



HAL
open science

DEM quality assessment with a 3D printed gravel bed applied to stereo photogrammetry

Stéphane Bertin, H Friedrich, P Delmas, E Chan, G Gimel'Farb

► **To cite this version:**

Stéphane Bertin, H Friedrich, P Delmas, E Chan, G Gimel'Farb. DEM quality assessment with a 3D printed gravel bed applied to stereo photogrammetry. *Photogrammetric Record*, 2014, 29, pp.241 - 264. 10.1111/phor.12061 . hal-03470769

HAL Id: hal-03470769

<https://hal.science/hal-03470769>

Submitted on 21 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DEM QUALITY ASSESSMENT WITH A 3D PRINTED GRAVEL BED APPLIED TO STEREOPHOTOGRAMMETRY

S. BERTIN (sber081@aucklanduni.ac.nz)

H. FRIEDRICH (h.friedrich@auckland.ac.nz),

P. DELMAS (p.delmas@auckland.ac.nz)

E. CHAN (ycha171@aucklanduni.ac.nz)

G. GIMEL'FARB (g.gimelfarb@auckland.ac.nz)

The University of Auckland, New Zealand

Abstract

Using stereophotogrammetry to obtain digital elevation models (DEMs) for surface topography analysis is becoming popular in hydraulic research, especially for coarse gravel beds. This paper assesses the DEM quality by using a realistic 3D printed gravel-bed model, with known elevations every 0.25 mm, as ground truth. Two Nikon D5100 cameras and non-proprietary photogrammetric software for camera calibration and DEM reconstruction are used for the study. A measured DEM is compared point by point with the ground truth and displays a very high measurement accuracy. The 3D printing of ground truths facilitates fast and versatile evaluation of both the DEM quality and the sensitivity of its errors to changes in surface topography and collection parameters. It has the potential to streamline evaluations of calibration and image quality, as well as error filtering strategies. Ultimately, 3D printed models will help in exploring stereomatching error reductions in occluded regions and defining the most suitable strategy for gravel-bed DEM collection, both in air and through water.

KEYWORDS: DEM accuracy, close range digital photogrammetry, gravel bed, 3D printing, ground truth

INTRODUCTION

THERE IS A GROWING DEMAND for high-resolution topographic data in hydraulic experiments, especially over natural water-worked gravel beds. Fine-scale representations of gravel beds with digital elevation models (DEMs) are becoming more and more

34 important for characterising the range of structures associated with water-worked
35 surfaces (Butler et al., 2001; Qin et al., 2012), and the links between microtopography
36 and interfacial hydraulics (Rice et al., 2014).

37 Digital stereophotogrammetry is capable of obtaining dense topographic data, in
38 both the laboratory and the field, and over dry and submerged riverbeds (Bertin et al.,
39 2012; Butler et al., 1998; 2002; Carbonneau et al., 2003; Chandler et al., 2001; 2003)
40 with lower cost, greater versatility and higher speed of data acquisition, than alternative
41 surveying techniques. However, as with all these methods, small-magnitude errors are
42 likely to propagate and significantly affect the quality of both the resulting DEM and
43 extracted topographic parameters (Lane et al., 2005), especially with low relief surfaces
44 (Hodge et al., 2009).

45 To ensure that recorded topographic models suit experimental requirements, it is
46 important to evaluate the overall performance of the recording environment. To do so, the
47 performance of the device and of the chosen methodological approach is represented in
48 terms of DEM quality and can be characterised by a global measure of error (Lane et al.,
49 2005).

50 Errors should also be investigated at each stage of the recording process, in order to
51 define a means of data collection improvement (Lane et al., 2005; Rieke-Zapp et al.,
52 2009). This is a particular concern for digital stereo-photogrammetric automated DEM
53 generation, as errors can arise from various parts of the process (Bertin et al., 2012;
54 Carbonneau et al., 2003; Lane et al., 2000).

55 RELATED WORK ON THE QUALITY OF TOPOGRAPHIC DATA

56 The topographic data quality in engineering surveying is described in terms of
57 accuracy, precision and reliability (Butler et al., 1998; Carbonneau et al., 2003; Cooper
58 and Cross, 1988; Hodge et al., 2009; Lane et al., 2000). The *accuracy* is the most
59 commonly employed descriptor (Lane et al., 2005) related to systematic errors in a DEM
60 and generally represented by the mean error (ME) between observed metrics and true
61 values. According to Cooper and Cross (1988), the inaccuracies occur in the
62 measurements due to an incorrect functional model, such as an inadequate lens distortion
63 model in stereo photogrammetry. The measurement *precision* refers to inconsistencies
64 (random errors) between repeated measurements under the same conditions. Such errors
65 cannot be eliminated by refining the functional model or applying corrections (Cooper
66 and Cross, 1988). Most importantly, global precision is traditionally computed using the
67 standard deviation of the errors (SDE). The root mean square error (RMSE) is normally
68 obtained through independent checks on the measured data (Brasington and Smart, 2003;
69 Butler et al., 2002; Carbonneau et al., 2003; Lane et al., 2000; Lane et al., 2003) and will
70 quantify random and systematic errors into a single measure of data quality (Butler et al.,
71 1998; Chandler et al., 2003; Lane et al., 2005). The RMSE is frequently labelled the
72 'accuracy' and assumes that the method of determining the independent checks is superior
73 to the method under consideration (photogrammetry in this instance). Gross errors or
74 blunders, which arise from an incorrect measuring process or malfunctioning equipment,
75 determine the *internal reliability* of the DEM. This descriptor is assessed by comparing
76 semi-independent metrics, such as DEMs obtained using different baselines (Butler et al.,
77 1998), or extracted from different imagery (Brasington and Smart, 2003), or obtained
78 through water and in air (Butler et al., 2002; Smith et al., 2012). The *external reliability*

79 is quantified by comparing a topographic parameter of interest, calculated from the DEM,
80 with a theoretical reference. Such external parameters can be slopes of the measured
81 surface for large-scale DEMs (Lane et al., 2000), or scaling properties of small-scale
82 surfaces (Carbonneau et al., 2003). Whilst tests of the external reliability can be
83 beneficial in investigating effects of changing DEM collection parameters, when
84 accuracy statistics are insensitive, the difficulty resides in defining the reference values
85 for the parameters of interest (Lane et al., 2000).

86 In practice, the topographic data quality is quantified in both the laboratory and field
87 by comparing metrics contained in a DEM for a small number of check points (generally,
88 less than 1% of all the DEM points), directly over the measured surface. Total stations or
89 alternative measuring devices are usually employed to locate and register the check
90 points within the DEM. Limitations of this approach are positioning errors, low density
91 and arbitrary distributions of the check points. Reliability of the data quality measured
92 with this approach cannot be ascertained, especially over rough surfaces, possibly leading
93 to wrong conclusions about the DEM quality (Bouratsis et al., 2013; Brasington and
94 Smart, 2003; Butler et al., 1998; 2002; Carbonneau et al., 2003; Chandler et al., 2003;
95 Gooch et al., 1999; Lane, 2000; Lane et al., 2000; Rieke-Zapp et al., 2009; Smith et al.,
96 2012).

97 To improve the check-point density, Chandler et al., (2001) measured a rough
98 flume-bed surface with both stereophotogrammetry and laser scanning and compared
99 point-by-point 5589 homogeneously distributed elevations. However, the laser-scanned
100 and photogrammetric data were obtained using different datums, and whilst data
101 alignment was said to be problematic (Lane, 2000), a detailed procedure was not
102 presented. Additionally, whilst laser scanning over rough surfaces is very accurate in
103 theory, it is not error-free in practice (Hodge et al., 2009). Therefore, each assessment
104 calls for both DEM reconstruction and error editing.

105 The measurement of objects of known size is another traditional approach to
106 evaluate the quality of topographic surveys (Bouratsis et al., 2013; Hodge et al., 2009;
107 Pollyea and Fairley, 2012). Wang et al. (2009) used a realistic 3D seafloor model of
108 known elevations to evaluate laser-scanning and photogrammetric capabilities in
109 measuring seabed roughness. The seafloor model built with a computer-assisted milling
110 machine achieved a 1 mm sampling distance between the check points and 0.1 mm
111 vertical accuracy. However, whilst the model mimicked the seafloor topography, its all-
112 white surface had unnatural appearance and lacked texture, resulting in adverse
113 photogrammetric measuring conditions. Additionally, Wang et al. (2009) had to align
114 check points with the observations, which hindered point-by-point comparisons and
115 computation of accuracy. Instead, external reliability measures were evaluated, with the
116 advantage that a reference value for an external parameter (roughness spectrum) could be
117 calculated from the known elevations of the seafloor model.

118 Whilst the above studies made significant progress on the DEM quality evaluation,
119 additional work is still needed to ensure that derived data-quality measures are sufficient
120 to evaluate a DEM collection strategy and allow for optimising the recording process. At
121 present, DEM evaluation strategy recommendations are twofold: confidence in check
122 points, their density and uniform distribution over the surface; and correct alignment of
123 measured and true values. The evaluation also needs to be improved for use in different
124 environments.

