle, and the limit is given by the
 643 camera frequency of acquisition. Based on Eq. (7), to perform the wave coupling and estimate
 644 the related derived quantities, it can be recommended to set the minimum value of the frequency
 645 of acquisition of a camera as follows:

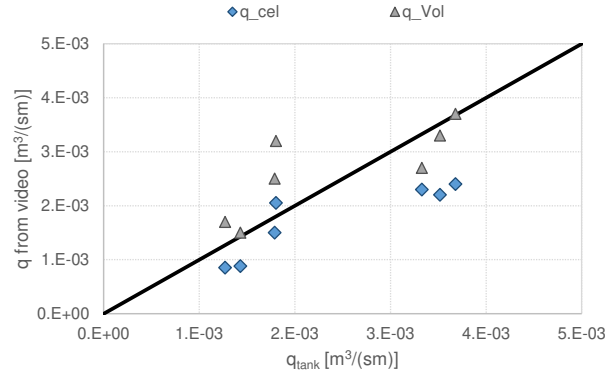
646
$$(\text{acquisition frequency})_{min} = 4 \cdot \frac{c_{max}}{diswg}. \quad (8)$$

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

647 For the Unibo tests, being the observed $c_{max} \approx 1.2$ m/s, (acquisition frequency) $_{min} \approx 50$ Hz for the
 648 most severe condition of $B=0.15$. In case the condition imposed by Eq. (8) cannot be addressed,
 649 it is recommended to prefer the estimations of q from q_{vol} rather than q_{cel} .

650



651

652 Figure 9. Values of q (ordinate) reconstructed from the cluster analysis of the video of the
 653 experiments and measured in the water tank of the lab (abscissa). The blue diamonds refer to
 654 the q_{cel} values, while the grey triangles refer to the q_{vol} values. Tests TD1-TD8.

655

656 4.4. Overtopping volumes

657 The results of the analyses of the overtopping volumes extracted with the videography are
 658 reported in Tables 4 and 5 and in Figure 10.

659 For each of the tests TD1-TD8, Table 4 provides the comparison among the q_{vol} -values at the
 660 OE and at the IE, the relative error between the quantities, the average relative error (μ) and the
 661 relative standard deviation ($\sigma\%$) characterizing the distribution.

662 The results of Table 4 show that the q_{vol} -values calculated at the OE and at the IE differ on
 663 average of -14%, with a standard deviation $\sigma\%=28\%$. In most cases the higher estimations of q
 664 are derived at the OE: this slight decay of q_{vol} is explained with the decay of the overtopping flow
 665 depths h along the berm, which is quantified of $\approx 30-40\%$ (see Figure 8) and which is in line with
 666 the expectations from the literature. The EurOtop (2018) manual indeed indicates that "the flow
 667 thickness h decreases of approximately 1/3 with respect to the value at the off-shore edge".
 668 Hence the results of the videography fulfill the continuity equation, because the variations of q_{vol}
 669 are due to the physical evolution of the h -values and not to a shortcoming of the methodology.

670

671 Table 4. Values of q_{vol} calculated from the videography at the OE and IE for the same tests and
 672 corresponding relative errors; average relative error (μ) and relative standard deviation ($\sigma\%$).

Test	q_{vol} OE [m ³ /(sm)]	q_{vol} IE [m ³ /(sm)]	rel. err. [-]
TD1	3.70E-03	2.70E-03	-0.27
TD2	3.30E-03	2.50E-03	-0.24
TD3	3.20E-03	3.30E-03	0.03
TD4	2.50E-03	2.40E-03	-0.04
TD5	2.70E-03	2.20E-03	-0.19
TD7	1.70E-03	1.50E-03	-0.12
TD8	1.50E-03	1.31E-03	-0.13
μ			-0.14
$\sigma\%$	0.28		

673

674 Table 5 and Figure 10 refer to the single test TD4 and compare the results from the videography
 675 to the elaborations of the data from the UVPs and (when possible) from the water tank.

676 Figure 10 reports in a bi-logarithmic chart the probability distributions of the V -values
 677 reconstructed from the cluster analysis (blue circles) and from the water-air interface as detected
 678 by UVPs (grey triangles). According to the literature, the distributions of the V -values can be
 679 approximated with a Weibull's function, characterized by the shape factor b . Following Formentin
 680 and Zanuttigh (2019b), the values of b were calculated as the slope of the linear fitting of the
 681 higher 20% overtopping. The 2 distributions present very similar shapes and almost the same
 682 changes of slope. In both cases, the 2 lower volumes are similarly detached from the main tend
 683 and the same consideration applies for the upper tail of the distributions.

684 Table 5 reports the number of overtopping waves (Now), the total and the mean volumes (V_{tot}
 685 and V_{bar}) and the Weibull's shape factors b as calculated from Figure 10 (1.17 for the UVPs, 0.93
 686 for the videography and not available for the water tank). The value of Now (not available for the
 687 water tank) differs of only 9 units (4-5%) among UVPs and videography, a deviation which is
 688 included in the measurement error of both the techniques. Similar considerations can be done
 689 for the values of V_{tot} and V_{bar} . In this case, the agreement among the videography and the
 690 traditional technique is even better, being the differences of 2 m³/m (2%) and of 8·10⁻⁵ m³/m (2%)
 691 for V_{tot} and V_{bar} respectively.

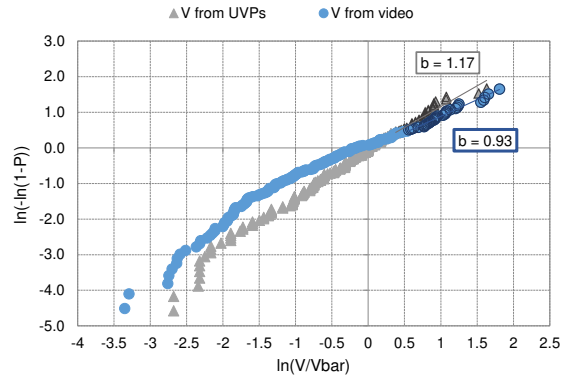
692

693 Table 5. Number overtopping waves, total and mean volumes and Weibull's shape factors b
 694 resulting from the distribution of the individual overtopping volumes of the test TD4.

	Now	V_{tot} [m ³ /m]	V_{bar} [m ³ /m]	Weibull's b factor
Values from the water tank	-	0.85	4.50e-3	-
Values from the UVPs	194	0.79	4.03e-3	1.17
Values from the video-cluster analysis	183	0.83	4.58e-3	0.93

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



695

696 Figure 10. Probability distributions of the individual overtopping volumes of the test TD4 as
 697 reconstructed from the cluster analysis of the video of the experiments (blue circles) and from
 698 the UVPs (grey triangles). The b -values refer to the linear fitting of the upper 20% volumes.

699

700 5. Application of the methodology to the detection of the air 701 entrainment

702 The detection of the flow areas subjected to turbulence and possible air entrainment would
 703 represent a key information for a number of applications, from the localization of the wave
 704 breaking, to the calibration of bi-phase numerical models (a.o., Gaeta and Lamberti, 2015), to
 705 the characterization of the wave impacts at walls, whose violence and intensity is directly
 706 connected to the aeration levels beneath the imping wave during the impact (a.o., Bullock et al.,
 707 2007).

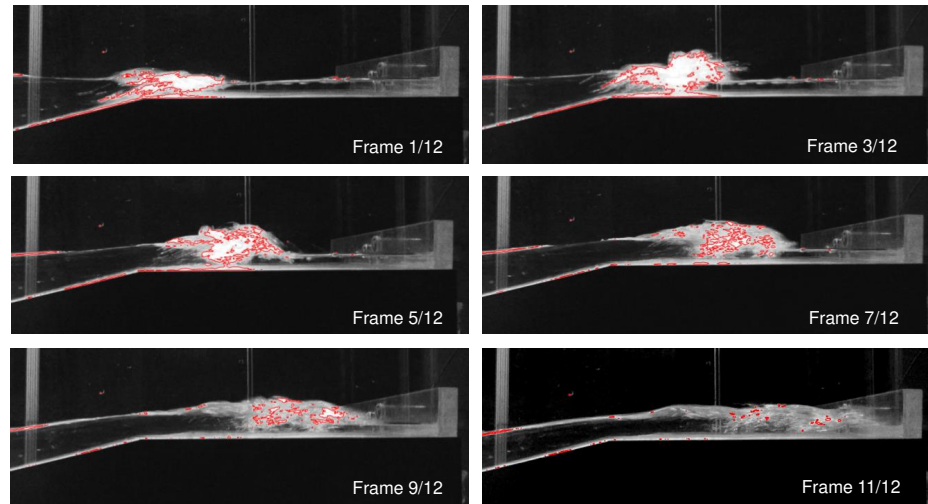
708 As illustrated in Sub-section 3.4.2, the video-cluster analysis of the experiments was here applied
 709 to the estimation of the percentage of the flow areas characterized by air entrainment, in order
 710 to derive maps of the spatial and time map distribution of the bi-phase flow zones during the
 711 wave overtopping. The derivation of the space-time air entrainment maps is illustrated in the
 712 following, with reference to an example overtopping event and an example overtopping
 713 experiment. For sake of simplicity, this analysis is limited to the flow along the slope and the
 714 berm. The analysis of the flow before and over the wall and the study of the wave impact are
 715 objects of separate research.

