

# Automated grain sizing from uncrewed aerial vehicles imagery of a gravel-bed river: Benchmarking of three object-based methods

Raphaël Miazza<sup>1</sup>  | Ivan Pascal<sup>1,2</sup>  | Christophe Ancey<sup>1</sup> 

<sup>1</sup>Laboratoire d'Hydraulique Environnementale, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

<sup>2</sup>Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Lausanne, Switzerland

## Correspondence

Raphaël Miazza, Laboratoire d'Hydraulique Environnementale, École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland.  
Email: [raphael.miazza@alumni.epfl.ch](mailto:raphael.miazza@alumni.epfl.ch)

## Funding information

Swiss National Science Foundation, Grant/Award Number: 200020 204108

## Abstract

Measuring grain sizes in gravel-bed rivers is crucial when studying river dynamics and sediment transport. Automated methodologies have been developed in recent years for detecting individual grains and measuring their size on digital imagery. These object-based methodologies have mainly been applied to handheld imagery. Low-cost and high-resolution orthoimages covering long river reaches are nowadays accessible with the improvements in uncrewed aerial vehicles (UAV) and structure-from-motion (SfM) photogrammetry. Applying object-based grain-sizing methodologies to such orthoimages may provide wide-scale information about the grain size spatial distribution along streambeds. We examined how accurate three object-based models (BASEGRAIN, PebbleCountsAuto and GALET) were, by comparing their outcomes to in-field manual measurements of grain sizes and manual grain labelling. We found that BASEGRAIN and PebbleCountsAuto underestimated grain sizes on average, whereas GALET generally overestimated grain size percentiles. Grain size measurements obtained by manually labelling grain features were consistent with in-field measurements.

## KEYWORDS

automated grain size measurements, grain size distribution, gravel-bed rivers, object detection, structure-from-motion photogrammetry, UAV

## 1 | INTRODUCTION

River beds are seldom composed of sediments of uniform size. Quantifying the size distribution of sediment mixtures is fundamental for understanding and modelling river flows, morphodynamics and sediment transport processes (Bunte & Abt, 2001). For instance, a number of flow resistance equations involve grain size quantiles (e.g., Ferguson, 2007; Hey, 1979; Keulegan, 1938; Rickenmann & Recking, 2011; Smart & Jaeggi, 1983; Strickler, 1923). Another example is provided by the methods for predicting the threshold of incipient motion and transport rates (e.g., Hodge, Richards, & Brasington, 2007; Parker, Klingeman, & McLean, 1982; Recking, 2013, 2016). When studying bedload transport in mountain rivers—whose sediment deposits are often prone to grain sorting processes, such as bed armouring and size-dependent selective transport—accurate quantification of local grain

size distributions is particularly important for assessing bedload transport mechanisms (e.g., Schlunegger, Delunel, & Garefalakis, 2020).

A variety of field sampling procedures have been developed for characterising river-bed grain sizes in quantitative terms. Mechanical sieving of volumetric samples and individual measurements of surface grains are commonly used techniques (Kellerhals & Bray, 1971). Nonetheless, these procedures can be laborious and entail limitations for the assessment of spatial variations of bed-material grain sizes (Bunte & Abt, 2001). To mitigate these issues, automated grain-sizing methods based on remote sensing technologies have emerged in recent decades. A number of research studies have been dedicated to utilizing images for measuring grain sizes (see Carrivick & Smith, 2019; Piégay et al., 2020). Parallely, recent improvements in uncrewed aerial vehicles (UAV) and structure-from-motion (SfM) photogrammetry software packages have allowed

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Earth Surface Processes and Landforms* published by John Wiley & Sons Ltd.

to produce easily and at relatively low-cost valuable topographic datasets, which typically comprise orthoimages, digital surface models (DSM) and dense point clouds. Combining UAV-SfM imagery and image-based grain sizing techniques can provide access to grain size estimates over large spatial scales, in a cost- and time-efficient manner (Carrivick & Smith, 2019).

The first image-based methodologies for measuring grain sizes in rivers relied on the visual interpretation of nadir photographs (e.g., Adams, 1979; Ibbeken & Schleyer, 1986). Grain size distributions (GSD) were estimated by measuring the projected intermediate axis (b-axis) of grains on photographs. Although less field time is required (only photographs need to be taken), this method still necessitates large processing times to measure grain features individually (Bunte & Abt, 2001). Since the early 2000s, numerous studies were dedicated to automatizing grain-sizing procedures on images (Butler, Lane, & Chandler, 2001; Graham, Rice, & Reid, 2005; McEwan et al., 2000). Two distinct approaches have been followed. The first approach involves deriving grain sizes from image statistics (Buscombe, 2013; Carbonneau, Lane, & Bergeron, 2004; Woodget & Austrums, 2017; Woodget, Fyffe, & Carbonneau, 2018). This methodology derives characteristic grain sizes using image texture metrics, autocorrelation or wavelet transformations. Such methods are based on regression between single parameters and grain sizes (e.g., Carbonneau, Lane, & Bergeron, 2004; Warrick et al., 2009; Woodget, Fyffe, & Carbonneau, 2018), or on more novel methods based on Convolutional Neural Networks (CNN; e.g., Buscombe, 2020; Lang et al., 2021). The second approach focuses on detection and measurement of individual grain features, and is thus referred to as 'object-based'. Individual surface grains can be identified by using image thresholding and segmentation processing (e.g., Detert & Weitbrecht, 2012; Graham, Rice, & Reid, 2005; Purinton & Bookhagen, 2019), or with most recent object detection algorithms based on CNN (e.g., Chen, Hassan, & Fu, 2022; Mair et al., 2023; Mörthl et al., 2022; Soloy et al., 2020). The advantages of object-based grain sizing over image statistics-based methods are (i) that the former does not require site-specific calibration and (ii) that one can derive more information from object-type data (e.g., grain arrangement). Object-type data can also be converted into grid-type data whose cell provide the local GSD.

The grain size distribution in gravel-bed rivers varies considerably at the local scale (Bluck, 1979). This variability has been studied by combining scattered manual measurements of grain sizes and visual analysis (e.g., Dietrich et al., 2006; Guerit et al., 2014; Lisle & Madej, 1992). The earliest object-based grain-sizing tools assumed that images were taken by handheld or pole-mounted cameras (e.g., Butler, Lane, & Chandler, 2001; Graham, Rice, & Reid, 2005). This imagery approach provides access only to small spatial scales (up to few tens of meters). Applying object-based grain-sizing methods to orthoimages issued from UAV surveys can convey information on the spatial variation in the grain size distribution, covering scales from decimetres to hectometres.

The quality of the grain size information derived from orthoimages and its potential for understanding geomorphological processes depend a great deal on the accuracy of the applied technique. Therefore, testing the latest object-based methodologies on high-

resolution orthoimages is key to identify specific limitations and biases. Mair et al. (2022) evaluated the uncertainties in grain size measurements on aerial imagery with regard to the UAV-SfM approach. A performance assessment of a set of existing grain-sizing routines has been conducted by Chardon, Piasny, & Schmitt (2022) for applications to handheld imagery. To the best of our knowledge, no study evaluated multiple object-based techniques for applications to high-resolution orthoimages.

In this context, we evaluated three object-based grain-sizing software routines (BASEGRAIN, Detert & Weitbrecht, 2012; PebbleCountsAuto, Purinton & Bookhagen, 2019; GALET, Mörthl et al., 2022) and compared their outcomes to in-field manual measurements of grain sizes and manual labelling on orthoimages. We wanted to answer the following questions. First, which automated grain-sizing software performs the best on orthoimages of a gravel-bed river? Second, which are the limitations of each tool?