125
126
127
128
129
130
131
132
133
134
135
136

137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167

PAPER OVERVIEW

In this work a novel methodology for assessing DEM quality was developed and applied to laboratory gravel-bed DEMs collected with digital stereo photogrammetry. Initially, a 3D printed model of known dimensions was produced, based on the experimental stereo-photogrammetric pre-measuring of spatial topography and the visual appearance of a water-worked gravel bed. The dense and uniformly-distributed grid of check points from the 3D printed gravel-bed model acted as ground truth for assessing the subsequent stereo-photogrammetric measurements. Adaptive surface de-trending and search means were implemented to align the measured and ground-truth data, thus enabling reliable point-by-point comparisons. The 3D printed model was then assessed as to its ability to reveal variations in DEM collection parameters, based on the parametric ‘DEM sampling distance’, and how representative the derived accuracy statistics were.

METHODOLOGY

Measurement Environment and General DEM Collection Process

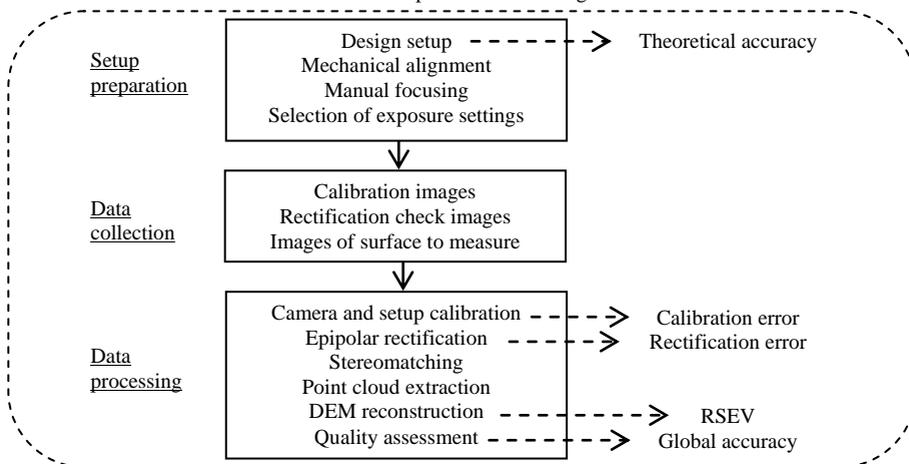
The DEM acquisition environment and the photogrammetric workflow (Table I) are identical to the earlier work by Bertin and Friedrich (2014) on monitoring gravel beds at the grain scale. A purpose-built hydraulic flume (19 m long, 0.5 m wide and 0.5 m deep, with a slope set to 0.5%) has been used for the measurements. A one metre long, full-width sediment recess 14 m downstream from the inlet was used to prepare the water-worked gravel bed. The evolving gravel-bed topography was recorded through air after the flume was drained.

A gantry-mounting system, sliding on a rail, allows the cameras to be rigidly attached horizontally above the test section. The stereo photogrammetric setup comprises two Nikon D5100 cameras with 16.2 Mpixel complementary metal oxide semiconductor (CMOS) sensors (4928 x 3264 pixels) and Nikkor 20 mm fixed-focus lenses. The baseline between the two cameras, and the setup ‘flying’ height, are adjustable within the ranges 200 to 400 mm, and 400 to 1000 mm, respectively. To minimise the effects of gravel protruding and shadowing the surroundings, the two cameras are mounted in a canonical vertical photogrammetric configuration, assuming parallel optical axes directed vertically toward the centreline of the flume.

The lighting consists of two pairs of 1 m long neon lights (58W with carbon dioxide), placed horizontally on each side of the flume, one light above the other. White Plexiglas sheets at the transparent flume sidewalls diffuse the lighting, creating a more homogenous illumination. The gravel-bed test section is thus illuminated by a ‘cool daylight’ colour. Two additional neon lights, which hang above the flume, are used during the acquisition of the calibration images. In order not to cast shadows on the surface, these latter lights are turned off for image acquisition of the gravel bed.

The general DEM collection process is presented in Table I and explained in detail below.

TABLE I. The workflow used for collecting, processing and assessing photogrammetric data. RSEV is the repeat stereomatching error value.



(1) The design of the stereo setup is the first step in a photogrammetric project. It constrains the theoretical accuracy achievable by the setup, and is thus critical to the success of the measurements (Lane et al., 2001; Lane, 2000). A MATLAB® routine, where the standard central perspective projection equations (pinhole camera model) are implemented, forms a systematic approach to determine the optimum camera-to-object distance (the distance resulting in the best measurement resolution), which complies with the experimental requirements of the measurement window size. Since DEM errors increase near the edges (Butler et al., 2002), a margin around the desired measurement window is commonly accounted for in the design and removed from the DEM during analysis.

(2) During image acquisition, each camera is connected to a computer, allowing remote control and live view with Nikon Control Pro 2® software. This helps to manually focus on the object to be measured and to mechanically align the cameras. After these steps, it is necessary to ensure no modification is made to the physical setup for the duration of the experiment. Adequate selection of the camera settings is important, as the crucial element to a successful close-range photogrammetric process is attaining ‘good photographs’ (Matthews, 2008), where the term ‘good’ refers to pictures that have uniform exposure with high-contrast sharp images. Since the lighting environment will constrain the potential camera settings, it is important that both the lighting environment and the exposure settings are optimised in common as an interdependent pair. To obtain the best quality photographs, the cameras are operated in manual mode. With the Nikkor 20 mm lenses, a combination of an aperture of f/8 (sometimes increased to f/10 or f/11) with a generic sensitivity value such as ISO 200, ensures a good depth of field, a reduced chance of vignetting and sharpness across the image. Once the aperture and ISO are set, these settings must remain constant throughout the acquisition of all images. The shutter speed can be changed to adapt to different lighting conditions.

(3) The camera and setup calibration is performed using the camera calibration toolbox for MATLAB®, developed by Bouguet (2010). This technique relies on the

210 calibration method of Zhang (1999), whereby calibration parameters (both intrinsic and
 211 extrinsic) are computed from a series of photographs of a planar chequerboard in
 212 different orientations. Radial distortion (up to the fourth order) and tangential distortion
 213 are also modelled during the calibration. Using the calibration results, all stereo pairs
 214 obtained with the setup in the calibrated configuration can be rectified to epipolar
 215 geometry, where corresponding pixels are on a same scan-line. Image rectification
 216 (Fusiello et al., 2000) is included in the calibration toolbox. It should be noted that the
 217 toolbox transforms images to greyscale equivalents during rectification, but the code can
 218 easily be modified to obtain rectified images in red/green/blue (RGB) format.

219 (4) Stereomatching is performed on the rectified RGB images using the symmetric
 220 dynamic programming stereo (SDPS) algorithm (Gimel'farb, 2002), providing both the
 221 grey-coded depth maps and ortho-images for a reconstructed DEM. The SDPS matches
 222 corresponding points by minimising pixel mismatches along each scan line in the stereo
 223 pairs. Compared with previous photogrammetric applications in hydraulic experiments
 224 (Bouratsis et al., 2013; Brasington and Smart, 2003; Butler et al., 1998; 2001; 2002;
 225 Carbonneau et al., 2003; Chandler et al., 2001; 2003; Gooch et al., 1999; Lane, 2000;
 226 Lane et al., 2000; 2003; 2005), the SDPS matching is not 'area-based'. Hence, the DEM
 227 smoothing will be less, and the sampling distance can be chosen, if necessary, as small as
 228 the pixel size at the object's distance. The SDPS search for the corresponding pixels is
 229 bounded by a disparity search range, calculated during the design by predetermining the
 230 surface elevation ranges. Post-processing by median filtering, a common practice in
 231 photogrammetric surveys (Carbonneau et al., 2003), is applied to eliminate blunders in
 232 the depth maps. Using the default filter width of 3 pixels and height of 11 pixels, each
 233 given pixel elevation is replaced by the median value over a 3 x 11 neighbourhood
 234 around the corresponding pixel position in the input image. In terms of the matching
 235 precision, the SDPS is able to distinguish disparity layers up to one pixel, which will give
 236 the theoretical depth resolution achievable by the setup after the pixel size is determined
 237 in the object space (see Table III).

238 From the depth maps and the calibration results, point cloud text files, containing the
 239 3D object coordinates of each pixel in the depth maps, are extracted. The points in these
 240 clouds are not regularly spaced, because they pertain to different disparity levels. Using
 241 the function 'gridfit' in MATLAB®, the point clouds are interpolated (using the default
 242 triangle interpolation) into regular grids, with adjustable spacing, and represented as 2.5D
 243 DEMs with normalised elevations.

244 *Preparation of a Realistic 3D Printed Gravel-Bed Model*

245 Grain sizes are frequently quoted as D_i , which is the grain size at which the specified
 246 percentage i of the grains are finer (so $D_{50} = 10$ mm means 50% of the grains are smaller
 247 than 10 mm). To ensure that the 3D printed model mimics the topography of gravel beds
 248 simulated in hydraulic laboratory experiments, a standard sediment mixture, with a
 249 median size of the intermediate axis $D_{50} = 7.3$ mm, a minimum size of 0.7 mm, a
 250 maximum size of 65 mm and a geometric standard deviation of the grain size
 251 distribution, calculated as $\sqrt{D_{84}/D_{16}} = 2.98$, was used to create a screeded gravel bed in
 252 the hydraulic flume. The sediment was water-worked with a flow rate of 60 l/s and a
 253 constant water depth of 200 mm, until no sediment movement was observed. The water-
 254 worked surface presented evidence of armouring, with a general coarsening of the

255 sediment. The grain size distribution was re-estimated using an orthophotograph of the
 256 surface and the image-based method developed by Detert and Weitbrecht (2012), called
 257 BASEGRAIN®. It was found to have a D_{50} of 13 mm, and $\sqrt{D_{84}/D_{16}} = 3.25$.