716 The clustering of a frame provides a classification map of the flow, where each pixel composing
 717 the frame is labelled as "water", "air entrainment" or other. Focusing on the air entrainment
 718 analysis, the labelling can be simplified to a "0-1" classification, where 1 is assigned to the pixels
 719 recognized as air entrainment and 0 otherwise. This information can be firstly processed to
 720 calculate the extent of the areas of air entrainment and localize their position in each frame.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

721 The analysis of multiple air entrainment maps derived from consecutive frames can provide
 722 further qualitative and quantitative information about the concentration of the entrapped-air
 723 areas. Figure 11 shows an example of the time-stack evolution (12 frames, 0.4 s) of the air
 724 entrainment areas detected during an example overtopping event, from the generation of the
 725 first bi-phase flow zones around the berm off-shore edge (frame 1), to the collapse and breaking
 726 of the wave above the berm (frames 3 and 5), to the development into the bore flow (frames 7
 727 and 9) and the dissipation of the last bi-phase flow areas just before the wall (frame 11). These
 728 images qualitatively show that the extent of the areas characterized by air entrainment tends to
 729 increase during the first part of the overtopping event, reaching a maximum at the wave breaking,
 730 while it starts decreasing during the bore flow. The images show also that before and during the
 731 wave breaking the air entrainment areas are relatively "compact" and their shape is influenced
 732 by the profile of the wave crest; on the contrary, after the wave breaking, the air entrainment
 733 areas are fragmented and tend to dissipate in many, small areas of irregular shape.



734
 735 Figure 11. Time-evolution of an overtopping event and detection of the areas of air entrainment
 736 (red contour) through the video-cluster analysis.

737

738 The air entrainment map areas of the frames 1 to 6 and of the frames 7 to 12 composing the
 739 whole overtopping event portrayed in Figure 11 have been summed up, obtaining 2 "air
 740 entrainment sum maps". The values of the pixels of such "sum maps" range between 0 and 6, in
 741 case the pixel is never or always affected by air entrainment, respectively. By normalizing the
 742 pixel values into the scale 0-1, it is possible to use the air entrainment sum maps to derive the
 743 spatial and time distribution of the air entrainment in the flow. Following this approach, the air
 744 entrainment concentration of each i -th pixel is defined as follows:

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

$$C_i = \frac{\text{pixel value from an air entrainment sum map}}{\text{maximum value of the air entrainment sum map}} \quad (9)$$

746 The 2 air entrainment sum maps relative to the frames 1 to 6 and 7 to 12 are represented in
747 Figure 12, in normalized values and in comparison to the envelope of the free-surface derived
748 for the corresponding 6 frames of each map. These maps provide a few quantitative information.

749 First, the extent and the concentration of the air entrainment is significantly higher during the
750 frames 1 to 6 (Figure 12-top) than during the frames 7 to 12 (Figure 12-bottom), in agreement
751 with the qualitative estimations given from Figure 11. This information is reported for each of the
752 12 frames in Table 6 in terms of ratio between the area of the flow affected by air entrainment
753 and the total flow area (second column). Indeed, the air entrainment area represents the 17-28%
754 of the total flow during the frames 1 to 6, while it is at maximum the 8.8% during the frames 7 to
755 12.

756 Second, Figure 12-top shows that the bi-phase flow is mostly localized close to the structure
757 edge and in the middle of the overtopping tongue (yellowish colour), where the air entrainment
758 concentration is equal to 1. On the contrary, in Figure 12-bottom the air entrainment
759 concentration is more uniform and significantly lower, ranging between 0 and 0.4 at maximum
760 (purple/magenta). The average values of air entrainment concentration calculated for each air
761 entrainment sum map are reported in the columns 3 to 8 of Table 6. While during the frames 1
762 to 6 the highest C-values (1 or 0.83) represent the 5 and the 16% of the total air entrainment
763 area, during the frames 7 to 12 the highest C-values are substantially null, and the lowest C-
764 values (0.33 and 0.17) represent respectively the 29.8% and the 43.4% of the total air
765 entrainment area.

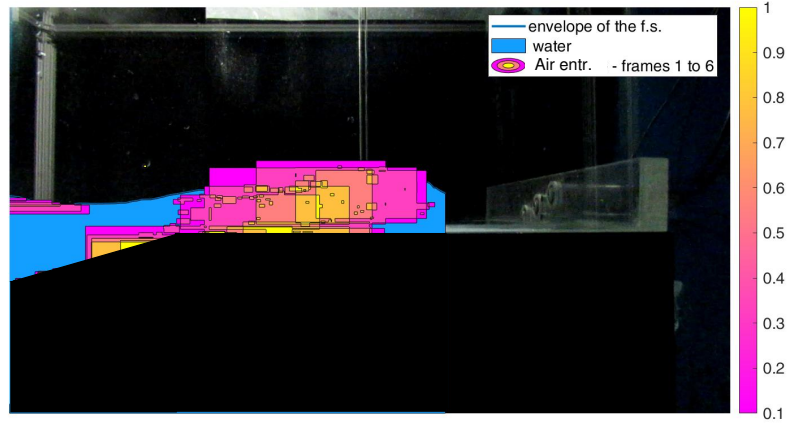
766 Air entrainment sum maps similar to the example ones represented in Figure 12, relative to an
767 overtopping event 12 frames-lasting, can be derived for the whole duration of the tests. In this
768 case, the pixel-values of the sum maps may range between 0 and – ideally- the total number of
769 frames of the test duration (typically, more than 16,000). However, the normalized concentration
770 as defined in Eq. (9) still ranges between 0 and 1, and the same quantitative information derived
771 for the example overtopping event can be derived for the whole test.

772

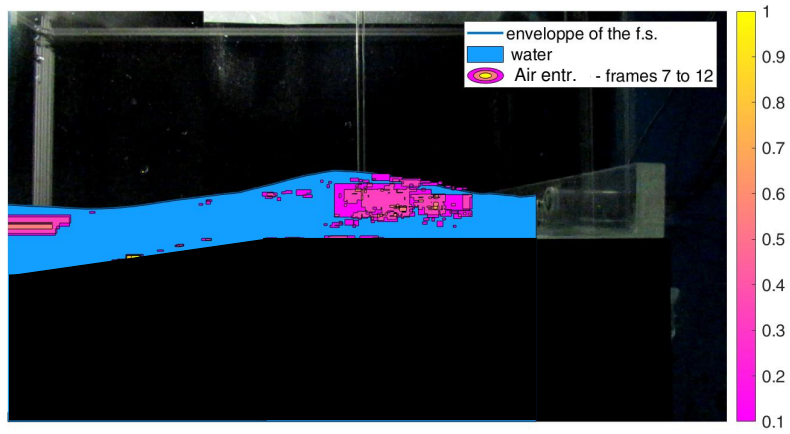
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

773



774



775 Figure 12. Air entrainment sum maps of the frames 1 to 6 (top) and 7 to 12 (bottom) of an
 776 overturning event. The maps are compared to the envelope of the free-surface obtained over
 777 the corresponding 6 frames.

778

779 Table 6. Frame-by-frame percentage of the area of the flow subjected to air entrainment (first
 780 column) and concentration of the air entrainment over the frames 1 to 6 (first 6 rows) and over
 781 the frames 7 to 12 (last 7 rows) for an example overtopping event. Values derived from the air
 782 entrainment sum maps.