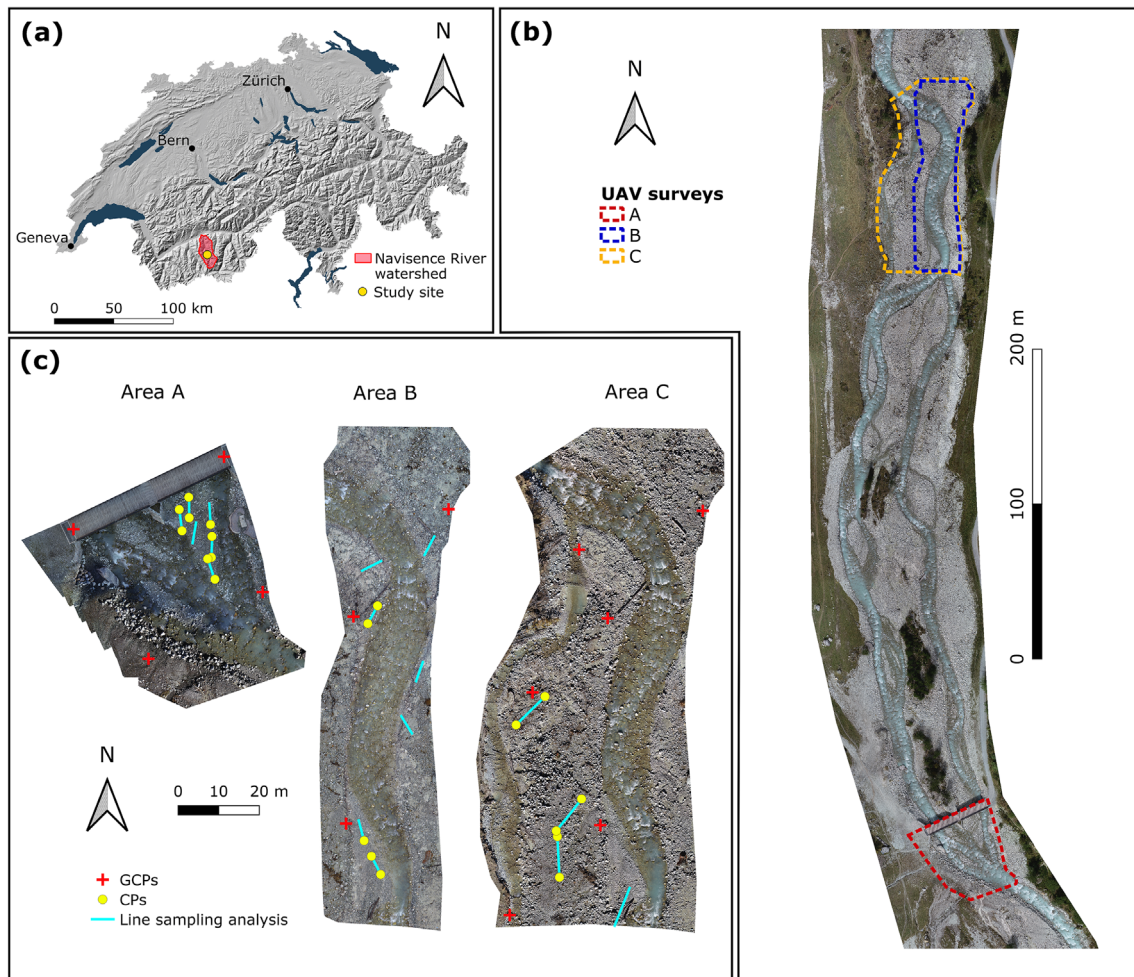
## 2 | METHODS

The surface grain size distribution of a mountain river (Section 2.1) was investigated by conducting line sampling (Section 2.2). We designed UAV surveys to reconstruct orthoimages of the study site based on SfM algorithms (Section 2.3). Grain sizes were measured digitally on these orthoimages by using manual labelling (Section 2.4) and three object-based grain-sizing methods (Section 2.5). These digitally measured grain sizes were compared with grain sizes measured manually in the field (Section 2.6).

### 2.1 | Study site

The Navisence is a mountain river located in the South-West Swiss Alps, tributary to the Rhône River (Figure 1a). This 23-km-long river drains a 257 km<sup>2</sup> catchment. Its main water source is the Zinal Glacier at 2300 m a.s.l. The river is hydrologically undisturbed upstream of the village of Zinal.

The study site is a 500-m-long and 60–90-m-wide river reach, located in the upstream part of a 2-km-long floodplain named 'Plats de la Lée' with an average slope around 3% (Figure 1b). The Navisence flows across this alpine floodplain and develops a braided network upstream of the village of Zinal (1650 m a.s.l.). There, the catchment area is 77 km<sup>2</sup>. A gauging station managed by a Walliser research institute (CREALP, Sion) is located downstream of the study reach, and has gathered data since 2011 (flow rates and bedload transport rates). The river has a glacio-nival hydrological flow regime, with very low flow rates in winter and high discharges in summer related to snow and glacier melting, with significant circadian variations (Travaglini et al., 2015). The typical low flow discharge is 1 m<sup>3</sup>/s in winter, while maximum hourly discharge can exceed 25 m<sup>3</sup>/s in summer. Over the last 5 years, the morphology of the braided network has been mostly impacted by a single major flood in July 2018. Regarding the sediment lithology, the surface alluvial deposits found in the river bed at the Plats de la Lée are mostly composed of metamorphic rocks (mainly orthogneiss). Therefore, sediments found on the study reach often



**FIGURE 1** (a) Location of the study site and the Navisence River watershed in Switzerland (© swisstopo). (b) Location of the UAV surveys over the study reach (the river flows northwise). The river image is an orthoimage obtained from a UAV survey carried out on 13 Sep 2022 (50-m-high flight). (c) Detailed view of the three orthoimages which were reconstructed from UAV collected images, with positions of the line sampling analysis, ground control points (GCPs) and check points (CPs). [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.com)]

exhibit variations of rock texture inside single grains, due to foliation or veins for example.

## 2.2 | Manual measurements of grain sizes

We used the line sampling procedure proposed by Fehr (1987). This procedure has been specially devised for mountain rivers, and has thus been used in numerous field studies of hydraulics and sediment transport in mountain rivers across the Alps (e.g., Konz et al., 2011; Ramirez et al., 2022; Rickenmann, 1997; Rickenmann & McCardell, 2007; Schneider et al., 2016). There are alternatives to Fehr's method; for instance, grid sampling protocols such as the Wolman pebble count (Wolman, 1954) and its variants are much more common than line sampling. The main advantage of Fehr's method over other techniques is that it is much easier to georeference a sample line than a grid (or a random walk) in the field and in orthophotographs. When comparing grain size estimates from field measurements and automated methods, it is thus possible to sample the same stones in the field and orthophotographs and thus, by doing so, we can benchmark methods on a fair basis. Another major advantage of Fehr's method is that by considering all the pebbles crossed

by the sampling line, it requires a smaller sampling space than methods based on areal or random particle sampling, which is of great interest for characterising spatial variations in grain size distribution when the riverbed involves patches of distinct grain sizes. A possible disadvantage of Fehr's method is that its accuracy has been much less studied than Wolman-based methods (Daniels & McCusker, 2010; Rice & Church, 1996). The recommendations drawn for Wolman-based methods should, however, apply to Fehr's method: the sampling line has to cross a large number of stones—as large as 400 according to Rice & Church (1996)—if sufficient precision in the quantile (or percentile) estimation is desired. In practice, this requirement is problematic for mountain rivers like the Navisence because of spatial variations in grain size over short distances. From a statistical viewpoint, the central limit theorem can help understand why accuracy of grain size quantiles varies as the sample size's square root as Rice & Church (1996) found in their field study, and why this result holds only when the sample is drawn from the same population, with the same mean and variance. When the grain size distribution exhibits substantial spatial variations, obtaining unbiased estimates becomes particularly difficult: (i) if the sample size is small, then it is probably representative of the local population, but estimates are inaccurate, (ii) if the sample size is large, then in principle, a higher accuracy is

expected in quantile estimation, but measurements are biased. The Navisence riverbed is characterised by a typical median stone size of around 10 cm and patch length of approximately 10 m. These features led us to consider that a sample size of about 150 pebbles—as proposed by Fehr (1987)—provided a suitable trade-off between precision and representativeness.

We collected samples over 17 lines distributed over areas A, B and C (Figure 1c). Following Fehr (1987), we stretched a string over the dry bed-material surface to be analysed. The b-axis of all pebbles underneath the string was measured. Pebbles with a b-axis smaller than 1 cm were not considered. The pebbles were divided into diameter classes and the number of pebbles falling in each grain size interval was computed (the corresponding diameter classes are specified in the Supporting Information, we used the sampling protocol provided in Fehr (1987)). Approximately 150 pebbles should be measured to ensure a good representativeness. This led us to choose sampling lengths of 5 or 10 m depending on the local grain size. The largest pebbles (over 10 cm) were often imbricated or clogged, which prevented a correct measurement of the b-axis. They were therefore manually extracted, which required significant effort and the use of a pickaxe. For one sample and a single operator, this procedure lasted about 1 hour. To georeference each line on geographic information systems, we measured the starting and endpoint positions using a pole-mounted GPS/GNSS system—Leica Zeno 20 coupled with a GG04P antenna, with real-time kinematic correction (Swipos-GIS/GEO network).

Fehr's method involves converting the line samples into approximate volumetric-sample equivalents of the subsurface grains via empirical relations between surface and subsurface grain sizes (Fehr, 1987). The frequency-by-number grain size distribution is converted into a frequency-by-weight distribution (describing the weight fraction of each grain size interval), so that the results are comparable with standard volumetric sampling. The conversion is based on the voidless cube model (Kellerhals & Bray, 1971), which was empirically adapted by Fehr (1987). As grains whose b-axis is smaller than 1 cm are neglected during the sampling process, the cumulative frequency of the components larger than 1 cm has to be corrected to take neglected finer components into account. According to Fehr (1987) observations based on field mechanical sieving in a large set of Swiss gravel-bed rivers, 20% to 30% of the subsurface layer volume is smaller than 1 cm in diameter. Finally, the GSD is extrapolated towards the finest grain sizes. Fehr (1987) observed that for the Swiss Alps, the distribution of the fine fraction of the bed and bedload material generally follows a Fuller curve. When predicting the proportion of fine material in the GSD, Fehr assumed that the final GSD follows a Fuller curve for the undersampled finest grain sizes (see Supporting Information for the detail). We consider this tail correction to be well suited to our field

site, whose bed is mostly structured by coarse particles and clogged by glacier flour.