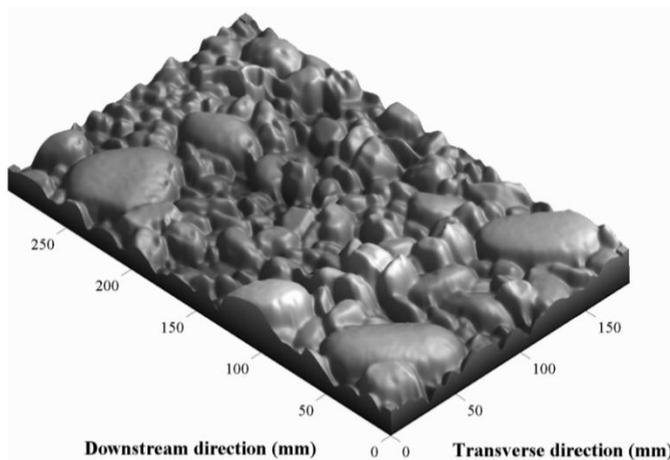
258 Digital stereophotogrammetry was chosen for data collection, as it was the available
 259 technique to allow the best compromise between measurement resolution and theoretical
 260 depth accuracy (Bertin and Friedrich, 2014). For obtaining the ground truth, two Nikon
 261 D90 cameras with Nikkor 18-105 mm lenses were experimentally set up with a baseline
 262 of 300 mm and a camera 'flying' height of 575 mm. The resulting pixel size on the gravel-
 263 bed surface was 0.18mm, the overlap was 60% and the theoretical depth resolution was
 264 0.35 mm.

265 Limitations due to the 3D printer restrained the size of the ground truth to
 266 296 mm and 184 mm along the flow and transverse directions, respectively. However,
 267 surfaces of this size and smaller have been shown to be suitable for characterising grain-
 268 scale surface properties of laboratory gravel beds (Marion et al., 2003; Ockelford and
 269 Haynes, 2013). Photogrammetric data were interpolated on a regular grid with a 0.25 mm
 270 sampling distance, generating 873 345 points of known elevations (Table II). To facilitate
 271 the assessment of future data using the 3D printed model, 5 mm squares at the four
 272 corners were made flat, and elevations were normalised to have a mean value of zero and
 273 stored in the *ground truth elevation matrix* (of data size 1185 x 737). Additionally, before
 274 sending it to the 3D printer, the DEM was rendered as a hollow object by replicating the
 275 surface 5 mm below the original and connecting the two with prolonged edges. The 3D
 276 DEM (Fig. 1) was finally transformed in a stereo lithography (.stl) file and sent to the 3D
 277 printer.

278 A ProJet® 3500 HDPlus from 3DSYSTEMS, available at The University of
 279 Auckland, was employed for the manufacturing. 3D printing is an additive process,
 280 where successive layers of material, VisiJet® Crystal (described as a transparent plastic),
 281 are laid down in different shapes. The gravel-bed model was built in approximately 15
 282 hours, with a manufacturer's specified accuracy range of 0.033 to 0.066 mm. After
 283 manufacturing, the model was left to dry. During the drying process, the edges of the
 284 model experienced a minimal upwards curvature due to thermal contraction of the
 285 material. The model was therefore mounted onto a flat Perspex plate, ensuring a planar
 286 3D datum.

287 TABLE II. Characteristics of the realistic 3D gravel-bed model (ground truth).

Size (mm)	296 x 184
Sampling distance (mm)	0.25
Number of points with known elevation	873 345
Standard deviation of surface elevation (mm)	5.5
Surface elevation span (mm)	32.6

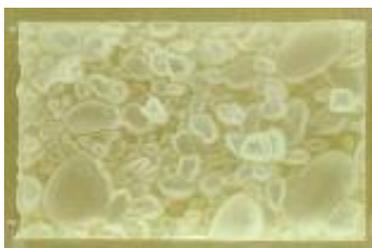


288
289
290

FIG.1. DEM of a representative patch of gravel bed, of size 296 x 184 mm, and 0.25 mm sampling distance, selected for 3D printing.

291 Preliminary tests showed that stereomatching of images of the 3D printed model
292 does not perform well because of the transparency of the material and weak texture, in
293 other words the lack of contrast and patterns on the surface (Fig. 2(a)). The surface was
294 subsequently finished (Fig. 2(b)), maximising the chance of correctly matched pixels.
295 This was an important step to allow the investigation of the joint effect of topography and
296 DEM collection strategy on the DEM quality, with minimal influence from surface-
297 dependent stereo correspondence problems. Similarly, images of natural gravel beds
298 contain fine texture for successful stereomatching (Bertin et al. 2012; Butler et al. 1998).

299 Surface finishing was achieved by painting the model in matt black to make the
300 object opaque. Very fine sand ($D_{50} < 0.2$ mm) was spread on the freshly painted surface so
301 that, due to the contact with the paint, a single layer of sand attached to the surface.
302 Theoretically, the addition of this real texture altered the initial topography contained in
303 the ground truth elevation matrix. However, the impact on the assessment of measured
304 data was neglected as: (i) the addition of sand can be approximated by a uniform shift in
305 bed elevation, which is suppressed when the measured DEM is vertically aligned with the
306 ground truth DEM; and (ii) the sand layer thickness is smaller than the minimum depth
307 measurable by the photogrammetric setup developed for the laboratory experiments
308 (Table III).



309
310

(a) (b)
FIG. 2. Photographs of the 3D gravel bed: (a) the raw model; (b) the rendered model.

311

EXPERIMENTS

312 *Implementation of the Measurement Technique*

313 This section reports on the particular photogrammetric setup that was evaluated with
 314 the 3D printed gravel bed. The setup was designed to comply with a desired measurement
 315 window of size 450 mm in the flow direction and 400 mm transversely. The baseline
 316 between the cameras was set experimentally as close as possible to 250 mm. This was
 317 confirmed after stereo calibration by a numerically estimated 3D distance between the
 318 two cameras of 250.3 mm. A 50 mm margin around the measurement window was
 319 accounted for and cropped before analysis. Table III summarises the resulting
 320 measurement specifications. JPEG (1:4) fine format was selected for the image recording.
 321 An aperture of $f/10$ and sensitivity of ISO 200, combined with a 1/15 s exposure time,
 322 allowed good contrast and dynamic range, while manual focusing ensured sharp images.

323 Data acquisition consisted in obtaining:

- 324 (i) 30 stereo pairs of a calibration chequerboard covering the whole cameras'
 325 common field of view (CFoV) and at distances close to the camera-to-object
 326 distance;
- 327 (ii) 10 additional calibration stereo pairs to evaluate the rectification error after
 328 calibration;
- 329 (iii) five stereo pairs of a gravel bed; and
- 330 (iv) five stereo pairs of the 3D printed model.

331 The number of images used for each case is independent of that employed in the other
 332 cases. For the camera and setup calibration, it has been shown that it is essential to have
 333 numerous check points (Carbonneau et al., 2003). Further tests, specific to the calibration
 334 method employed, revealed that the calibration error reaches a minimum with 20 or more
 335 calibration images. The rectification error was spatially determined over the whole CFoV
 336 by using 10 additional calibration images. Finally, five stereo pair images of the
 337 measured surfaces (3D printed model and gravel bed) allowed an investigation into the
 338 effect of different imagery on stereomatching results (see the section '*Assessment of*
 339 *Stereomatching Reliability*' for information on the images). In addition, restricting the
 340 number of images to five allowed for a time-efficient process to collapse the individually-
 341 extracted DEMs into one, with the aim of reducing stereo correspondence gross errors.

342 TABLE III. Summary of the photogrammetric setup designed for the test. All values were theoretically
 343 determined using the design equations and the rounded camera-to-object distance. Actual values may deviate
 344 because of the difficulty, in practice, to accurately set the cameras' 'flying' height and baseline.

Rounded camera-to-object distance (mm)	636
CFoV (mm)	500 x 498
Baseline (mm)	250
Range of disparity for 50 mm elevation span (pixels)	1580 to 1710
Overlap (%)	67
Sampling distance (mm) / resolution (pixel/mm ²)	0.15 / 43
Theoretical vertical accuracy or minimum measurable depth (mm)	0.39
Number of pixels in the 450 x 400 mm measurement window	≈7 800 000

345 *Variation of DEM Sampling Distance*

346 The visual aspect and quality of a DEM is strongly impacted by the sampling
 347 distance with which the DEM is represented (Lane et al., 2000). Too coarse a point

348 spacing results in a loss of topographic information. At the same time, grid size directly
 349 influences the size of numerical dataset that needs to be handled. A point spacing which
 350 is too fine will slow down the calculations. Although, the stereo-photogrammetric system
 351 allowed sampling distances as small as 0.15 mm (Table III), the default value for grid
 352 spacing was set to 0.25 mm, which is the same sampling distance used for the 3D printed
 353 model. The sensitivity of the DEM quality to changes in the collection parameters was
 354 evaluated by re-collecting the DEMs using different values for the parameter (sampling
 355 distance of 0.5, 0.75, 1, 1.5 and 3 mm).