Frame	(Air entr. area)/ (water area)	Air entr. area with C=1 (6/6)	Air entr. area with C=0.83 (5/6)	Air entr. area with C=0.67 (4/6)	Air entr. area with C=0.5 (3/6)	Air entr. area with C=0.33 (2/6)	Air entr. area with C=0.17 (1/6)
1	27%	5.0%	16%	21%	17%	19%	22%
2	21%						
3	28%						
4	24%						
5	20%						
6	17%						
7	8.8%	0.48%	0.66%	6.7%	19.1%	29.8%	43.4%
8	7.2%						
9	4.3%						
10	1.4%						
11	1.0%						
12	0.8%						

783

784 6. Conclusions

785 This contribution presented the application of a video-cluster-based methodology to the
 786 modelling of the wave-structure interaction processes from the video recording of laboratory
 787 experiments of wave overtopping at sea-dikes with walls. The novelty of the contribution consists
 788 in the pre- and post-clustering procedures set up to improve the automatic detection of features
 789 from the images and in the direct and indirect outcomes of the application.

790 The pre-clustering techniques operate image subtraction, light and color adjustment and noise
 791 addition to optimize the pictures as taken from the camera for the automatic clustering phase.
 792 The post-clustering techniques handle the clustered maps as returned by the cluster analysis
 793 with morphological filters and iterative procedures to extract and elaborate the data embedded
 794 in the maps themselves.

795 The direct outputs of the whole videography methodology are the time series of the free-surface
 796 elevations along the dike slopes and berm and the spatial-time distribution of the areas of the
 797 flow affected by air entrainment. The derived outputs are the overtopping flow depths above the
 798 structure berm, the individual and total overtopping volumes and discharges and the time series
 799 of the flow celerities derived from the elaboration of the signals of the free-surface elevations
 800 “virtually” recorded through the video-analysis. These virtual quantities were derived for the first
 801 time from the results of a cluster analysis through the introduction of “virtual gauges” and the

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

802 use of a procedure for the automatic detection and coupling of the individual overtopping events.
803 The virtual quantities and compared to the corresponding quantities obtained from traditional
804 measurements in the lab to verify and assess the accuracy of the methodology. The agreement
805 among virtual and laboratory quantities is synthesized by the following relative errors: +36% for
806 the extreme flow depths values ($h_{2\%}$); +19% and +25% for the average overtopping discharge
807 estimated from the overtopping volumes and flow celerities, respectively; +5% for the estimation
808 of the number of overtopping waves; +2% for the average and total overtopping volumes. The
809 positive sign of the errors indicates that videography gave always conservative estimations of
810 the virtual quantities. Such level of accuracy was achieved thanks to the introduction of the pre-
811 and post-clustering techniques, whose ensemble represents a best practice to be recommended
812 for other applications.

813 Another novel result of the research is represented by the maps of the concentration of the air
814 entrainment related to the wave breaking and the turbulent overtopping flow. The eye-
815 examination of these maps suggests that the areas of the flow more frequently subjected to the
816 formation of air bubbles are accurately caught by the methodology. The time-stack values of the
817 air entrainment concentration was reconstructed for an example overtopping event. This
818 application allowed a detailed analysis of the spatial-time evolution of the concentration of air in
819 the liquid phase. Specifically, it was found that the air entrainment is maximum around the berm
820 off-shore edge and decreases radially towards the middle of the flow and the berm width (spatial
821 distribution); it is maximum during the wave breaking, and starts decreasing during the bore flow
822 (time-stack analysis). It should be noted that the air entertainment becomes 3D during the later
823 stage of breaking since the flow is very turbulent. Therefore, the indications about the air
824 entrainment mapping are valid for 2D conditions only and cannot be extended to a 3D domain,
825 since just one camera was used to record the experiments.

826

827 List of notations

B	Berm width
c	Instantaneous wave celerity (front velocity of the overtopping tongue)
C	Air entrainment concentration
$diswg$	Distance between D1 and D3, corresponding approximately to the positions of the off-shore and in-shore edges of the berm, respectively
D1, D2, D3	Reference to the UVP 1, 2 and 3 installed along the berm
g	Acceleration due to gravity
h	Overtopping flow depths above the berm

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

$h_{2\%}$	Upper 2% percentile of the distribution of the h -values
h_b	Berm submergence ($h_b < 0$ and $h_b > 0$ respectively for emerged and submerged berm)
h_c	Elevation of the structure berm with respect to the bottom of the channel, excluding the crown wall
h_w	Height of the crown wall
H_{m0}	Spectral wave height
IE	Abbreviation of "in-shore edge"
K_r	Bulk wave reflection coefficient
$L_{m-1,0}$	Wave length from spectral analysis
N_{ow}	Number of overtopping waves
OE	Abbreviation of "off-shore edge"
q	Average specific wave overtopping discharge
q_{cel}	Value of q obtained from the integration of the overtopping wave celerities (c) with the overtopping flow depths (h)
q_{tank}	Value of q measured in the laboratory from the water tank
q_{Vol}	Value of q obtained from the sum of the V-values (V_{tot}) to the duration of the test (t_{tot})
swl	Acronym of "still water level"
R^2	Coefficient of determination
R_c	Structure freeboard with the respect to the still water level ($R_c = h_w - h_b$)
$s_{m-1,0}$	Wave steepness calculated based on the spectral wave period
t	Time
t_{tot}	Total duration of an experiment (corresponding to 480 s)
$T_{m-1,0}$	Spectral wave period
T_p	Peak wave period
u	Instantaneous flow velocity
UVP(s)	Acronym of "Ultrasonic Velocity Profiler(s)"

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

V	Instantaneous wave overtopping volume obtained from the integration of the h -values in the time domain
V_{bar}	Average wave overtopping volume of a test
V_{tot}	Total wave overtopping volume of a test
wgs	Acronym of "wave gauges"
α_d	Dike off-shore slope below the berm
γ_f	Structure roughness factor according to the EurOtop (2018) manual
μ	Mean
$\xi_{m-1,0}$	Iribarren-Battjes breaker parameter
$\sigma\%$	Relative standard deviation (or coefficient of variation)

828

829 **Author's contributions**

830 Conceptualization, S.M.F and B.Z.; formal analysis, S.M.F., M.G.G. and M.G.; investigation,
 831 S.M.F., R.D.V. and M.G.G; data curation, S.M.F and R.D.V.; writing—original draft preparation,
 832 S.M.F.; writing—review and editing, M.G.G., M.G. and B.Z.; supervision, B.Z.

833 **Acknowledgments**

834 The support of the European Commission through the Horizon 2020 project BRIGAID ("BRIdging
 835 the GAp for Innovations in Disaster resilience," [www.brigaid.eu](http://www brigaid.eu)) is gratefully acknowledged.

836 **Data availability statement**

837 The data that support the findings of this study are available from the corresponding author upon
 838 reasonable request.

