RESEARCH ARTICLE

Measuring the grain-size distributions of mass movement deposits

Erin L. Harvey[1] 🟢 | Tristram C. Hales[1] 🟢 | Daniel E. J. Hobley[1,2] 🟢 | Jie Liu[3] |
Xuanmei Fan[3]

[1]School of Earth and Environmental Science, Cardiff University, Cardiff, UK

[2]RSK ADAS Ltd, Helsby, Cheshire, UK

[3]State Key Laboratory for Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, China

Correspondence
Erin L. Harvey, School of Earth and Environmental Science, Cardiff University, Cardiff CF10 3AT, UK.
Email: harveyel1@cardiff.ac.uk

Funding information
NERC GW4+ Doctoral Training Partnership studentship, Grant/Award Number: NE/L002434/1

Abstract

Mass movement deposit grain-size distributions (GSDs) record initiation, transport and deposition mechanisms, and contribute to the rate at which sediment is exported from hillslopes to channels. Defining the GSD of a mass movement deposit is a significant challenge because they are often difficult to access, are heterogeneous in planform and with depth, contain grain sizes from clay (<63 μm) to boulders (>1 m), and require considerable time to calculate accurately. There are numerous methods used to measure mass movement GSDs, but no single method alone can measure the entire range of grain sizes. This paper compares five common methods for determining mass movement deposit GSDs to assess how their accuracy may affect their applicability to different research areas. We applied an automated wavelet analysis (pyDGS), Wolman pebble counts, survey tape counts, manual photo counts and sieving to three different mass movement deposits (two debris flows, one rockslide) in Tredegar, Wales and the Longmen Shan, China. We found that pyDGS and survey tape counts produced comparable GSDs to sieving over a single order of magnitude. PyDGS required calibration to achieve accurate results, limiting its use for many applications. In Tredegar, Wolman pebble counts over-estimated grain sizes in the lower 80% of the distribution relative to the other four methods used. We demonstrate that method choice can introduce significant uncertainties, particularly at the edges of the distributions, such that D_{16} values differ by up to a factor of five. These methodological uncertainties limit GSD comparisons across studies, particularly where these are used to infer processes within deposits. To minimize these challenges, the methods chosen should be both carefully reported and consistent with the research question.

KEYWORDS
debris flows, grain size, grain-size distributions, mass movements, rockslides

1 | INTRODUCTION

Mass movement deposit grain-size distributions (GSDs) can constrain the source of the material eroded (Dunning, 2006; Marc et al., 2021), the transport and emplacement mechanisms of the deposit (de Haas et al., 2015; Makris et al., 2020) and moderate sediment transport rates in fluvial systems (Neely & DiBiase, 2020; Sklar et al., 2017, 2020). Mass movement deposit GSDs are typically heterogeneous,

can extend up to eight orders of magnitude (from <1 μm to >10 m), and vary spatially and with depth. These GSD characteristics reflect source properties, such as lithology or fracture spacing (Attal & Lavé, 2006; Marc et al., 2021) and processes occurring during transit, including winnowing (Crosta et al., 2007; Dufresne & Dunning, 2017; Locat et al., 2006). There remains no single method that can record GSDs over the range and scale of most mass movement deposits (Table 1). Hence, different approaches or combinations of approaches

TABLE 1 A summary of the key advantages, limitations and resolutions of the methods used to measure mass movement deposit GSDs and key references for the use of these methods