356 *Quantitative Assessment of DEM Quality Using the 3D Printed Gravel Bed*

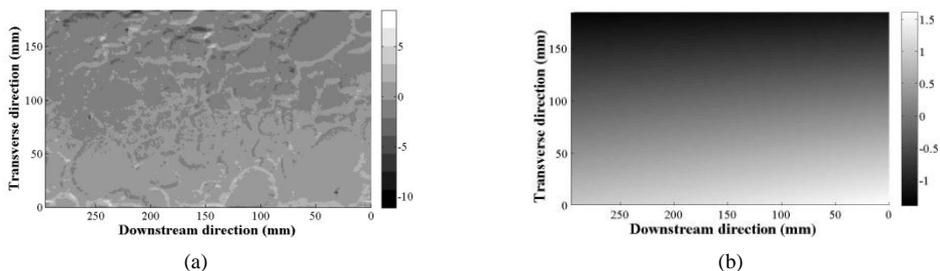
357 For assessing the measurement quality, the measured DEM of the 3D printed gravel
 358 bed needed to be aligned with the ground truth DEM. Data alignment had to be
 359 performed separately to the DEM reconstruction process, since no means was used to
 360 register measured DEMs within a common coordinate system with the ground truth. The
 361 data was aligned as follows:

- 362 (i) the 3D printed model was placed horizontally in the hydraulic flume at the
 363 location where the gravel bed is normally recorded, with its long axis parallel to
 364 the cameras' baseline horizontal axis;
- 365 (ii) horizontal alignment is performed with a search algorithm to find the region
 366 where the 3D printed model is represented in the measured DEM; and
- 367 (iii) vertical alignment makes use of surface de-trending algorithms to remove any
 368 tilt between the measured and ground truth DEMs.

369 In cases where the sampling distance of the measured DEM is not 0.25 mm, elevations
 370 are interpolated onto a 0.25 mm grid to allow comparison with all check points contained
 371 in the 3D printed model. Finally, measured DEMs of the 3D printed model are
 372 quantitatively assessed after point-by-point elevation comparison with the ground truth
 373 DEM, providing measures of DEM quality.

374 Any vertical misalignment of the cameras, with the baseline not set parallel to the
 375 mean surface of the 3D printed model and/or when the cameras' optical axes are not
 376 perpendicular to it, results in systematic comparison errors. This translates to a DEM of
 377 difference (DoD) between the measured and ground-truth elevation maps that exhibits a
 378 linear trend when the cameras are tilting along one (nominally) horizontal axis and a
 379 bilinear trend when they are tilting along both axes (Fig. 3(a)). The trend corresponding
 380 to the deviation from the correct position is estimated by fitting a bilinear surface to the
 381 DoD in a least squares method (Fig. 3(b)). This surface is then removed from the
 382 measured DEM to perform the vertical alignment. The scale in Fig. 3(b) indicates that the
 383 vertical misalignment is responsible for systematic elevation discrepancies of up to
 384 several millimetres. It is thus fundamental to remove the misalignment before a realistic
 385 assessment of the measured DEM can be undertaken.

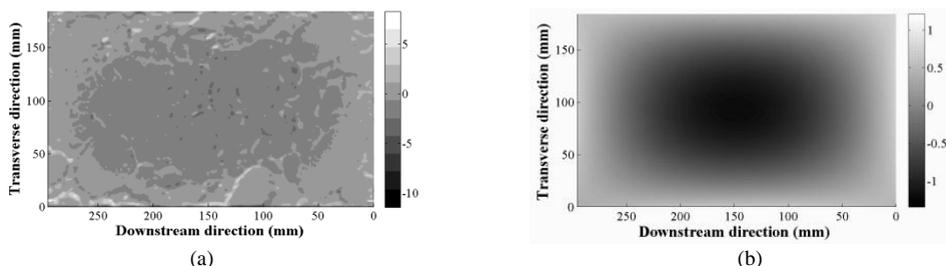
386 The reliability of the vertical alignment process was assessed by comparing the
 387 trends extracted from DEMs collected using a 0.25 mm (Fig. 3) with another employing
 388 a 3 mm sampling distance. The latter is the coarsest grid used in this study and thus it was
 389 expected that differences between the two trends would be maximised for this case. The
 390 vertical alignment was reliable, as even for the coarsest grid the mean unsigned (absolute)
 391 difference between the two estimated trends was 24 μm , the standard deviation was
 392 29 μm and the maximum unsigned difference was 88 μm .



393 (a) (b)
 394 FIG. 3. (a) DoD before vertical alignment of the measured DEM collected with a 0.25 mm sampling distance
 395 with the ground truth DEM. (b) Fitted bilinear trend representing the vertical misalignment of the stereo setup.
 396 Elevations are displayed as gradient of greys; values are in millimetres.

397 Another trend was observed in the DoDs, after alignment and differentiation of the
 398 measured DEMs with the ground-truth DEM (Fig. 4). The trend is comparable to the
 399 material thermal contraction which caused the curvature observed during the drying
 400 process for the 3D printed model. The deviation of the final 3D printed model from its
 401 initial topography was measured with high precision by fitting a biquadratic surface to
 402 the DoDs using least squares. The magnitude of the deviation was found to be consistent
 403 throughout the various usages of the 3D printed model, exhibiting a mean absolute
 404 difference between two estimated trends (obtained using different setup geometry) of
 405 0.02 mm, a standard deviation of 0.03 mm and a maximum unsigned difference of
 406 0.24 mm, suggesting that the curvature originated during manufacturing and remained
 407 constant afterwards. The shape (Fig. 4(a)), inherent to the 3D printed model, was
 408 obtained by averaging two estimated trends and was incorporated in the ground-truth
 409 elevation matrix. The ground-truth DEM now represented the actual 3D printed model.

410 To confirm that the curvature trend observed in the DoDs did not originate from the
 411 measurements, but was indeed inherent to the 3D printed model, a stereo pair of the flat
 412 calibration chequerboard was matched with the SDPS. After alignment of the
 413 chequerboard's surface with the cameras' image plane using bi-linear de-trending, a
 414 range of elevations of 0.4 mm (which is exactly the theoretical accuracy of the setup and
 415 about 15% of the range of elevations found in Fig. 4) characterised the chequerboard
 416 surface and no clear trend was observed.



417 (a) (b)
 418 FIG. 4. (a) DEM of difference (DoD) before removal of the manufacturing-induced trend. (b) Fitted
 419 manufacturing-induced curvature that could not be eliminated by fixing the 3D printed model to a flat Perspex
 420 plate. Elevations are displayed as gradient of greys and values are in millimetre.

421 *Assessment of Stereomatching Reliability*

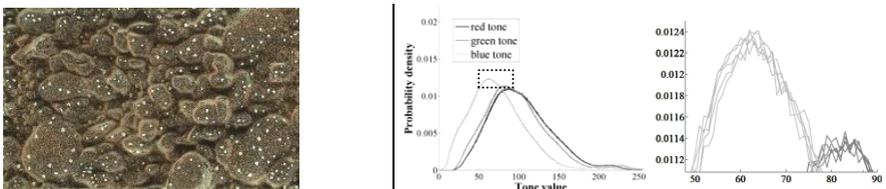
422 Quantitative evaluation and comparison of matching errors in DEMs of the 3D
 423 printed model and DEMs of the gravel bed was performed using the five DEMs that were
 424 collected for each case. It is important to note that the gravel-bed surface differs from the
 425 3D printed model, although they share some common properties (the same sediment
 426 mixture and the same armouring process). The motivation of this test was to examine
 427 whether the 3D printed model is suitable as a real gravel bed substitute in terms of
 428 stereomatching.

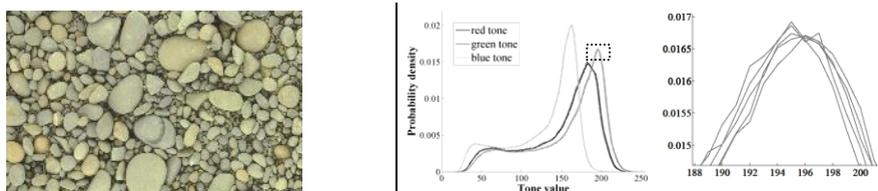
429 The five DEMs were the result of matching stereo pairs taken continuously (Fig. 5),
 430 and rectified with the same calibration results. Matching performance between stereo
 431 pairs will thus only differ because of differences in the images themselves. Fig. 5 (b)
 432 reveals that image data acquired with a 2 s interval timer differ, although variations are
 433 small (the standard deviation in overall pixel intensity (averaged over a given image)
 434 between the five images was of the order of 0.01). However, intensity variations between
 435 repeat photographs at some pixel locations are larger than one, and influenced the
 436 stereomatching. Pixel intensity is defined as the maximum tonal value between the red,
 437 green and blue channels, expressed on an 8-bit (0-255) scale. The neon lights were
 438 identified as the cause of these differences, as they consist of periodic light stripes.
 439 Because the left-hand images are different (compared with the right-hand images), the
 440 similarity in intensity (area-based) between the left and the right images may change
 441 between stereo pairs (Table IV).