839 **References**

840 Aleixo R., Soares-Frazão S. and Zech Y., 2018. Statistical analysis methods for transient flows—
 841 the dam-break case. *J. Hydraul. Res.*57, 688–701.
 842 Almar R., Cienfuegos R., Catalán P.A., Michallet H., Castelle B., Bonneton P. and Marieu V.,
 843 2012. A new breaking wave height direct estimator from video imagery, *Coastal Engineering* 61,
 844 42-48.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 845 Almar, R., Blenkinsopp, C., Almeida, L.P., Cienfuegos, R., Catalán, P.A., 2017. Wave runup
846 video motion detection using the radon transform. *Coastal Engineering* 130, 46–51.
847 <http://dx.doi.org/10.1016/j.coastaleng.2017.09.015>.
- 848 Bechle, A.J., Wu, C.H., 2011. Virtual wave gauges based upon stereo imaging for measuring
849 surface wave characteristics. *Coastal Engineering* 58 (4), 305–316.
850 <http://dx.doi.org/10.1016/j.coastaleng.2010.11.003>.
- 851 Benetazzo A., Barbariol F., Bergamasco F., Torsello A., Carniel S., Sclavo, M., 2015.
852 Observation of extreme sea waves in a space–time ensemble, *J. Phys. Oceanogr.* 45, 2261.
- 853 Blenkinsopp, C. E., and J. R. Chaplin (2007), Void fraction measurements in breaking waves,
854 *Proc. R. Soc. A*, 463(2088), 3151–3170.
- 855 Blenkinsopp, C.E., Turner, I.L., Allis, M.J., Peirson, W.L., Garde, L.E., 2012. Application of
856 LiDAR technology for measurement of time-varying free-surface profiles in laboratory wave
857 flume. *Coastal Engineering* 68, 1–5.
- 858 Bouguet, J.Y. Camera Calibration Toolbox for Matlab. 2015. Available online:
859 http://www.vision.caltech.edu/bouguetj/calib_doc/index.html (accessed on 1 September 2020).
- 860 Buscombe, D., Carini, R.J., Harrison, S.R., Chickadel, C.C., Warrick, J.A., 2020. Optical wave
861 gauging using deep neural networks. *Coastal Engineering* 155, 103593.
862 <http://dx.doi.org/10.1016/j.coastaleng.2019.103593>.
- 863 Buscombe, D., Carini, R.J., 2019. A Data-Driven Approach to Classifying Wave Breaking in
864 Infrared Imagery, *Remote Sensing* 11, 859. <http://dx.doi.org/10.3390/rs11070859>.
- 865 Bullock, G.N., Obhrai, C., Peregrine, D.H. and Bredmose, H. (2007). Violent breaking wave
866 impacts. Part 1: Results from large-scale regular wave tests on vertical and sloping walls,
867 *Coastal Engineering*, 54, 602–617.
- 868 Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach.*
869 *Intell.* 8, 679–714.
- 870 Carini, R.J., Chickadel, C.C., Jessup, A.T., Thomson, J., 2015. Estimating wave energy
871 dissipation in the surf zone using thermal infrared imagery. *J. Geophys. Res. Oceans* 120, 3937–
872 3957.
- 873 Chanson, H., 2002. Air–water flow measurements with intrusive phase-detection probes: Can
874 we improve their interpretation, *J. Hydraul. Eng.*, 128(3), 252–255.
- 875 Cox, D. T., and S. W. Shin (2003), Laboratory measurements of void fraction and turbulence in
876 the bore region of surf zone waves, *J. Eng. Mech.*, 129(10), 1197–1205
- 877 Deane, G. B. and Stokes, (2002), Scale dependence of bubble creation mechanisms in breaking
878 waves, *Nature*, 418, 839–844.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

879 Dempster, A. P., Laird, N. M. and Rubin, D. B., 1977. Maximum likelihood from incomplete data
 880 via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1),
 881 1–38.

882 Den Bieman J.P., de Ridder M.P., van Gent M.R.A., 2020. Deep learning video analysis as
 883 measurement technique in physical models. *Coastal Engineering* 158, 103689.

884 De Vries, S., Hill, D.F., de Schipper, M.A., Stive, M.J.F., 2011. Remote sensing of surf zone
 885 waves using stereo imaging. *Coastal Engineering* 58 (3), 239–250.
 886 <http://dx.doi.org/10.1016/j.coastaleng.2010.10.004>.

887 Drazen, D. A., and W. K. Melville (2009), Turbulence and mixing in unsteady breaking surface
 888 waves, *J. Fluid Mech.*, 628, 85–119.

889 Duin, R.P.W.; Pekalska, E., 2015. *Pattern Recognition: Introduction and Terminology*; Delft
 890 University of Technology: Delft, The Netherlands. Available online:
 891 <http://resolver.tudelft.nl/uuid:f5c560ed-5fc7-4320-84b4-a20614012bc7> (last accessed July 1, 2021).

892 Duz, B., Scharnke, J., Hallmann, R., Tukker, J., Khurana, S. and K. Blanchard, 2020.
 893 Comparison of the CFD results to PIV measurements in kinematics of spilling and plunging
 894 breakers, *Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and*
 895 *Arctic Engineering*, OMAE2020-19268

896 EurOtop, 2018. Manual on wave overtopping of sea defences and related structures. An
 897 overtopping manual largely based on european research, but for worldwide application. N.W.H.
 898 Allsop, T. Bruce, J. DeRouck, A. Kortenhaus, T. Pullen, H. Schuttrumpf, P. Troch, J.W. van der
 899 Meer and B. Zanuttigh. www.overtoppingmanual.com.

900 Formentin, S.M., Zanuttigh, B., 2018. A new method to estimate the overtopping and overflow
 901 discharge at over-washed and breached dikes, *Coastal Engineering* 140, 240-256.

902 Formentin S.M. and Zanuttigh B., 2019a. A Genetic Programming based formula for wave
 903 overtopping by crown walls and bullnoses, *Coastal Engineering* 152, 103529, 17 pp.

904 Formentin S.M. and Zanuttigh B., 2019b. Semi-automatic detection of the overtopping waves
 905 and reconstruction of the overtopping flow characteristics at coastal structures, *Coastal*
 906 *Engineering* 152, 103533, 18 pp.

907 Formentin S.M., Gaeta M.G., Palma G, Zanuttigh B., and Guerrero M., 2019a. Flow Depths and
 908 Velocities across a Smooth Dike Crest, *Water* 2019, 11, 2197; doi:10.3390/w11102197.

909 Formentin S.M., Zanuttigh B., Palma G, Gaeta M.G. and Guerrero M., 2019b. Experimental
 910 analysis of the wave loads on dike crown walls with parapets, *Proc. of Coastal Structure*
 911 *Conference, Hannover (D)*, 2019.

912 Gaeta M.G., Guerrero M., Formentin S.M., Palma G. and Zanuttigh B., 2020. Non-Intrusive
 913 Measurements of Wave-Induced Flow over Dikes by Means of a Combined Ultrasound Doppler
 914 Velocimetry and Videography, *Water* 2020, 12, 3053; doi:10.3390/w12113053.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 915 Gaeta, M.G., Lamberti, A., 2015. The role of air modeling on numerical investigation of coastal
 916 dynamics and wave-structure interactions, *Computers & Fluids*, 111, 114-126.
- 917 Hofland B., Diamantidou E., van Steeg P., Meys P., 2015. Wave runup and wave overtopping
 918 measurements using a laser scanner. *Coastal Engineering* 106, 20–29.
- 919 Hoque, A., and S. I. Aoki (2005), Distributions of void fraction under breaking waves in the surf
 920 zone, *Ocean Eng.*, 32, 1829–1840
- 921 Holman, R. and J. Stanley, 2007. The history and technical capabilities of Argus. *Coastal*
 922 *Engineering* 54, 477-491.
- 923 Jessup, A., Zappa, C., Loewen, M., Hesany, V., 1997a, Infrared remote sensing of breaking
 924 waves. *Nature* 385, 52.
- 925 Jessup T., Zappa C.J., Yeh H., 1997b. Defining and quantifying microscale breaking waves with
 926 infrared imagery, *J. Geophys. Res.* 102, 23145.
- 927 Kimmoun, O. and H. Branger, 2007. A particle image velocimetry investigation on laboratory
 928 surf-zone breaking waves over a sloping beach, *Journal of Fluid Mechanics*, Vol.588, 353 – 397.
- 929 Li, H., Fernex, D., Semaan, R., Tan, J., Morzyński, M., & Noack, B., 2021. Cluster-based network
 930 model. *Journal of Fluid Mechanics*, 906, A21. doi:10.1017/jfm.2020.785
- 931 Lim, H. J., K. A. Chang, C. B. Su, and C. Y. Chen, 2008. Bubble velocity, diameter, and void
 932 fraction measurements in a multiphase flow using fiber optic reflectometer, *Rev. Sci. Instrum.*,
 933 79, 125105.
- 934 Lim, H.-J., K.-A. Chang, Z.-C. Huang, and B. Na (2015), Experimental study on plunging breaking
 935 waves in deep water, *J. Geophys. Res. Oceans*, 120, 2007–2049, doi:10.1002/2014JC010269
- 936 Mori, N., T. Suzuki, and S. Kakuno (2007), Experimental study of air bubbles and turbulence
 937 characteristics in the surf zone, *J. Geophys. Res.*, 112, C05014, doi:10.1029/2006JC003647.
- 938 Mori, N., and S. Kakuno (2008), Aeration and bubble measurements of coastal breaking waves,
 939 *Fluid Dyn. Res.*, 40, 616–626.
- 940 Na B., Chang K., Huang Z. and H. Lim, 2015. Turbulent flow field and air entrainment in
 941 laboratory plunging breaking waves, *Journal of Geophysical Research: Oceans*,
 942 10.1002/2015JC011377,
- 943 Oh, S., Mizutani, N. and K. Suh, 2008. Laboratory observation of coherent structures beneath
 944 microscale and large-scale breaking waves under wind action, *Experimental Thermal and Fluid*
 945 *Science*, 32 (6), 1232-1247.
- 946 Perugini E., Soldini L., Palmsten M.L., Calantoni J. And Brocchini M., 2019. Linear depth
 947 inversion sensitivity to wave viewing angle using synthetic optical video, *Coastal Engineering*
 948 152, 103535. <https://doi.org/10.1016/j.coastaleng.2019.103535>.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