Method	Advantages	Limitations	Sampling range and size	Key references
Volumetric sieving Frequency/volume	• Can collect subsurface GSDs by digging pits or using vertical exposures. • Can constrain the proportion of grains <1 mm.	• Limit to the maximum grain size obtained using sieving directly (typically 80 mm). • Time consuming, which limits its application to detailed, small sections of the deposit. • Difficult to apply to questions of spatial variability in deposits. • May require some larger grains to be measured by hand to obtain a full GSD for each pit.	<0.063–80 mm. We use a maximum limit of 80 mm and record grains >80 mm by hand in the field. We found a 1 m × 1 m × 0.5 m pit took ~6–8 h to dig and sieve. As such, only a small proportion of the deposit can be sampled. Here, we sieved 1000 kg per pit. Church et al. (1987) recommend that the largest particles should represent no more than 5% of the total sample mass. However, this approach is often unachievable in mass movement deposits, where extremely large boulders are present. For example, if grains >50 kg are present, >950 kg of sediment must be sieved. As a result, mass movement GSDs are often generated from smaller than ideal sample sizes, without the rigorous reporting of sampling that is common in fluvial geomorphology.	Attal and Lavé (2006), Bunte and Abt (2001a, b), Casagli et al. (2003), Chen et al. (2001), Dunning (2006), Genevois et al. (2001), Hubert and Filipov (1989), Ibbeken et al. (1998), Major and Voight (1986), Sosio et al. (2007), Whipple and Dunne (1992), Zhang et al. (2011, 2015)
Pebble count and survey tape Frequency/number	• Can record all three axes of a grain, which is useful when working with non-spherical grains. • Sampling typically involves >100 grains. This is quick relative to other methods (~1 h)	• Only used to collect surface GSDs. • Field intensive. • Bias towards sampling only visible grains.	The smallest detectable grain size is typically gravel as this is easily visible. Studies typically give a minimum grain size of 4–5 mm (Casagli et al., 2003; Sklar et al., 2020). However, when survey tapes are used, the minimum detectable grain size is thought to be lower (~2 mm) (Bunte & Abt, 2001a). The number of grains measured can be as low as 100, however this value increases with more heterogeneous deposits. We used a sample size of 300 for a small, heterogeneous landslide deposit.	Attal and Lavé (2006), Casagli et al. (2003), Hubert and Filipov (1989), Kim and Lowe (2004), Major and Voight (1986), Vallance and Scott (1997), Zhang et al. (2011)
Manual photo analysis Frequency/number	• Requires considerably less time in the field in comparison to other methods.	• Only used to collect surface GSDs. • Bias towards sampling coarser grains.	The minimum grain size depends on the resolution of the image and the maximum grain	Attal and Lavé (2006), Casagli et al. (2003), Crosta et al. (2007), Genevois et al. (2001),

(Continues)

TABLE 1 (Continued)

Method	Advantages	Limitations	Sampling range and size	Key references
	• Does not disturb the surface of the deposit (this allows the method to be compared directly to sieving for the same area). • UAV imagery can be used in less accessible locations. • The results can be reproduced.	• Can only measure visible axes and some grains may overlap and therefore the *b*-axis will be measured incorrectly.	size depends on the extent of the photo. This technique can be used across large surface areas, for example by using UAVs. In this study, photos were taken with a 50 cm × 50 cm frame for reference with resolution >0.12 mm pi^{-1}.	Ibbeken et al. (1998), Kellerhals and Bray (1971), Zhang et al. (2015)
pyDGS *A texture-based approach which uses the spatial and spectral properties of an image to generate a GSD* Frequency/number	• Does not require calibration, though our results suggest some calibration is necessary when choosing the *shape parameter*. • Requires considerably less time in the field in comparison to other methods. • UAV imagery can be used in less accessible locations. • Does not disturb the surface of the deposit.	• Only used to collect surface GSDs. • Similar to manual photo counts, can only detect visible grains. • Coarsest grain size is determined by photo window size. • Errors may be generated if grains are wet, imbricated or similar in colour.	Minimum grain size is determined by photo resolution and number of pixels required to clearly identify a grain. Based on v4.0, pyDGS detects grains ∼6 pixels in length. Maximum grain size is dependent on *maxscale* parameter. Whilst UAVs can be used to survey large surface areas, to measure grains 10 mm in size, photo resolution would need to be 1.66 mm pi^{-1}. Here, we use photos with at least 0.12 mm pi^{-1} resolution.	Buscombe (2013) **Beach GSDs** (Prodger et al., 2017) **Dryland basin GSDs** (Michaelides et al., 2018)

have been used (e.g. Attal & Lavé, 2006; Attal et al., 2015; Casagli et al., 2003; Crosta et al., 2007; Dunning, 2006; Zhang et al., 2015). These approaches typically involve measuring the *b*-axis of grains using two methods, one of which can capture the finest grains, such as sieving, and another for the coarsest grains, such as Wolman pebble counts or photo-based techniques, and combining these to produce a new distribution (Attal & Lavé, 2006; Casagli et al., 2003; Fripp & Diplas, 1993).