## 2.3 | UAV surveys and structure-from-motion photogrammetry

### 2.3.1 | Areas A, B and C: data acquisition

Three UAV surveys were carried out over different sectors of the study reach. The covered areas were named A, B, and C and their location is shown in Figure 1b,c. We conducted these surveys in order to evaluate the accuracy of digital object-based grain-sizing tools on orthoimages, compared with in-field line sampling. Nine manually sampled lines were located in area A, seven in area B and four in area C. We used a DJI Phantom 4 pro and a DJI Phantom 4 pro v2 UAVs. These rotary-wing quadcopters are equipped with a GPS for automated flights. They have an integrated camera with a 20-mega-pixel resolution. The automated flights were planned using Pix4Dcapture software (v. 4.13.1; developed by Pix4D, Lausanne, Switzerland). Images were taken vertically on a predefined trajectory, with a frontal and lateral overlap between individual images in the order of 70%. During image acquisition, the UAV stayed stationary to avoid motion blur. It then moved to the next predefined position along the grid line map. In order to obtain the best compromise between image resolution and spatial coverage, we conducted our flights at an elevation of approximately 10 m above the take-off position. As the UAVs flew horizontally, the effective flight height varied depending on ground slope and local topographic features. Ground resolution ranged from 2.9 mm/px to 3.7 mm/px. The survey C was conducted under sunny conditions, whereas the surveys over area A and B were conducted under shaded conditions, on clear days and before the sun illuminated the study reach (see Table 1). These three UAV surveys were always performed before manual line sampling. Further information about the camera parameters used can be found in Tables S1 and S2.

### 2.3.2 | Auxiliary georeferenced points

Ground control points (GCP) were distributed over the surveyed areas to constrain more accurately the SfM photogrammetric reconstruction and to assess errors. The GCP were marked with paint, and their position was measured using the same GPS/GNSS system described in Section 2.2. This provided a horizontal positioning accuracy close to 1 cm and a vertical accuracy in the 2- to 4-cm range. The GCP coordinates were measured in the CH1903+/LV95 coordinate system (EPSG 2056). During the SfM reconstruction preprocessing steps, we located the GCP position on the images. The start and endpoints of

**TABLE 1** Summary of the UAV surveys.

Area name	Area size (m <sup>2</sup> )	Date	Weather	UAV model	Number of images	Number of tie points	Number of GCP/CP	Ground sampling distance (mm/px)
A	2500	5 Oct 2022	Shady, clear sky	DJI Phantom 4 pro	232	6 070 479	4/9	2.9
B	3450	12 Oct 2022	Shady, clear sky	DJI Phantom 4 pro	223	6 053 898	3/5	3.7
C	5850	26 Oct 2022	Sunny	DJI Phantom 4 pro v2	419	12 511 931	6/6	3.3

Abbreviations: CP, check points; GCP, ground control points.



the manually sampled lines served as independent check points to evaluate the accuracy of the reconstructed orthoimages. Some start and endpoints could not be located with certainty on the basis of the photographs taken in the field (14 out of 34), and were therefore not considered check points. The resulting orthoimage positioning errors are given in Table 2.

### 2.3.3 | Data processing

Geo-referenced orthoimages were obtained by processing the images with the Pix4Dmapper software (v. 4.8.0; Pix4D, Lausanne, Switzerland). By combining SfM photogrammetry and multi-stereo view algorithms, the software reconstructs the three-dimensional (3D) surface topography. In the SfM framework, the 3D positions of a large set of features automatically extracted from images are retrieved, simultaneously with camera positions and orientations by iteratively solving a highly redundant system of triangulation equations (Westoby et al., 2012). This method provides a point cloud, which can then be converted into a DSM and an orthoimage. The main parameters used in Pix4Dmapper can be found in Table S3.

## 2.4 | Digital manual labelling

Manual labelling was performed by a single operator on the orthoimages using the QGIS software (v. 3.22). We manually drew a polygon on all visible grain features intersected by the georeferenced lines. This labelling operation took approximately 10–15 min per line. We measured the b-axis of all labelled grains by automatically fitting an ellipse to each feature (fitting based on the second central moment of the object geometry). The detailed ellipse fitting procedure is described in the Supporting Information. The b-axis corresponds to the minor axis of the fitted ellipse. This procedure for extracting the b-axis of each labelled grain is similar to the ellipse fitting procedure used by the object-based grain-sizing methods described in Section 2.5. Finally, we derived the GSD from the b-axis of the identified grain features for each line using Fehr's (Fehr, 1987) method implemented in a Python script.

## 2.5 | Description of selected object-based grain-sizing tools

In this section, we describe the internal frameworks of the three object-based grain-sizing tools under investigation. We specify how

each tool was implemented to derive grain size distributions that are comparable with those based on in-field manual samples. Figure 2 shows an example of detected grain features along a line using the different methodologies.

### 2.5.1 | BASEGRAIN

We used BASEGRAIN (v.2.3), which is a free access MATLAB-based method developed by Detert & Weitbrecht (2012). It performs individual grain segmentation on digital top-view photographs in five preprocessing steps. Three out of five steps require supervised parameter tuning to optimise performance. Ellipses are then fitted to the detected objects, and the minor axis is considered as the b-axis of the grain feature.

Orthoimages were cut out into image tiles corresponding to the bed surface patches where the line sampling was conducted. This splitting was required because BASEGRAIN is not able to handle georeferenced orthoimages. The image tiles were rotated in order to position the line vertically. The BASEGRAIN processing was much influenced by the variations in colour and texture of the Navisence sediments, thus no unique set of parameters allowed for optimised object detection for all image tiles. Therefore, object detection had to be performed individually for each image tile when using BASEGRAIN. We tuned the different parameters until a visually optimal segmentation of the grains was obtained. No postprocessing was applied to the detected objects. The virtual sampling line implemented in BASEGRAIN was placed so as to match the position of each field sampling line. We extracted the dimensions of all the grains detected and intersected by the line in BASEGRAIN. We then tallied grains using the same grain size intervals as those used for manual sampling, and computed the GSD according to Fehr's (Fehr, 1987) method.