- 949 Raby A., Jayaratne R., Bredmose H. and Bullock G., 2028. Individual violent wave-overtopping
950 events: Behaviour and estimation. *J. Hydraul. Res.* 58, 34-46.
- 951 Robinson I.S., Wells N.C., Charnock H., 1984. The sea surface thermal boundary layer and its
952 relevance to the measurement of sea surface temperature by airborne and spaceborne
953 radiometers, *Int. J. Remote Sens.* 5, 19.
- 954 Rojas, G., and M. R. Loewen (2010), Void fraction measurements beneath plunging and spilling
955 breaking waves, *J. Geophys. Res.*, 115, C08001, doi:10.1029/2009JC005614
- 956 Siddiqui M.H.K, Loewen M.R., Richardson C., Asher W.E., Jessup A., 2001. Simultaneous
957 particle image velocimetry and infrared imagery of microscale breaking waves, *Physics of Fluids*
958 13, 1891. <https://doi.org/10.1063/1.1375144>
- 959 Streicher M., Hofland B., Lindenbergh R.C., 2013. Laser Ranging for Monitoring Water Waves
960 in the new Deltares Delta Flume. *ISPRS Annals of the Photogrammetry, Remote Sensing and*
961 *Spatial Information Sciences*, Volume II-5/W2, 2013.
- 962 Stringari, C.E., Harris, D.L., Power, H.E., 2019. A novel machine learning algorithm for tracking
963 remotely sensed waves in the surf zone. *Coastal Engineering* 147, 149–158.
964 <http://dx.doi.org/10.1016/j.coastaleng.2019.02.002>.
- 965 Stringari, C.E., Veras Guimarães P., Filipot J.F, Leckler F., Duarte R., 2021. Deep neural
966 networks for active wave breaking classification, *Scientific Reports* 11:3604.
967 <https://doi.org/10.1038/s41598-021-83188-y>
- 968 Techet, A.H. and A.K. McDonald, 2005. High Speed PIV of Breaking Waves on Both Sides of
969 the AirWater Interface, 2005. *Proc. 6th International Symposium on Particle Image Velocimetry*,
970 Pasadena, California, USA, September 21-23, 2005.
- 971 Vousdoukas M.I., Kirupakaramoorthy T., de la Torre M., Wübbold F., Wagner W., Schimmels S.,
972 Oumeraci H., 2014. The role of combined laser scanning and video techniques in monitoring
973 wave-by-wave swash zone processes. *Coastal Engineering* 83, 150–165.
- 974 Wang, B. and Q. Liao, 2016. Field observations of turbulent dissipation rate profiles immediately
975 below the air-water interface, *Journal of Geophysical Research: Oceans*, 4377-4391.
- 976 Waniewski, T. A., Brennen, C. E. and F. Raichlen, 2001. Measurements of Air Entrainment by
977 Bow Waves, *Journal of Fluids*, 123, 57-63.
- 978 Zavadsky A., Benetazzo A., Shemer L., 2017. On the two-dimensional structure of short gravity
979 waves in a wind wave tank, *Physics of Fluids* 29, 016601. <https://doi.org/10.1063/1.4973319>

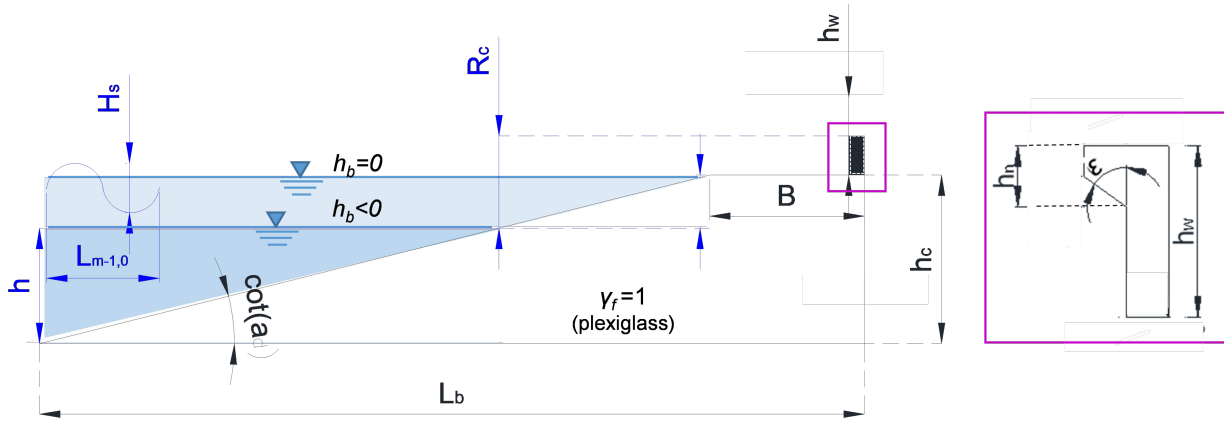
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



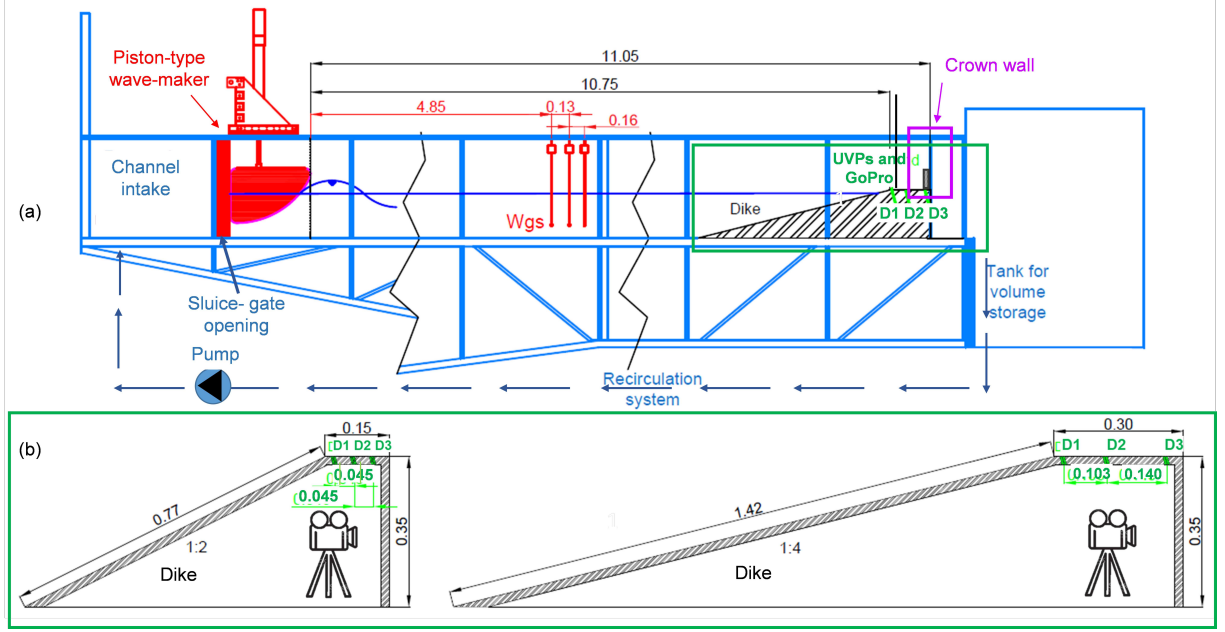
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0065019