Mass movement deposits often segregate by grain size during transit, which results in the development of facies within deposits (Attal & Lavé, 2006; Dufresne & Dunning, 2017; Dunning, 2006; Marc et al., 2021; Vallance & Savage, 2000). For example, flowing mass movements create coarse fronts and levees within their deposits due to kinetic sieving and shear driven size segregation (Attal & Lavé, 2006; Johnson et al., 2012; Marc et al., 2021). Similarly, rock avalanche deposits often have a finer core and coarse carapace associated with high shear at the landslide base (Dufresne & Dunning, 2017). As a result, bimodal and multimodal distributions are often observed in landslide dams and rock avalanches where comminution and shear are prevalent (Casagli et al., 2003; Crosta et al., 2007). In these cases, process-based studies of the emplacement of these deposits may require high spatial sampling resolutions across all three dimensions (Casagli et al., 2003; Makris et al., 2020). It is therefore important to ensure methods can readily characterize spatial and vertical changes across a range of grain sizes from clay to boulders.

Wide and multimodal GSDs also limit the applicability of single grain size metrics like D_{50} to characterize a mass movement deposit (Casagli et al., 2003). A full GSD is useful for inferring processes that involve multiple different grain sizes, such as comminution and kinetic sieving (Dufresne & Dunning, 2017; Makris et al., 2020), and also provides insight into the textural properties of deposits (Casagli et al., 2003). It is therefore more useful to use several quartiles, such as D_5, D_{16}, D_{50}, D_{84} and D_{95}, as opposed to a single metric to characterize the entire GSD for mass movement deposits (Folk & Ward, 1957; Purinton & Bookhagen, 2021). The higher percentiles, such as D_{95} and D_{99}, are prone to larger uncertainty, which arises because of the difficulties associated with sampling the coarsest grains and the often heavy-tailed nature of the distributions. This uncertainty can be mitigated by increasing the sample size, to include as much of the coarser grains as possible (Eaton et al., 2019; Guerit et al., 2018; Purinton & Bookhagen, 2021). However, increasing sample size subsequently results in increased sampling time per site.

Automated and semi-automated techniques that obtain GSDs from static photos may mitigate the large sample sizes required for wide, multimodal GSDs (Table 2). Photo-based methods are also less invasive, typically require less field work and can measure surface GSDs across larger areas over a shorter time period (Table 2; Purinton & Bookhagen, 2021). These methods include both image segmentation and texture-based approaches (Table 2). Image

TABLE 2 The two main types of automated and semi-automated methods for measuring GSDs from photos; key references refer to the use of these methods in mass movement deposits as well as other deposits