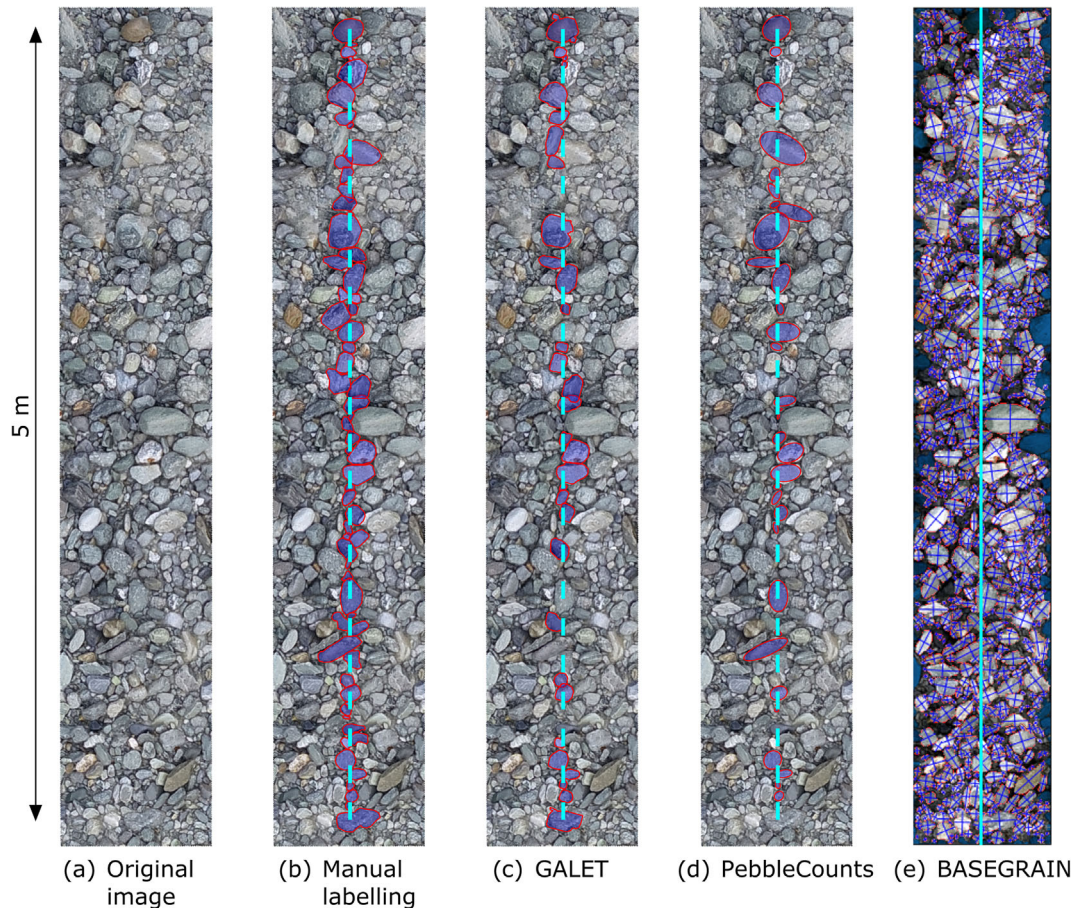
### 2.5.2 | GALET

GALET is a deep-learning image segmentation model for grain size analysis developed by Styx4D (Bourget-du-Lac, France) and presented in Mörtl et al. (2022). The CNN model implemented in GALET was trained using a dataset generated by a technique based on automated image creation. Mörtl et al. (2022) used manually cropped grain images and synthesised artificial grain images to generate labelled training images. During the grain detection and segmentation steps in GALET, orthoimages are split into 512 or 1024 pixel-large tiles. The software performs grain detection and estimates the shapes of

**TABLE 2** Quality assessment of structure-from-motion photogrammetry results.

Area name	Ground control points						Check points					
	X (cm)		Y (cm)		Z (cm)		X (cm)		Y (cm)		Z (cm)	
	ME	STDEV	ME	STDEV	ME	STDEV	ME	STDEV	ME	STDEV	ME	STDEV
A	0.0	0.8	0.0	0.9	0.0	1.0	0.2	1.5	-2.8	2.3	0.3	2.8
B	0.0	0.8	0.0	0.5	0.0	0.0	-0.9	1.5	-1.8	2.1	-4.2	2.0
C	0.0	1.0	0.0	0.9	0.0	1.2	0.8	1.8	0.7	1.7	2.8	1.4

Note: Mean error (ME) and standard deviation of error (STDEV) on ground control points and check points after bundle block adjustment.



**FIGURE 2** Example of digital line sampling with manual labelling and the software routines investigated. The blue line corresponds to the location of the georeferenced line. The grain features displayed for manual labelling and GALET are the original grain feature polygons and not the fitted ellipses. The BASEGRAIN image originates from the software GUI. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/esp.5782)]

overlapping grains. A final shapefile is produced in which all detected instances are vectorised.

The entire GALET segmentation process was applied on each orthoimage. The routine detected the grain features and provided the corresponding shapefiles. We measured the b-axis of all detected grains by automatically fitting an ellipse using the same method as the one utilised for manual labelling. Digital line sampling was performed on these shapefiles in QGIS, as the positions of the field sampling lines were georeferenced. We classified the vectorised grains that intersected the georeferenced lines according to their b-axis. The same grain size intervals used for manual sampling were considered and the GSDs were computed by using the Fehr (1987) method.

### 2.5.3 | PebbleCountsAuto

PebbleCounts is an open-source Python-based algorithm, developed by Purinton & Bookhagen (2019). Here, we used its highly automated version named PebbleCountsAuto. It is an image segmentation method that performs individual grain detection. Resulting grain features are measured via ellipse fitting.

As PebbleCountsAuto requires significant computing time to process entire orthoimages (Purinton & Bookhagen, 2021), grain feature detection was performed on image tiles cut out from the orthoimages (similarly to the procedure used with BASEGRAIN). These tiles corresponded to the locations where in-field line sampling was

performed. PebbleCountAuto only required us to manually tune the threshold level of Otsu's threshold matrix. This parameter was tuned for each processed image tile in order to obtain a visually optimal segmentation of grain features. The default parameter defining the minimum area in pixel for a feature to be considered a grain was modified and set at 23 pixels in order to be consistent with the same parameter defined in BASEGRAIN—value based on the limit of grain feature detectability in images (see Graham, Rice, & Reid, 2005). The model uses a size cut-off criterion to discard grains whose b-axis is too small. By default, the cut-off value is set at 20 px, but we decreased it to 3 px so that all grains with a b-axis exceeding 1 cm could be considered. PebbleCountsAuto allows one to work with georeferenced orthoimages. Therefore, information about detected grain features such as northing and easting coordinates in an UTM coordinate system, major and minor axis of the fitted ellipses and their orientation could be exported as text files. These data were used to reconstruct the detected features as georeferenced ellipses in QGIS. Digital line sampling was then performed by computing the GSD from grain features intersected by each georeferenced line.

## 2.6 | Accuracy evaluation

The grain size percentiles obtained by manually labelling images and those obtained by applying the three grain-sizing tools were compared with the grain sizes retrieved by on-field manual line sampling.

The comparisons were done by normalising all digitally obtained grain size percentiles by their corresponding in-field manually measured grain size percentiles as follows:

$$d_{\text{norm}} = \frac{d_{\text{digital}}}{d_{\text{manual}}}$$

where  $d_{\text{digital}}$  corresponds to the digitally obtained grain size percentiles, while  $d_{\text{manual}}$  corresponds to the grain size percentiles obtained from in-field manual sampling. Therefore, if the digital measurements were accurate,  $d_{\text{norm}}$  should be close to unity.

We considered the normalised root mean square error (NRMSE) to quantify the errors of digitally based grain size percentiles in terms of the corresponding fraction of the mean grain size percentiles obtained from in-field manual sampling. This error metric allowed us to directly compare the accuracy of the model estimates for different grain size percentiles, even if the grain scales were different (like for  $d_{16}$  and  $d_{84}$  values, which may not be of the same order of

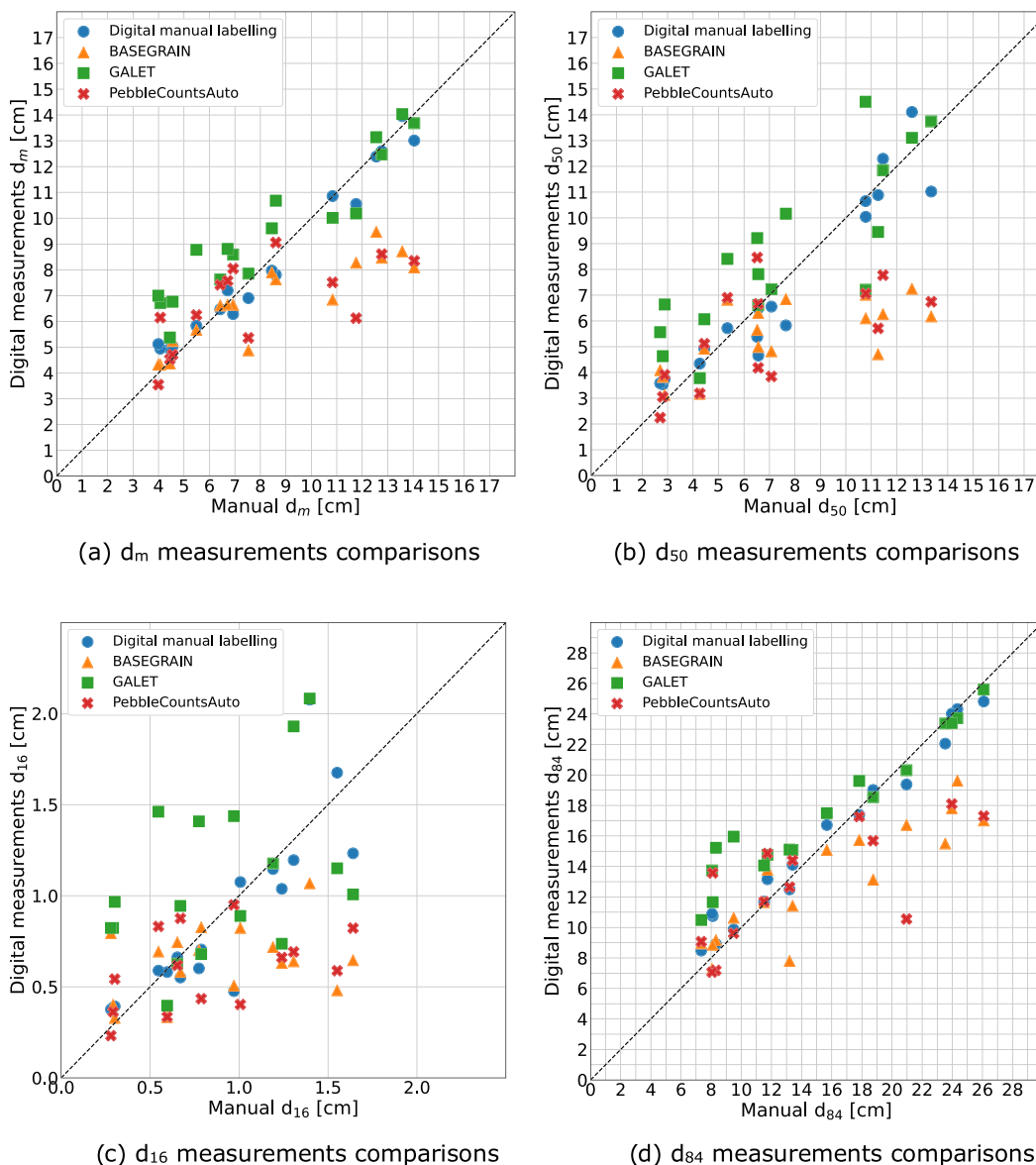
magnitude). For each grain size percentile, the NRMSE was calculated as follows:

$$\text{NRMSE} = \frac{\sqrt{\sum_{i=1}^n (y_i - x_i)^2}}{\sqrt{n} \cdot x_{\text{mean}}}$$

where  $y_i$  is the value of the digitally estimated grain size percentile on sample  $i$ ,  $x_i$  is the value of the manually measured grain size percentile on sample  $i$ ,  $x_{\text{mean}}$  is the mean value of the manually measured grain size percentiles and  $n$  is the number of line samples— $n = 17$  in this study.

### 3 | RESULTS

Pairs of digitally and manually measured characteristic grain sizes are presented in Figure 3 for the set of Fehr line samples collected during



**FIGURE 3** Characteristic grain size values from digital measurement procedures plotted against in-field manually measured characteristic grain sizes, for each of the 17 line sampling analysis performed. The black dashed line corresponds to the 1:1 trend. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

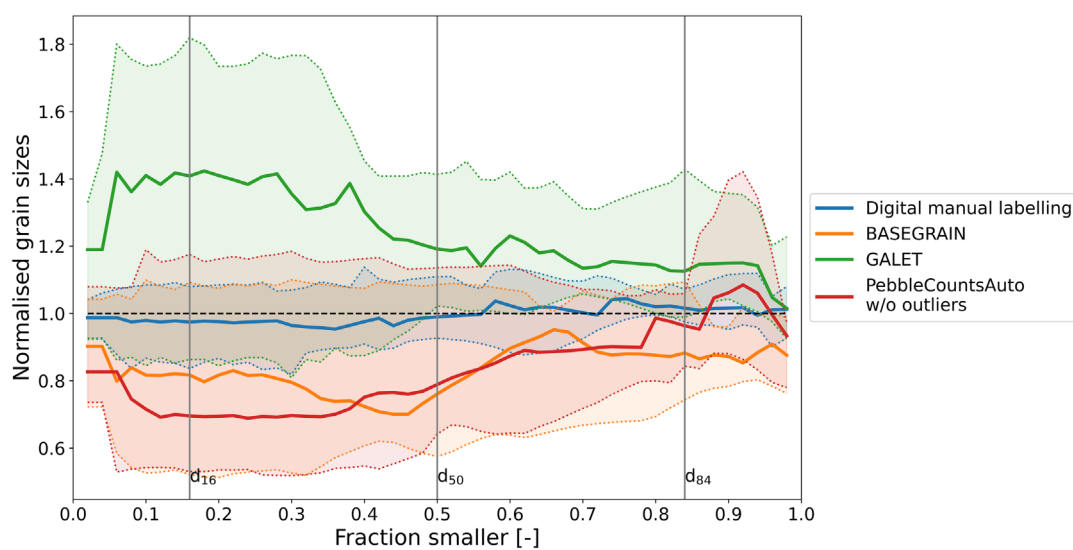
the field campaign. Grain size estimates from manual in-field line samples were regarded as the reference 'ground-truth' values when comparing the performance of the different methods. Overall, the data pairs corresponding to digital manual labelling and their associated in-field manual measurements proved to be mutually consistent. Grain size estimates from digital manual labelling closely follow the 1:1 trend when plotted against their in-field ground-truth counterparts (Figure 3). The grain size estimates issued from GALET follow the 1:1 trendline with a particularly good agreement on the highest half of the grain size domain for the  $d_m$ ,  $d_{50}$  and  $d_{84}$  cases. However, the GALET estimates were often larger than the manual ones on the lowest half of the grain size domain (Figure 3a,b,d).

PebbleCounts grain size estimates included three outliers. For the sake of readability, the outliers were not plotted in Figure 3, as they differed strongly from manually measured grain sizes for the  $d_m$  and

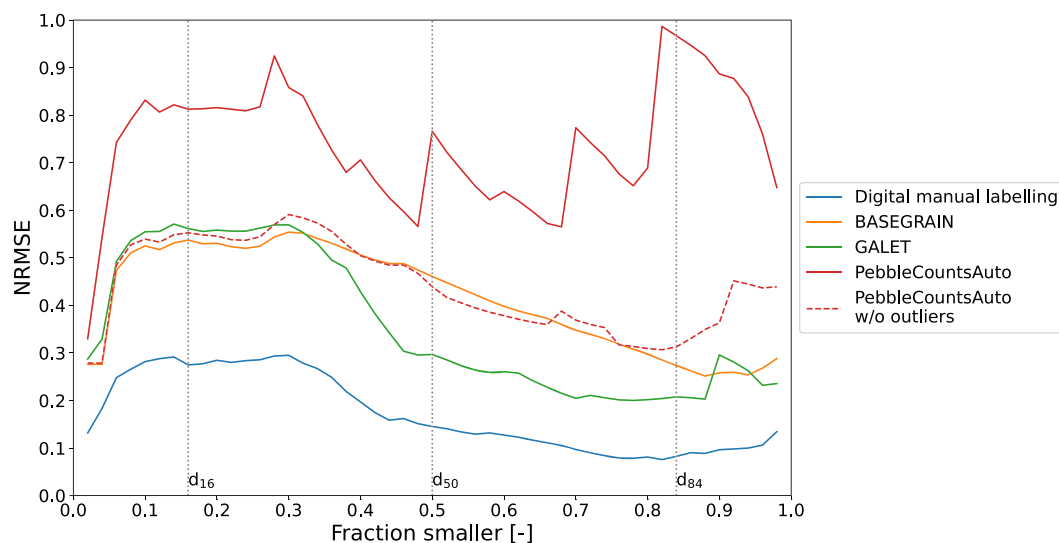
$d_{84}$  grain sizes (their values ranged from 20 to 28 cm for the  $d_m$  grain size, and from 42 to 62 cm for the  $d_{84}$  grain size, see Figure S2). These outliers arose because a large grain feature (b-axis > 40 cm) was detected in each of the three concerned samples. Such large features resulted from the undersegmentation of sediment patches in PebbleCounts.

When ignoring the above outliers, we found that PebbleCounts and BASEGRAIN grain size estimates were often located under the identity line in Figure 3. Particularly, the largest manually measured values were systematically underestimated by BASEGRAIN and PebbleCounts routines for all characteristic grain sizes.

The normalised grain size percentiles are presented in Figure 4. The median normalised grain size percentiles derived from manual labelling are close to unity for all grain size percentiles. This indicates a good match with grain sizes issued from in-field manual sampling.



**FIGURE 4** Median normalised grain size percentile of each digital measuring procedure (thick line). The grain size percentiles were normalised by the in-field manually derived grain size percentiles ( $d_{\text{digital}}/d_{\text{manual}}$ ). The  $q_{25}$  and  $q_{75}$  quartiles of the normalised digital estimates are represented by the dotted lines, meaning that 50% of the normalised estimates are located within the coloured area. Key grain size percentiles (i.e.,  $d_{16}$ ,  $d_{50}$  and  $d_{84}$ ) are indicated by vertical lines. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]



**FIGURE 5** Normalised root mean squared error (NRMSE) of each digital measuring procedures compared with in-field manual sampling, for each grain size percentile. Key grain size percentiles (i.e.,  $d_{16}$ ,  $d_{50}$  and  $d_{84}$ ) are indicated by vertical dotted lines. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]



The normalised grain size percentiles computed by GALET were frequently in excess of the grain size values derived from in-field manual measurements (median value above unity). This overestimation was particularly pronounced for grain size percentiles smaller than  $d_{40}$ . Normalised grain size percentiles obtained from BASEGRAIN and PebbleCounts (whose outliers were not considered) underestimated grain size percentiles on average. This underestimation was less severe for larger grain size percentiles. If we look at the  $q_{25}$  and  $q_{75}$  quartiles in Figure 4, the object-detection software routines provide normalised grain size percentiles that are more scattered in the lower half of the grain size percentiles (i.e., between  $d_5$  and  $d_{50}$ ) than in the upper half, with a minimum scatter reached around the  $d_{80}$  grain size. This trend was also observed for manual labelling, but it was less pronounced—the interquartile range of the normalised grain size percentiles was relatively low compared with the object detection software estimates.