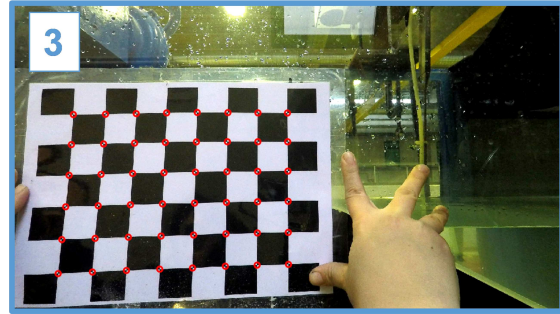
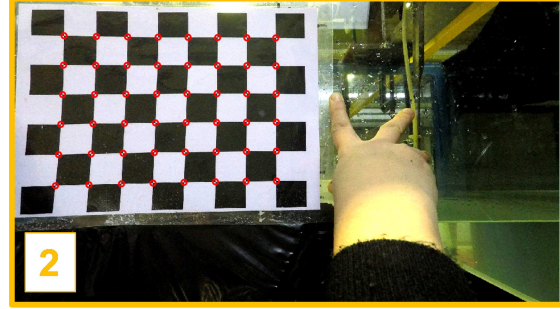
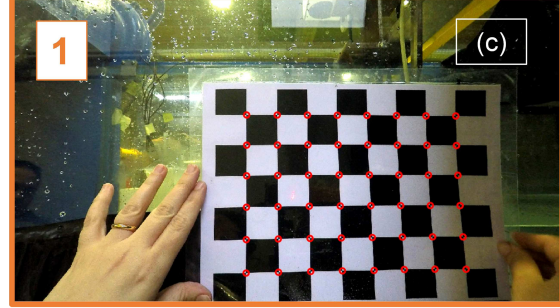
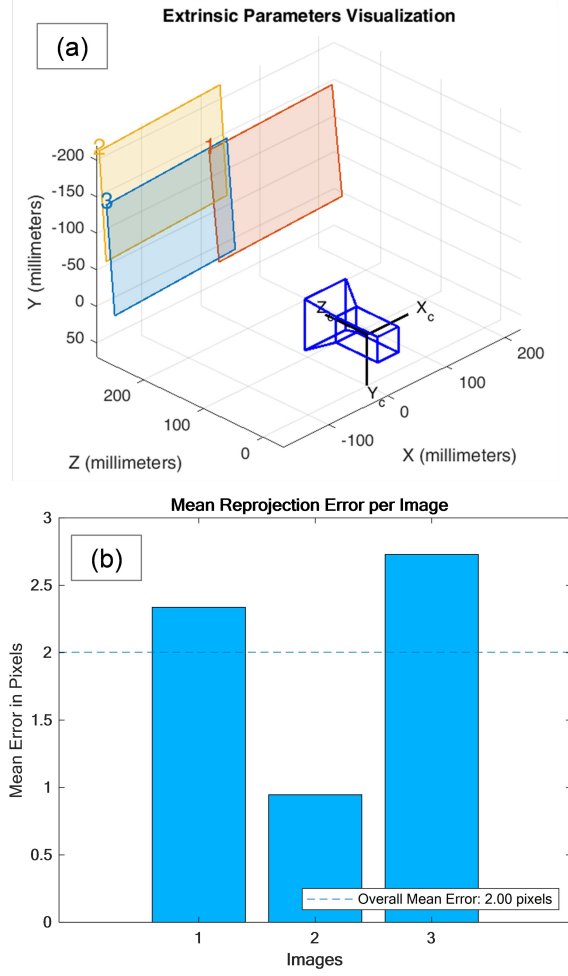
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



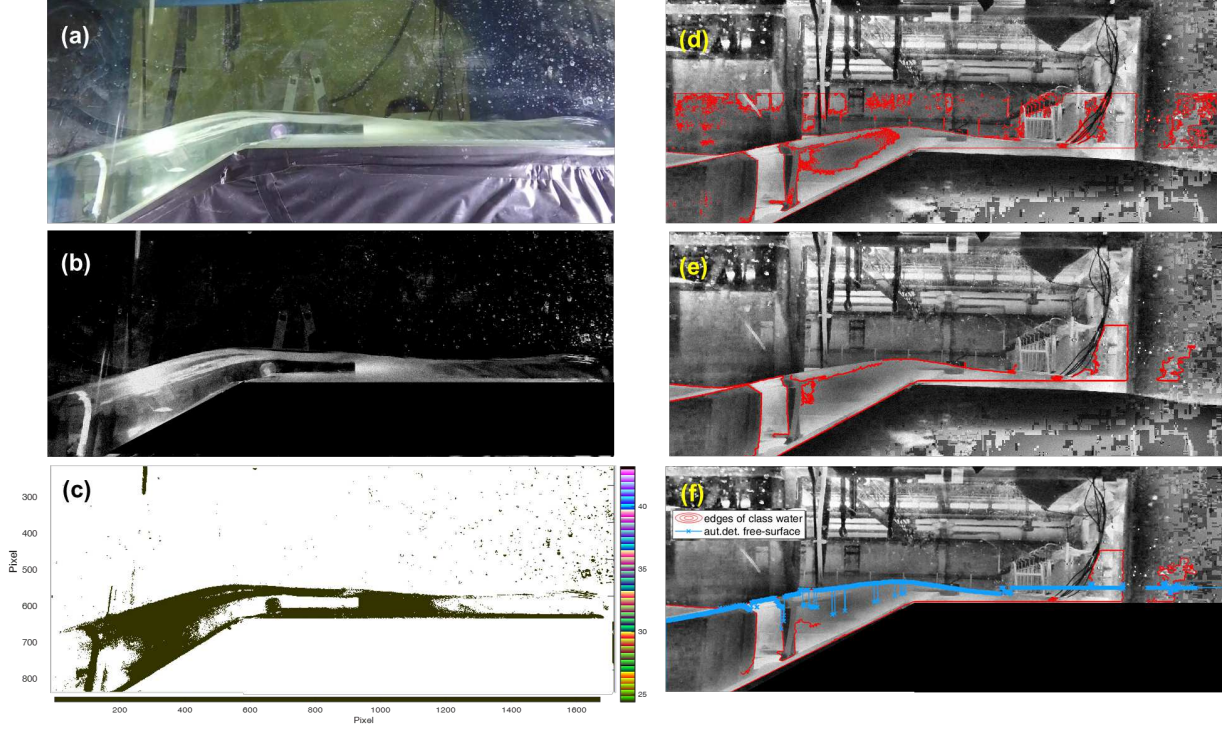
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0065019

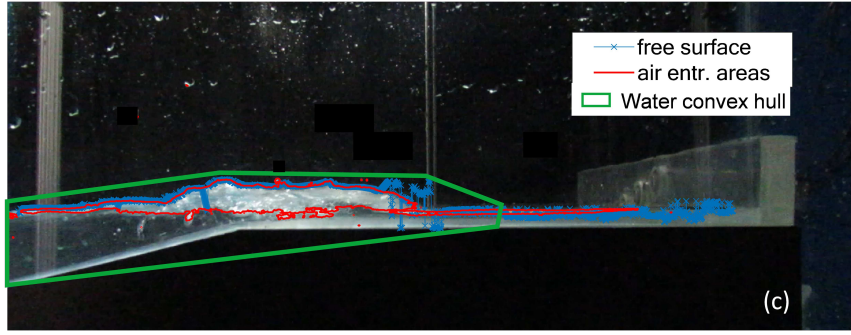
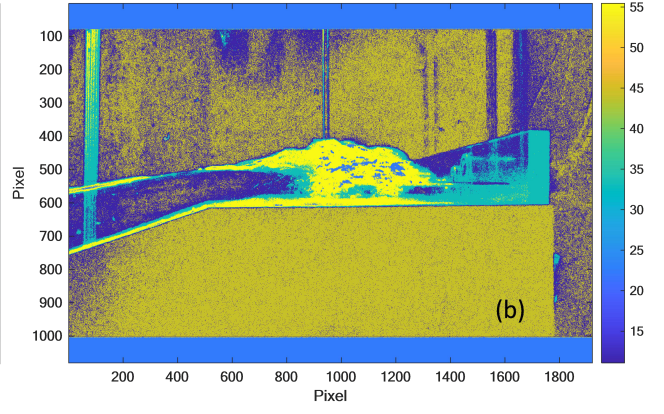
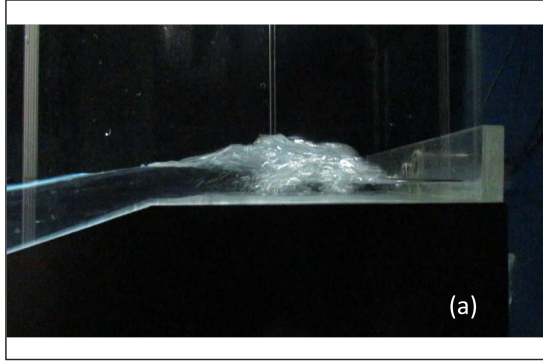