Method	Advantages	Limitations	Sampling range and size	Key examples and references
Image segmentation *These techniques aim to isolate and measure the visible axes of the individual grains in an image*	• Requires considerably less time in the field in comparison to other methods. • Does not disturb the surface of the deposit. • Methods may require extensive calibration and time to understand, however once running they can readily be applied to large areas. • UAV imagery can be used in less accessible locations. • Do not require user intervention, which reduces operator bias. • Obtain measurements for each individual grain, from which interpolation is not required to obtain percentiles.	• Only used to collect surface GSDs. • Most methods need calibrated detection algorithms to isolate individual grains or large calibration datasets. Therefore, not universally applicable. • Coarsest grain size is determined by photo window size, camera height and resolution. • Can only measure clearly visible grains. • Limited by image complexity, such as vegetation, variations in colour and texture (e.g. veins) and imbrication.	The minimum grain size depends on the resolution of the image and the maximum grain size depends on the extent of the photo. The PebbleCounts algorithm can detect grains ≥20 pixels (Purinton & Bookhagen, 2019). Purinton and Bookhagen (2021) use automated PebbleCounts with minimum grain size of 2.5 cm. Automated image segmentation methods can use cameras on tripods or UAVs to survey large areas. For example, Purinton and Bookhagen (2021) surveyed areas between 944 and 3470 m² for sand- and gravel-bed rivers in the South-Central Andes.	Graham et al. (2005, 2010), Storz-Peretz and Laronne (2013) Two examples of semi- and automated image segmentation techniques. Both were developed using fluvial GSDs. **Basegrain** (Detert & Weitbrecht, 2012) **PebbleCounts** (Purinton & Bookhagen, 2019, 2021)
Texture-based approaches *These techniques generate grain-size distributions and grain-size percentiles using statistics based on the texture of a 2D or 3D image*	• UAV imagery can be used in less accessible locations. • Does not necessarily require fieldwork, though often needs to be calibrated. • UAVs have been used successfully in homogeneous fluvial environments to obtain GSDs. • Does not disturb the surface of the deposit. • Once the photos are taken and the model is running, larger areas can be sampled quicker.	• Only used to collect surface GSDs. • Requires very high-resolution DEMs or images to measure the finest grains. • UAVs to generate 3D point clouds have not worked well in poorly sorted environments (Westoby et al., 2015). • Typically requires high computer processing power. • Bias towards coarse grains. • Coarsest grain size is determined by photo window size. • Generally, requires site-specific calibration to establish relationships between texture and grain size in each location. • May require extensive measurements of GSDs in the field to compare or train the algorithm.	The minimum grain size depends on the resolution of the image and the maximum grain size depends on the extent of the photo. Can be used to survey larger areas than sieving and Wolman pebble counts, based on the reduced field time required. Though high-resolution imagery is required to detect the finest grains.	**Semivariance** (Carbonneau et al., 2004, 2005) **Autocorrelation** (Rubin, 2004) **Wavelet transforms** (Buscombe, 2013) **Convolutional neural networks** (Buscombe, 2020; Lang et al., 2021) **3D-based roughness** (Brasington et al., 2012; Neverman et al., 2019; Vázquez-Tarrío et al., 2017; Westoby et al., 2015) **Structure from motion and TLS to produce DEMs of mass movement deposits** (Bitelli et al., 2004; Cucchiaro et al., 2018; Dunning et al., 2009; Gupta & Shukla, 2018; Saunders, 2014)

segmentation techniques isolate and measure the visible axes of individual grains (e.g Graham et al., 2005; Purinton & Bookhagen, 2019), whereas texture-based techniques are statistical approaches which produce GSDs using information about how intensity and colour vary within 2D and 3D images (Buscombe, 2013; Lang et al., 2021), for example a high-resolution digital elevation model (DEM).

Traditional methods used to measure GSDs are often limited by sampling size, inaccessibility and time constraints, as described in Table 1. The disadvantages associated with using each method are likely to introduce uncertainty into the measured GSDs, for example by excluding fine grain sizes or using small sample sizes (Casagli et al., 2003). Uncertainty in measured GSDs may affect our ability to compare across different studies. Whilst the uncertainty associated with comparing different methods has been widely discussed for fluvial GSDs (e.g. Bunte & Abt, 2001a,b; Wohl et al., 1996), the effect of method choice on comparisons of mass movement deposit GSDs has been less well explored. The uncertainties associated with different methods may be more pronounced in mass movement deposits, which have wider GSDs, greater angularity and grains in excess of 1 m. GSDs may differ in terms of methodological uncertainty, sample size and sample type, which can affect our ability to accurately develop process-based conclusions regarding transport and depositional mechanisms in mass movement deposits. Methodological uncertainties refer to how much the GSD varies depending on the method chosen, sample size refers to the number of grains measured and sample type refers to the region of the deposit considered by each method (i.e. surface or subsurface). Here we compare and combine GSDs generated for three different deposits using five different methods. We compare the methods using D_{16}, D_{50} and D_{84} percentiles as well as statistically using chi-square tests.