442 The five DEMs were differentiated one by one to provide a measure of the repeat
 443 stereomatching error value (RSEV) over the investigated surface. This procedure is
 444 similar to the use of the ‘repeat scan error value’ by Hodge et al. (2009). Here, RSEV is a
 445 matrix, the same size as the measured data, containing the mean absolute height
 446 discrepancies at the grid nodes between all repeated measurements. In the case of five
 447 repeated measurements, this leads to 10 independent comparisons (Fig. 5). The RSEV
 448 was calculated as:

$$449 \quad \text{RSEV}(x,y) = \frac{\sum_{i=1}^5 \sum_{j=1}^5 |z_i(x,y) - z_j(x,y)|}{10}, \forall (i,j) \in [1,5] \cdot \text{with } i \neq j$$

450 where $z_i(x,y)$ is the elevation at grid node with coordinates (x,y) in scan or stereomatch
 451 i . Both the mean RSEV (after averaging the height differences over the grid) and the
 452 maximum RSEV were calculated over the 3D printed model and the gravel bed for
 453 various DEM sampling distances (from 0.25 to 3 mm), providing an estimate of
 454 stereomatching reliability over the two surfaces.





455 FIG. 5. Example of photographs (left), and image histograms (middle) in the red, green and blue channels, for
 456 the five left-hand images, reduced to the zones of interest shown in the photographs of (top) the 3D printed
 457 model and (bottom) the gravel bed. The enlargements on the right show the small differences between images.

458 TABLE IV. Area-based differences in pixel intensity between the two images forming a stereo pair, for the five
 459 stereopairs taken in a row, with evaluation limited to the zones of interest shown in Fig. 5.

<i>Difference in the mean pixel intensity</i>		<i>Difference in the median pixel intensity</i>	
<i>3D printed model</i>	<i>Gravel bed</i>	<i>3D printed model</i>	<i>Gravel bed</i>
0.45	0.80	0	0
0.44	0.79	1	0
0.45	0.77	0	0
0.44	0.79	1	0
0.45	0.80	1	0

460 **RESULTS**

461 *Qualitative Evaluation of DEM Generation Using Default Sampling Distance*

462 Fig. 6 presents DEMs obtained from recorded data of the 3D printed model and
 463 gravel bed, collected with a 0.25 mm sampling distance. The associated probability
 464 density functions (PDFs) of surface elevations are shown. Visually, both DEMs correctly
 465 represent the surface investigated. However, the enriched texture and greater colour range
 466 in images of the gravel bed are thought to contribute to a ‘sharper’ graphical
 467 representation of the gravel-bed DEM.

468 Fig. 6 also shows that the recorded surface of the 3D printed model has a larger
 469 range of elevations with a narrower distribution around the zero-mean than the measured
 470 surface of the ‘real’ gravel bed. This infers that the effect of water-working was stronger
 471 on the surface replicated by 3D printing, which agrees well with the visual observation of
 472 coarser exposed particles and thus deeper troughs between particles, compared to the
 473 ‘real’ gravel bed.

474

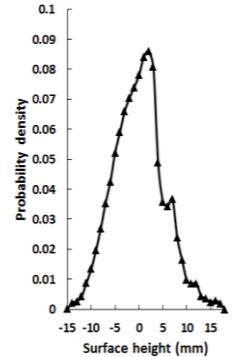
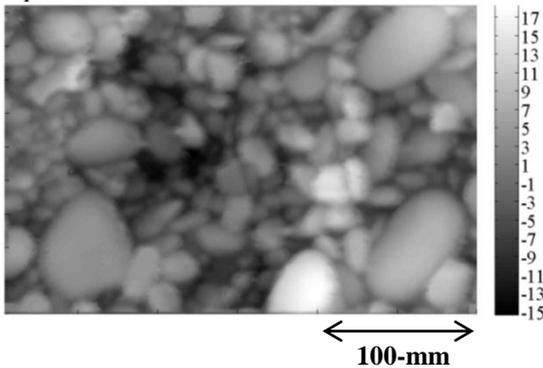
475

476

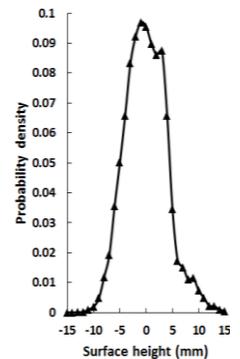
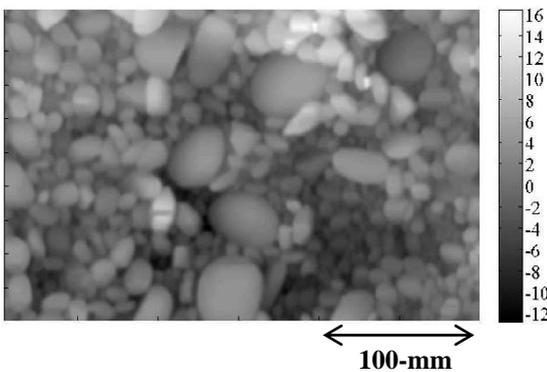
477

478

3D printed model



Gravel bed



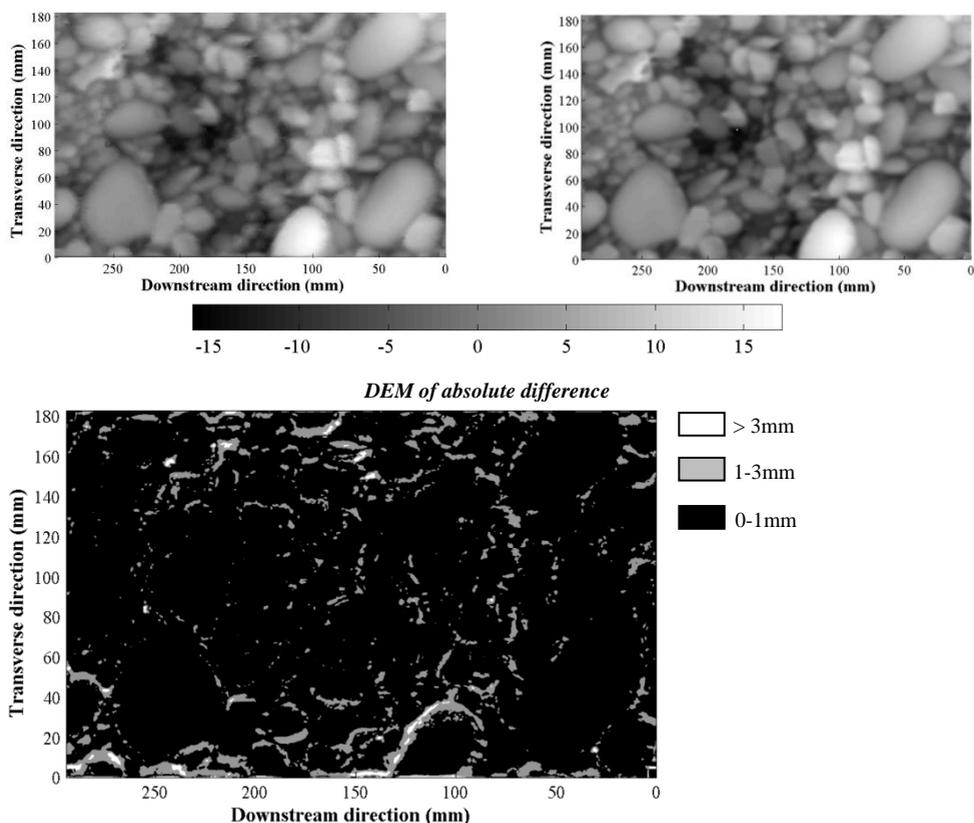
479 FIG.6. 2.5D DEMs collected with 0.25 mm sampling distance of (top) the 3D printed model and (bottom) the
 480 gravel bed. Elevations are in millimetres, and are represented with a gradient of greys.
 481 On the right are the associated PDFs of the DEM elevations.

482 *Quantitative Assessment of DEM Quality*

483 Fig. 7 shows the DoD (absolute differences are represented for clarity) obtained
 484 from the differentiation of the measured DEM (top left), after horizontal and vertical
 485 alignment, with the ground-truth DEM (top right), over the full dimensions of the 3D
 486 printed model. Visually, the two DEMs do not display obvious differences and analysis
 487 of the DoD is required. The DoD shows that surface errors are located at the grains'
 488 edges and gaps between grains. This observation relates well with findings obtained by
 489 Chandler et al. (2001), where stereomatching outliers were found near deep crevices
 490 between grains at the surface of a gravel bed. The reason is that occluded regions in the
 491 imagery, where pixels are replaced by values interpolated from neighbouring data, based
 492 on the assumption of a continuous surface, are smoothed in the DEMs and introduce
 493 surface errors.

Measurement

Ground truth



494 FIG.7. Top left: measured DEM of the 3D printed model collected with a 0.25 mm sampling distance,
 495 horizontally and vertically aligned with the ground truth. Top right: ground-truth DEM. Bottom: DoD over the
 496 full dimensions of the 3D printed model. The same colour representation, with elevations in millimetres, is used
 497 to represent both the measured and the ground-truth DEMs.

498 Table V shows statistics on the surface error, obtained from the comparison of the
 499 elevations contained in the DEM collected with default parameters with the ground truth
 500 elevations, over all check points (n=873 345). There is little systematic error in the
 501 measured surface, witnessed by a mean error (ME) which is close to zero. In this study,
 502 computations of the mean height discrepancies between the measurements and ground
 503 truth confirm that the vertical alignment was correctly performed. In the case of DEMs
 504 registered with an external coordinate system, such as by using total stations, evaluation
 505 of the mean error can provide information on the presence of systematic shifts in the
 506 recorded elevations (Butler et al., 1998; 2002; Carbonneau et al., 2003; Lane et al., 2000).