The evolution of the NRMSE as a function of the grain size percentile is presented in Figure 5. This error metric indicates that the digital measurement procedures were most accurate around the  $d_{84}$  grain size percentile, except for PebbleCounts when its outliers were considered. Grain sizes computed from digital manual labelling showed the lowest NRMSE. Concerning the software routines, they exhibited mutually similar error values for grain size percentiles between  $d_5$  and  $d_{40}$ . Between the  $d_{40}$  and  $d_{90}$  grain sizes percentiles, GALET showed the lowest errors, whereas BASEGRAIN and PebbleCounts (without outliers) exhibited comparable higher NRMSE values. The NRMSE of grain size estimates of PebbleCounts was significantly reduced when removing the outliers from the error metric computation.

## 4 | DISCUSSION

The accuracy of digital manual labelling and of object-based grain-sizing methods compared with in-field line sampling is discussed in Sections 4.1 and 4.2, respectively.

### 4.1 | Digital manual labelling accuracy

The characteristic grain size values derived from manual labelling of orthoimages showed a great similarity with the values derived from in-field manual sampling (Figures 3 and 4). This similarity contrasts with previous studies on manual labelling of individual grains in photographs, which found that grain sizes were generally underestimated compared with in-field sampling results (e.g., Adams, 1979; Church, McLean, & Wolcott, 1987; Garefalakis et al., 2023; Ibbeken & Schleyer, 1986). These authors linked this underestimation to the partial information accessible in photography-based grain-sizing methods, because only the exposed part of grains is visible in nadir photographs. Partial burying of grains, grain imbrication or foreshortening of grains due to the angle of the photograph can lead to underestimate the true grain sizes (Graham et al., 2010).

In our study, grain size percentiles derived from digital manual labelling did not suffer from underestimation, most likely because it was difficult to identify smallest particles (b-axis < 2 cm approximately) owing to the orthoimage resolution or to their location in-

between coarser particles. The weak detection of finest particles resulted in different calibrations of the Fuller curves that describe the lower end of the GSDs and probably led to higher NRMSE values for grain size percentiles in the  $d_{10}$ – $d_{40}$  range (Figure 5). In our samples, percentiles smaller than  $d_{10}$  corresponded to grain sizes smaller than 1 cm and no grain-sizing method (including in-field sampling) provided direct measurements for such small grains. Therefore, under  $d_{10}$ , the Fuller interpolation in the Fehr method completely determined the grain size percentiles, made the GSD tails mutually similar and provided lower NRMSE values. The undersampling of fine grains likely counterbalanced the size underestimation for the largest particles, thus providing average grain size estimates that were similar to those derived from in-field manual sampling. It is worth mentioning that manual labelling was performed by a single operator and we did not investigate how the results may differ depending on the operator.

### 4.2 | Accuracy of object-based grain-sizing methods

Comparing the three software routines (BASEGRAIN, GALET and PebbleCountsAuto) revealed differences in accuracy and limitations. GALET tends to overestimate in-field grain sizes issued from line sampling. The main explanation for this phenomenon is that GALET did not detect the smallest grains (b-axis < 2–3 cm). Mörtl et al. (2022) noted that the resolution of orthoimages determines the smallest detectable grain size by GALET. The resolution of the generated orthoimages (approximately 0.3 cm/px) was likely too low for the detection of the smallest grains. Fine grained samples (with  $d_m$  smaller than 7 cm, Figure 3a) were therefore particularly affected by the absence of small grains in GALET grain size estimates and thus led to overestimated values. To reduce the detection limit for small grains, shorter ground sampling distances would be required.

Visual inspection of the grain features detected by GALET suggested that the software performance was not affected by different rock texture patterns inside individual grains. The CNN training dataset used by Mörtl et al. (2022) is probably well suited to applying the routine to the Navisence bed images. A significant number of grains was not detected in the GALET routine. Overall, GALET detected 40% less grain features, regardless of their size, along the lines compared with the manual labelling conducted by a human operator. The largest grains (b-axis > 20 cm) sometimes appeared over-segmented by a vertical or horizontal line. This feature splitting was caused by the edges of the finite-size moving window used for grain feature detection in GALET. This issue did not arise with large grain features along the lines, but it could lead to a size underestimation for some of the largest pebbles present in the river bed.

BASEGRAIN and PebbleCounts produced similar results, as both software routines generally underestimated characteristic grain sizes. For three sampling lines, PebbleCounts outcomes were affected by several feature merging occurrences and missed grain detections (the outliers mentioned in Section 3). These errors were likely caused by glacial flour, which partially covered pebbles. This led to a large overestimation of grain size percentiles for these lines. The grain sizes of the other lines were generally underestimated because the largest stones were often not detected, which may be due to the presence of intergranular textures that prevented optimal edge detection. In

addition, direct sun illumination on orthoimage C caused size underestimation of the detected stones, as the shaded grain faces were not included in the detected object boundaries. Finally, we observed that the grain masks identified by PebbleCounts were generally smaller than the apparent size of stones in the orthoimages. This shortcoming in PebbleCounts led to grain size underestimation. PebbleCounts showed a poor detection rate along the lines on the orthoimages, as the number of grain features detected is 62% smaller than that resulting from manual labelling.

Concerning BASEGRAIN, rock-texture variations inside single grains (e.g., due to foliation or veins) can be detected as grain edges during the segmentation procedure. Therefore, large particles often appeared over-segmented. This resulted in the detection of several smaller particles instead of a single large particle. These over-segmentation errors were less frequent for small particles, whose detection was less influenced by rock texture details—apparently because of the limited image resolution. BASEGRAIN suffered from the same problem identified in PebbleCounts: it often did not merge shaded grain-surfaces into the object boundaries. This resulted in underestimation of the real size of grain features. Rock texture variations induced BASEGRAIN to detect fictitious edge patterns. Non-detection of large particles might arise when BASEGRAIN was unable to reconstruct a closed grain boundary from these edge patterns. The large number of tunable parameters in BASEGRAIN makes the performance highly dependent on the operator's choices. Meticulous parameter tuning and visual checks of the quality of the object detection were performed in the present study, but it cannot be excluded that different operators could have obtained different results.

Systematic grain size underestimation from automated image-based methods induced by rock texture is an issue pinpointed by earlier studies (e.g., Strom, Kuhns, & Lucas, 2010). We found that PebbleCounts results were relatively less influenced by rock texture patterns than BASEGRAIN ones. Over-segmentation errors with BASEGRAIN may be even more severe when no parameter tuning is performed (see Chardon, Piasny, & Schmitt, 2022).

Among the routines considered here, GALET, based on deep learning for object detection, emerged as the best-suited tool for grain size analysis when studying gravel bars from orthoimages. GALET was designed to fit the wide range of rock texture found in gravel-bed rivers and to conduct grain feature detection on long stream reaches. The object-detection performances of deep-learning methods (see Zhao et al., 2019), the recent implementation of deep learning in object-based grain-sizing techniques (e.g., Chen, Hassan, & Fu, 2022; Soloy et al., 2020) and the results of the present study indicate that the deep-learning technology may enable a step forward in automated optical granulometry. We believe that the use of novel automated grain-sizing techniques applied to large scale orthoimages will give a new impetus to understanding the processes that drive sediment sorting in gravel-bed rivers. In the Supporting Information, we illustrate by way of an example how grain size data obtained by object-based segmentation on orthoimages can be used to study the grain size distribution's spatial variation along a mountain river. This example involves statistical tools originally developed in other scientific fields—such as Moran's index (Moran, 1950) and Local Indicator of Spatial Association (Anselin, 1995)—to analyse the spatial variability of riverbed deposits according to their grain sizes.

## 5 | CONCLUSIONS

We presented a benchmarking study of three object-based grain-sizing models (BASEGRAIN, PebbleCountsAuto, and GALET) on a mountain river bed. The main difference between them was that GALET uses deep-learning technology whereas the two others are based on image thresholding for grain segmentation. The three methods were applied to orthoimages obtained from UAV surveys and Structure-from-Motion photogrammetry. In-field estimates of grain sizes obtained using Fehr's line sampling technique served as a reference dataset to evaluate the accuracy of each grain-sizing method. We supplemented the comparison by manually labelling grain features on the same orthoimages. By computing the grain size distributions within Fehr's (Fehr, 1987) framework, we ensured that all methods were comparable on the same footing.

Manual labelling provided estimates that were fully consistent with field measurements. BASEGRAIN and PebbleCountsAuto underestimated grain sizes on average, whereas GALET generally overestimated grain size percentiles.