2 | STUDY SITES

We chose mass movements from three field sites, a rockslide deposit south of Tredegar, Wales and two debris flows in the Longmen Shan, China (Figure 1).

In Tredegar the rockslide was triggered within the Carboniferous Deri Formation, a sedimentary unit of interbedded sandstone, mudstone and siltstone (Barclay et al., 1989; George, 2015). The rockslide was triggered in a former quarry face during a winter storm on a 26° slope and measured as 26.5 m long and 15 m at the widest point (Figure 1A).

We sampled two large post-seismic debris flow deposits triggered during a monsoonal storm in 2019 in the Luoquan gully and Liusha gully, Longmen Shan, China (Figures 1B and C). The abundance of sediment in the channel prior to these events was a result of previous debris flows and landslides generated during and after the 2008 Wenchuan earthquake. The Luoquan gully consisted mainly of Mesoproterozoic granitoids of the Penguan massif (Figures 1C and D). The Luoquan debris flow had a slope of 9° and average width of 42 m when averaged across the length of the debris flow (∼8 km). The Liusha gully includes both Mesoproterozoic granitoid material and Palaeozoic greywacke and shale (Figures 1B and D; Ma, 2002). The Liusha debris flow had a slope of 23° and average width of 8 m across the 1.5 km length of the debris flow.

3 | METHODS

We applied five methods (Figure 2) to the three different field sites: sieving, survey tape counts, Wolman pebble counts, pyDGS and manual photo counts.

3.1 | Volumetric sieving

We sieved each deposit using a protocol previously used for fluvial sediments, landslide deposits and debris flow deposits (Attal & Lavé, 2006; Attal et al., 2015; Bunte & Abt, 2001a; Zhang et al., 2014). We measured a 1 m × 1 m pit in the centre of the deposit and excavated material at 10 cm increments to a depth of 30 cm in Tredegar and 50 cm in Longmen Shan. The shallower depth in Tredegar was due to the steeper slope and smaller apparent grain size and failure. We used square sieves to separate the remaining sediment into the following size fractions: >4 cm, 2–4 cm, 1–2 cm and <1 cm. We weighed all fractions in the field using fishing scales and separated 1 kg of sediment from the fraction of sediment <1 cm to analyse in the laboratory (Attal & Lavé, 2006; Hubert & Filipov, 1989). We sieved approximately 1000 kg of sediment per pit to fulfil the 5% of total weight limit for the largest grain set out by Church et al. (1987). The coarser sediment was not air dried in the field as the difference in weight for large gravels is negligible (Bunte & Abt, 2001b). We weighed and measured all three axes for grains >8 cm in diameter, which accounted for up to 35% of grains by weight. By measuring multiple axes for these grains, we were able to quantify grain shape as well as size. Where large grains covered multiple layers (e.g. >10 cm on at least one axis), we consistently sampled the grain from the lowest layer to avoid disturbing layers unnecessarily. We adjusted our sieving GSDs to account for this by averaging the weight of grains with a b-axis >10 cm across the appropriate number of layers. In the lab, we wet sieved the 8, 4, 2, 1, 0.5, 0.25, 0.125 and 0.063 mm fractions. For samples containing a large proportion of gravels (>2 mm), we used a sieve shaker to separate the first four fractions. Manual endpoint tests were carried out to ensure all grains had passed through each sieve (Dufresne & Dunning, 2017). The tests involved briefly shaking the sieve into a clean, dry sieve pan to see if any grains still passed through. We noticed that there were still grains passing through the five smallest sieves, so we also wet sieve the fraction <1 mm.

A square sieve correction (Attal & Lavé, 2006) of the form

$$b = \frac{2km}{\sqrt{1+k^2}} = \frac{k\sqrt{2}}{\sqrt{1+k^2}} S \qquad (1)$$

where S is the sieve mesh size and k is the ellipse eccentricity, or the ratio between the b-axis and the c-axis of grains, was applied to our data. There were a large range of values obtained for k within both pits, with the sieves potentially over-estimating the b-axis of each grain by a maximum of 41% in Liusha and 35% in Luoquan. The mean b-axis over-estimate in Liusha and Luoquan was 21% and 17%, respectively. This equated to an approximately 0.7 cm difference between the adjusted size of a 4 cm sieve.