507 The evaluation of the average surface error, represented by a mean unsigned, or
 508 absolute, error (MUE) of 0.43 mm, is needed to characterise the DEM accuracy. The
 509 global surface precision is characterised by a standard deviation of 0.62 mm ($\approx 5\%$
 510 of surface D_{50}). Most DEM points (90%) are within ± 1 mm of check points, and 0.5% of
 511 DEM points exhibit errors above 3 mm (Fig. 7, bottom).

512

TABLE V. Quantitative assessment of the DEM collected with 0.25 mm sampling distance.

ME (mm)	0.04
MUE (mm)	0.43
SDE (mm)	0.62
Maximum absolute error (mm)	8.16
DEM points within ± 0.5 mm of check points (%)	71.4
DEM points within ± 1 mm of check points (%)	90.6
DEM points within ± 3 mm of check points (%)	99.5

513 The quality of the 'default' DEM is encouraging, as the actual accuracy of the
 514 measurements, represented by a MUE of 0.43 mm, is not severely degraded compared
 515 with the theoretical vertical accuracy of 0.39 mm, and is constrained by the image
 516 resolution. Fig. 7 confirms again that the loss in accuracy is associated with occlusions in
 517 the imagery, suggesting that the process steps undertaken to reconstruct the DEM were
 518 correctly executed, with minimal error propagation.

519 The quality of the DEM obtained in this study improves on previous
 520 photogrammetric measurements of gravel beds. However, any comparison is rather
 521 ambiguous, since the success rate of a photogrammetric survey essentially depends on the
 522 photogrammetric design and on the surface investigated, which differ between studies. In
 523 Carbonneau et al. (2003), mean errors from -1.5 to +3.6 mm, with surface precisions
 524 (SDE) ranging from 2.1 mm to 8.5 mm, were associated with the measurement of a dry
 525 natural gravel bed in the field, with a camera-to-object distance of 1.1 m. The surveyed
 526 surfaces were made of larger grains, with a D_{50} ranging from 18 mm to 61 mm, compared
 527 with a D_{50} of 13 mm in the present study. In the laboratory, a RMSE of 1.7 mm was
 528 obtained with a camera 'flying' height of 2 m by Chandler et al. (2001).

529 *Effect of DEM Grid Size on DEM Quality*

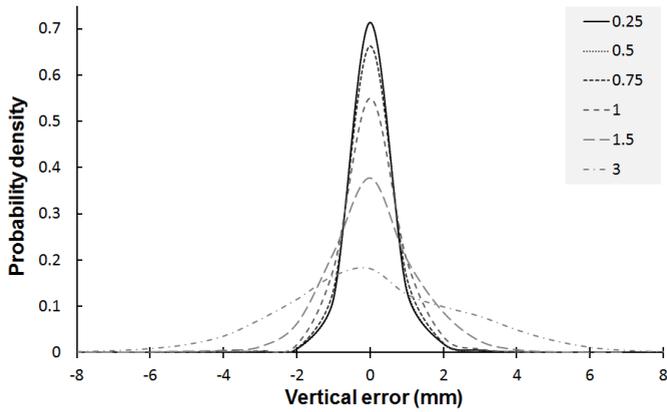
530 Fig. 8 presents the probability density functions of surface error for various grid
 531 spacings. No difference is visible when the sampling distance is increased from 0.25 mm
 532 to 0.5 mm. Above 0.5 mm, there is a consistent increase in the vertical error with
 533 increasing sampling distance (Fig. 8). Thus, increasing DEM grid size significantly
 534 affects the DEM quality.

535 Fig. 9 shows the effect of DEM grid spacing on statistical quality measures of the
 536 DEMs obtained from a single stereo pair and DEMs obtained from averaging five stereo
 537 pairs. It can be seen that the surface errors do not vary with DEM averaging. Fig. 9 agrees
 538 well with Fig.8 in that changing the sampling distance from 0.25 mm to 0.5 mm does not
 539 vary the DEM quality. When the sampling distance is increased beyond 0.5 mm, the
 540 DEM quality is reduced. In addition, the flattening of the graphs, with decreasing grid
 541 spacing, suggests that no improvement will result from reducing the sampling distance
 542 below 0.25 mm. This cannot be verified since check points for this study are sampled
 543 every 0.25 mm, and a denser grid of check points is needed to confirm this statement.

544 There is a direct correlation between the different statistical parameters and DEM
 545 grid size. A very small change is quantified when grid size is increased from 0.25 mm to
 546 0.5 mm, with an improvement or deterioration of the parameters of less than 1%. An
 547 average deterioration of 9% is associated with an increase of the grid size from 0.25 mm
 548 to 0.75 mm (the deterioration is 10.6% in the MUE and 7.7% for the percentage of DEM

549 points within ± 0.5 mm of the reference ground truth elevations).The deterioration is, on
 550 average, 34%, 102% and 308% (with a standard deviation between parameters of 4.3%,
 551 13.4% and 33.8%), when the grid size is increased to 1 mm, 1.5 mm and 3 mm,
 552 respectively.

553

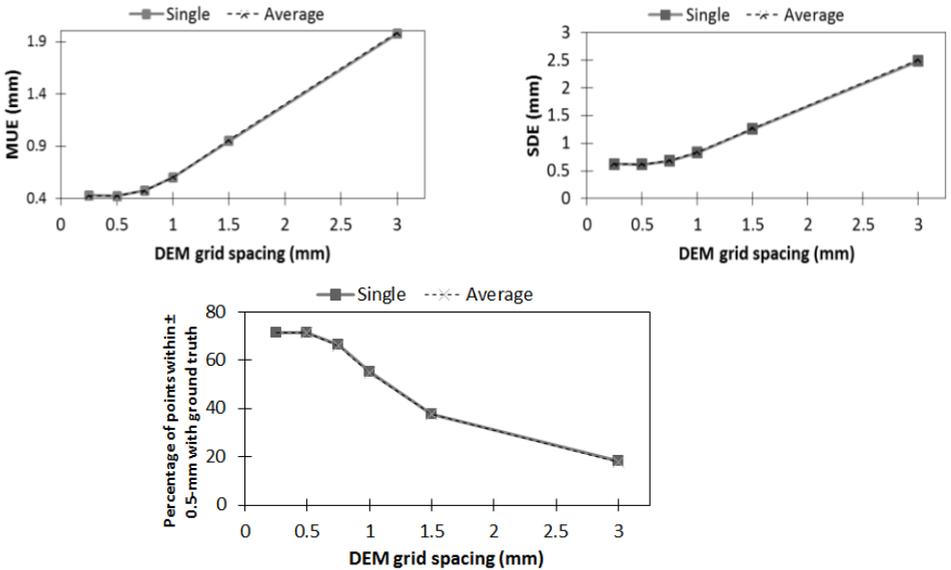


554

555 FIG.8. PDFs of vertical error for different grid spacing. Surface error is normally distributed, resulting in mean
 556 error close to zero. The PDFs for 0.25 and 0.5 mm grid spacings are virtually indistinguishable.

557

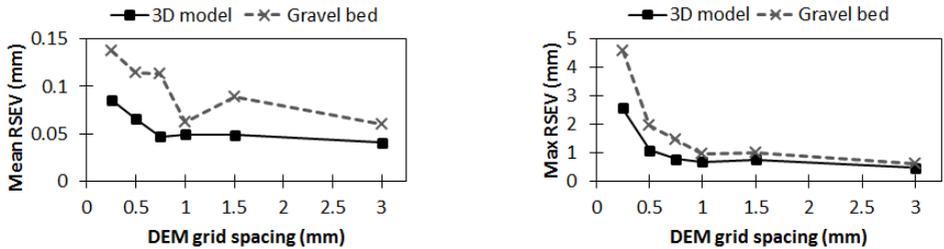
558



559 FIG.9. Effects of DEM grid spacing and averaging DEMs collected from different imagery on statistical
 560 measures of DEM quality. *Single* refers to a single DEM whereas *average* corresponds to the gridded elevations
 561 of the five DEMs averaged to produce one DEM.

562 *Internal Reliability of Stereomatching*

563 Fig. 10 presents the mean and maximum repeat stereomatching error values
 564 (RSEVs) for the DEMs of both the 3D printed model and the gravel bed, relating to
 565 different sampling distances. Overall, similar trends are observed, with reduced RSEVs
 566 for increased grid size. This is caused by smoothing DEM matching errors, due to more
 567 point cloud data being used for interpolating elevations at grid nodes.



568 FIG.10. Mean (left) and maximum (right) repeat stereomatching error value (RSEV) computed with 10
 569 independent DEM comparisons for the 3D printed model and the gravel bed.

570 Figs. 10 and 11 and Table VI confirm that stereomatching is more consistent for
 571 data of the 3D printed model, compared to data of the gravel bed, with a mean RSEV
 572 reduced by 35%. Similarly, the maximum RSEV is larger for the gravel bed (4.59 mm)
 573 than for the 3D printed model (2.57 mm).