We identified some limitations in the three models. BASEGRAIN often led to over-segmentation in grain features due to the rock-texture influence on object detection. PebbleCountsAuto outcomes were often affected by missed detections of large grains. Shaded grain faces and glacier flour also influenced grain detection in PebbleCountsAuto. The available image resolution prevented the detection of the smallest grain features with GALET.

Our study showed that GALET was generally the most accurate automated grain-sizing method. Our results suggest that object-based methodologies based on deep learning may become the new cornerstone of optical granulometry for monitoring river-surface grain sizes with high spatial coverage and accuracy.

### AUTHOR CONTRIBUTIONS

**Raphaël Miazza:** Data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); visualisation (lead); writing—original draft preparation (lead); writing—review & editing (equal).  
**Ivan Pascal:** Conceptualisation (lead); formal analysis (supporting); investigation (supporting); methodology (equal); project administration (lead); supervision (lead); writing—review & editing (equal).  
**Christophe Ancey:** Conceptualisation (supporting); funding acquisition (lead); supervision (supporting); writing—review & editing (equal).

### ACKNOWLEDGEMENTS

The authors acknowledge the support of the Swiss National Science Foundation (grant No. 200020 204108). We kindly thank the Styx4D company (Le-Bourget-du-Lac, France) for performing the GALET grain-detection processing. We also thank Bob de Graffenried for his assistance in the field and with the Pix4Dmapper software and for providing the orthoimage produced from 50-m-high flights. Open access funding provided by Ecole Polytechnique Federale de Lausanne.

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available from Zenodo at <https://doi.org/10.5281/zenodo.8020652>.

## ORCID

Raphaël Miazza  <https://orcid.org/0009-0008-2242-7888>

Ivan Pascal  <https://orcid.org/0000-0001-6599-6222>

Christophe Ancey  <https://orcid.org/0000-0003-3354-7398>

## REFERENCES

- Adams, J. (1979) Gravel size analysis from photographs. *Journal of the Hydraulics Division*, 105(10), 1247–1255. Available from: <https://doi.org/10.1061/JYCEAJ.0005283>
- Anselin, L. (1995) Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2), 93–115. Available from: <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Bluck, B.J. (1979) Structure of coarse grained braided stream alluvium. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 70(10–12), 181–221. Available from: <https://doi.org/10.1017/S0080456800012795>
- Bunte, K. & Abt, S.R. (2001) Sampling surface and subsurface particle-size distributions in wadable gravel- and cobble-bed streams for analyses in sediment transport, hydraulics, and streambed monitoring. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station (General Techn. Report RMRS-GTR-74).
- Buscombe, D. (2013) Transferable wavelet method for grain-size distribution from images of sediment surfaces and thin sections, and other natural granular patterns. *Sedimentology*, 60(7), 1709–1732. Available from: <https://doi.org/10.1111/sed.12049>
- Buscombe, D. (2020) SediNet: a configurable deep learning model for mixed qualitative and quantitative optical granulometry. *Earth Surface Processes and Landforms*, 45(3), 638–651. Available from: <https://doi.org/10.1002/esp.4760>
- Butler, J.B., Lane, S.N. & Chandler, J.H. (2001) Automated extraction of grain-size data from gravel surfaces using digital image processing. *Journal of Hydraulic Research*, 39(5), 519–529. Available from: <https://doi.org/10.1080/00221686.2001.9628276>
- Carbonneau, P.E., Lane, S.N. & Bergeron, N.E. (2004) Catchment-scale mapping of surface grain size in gravel bed rivers using airborne digital imagery. *Water Resources Research*, 40(7), W07202. Available from: <https://doi.org/10.1029/2003WR002759>
- Carrivick, J.L. & Smith, M.W. (2019) Fluvial and aquatic applications of structure from motion photogrammetry and unmanned aerial vehicle/drone technology. *WIREs Water*, 6(1), e1328. Available from: <https://doi.org/10.1002/wat2.1328>
- Chardon, V., Piasny, G. & Schmitt, L. (2022) Comparison of software accuracy to estimate the bed grain size distribution from digital images: a test performed along the Rhine River. *River Research and Applications*, 38(2), 358–367. Available from: <https://doi.org/10.1002/rra.3910>
- Chen, X., Hassan, M.A. & Fu, X. (2022) Convolutional neural networks for image-based sediment detection applied to a large terrestrial and airborne dataset. *Earth Surface Dynamics (Pembroke, Ont.)*, 10(2), 349–366. Available from: <https://doi.org/10.5194/esurf-10-349-2022>
- Church, M., McLean, D.G. & Wolcott, J.F. (1987) 'River bed gravels: sampling and analysis.' In *Sediment transport in gravel-bed rivers*, 43–88.
- Daniels, M.D. & McCusker, M.H. (2010) Operator bias characterizing stream substrates using Wolman pebble counts with a standard measurement template. *Geomorphology*, 115(1), 194–198. Available from: <https://doi.org/10.1016/j.geomorph.2009.09.038>
- Detert, M. & Weitbrecht, V. (2012) 'Automatic object detection to analyze the geometry of gravel grains - A free stand-alone tool'. *River Flow 2012 - Proceedings of the International Conference on Fluvial Hydraulics*, 1, 595–600.
- Dietrich, W.E., Nelson, P.A., Yager, E., Venditti, J.G., Lamb, M.P. & Collins, L. (2006) Sediment patches, sediment supply, and channel morphology. In: Parker, G. & Garcia, M.H. (Eds.) *River coastal and estuarine Morphodynamics*. London: Taylor & Francis, pp. 79–90. Available at: <https://resolver.caltech.edu/CaltechAUTHORS:20130307-142351272>
- Fehr, R. (1987) 'Geschiebeanalysen in Gebirgsflüssen: Umrechnung und vergleich von verschiedenen Analyseverfahren'. *Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie an der Eidgenössischen Technischen Hochschule Zürich*, 92.
- Ferguson, R. (2007) Flow resistance equations for gravel- and boulder-bed streams. *Water Resources Research*, 43(5), W05427. Available from: <https://doi.org/10.1029/2006WR005422>
- Garefalakis, P., do Prado, A.H., Mair, D., Douillet, G.A., Nyffenegger, F. & Schlunegger, F. (2023) Comparison of three grain size measuring methods applied to coarse-grained gravel deposits. *Sedimentary Geology*, 446, 106340. Available from: <https://doi.org/10.1016/j.sedgeo.2023.106340>
- Graham, D.J., Rice, S.P. & Reid, I. (2005) A transferable method for the automated grain sizing of river gravels. *Water Resources Research*, 41(7), W07020. Available from: <https://doi.org/10.1029/2004WR003868>
- Graham, D.J., Rollet, A.J., Piégay, H. & Rice, S.P. (2010) Maximizing the accuracy of image-based surface sediment sampling techniques. *Water Resources Research*, 46(2), W02508. Available from: <https://doi.org/10.1029/2008WR006940>
- Guerit, L., Barrier, L., Narteau, C., Métivier, F., Liu, Y., Lajeunesse, E., et al. (2014) The grain-size patchiness of braided gravel-bed streams – example of the Urumqi River (Northeast Tian Shan, China). *Advances in Geosciences*, 37, 27–39. Available from: <https://doi.org/10.5194/adgeo-37-27-2014>
- Hey, R.D. (1979) Flow resistance in gravel-bed rivers. *Journal of the Hydraulics Division*, 105(4), 365–379. Available from: <https://doi.org/10.1061/JYCEAJ.0005178>
- Hodge, R., Richards, K. & Brasington, J. (2007) A physically-based bedload transport model developed for 3-D reach-scale cellular modelling. *Geomorphology*, 90(3), 244–262. Available from: <https://doi.org/10.1016/j.geomorph.2006.10.022>
- Ibbeken, H. & Schleyer, R. (1986) Photo-sieving: a method for grain-size analysis of coarse-grained, unconsolidated bedding surfaces. *Earth Surface Processes and Landforms*, 11(1), 59–77. Available from: <https://doi.org/10.1002/esp.3290110108>
- Kellerhals, R. & Bray, D.I. (1971) Sampling procedures for coarse fluvial sediments. *Journal of the Hydraulics Division*, 97(8), 1165–1180. Available from: <https://doi.org/10.1061/JYCEAJ.0003044>
- Keulegan, G.H. (1938) Laws of turbulent flow in open channels. *Journal of Research of the National Bureau of Standards*, 21(6), 707–741. Available from: <https://doi.org/10.6028/jres.021.039>
- Konz, M., Chiari, M., Rimkus, S., Turowski, J.M., Molnar, P., Rickenmann, D., et al. (2011) Sediment transport modelling in a distributed physically based hydrological catchment model. *Hydrology and Earth System Sciences*, 15(9), 2821–2837. Available from: <https://doi.org/10.5194/hess-15-2821-2011>
- Lang, N., Irniger, A., Rozniak, A., Hunziker, R., Wegner, J.D. & Schindler, K. (2021) GRAINet: mapping grain size distributions in river beds from UAV images with convolutional neural networks. *Hydrology and Earth System Sciences*, 25(5), 2567–2597. Available from: <https://doi.org/10.5194/hess-25-2567-2021>
- Lisle, T.E. & Madej, M.A. (1992) Spatial variation in armouring in a channel with high sediment supply. In: Billi, P., Hey, R. D., Thorne, C. R. & Tacconi, P. (Eds.) *Dynamics of gravel-bed rivers*. Chichester: Wiley, pp. 277–293.
- Mair, D., do Prado, A.H., Garefalakis, P., Lechmann, A., Whittaker, A. & Schlunegger, F. (2022) Grain size of fluvial gravel bars from close-range UAV imagery – uncertainty in segmentation-based data. *Earth Surface Dynamics (Pembroke, Ont.)*, 10(5), 953–973. Available from: <https://doi.org/10.5194/esurf-10-953-2022>
- Mair, D., Witz, G., Do Prado, A.H., Garefalakis, P. and Schlunegger, F. (2023) 'Automated finding, segmenting, and measuring of grains in images of fluvial sediments – the potential of transfer learning in deep neural networks'. Available at: <https://eartharxiv.org/repository/view/5523/> (Accessed: 23 December 2023).
- McEwan, I.K., Sheen, T.M., Cunningham, G.J. & Allen, A.R. (2000) Estimating the size composition of sediment surfaces through image analysis. *Proceedings of the Institution of Civil Engineers - Water and Maritime Engineering*, 142(4), 189–195. Available from: <https://doi.org/10.1680/wame.2000.142.4.189>