3.2 | Wolman pebble counts and survey tape counts

We conducted a Wolman (1954) pebble count and survey tape pebble count across the surface of the deposit in Tredegar. Typically, at least 100 grains are required for a Wolman pebble count (Wolman, 1954). Due to the heterogeneity of landslide deposits, we
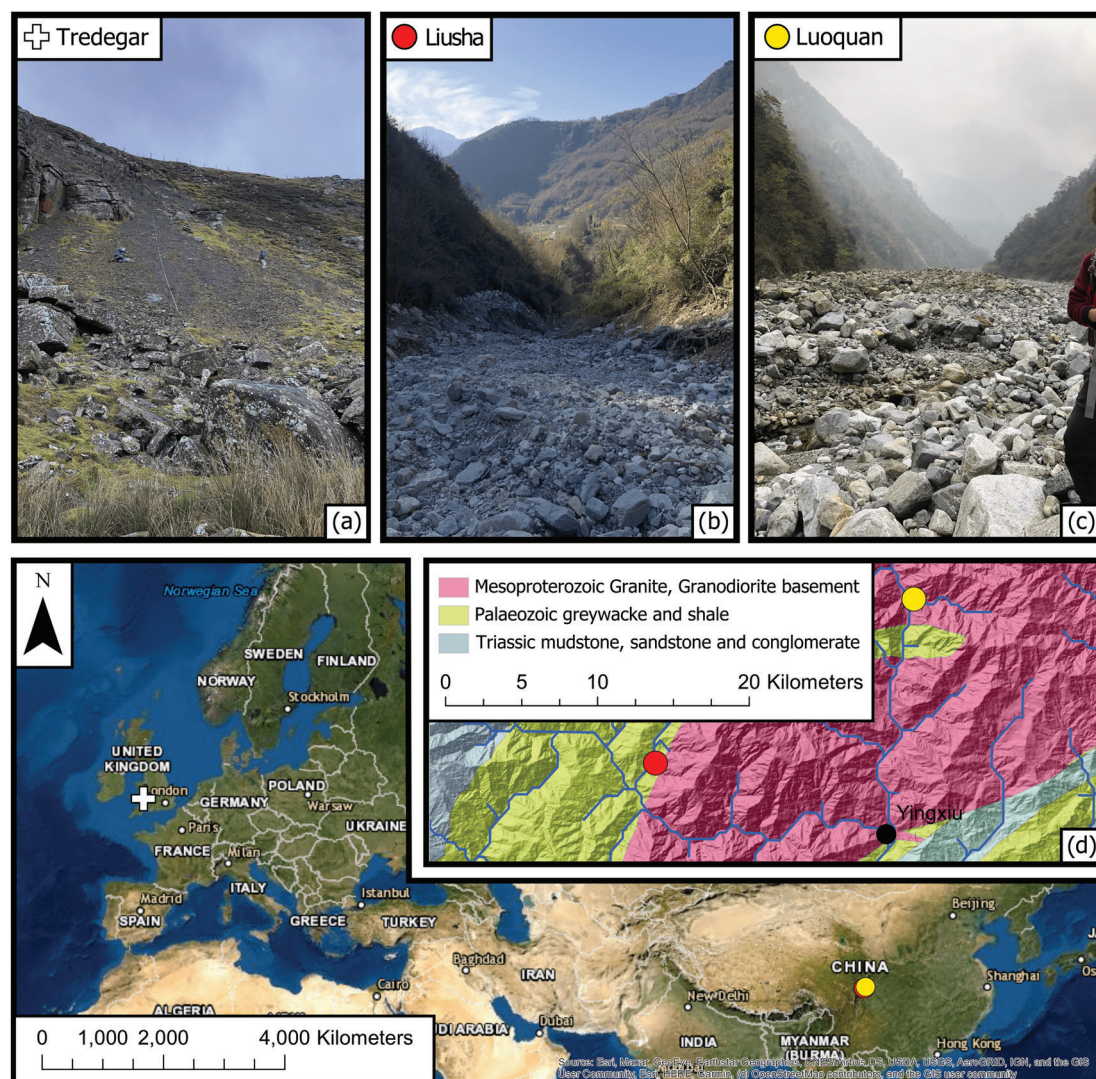
FIGURE 1 Map showing the three locations studied. Inset (a) shows the Tredegar landslide in South Wales. Insets (b) and (c) show the Liusha and Luoquan debris flows in the Longmen Shan, respectively. Inset (d) provides a closer location map for the debris flows in the Longmen Shan with the geology for the region also shown