574 Finishing the surface of the 3D printed model maximised stereomatching
 575 performance, which is exemplified by a mean RSEV of 0.09 mm. However, the
 576 assessment of gravel-bed DEM quality using the 3D printed model results in accuracy
 577 statistics that may be understated (which means that DEM quality of a natural gravel bed
 578 is to be revised downwards), since errors due to stereomatching are globally lower in the
 579 DEMs of the 3D printed model.

580 TABLE VI. Summary of RSEV for DEMs of 3D printed model and gravel bed collected with a 0.25 mm
 581 sampling distance.

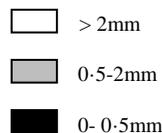
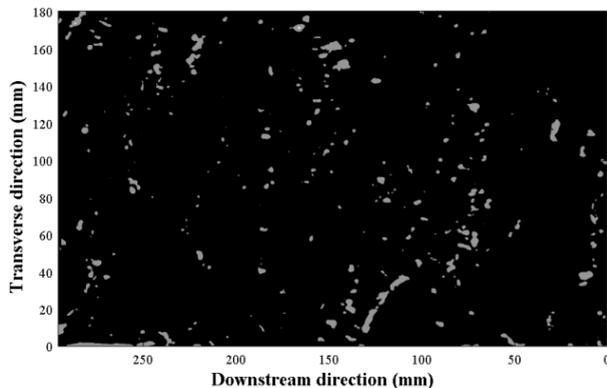
	<i>3D printed model</i>	<i>Grave l bed</i>
Mean (mm)	0.09	0.14
Maximum (mm)	2.57	4.59
DEM points with RSEV ≤ 0.5-mm (%)	96.1	89.08
DEM points with RSEV ≤ 2-mm (%)	>99.99	99.93
DEM points with RSEV > 2-mm (%)	<0.01	0.07

582 The ratio of the mean RSEV (0.09 mm) to the MUE (0.43 mm) is approximately 1:5
 583 for the DEMs of the 3D printed model. This confirms the observations made in previous
 584 studies (Butler et al., 1998), which suggested that image quality is an essential parameter
 585 in determining the success of photogrammetric surveys. However, image quality is often
 586 reduced by the texture contained in the imagery. Results of the present study, however,
 587 where DEMs are collected from different images of similar quality, suggest that intensity

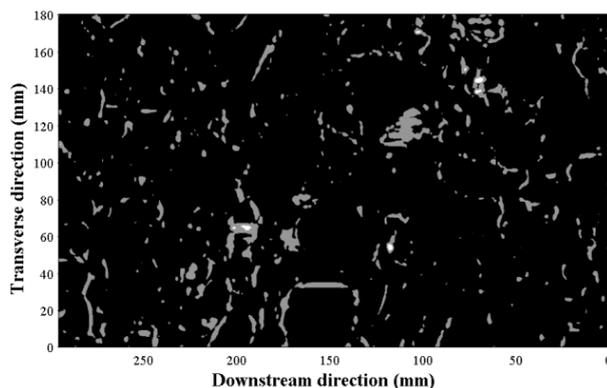
588 similarity between the two images forming a stereopair, on which stereo correspondence
 589 establishment relies, does affect DEM quality.

590 Whilst RSEVs cannot be ignored, especially at small sampling distances, averaging
 591 several DEMs collected from different imagery to mitigate stereomatching gross errors,
 592 proved to be of negligible interest in terms of measurement quality (Fig. 9). As an
 593 alternative, a filter, which removes and interpolates points in the average DEM, when
 594 RSEVs greater than 1 mm are detected, was tested on the DEM collected with default
 595 parameters. No significant change was observed. This is explained by the fact that
 596 stereomatching errors are essentially located in occluded regions of the DEM (Fig. 11),
 597 where elevations are already interpolated during stereomatching. Hence, a means to
 598 mitigate stereomatching errors in occluded regions of the DEMs has yet to be found,
 599 although the potential of multi-view stereo (MVS) methods might be investigated in this
 600 context.

3D printed model



Gravel bed



601 FIG. 11. Representation of the RSEV, which is the absolute difference between DEM points collected using
 602 different imagery, averaged over all ten independent comparisons (five repeat DEMs), for (top) the 3D printed
 603 model and (bottom) the gravel bed. The sampling distance is 0.25 mm.
 604

605

DISCUSSION

606 Methods for assessing DEM quality, using independent check points, can easily
607 result in misleading statistics, when the net of check points employed for the assessment
608 is of low density and/or of poor repartition (Lane, 2000). Moreover, accuracy statistics
609 might be insensitive to changes in DEM collection parameters. Another potential issue
610 relates to the confidence with which check points are estimated and registered within the
611 DEM. When measuring devices, such as total stations and laser scanners, are employed,
612 errors can propagate so that the actual survey precision of check points is decreased and
613 is generally not known.

614 The use of a 3D printed model to assess DEM quality improves greatly on previous
615 approaches, which have been characterised by the need for an additional instrument to
616 locate and register check points within the DEM. Furthermore, previously the preparation
617 of ground truth data has been repeated for each assessment and has required DEM
618 reconstruction and error editing, which is time consuming. This has been replaced, in the
619 current method, by the preparation of a reliable and practical 3D printed model. Thus,
620 future studies are not restricted by the common characteristics of check point measuring
621 devices, such as bulkiness and long recording times, which not only limits field
622 applications, but also the ability to record underwater check points. In future, a 3D
623 printed model provides a dense and accurate set of check data, where resolution is only
624 limited by the 3D printer capabilities. In this study, the 3D printer would have allowed
625 check points to be spaced with distances as small as 0.07 mm. This is beyond what is
626 achievable with the measurement techniques currently used in hydraulic applications.
627 Producing and handling such high-resolution data would still be difficult at present. In
628 addition, the depth accuracy with which the 3D model was printed, between 0.033 and
629 0.066 mm and with no error slip, ensured increased confidence in check point location.
630 The size of the 3D printed model, 296 x 184 mm, which is the maximum size that the 3D
631 printer available for the study allowed, might not be suitable for all applications. DEM
632 quality assessment solutions over larger measurement windows are either: (i) to move the
633 ground truth to determine the spatial distribution of DEM errors; or (ii) to produce several
634 ground-truth models that can be merged. However, the relatively small size of the model
635 allows it to be used easily at various locations in future, for both laboratory and field
636 setups.

637 The main uncertainties in the assessment using a 3D printed model result from the
638 alignment procedure necessary to superpose the measurements with the ground truth.
639 These uncertainties were present when metrics obtained with stereo photogrammetry
640 were compared with laser-scanned data (Chandler et al., 2001; Lane, 2000). In the
641 present study, the alignment was performed using search and surface de-trending
642 algorithms. The reliability of the procedure was assessed semi-independently by
643 comparing the setup vertical misalignment trends with different sampling distances. This
644 demonstrated that the alignment procedure is able to align data with great precision,
645 characterised in Table V by a mean error of 0.04 mm.

646 The present work also highlights the challenges that larger rigid 3D printed models
647 might face, due to the material's thermal contraction. There is limited information
648 available on preparation that ensures minimal thermal contraction during the printing and
649 drying phase. The observed deformation is largely dependent on the 3D printer and
650 material used, and care needs to be taken to design a suitable model. In this study, the

651 small warping effect the 3D print endured was corrected by mounting it on a flat and
 652 rigid Perspex sheet. The remaining trend of deformation was recorded by the setup, and
 653 was confirmed by multiple recordings and corrected during post-processing.

654 The inherent difficulty related to the use of a 3D printed model to assess surface
 655 error is the need to produce a realistic model of the surface to survey, since a global
 656 measure of error can be unreliable if it is based upon test sites that have little resemblance
 657 with the site of interest (Lane et al., 2005; Wang et al., 2009). Topography has an
 658 essential impact on measurement quality, especially with photogrammetric surveys, with
 659 occluded regions accounting for most of the errors detected on the surface (Chandler et
 660 al., 2001). A realistic topography was acquired by measuring a patch of gravel bed,
 661 representative of simulated riverbeds in laboratory flumes. However, because the ground
 662 truth was acquired with imagery, the interstices between particles are not always
 663 represented accurately. A technique which allows vertical recording, such as laser
 664 scanning, if available, may theoretically improve the details at the interstices between
 665 particles of the ground truth. To obtain a natural look for the transparent 3D printer
 666 material, the gravel-bed model was painted and sand was spread on the wet paint to
 667 ensure reliable stereomatching on images of the 3D printed model (mean and maximum
 668 RSEV of 0.09 mm and 2.57 mm, respectively). The difference in stereomatching
 669 performance between images of the 3D printed model and those of a laboratory gravel
 670 bed was quantified in terms of the repeat stereomatching error value (RSEV). This is
 671 proposed as an internal reliability test, addressing the stereomatching procedure. Larger
 672 RSEVs were obtained over the gravel bed (characterised by a 35% increase in mean
 673 RSEV), suggesting that stereomatching performs better, globally, over images of the 3D
 674 printed model. Hence, it is expected that DEM quality evaluated using the 3D printed
 675 gravel-bed model overestimates the DEM quality of natural gravel beds. For the 3D
 676 printed gravel-bed model, a ratio of 1 to 5 was obtained between the mean RSEV and the
 677 MUE. Assuming this ratio is valid for DEMs of natural gravel beds, this corresponds to
 678 an increase in the MUE from 0.43 mm for the 3D printed model to 0.7 mm for the
 679 natural gravel bed.