- Moran, P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17–23. Available from: <https://doi.org/10.2307/2332142>
- Mörtl, C., Baratier, A., Berthet, J., DuVillard, P.A. & De Cesare, G. (2022) 'GALET: A deep learning image segmentation model for drone-based grain size analysis of gravel bars'. In Proceedings of the 39th IAHR World Congress—From Snow To Sea. <https://doi.org/10.3850/IAHR-39WC252171192022895>
- Parker, G., Klingeman, P.C. & McLean, D.G. (1982) Bedload and size distribution in paved gravel-bed streams. *Journal of the Hydraulics Division*, 108(4), 544–571. Available from: <https://doi.org/10.1061/JYCEAJ.0005854>
- Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carboneau, P., et al. (2020) Remotely sensed rivers in the Anthropocene: state of the art and prospects. *Earth Surface Processes and Landforms*, 45(1), 157–188. Available from: <https://doi.org/10.1002/esp.4787>
- Purinton, B. & Bookhagen, B. (2019) Introducing PebbleCounts: a grain-sizing tool for photo surveys of dynamic gravel-bed rivers. *Earth Surface Dynamics (Pembroke, Ont.)*, 7(3), 859–877. Available from: <https://doi.org/10.5194/esurf-7-859-2019>
- Purinton, B. & Bookhagen, B. (2021) Tracking downstream variability in large grain-size distributions in the south-Central Andes. *Journal of Geophysical Research - Earth Surface*, 126(8), e2021JF006260. Available from: <https://doi.org/10.1029/2021JF006260>
- Ramirez, J.A., Mertin, M., Peleg, N., Horton, P., Skinner, C., Zimmermann, M., et al. (2022) Modelling the long-term geomorphic response to check dam failures in an alpine channel with CAESAR-Lisflood. *International Journal of Sediment Research*, 37(5), 687–700. Available from: <https://doi.org/10.1016/j.ijsrc.2022.04.005>
- Recking, A. (2013) Simple method for calculating reach-averaged bed-load transport. *Journal of Hydraulic Engineering*, 139(1), 70–75. Available from: [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000653](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000653)
- Recking, A. (2016) A generalized threshold model for computing bed load grain size distribution. *Water Resources Research*, 52(12), 9274–9289. Available from: <https://doi.org/10.1002/2016WR018735>
- Rice, S. & Church, M. (1996) Sampling surficial fluvial gravels; the precision of size distribution percentile sediments. *Journal of Sedimentary Research*, 66(3), 654–665. Available from: <https://doi.org/10.2110/jsr.66.654>
- Rickenmann, D. (1997) Sediment transport in Swiss torrents. *Earth Surface Processes and Landforms*, 22(10), 937–951. Available from: [https://doi.org/10.1002/\(SICI\)1096-9837\(199710\)22:10<937::AID-ESP786>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1096-9837(199710)22:10<937::AID-ESP786>3.0.CO;2-R)
- Rickenmann, D. & McArdell, B.W. (2007) Continuous measurement of sediment transport in the Erlenbach stream using piezoelectric bedload impact sensors. *Earth Surface Processes and Landforms*, 32(9), 1362–1378. Available from: <https://doi.org/10.1002/esp.1478>
- Rickenmann, D. & Recking, A. (2011) Evaluation of flow resistance in gravel-bed rivers through a large field data set. *Water Resources Research*, 47(7), W07538. Available from: <https://doi.org/10.1029/2010WR009793>
- Schlunegger, F., Delunel, R. & Garefalakis, P. (2020) Short communication: field data reveal that the transport probability of clasts in Peruvian and Swiss streams mainly depends on the sorting of the grains. *Earth Surface Dynamics (Pembroke, Ont.)*, 8(3), 717–728. Available from: <https://doi.org/10.5194/esurf-8-717-2020>
- Schneider, J.M., Rickenmann, D., Turowski, J.M., Schmid, B. & Kirchner, J.W. (2016) Bed load transport in a very steep mountain stream (Riedbach, Switzerland): measurement and prediction. *Water Resources Research*, 52(12), 9522–9541. Available from: <https://doi.org/10.1002/2016WR019308>
- Smart, G.M. & Jaeggi, M.N.R. (1983) Sediment transport on steep slopes. *Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie Und Glaziologie*, 64, 91–191.
- Soloy, A., Turki, I., Fournier, M., Costa, S., Peuziat, B. & Lecoq, N. (2020) A deep learning-based method for quantifying and mapping the grain size on pebble beaches. *Remote Sensing*, 12(21), 3659. Available from: <https://doi.org/10.3390/rs12213659>
- Strickler, A. (1923) *Beiträge zur Frage der Geschwindigkeitsformel und der Rauheitszahlen für Ströme, Kanäle und geschlossene Leitungen*. Amt für Wasserwirtsch: Eidg.
- Strom, K.B., Kuhns, R.D. & Lucas, H.J. (2010) Comparison of automated image-based grain sizing to standard pebble-count methods. *Journal of Hydraulic Engineering*, 136(8), 461–473. Available from: [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000198](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000198)
- Travaglini, E., Bardou, E., Ancey, C. & Bohorquez, P. (2015) Analysis of Sediment Transport from Recorded Signals of Sediments in a Gravel-Bed River: Role of Sediment Availability. In: Lollino, G., et al. (Eds.) *Engineering geology for society and territory - volume 3*. Cham: Springer International Publishing, pp. 477–481 [https://doi.org/10.1007/978-3-319-09054-2\\_98](https://doi.org/10.1007/978-3-319-09054-2_98)
- Warrick, J.A., Rubin, D.M., Ruggiero, P., Harney, J.N., Draut, A.E. & Buscombe, D. (2009) Cobble cam: grain-size measurements of sand to boulder from digital photographs and autocorrelation analyses. *Earth Surface Processes and Landforms*, 34(13), 1811–1821. Available from: <https://doi.org/10.1002/esp.1877>
- Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J. & Reynolds, J.M. (2012) 'Structure-from-motion' photogrammetry: a low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314. Available from: <https://doi.org/10.1016/j.geomorph.2012.08.021>
- Wolman, G. (1954) A method of sampling coarse river-bed material. *Eos, Transactions American Geophysical Union*, 35(6), 951–956. Available from: <https://doi.org/10.1029/TR035i006p00951>
- Woodget, A.S. & Austrums, R. (2017) Subaerial gravel size measurement using topographic data derived from a UAV-SfM approach. *Earth Surface Processes and Landforms*, 42(9), 1434–1443. Available from: <https://doi.org/10.1002/esp.4139>
- Woodget, A.S., Fyffe, C. & Carboneau, P.E. (2018) From manned to unmanned aircraft: adapting airborne particle size mapping methodologies to the characteristics of sUAS and SfM. *Earth Surface Processes and Landforms*, 43(4), 857–870. Available from: <https://doi.org/10.1002/esp.4285>
- Zhao, Z.-Q., Zheng, P., Xu, S.T. & Wu, X. (2019) Object detection with deep learning: a review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232. Available from: <https://doi.org/10.1109/TNNLS.2018.2876865>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Miazza, R., Pascal, I. & Ancey, C. (2024) Automated grain sizing from uncrewed aerial vehicles imagery of a gravel-bed river: Benchmarking of three object-based methods. *Earth Surface Processes and Landforms*, 49(5), 1503–1514. Available from: <https://doi.org/10.1002/esp.5782>