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

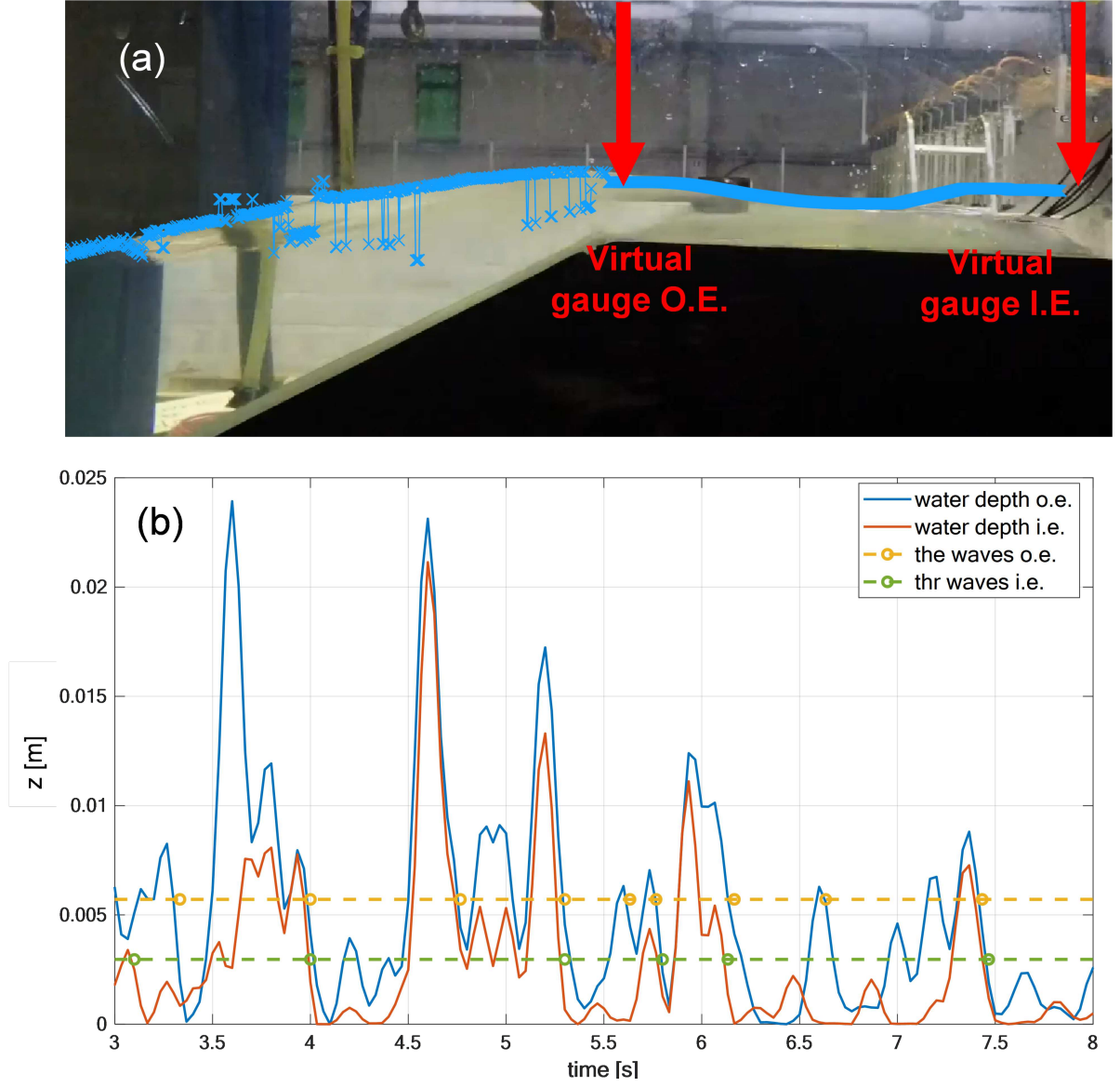


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
 PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