decided that rather than choose a particular number of samples, we would measure grains until the mean value converged (i.e. any additional grain measured did not change the mean beyond 0.1 mm). We found the mean, D_{50} and D_{84} converged when measuring 300 grains, while D_{90} did not.

The survey tape method involved placing three 50 m tapes horizontally across the deposit and one tape from the scar of the failure to the toe. We measured the b-axis of the grain directly below the tape every 0.25 m interval. This spacing was decided based on the size of grains in the deposit, to ensure no grain covered two points on the tape (Kellerhals & Bray, 1971). If grains were too small to be measured, the nearest grain was chosen instead (de Scally & Owens, 2005; Hubert & Filipov, 1989). If grains were too large the same protocol would apply, however we did not encounter this in Tredegar. For this method, we sampled 174 grains in total and obtained a mean of 17.1 mm, which was 0.7 mm larger than the mean obtained using random Wolman pebble counts. We included grains as small as 1 mm (survey tape) and 3 mm (Wolman count) as they were visible in the field.

3.3 | Manual photo counts

Manual photo counts involved measuring the apparent b-axis of grains using photos taken parallel to the surface (Attal & Lavé, 2006; Casagli et al., 2003; Crosta et al., 2007; Genevois et al., 2001; Ibbeken et al., 1998; Kellerhals & Bray, 1971; Zhang et al., 2011, 2015). We conducted manual photo counts in all three locations by taking photos using a smartphone camera (Figure S4, image resolutions ranged from 0.12 mm pi^{-1} in Tredegar to 0.39 and 0.46 mm pi^{-1} in Liusha and Luoquan). We used a tape measure in Tredegar and a 50 cm × 50 cm frame in the Longmen Shan to determine the resolution of the image. The tape measure and frame also helped to identify when photos were not taken parallel to the slope. These images were discarded alongside photos with inconsistent resolutions and photos of the same surface to ensure no grains were counted multiple times. We conducted manual photo counts on six images in Tredegar, measuring a total of 300 grains. In Longmen Shan, we took photos of the surface of the pit and used these photographs to conduct a grid-by-number analysis (Figures 5 and 6). The width and height of the grid were
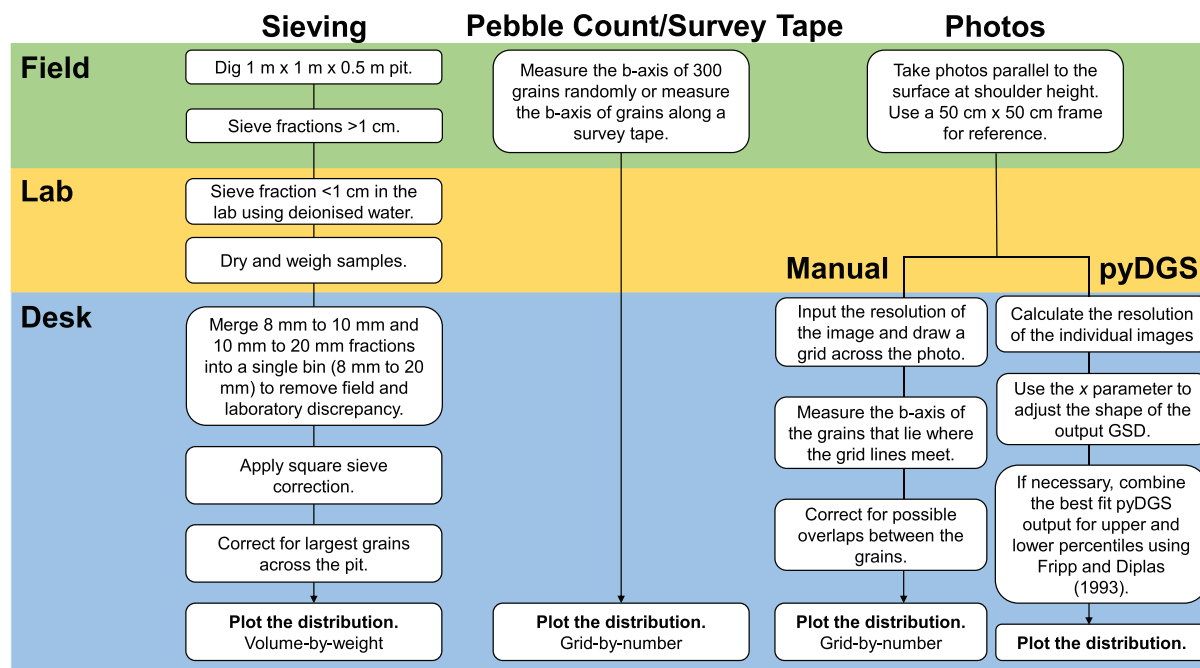
	Sieving	Pebble Count/Survey Tape	Photos
Field	Dig 1 m x 1 m x 0.5 m pit. → Sieve fractions >1 cm.	Measure the b-axis of 300 grains randomly or measure the b-axis of grains along a survey tape.	Take photos parallel to the surface at shoulder height. Use a 50 cm x 50 cm frame for reference.