680

CONCLUSIONS

681 A 3D printed gravel-bed model was produced and subsequently used to evaluate the
 682 overall performance of a stereophotogrammetric setup developed for hydraulic
 683 experiments. The versatility and ease of use of the assessment method make it suitable to
 684 a large range of research areas, where small-scale and accurate topographic models are
 685 needed. The key advantages over traditional DEM evaluation methods can be
 686 summarised as follows:

- 687 (i) High density and uniform repartition of check points, which makes the DEM
 688 assessment more realistic and allows for a better quantification of errors due to
 689 changes in the DEM collection parameters, and as such allows optimisations to be
 690 performed on the photogrammetric workflow.
- 691 (ii) Strong (and measurable) confidence in check point locations and registration with
 692 the measured DEM.
- 693 (iii) Easier implementation of the assessment routine in a DEM acquisition campaign,
 694 with significant improvements in time and efficiency, and a larger range of
 695 situations where the assessment can be performed.

696 The results in this paper will be useful for setting up the optimum strategy for
 697 gravel-bed DEM collection, both in air and through water. Experiments have confirmed
 698 that the grid spacing should be chosen with care, in order to ensure correct surface
 699 representation. In particular, a sampling distance larger than 0.5 mm resulted in a
 700 significant loss of geometric information and degraded the DEM quality. Furthermore, it
 701 has been shown that the ground truth needs to be representative of the surface under
 702 consideration, in order to obtain the most realistic assessment. 3D printing is suitable, and
 703 at present the most advanced method, for obtaining realistic ground truth for known
 704 surfaces, although printed models can be affected by the material's thermal contraction.
 705 The acquisition of ground truth with imagery affects the surface details at interstices
 706 between particles.

707 The assessment of using off-the-shelf processing and consumer-grade digital
 708 cameras for a stereophotogrammetric setup showed that high accuracy and precision, in
 709 terms of sub-millimetre MUE and SDE, can be obtained with the presented default DEM
 710 collection parameters. Finally, this research confirmed the common expectation that most
 711 of the errors are caused by occluded regions in the imagery, a recurring stereo-
 712 photogrammetric measurement problem on rough surfaces. This suggests that new means
 713 for capturing occlusion, beyond traditional uniform surface smoothing, must be
 714 envisaged and the potential of employing multi-view stereo (MVS) methods to assist in
 715 this regard may be worthy of investigation. The ultimate goal is to obtain measurements
 716 with local vertical accuracy as close as possible to the optimal accuracy allowable by the
 717 setup.

718

REFERENCES

- BERTIN, S. and FRIEDRICH, H., 2014. Measurement of gravel-bed topography: an evaluation study applying statistical roughness analysis. *Journal of Hydraulic Engineering*, 140(3): 269-279.
- BERTIN, S., FRIEDRICH, H., CHAN, E. and DELMAS, P., 2012. The development and internal assessment of a high-resolution, non-proprietary, stereo-photogrammetric setup for hydraulic experiments. *Proceedings of 27th Conference on Image and Vision Computing New Zealand*, Dunedin, New Zealand: 515-520.
- BOUGUET, J.-Y., 2010. http://www.vision.caltech.edu/bouguetj/calib_doc/ [Accessed: 1st July 2013].
- BOURATSIS, P., DIPLAS, P., DANCEY, C. L. and APSILIDIS, N., 2013. High-resolution 3-D monitoring of evolving sediment beds. *Water Resources Research*, 49(2): 977-992.
- BRASINGTON, J. and SMART, R. M. A., 2003. Close range digital photogrammetric analysis of experimental drainage basin evolution. *Earth Surface Processes and Landforms*, 28(3): 231-247.
- BUTLER, J. B., LANE, S. N. and CHANDLER, J. H., 1998. Assessment of DEM quality for characterising surface roughness using close range digital photogrammetry. *Photogrammetric Record*, 16(92): 271-291.
- BUTLER, J. B., LANE, S. N. and CHANDLER, J. H., 2001. Characterization of the structure of river-bed gravels using two-dimensional fractal analysis. *Mathematical Geology*, 33(3): 301-330.
- BUTLER, J. B., LANE, S. N., CHANDLER, J. H. and PORFIRI, E., 2002. Through-water close range digital photogrammetry in flume and field environments. *Photogrammetric Record*, 17(99): 419-439.
- CARBONNEAU, P. E., LANE, S. N. and BERGERON, N. E., 2003. Cost-effective non-metric close-range digital photogrammetry and its application to a study of coarse gravel river beds. *International Journal of Remote Sensing*, 24(14): 2837-2854.
- CHANDLER, J., SHIONO, K., RAMESHWARAN, P. and LANE, S., 2001. Measuring flume surfaces for hydraulics research using a Kodak DCS460. *Photogrammetric Record*, 17(97): 39-61.
- CHANDLER, J. H., BUFFIN-BÉLANGER, T., RICE, S., REID, I. and GRAHM, D. J., 2003. The accuracy of a river bed moulding/casting system and the effectiveness of a low-cost digital camera for recording river bed fabric. *Photogrammetric Record*, 18(103): 209-223.
- COOPER, M. A. R. and CROSS, P. A., 1988. Statistical concepts and their application in photogrammetry and surveying. *Photogrammetric Record*, 12(71): 637-663.

- DETERT, M. and WEITBRECHT, V., 2012. Automatic object detection to analyze the geometry of gravel grains. *Proceedings of River Flow 2012*, San Jose, Costa Rica. Pages 595-600.
- FUSIELLO, A., TRUCCO, E. and VERRI, A., 2000. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12(1): 16-22.
- GIMEL'FARB, G., 2002. Probabilistic regularisation and symmetry in binocular dynamic programming stereo. *Pattern Recognition Letters*, 23(4): 431-442.
- GOOCH, M. J., CHANDLER, J. H. and STOJIC, M., 1999. Accuracy assessment of digital elevation models generated using the Erdas Imagine Orthomax digital photogrammetric system. *Photogrammetric Record*, 16(93): 519-531.
- HODGE, R., BRASINGTON, J. and RICHARDS, K., 2009. In situ characterization of grain-scale fluvial morphology using terrestrial laser scanning. *Earth Surface Processes and Landforms*, 34(7): 954-968.
- LANE, S. N., 2000. The measurement of river channel morphology using digital photogrammetry. *Photogrammetric Record*, 16(96): 937-961.
- LANE, S. N., JAMES, T. D. and CROWELL, M. D., 2000. Application of digital photogrammetry to complex topography for geomorphological research. *Photogrammetric Record*, 16(95): 793-821.
- LANE, S., CHANDLER, J. and PORFIRI, K., 2001. Monitoring river channel and flume surfaces with digital photogrammetry. *Journal of Hydraulic Engineering*, 127(10): 871-877.
- LANE, S. N., WESTAWAY, R. M. and HICKS, D. M., 2003. Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing. *Earth Surface Processes and Landforms*, 28(3): 249-271.
- LANE, S. N., REID, S. C., WESTAWAY, R. M. and HICKS, D. M., 2005. Remotely sensed topographic data for river channel research: the identification, explanation and management of error. In: *Spatial Modelling of the Terrestrial Environment* (Eds. R.E.J. Kelly, N.A. Drake and S.L. Barr), chapter 6. John Wiley & Sons, Ltd, Hoboken, NJ, USA. 290 pages: 113-136.
- MARION, A., TAIT, S. J. and MCEWAN, I. K., 2003. Analysis of small-scale gravel bed topography during armoring. *Water Resources Research*, 39(12): 1334-1345.
- MATTHEWS, N. A., 2008. Aerial and close-range photogrammetric technology: providing resource documentation, interpretation, and preservation. *U.S. Department of the Interior, Bureau of Land Management, Technical Note 428*. National Operations Center, Denver, Colorado, U.S.A. 42 pages.
- OCKELFORD, A.-M. and HAYNES, H., 2013. The impact of stress history on bed structure. *Earth Surface Processes & Landforms*, 38(7): 717-727.
- POLLYEA, R. M. and FAIRLEY, J. P., 2012. Experimental evaluation of terrestrial LiDAR-based surface roughness estimates. *Geosphere*, 8(1): 222-228.
- QIN, J., ZHONG, D., WANG, G. and NG, S. L., 2012. On characterization of the imbrication of armored gravel surfaces. *Geomorphology*, 159-160: 116-124.
- RICE, S. P., BUFFIN-BÉLANGER, T. and REID, I., 2014. Sensitivity of interfacial hydraulics to the microtopographic roughness of water-lain gravels. *Earth Surface Processes and Landforms*, 39(2): 184-199.
- RIEKE-ZAPP, D. H., ROSENBAUER, R. and SCHLUNEGGER, F., 2009. A photogrammetric surveying method for field applications. *Photogrammetric Record*, 24(125): 5-22.
- SMITH, M., VERICAT, D. and GIBBINS, C., 2012. Through-water terrestrial laser scanning of gravel beds at the patch scale. *Earth Surface Processes and Landforms*, 37(4): 411-421.
- WANG, C.-C., HEFNER, B. T. and TANG, D., 2009. Evaluation of laser scanning and stereo photography roughness measurement systems using a realistic model seabed surface. *IEEE Journal of Oceanic Engineering*, 34(4): 466-475.
- ZHANG, Z., 1999. Flexible camera calibration by viewing a plane from unknown orientations. Proceedings of 7th IEEE International Conference on Computer Vision. Pages 666-673.

Résumé

Zusammenfassung