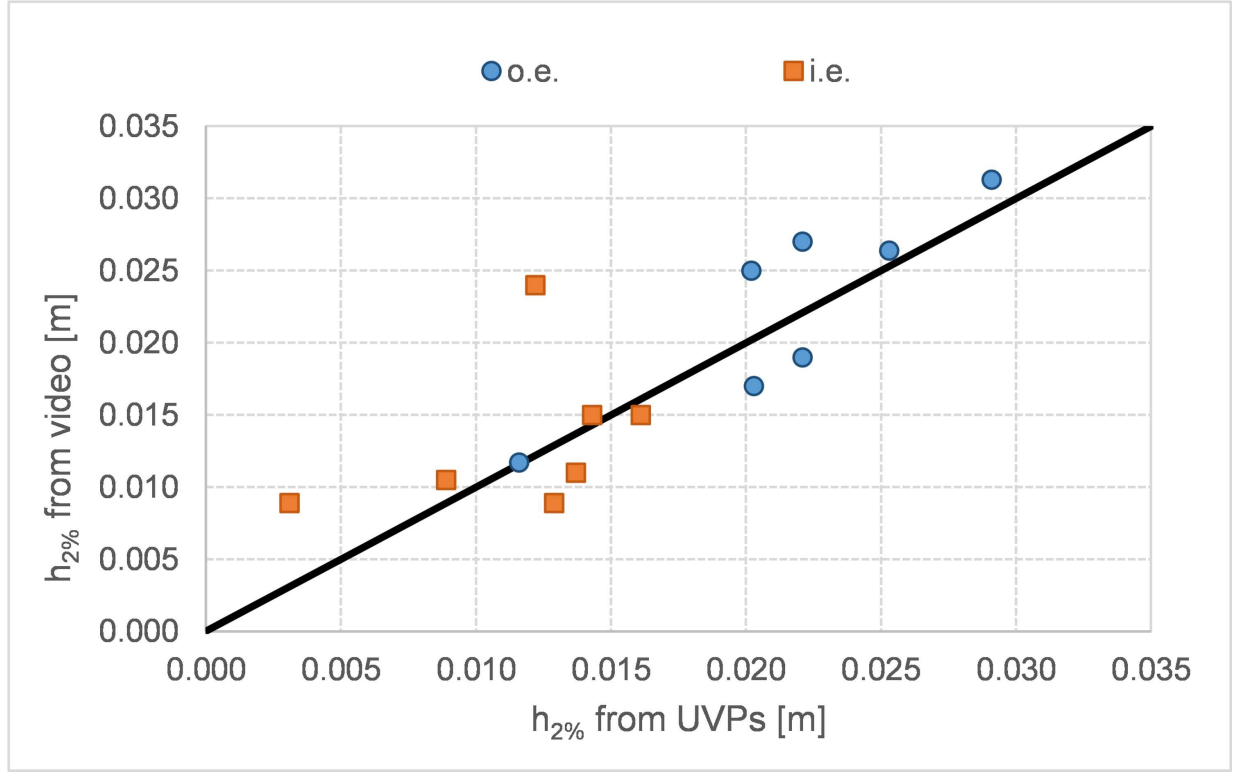


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

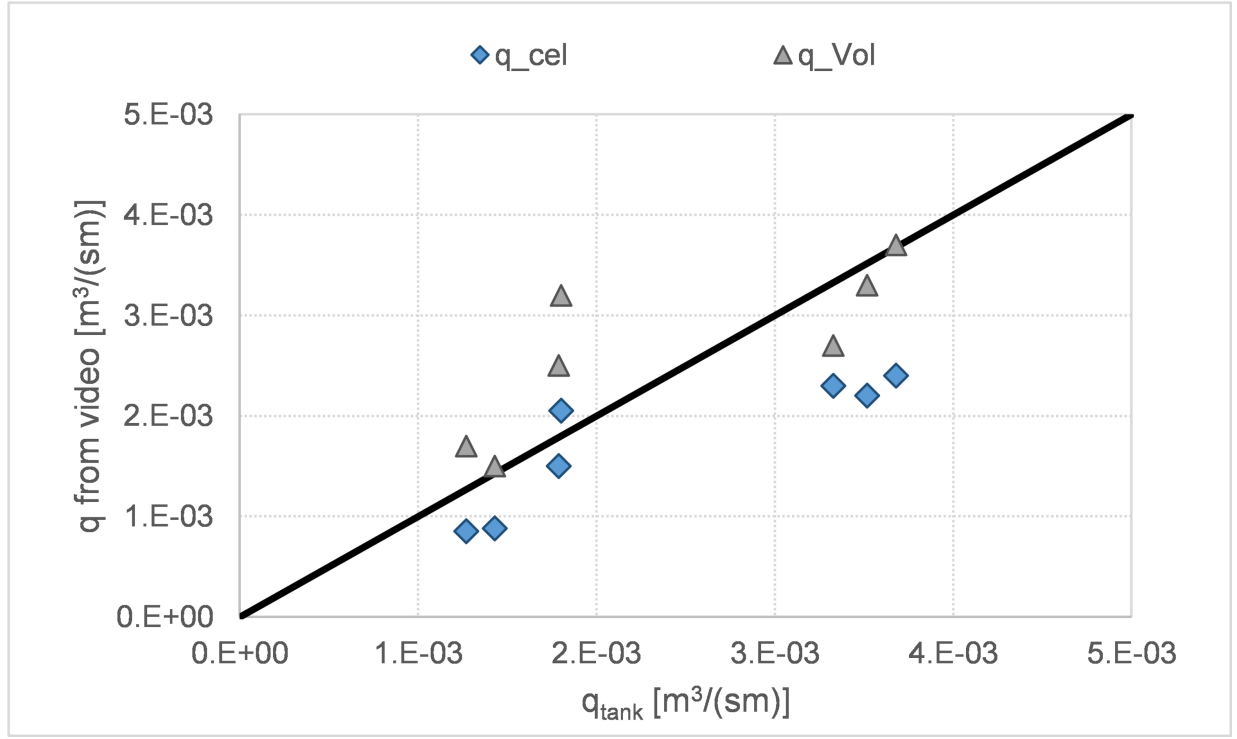
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
 PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

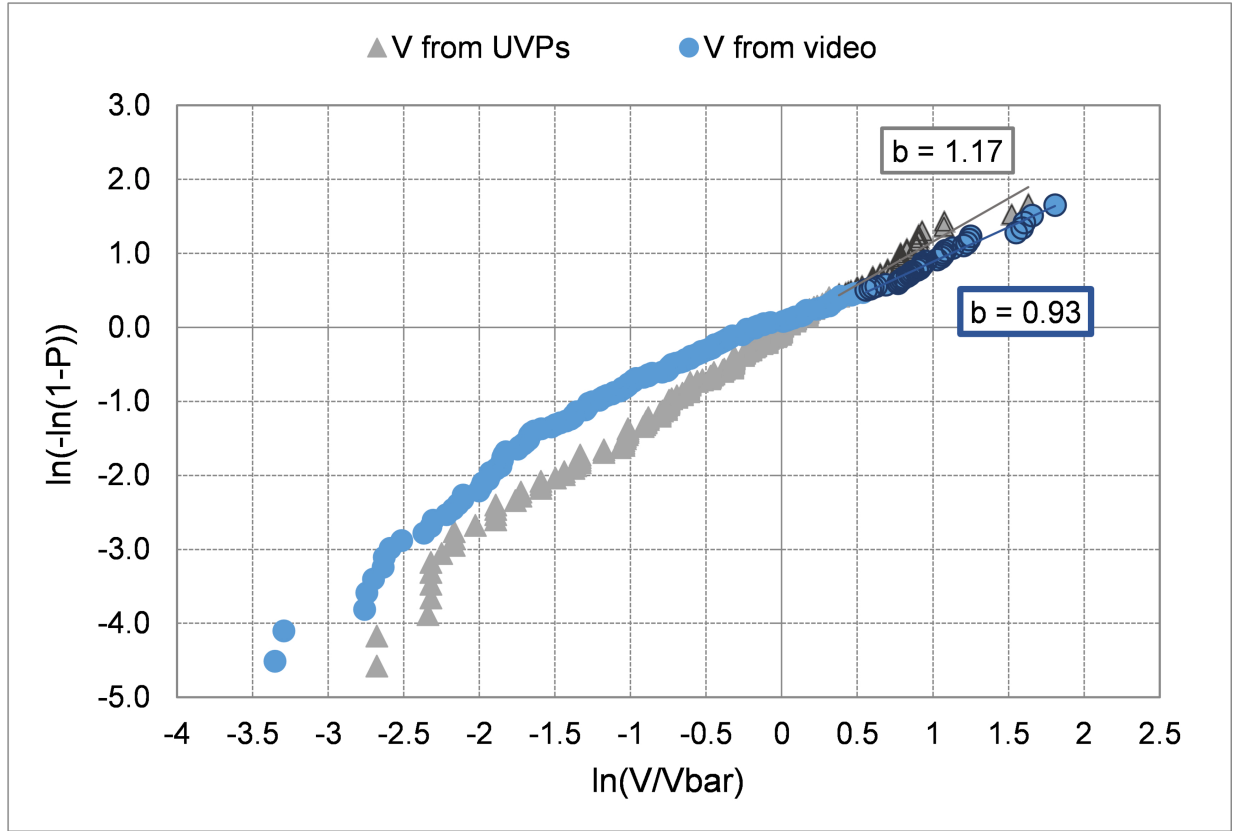


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
 PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



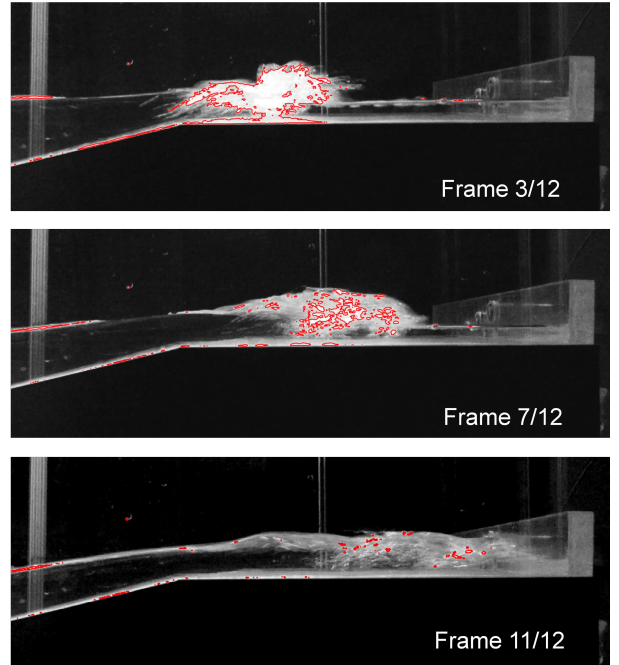
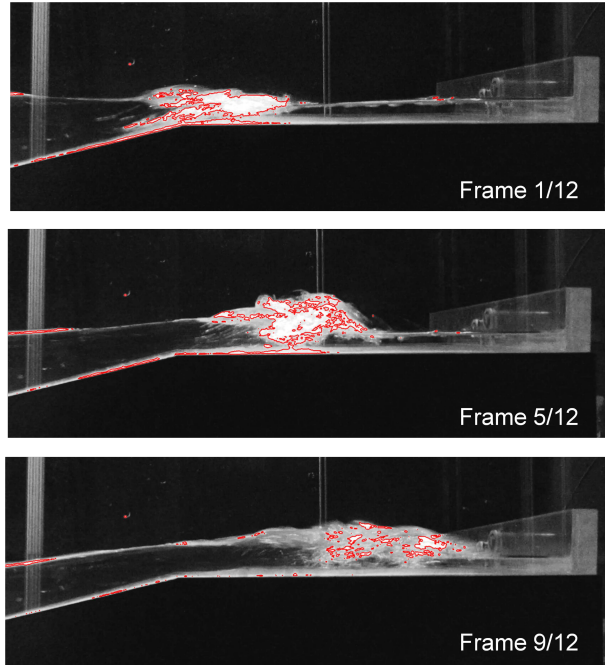
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/1.50065019