FIGURE 2 A flow diagram detailing the key steps taken for the methods used in this study. Sieving required field, laboratory and desk work. The second method, Wolman pebble counts, was split into two approaches, survey tape measurements and more randomized Wolman pebble counts. The final methods required photos taken in the field. These photos were then analysed using two different methods, automatic grain-size analysis using pyDGS and manual photo counts

determined by the largest grain in the photo to ensure no grain was counted twice (Bunte & Abt, 2001b).

3.4 | Automated photo analysis (pyDGS)

We applied a texture-based approach, pyDGS (v4.0), as it allows for the rapid identification of GSDs from photos and is beneficial for obtaining a GSD for a large surface area. PyDGS has been successfully applied to dryland basins (Michaelides et al., 2018), beaches (Prodger et al., 2017) and bioclastic sediments (Cuttler et al., 2017), as well as a range of sorted and poorly sorted sediments (Buscombe, 2013). The algorithm requires minimal calibration and can detect grains ~6 pixels in length (fine gravel) from photos taken using a smartphone camera. The algorithm works best for coarse, well-sorted grains, where the brightness of the grains is not positively correlated with size and there are >100 grains in each image (Buscombe, 2013).

There are three key parameters in pyDGS (July 2020 version); x, $maxscale$ and $resolution$. x varies from 1 to -1 and is an exponent that converts the area-based pyDGS output to a volume-based GSD (Buscombe, 2013; Cuttler et al., 2017). The x $exponent$ (hereafter referred to as the shape parameter) relates to the size of the grains, their porosity and sorting (Bunte & Abt, 2001b; Cuttler et al., 2017; Diplas & Fripp, 1992; Diplas & Sutherland, 1988). For example, a negative value of x (-1) can represent poorly sorted coarse gravels with low porosity and a high sand content, whereas a value of 0 is indicative of well-sorted gravel (Bunte & Abt, 2001b). We tune the shape parameter in this paper based on our sieving data. In Tredegar, a single shape parameter consistently represented the GSD obtained using sieving (Figure 3). However, in the Longmen Shan, a single pyDGS shape parameter did not fit the GSD obtained using sieving or manual
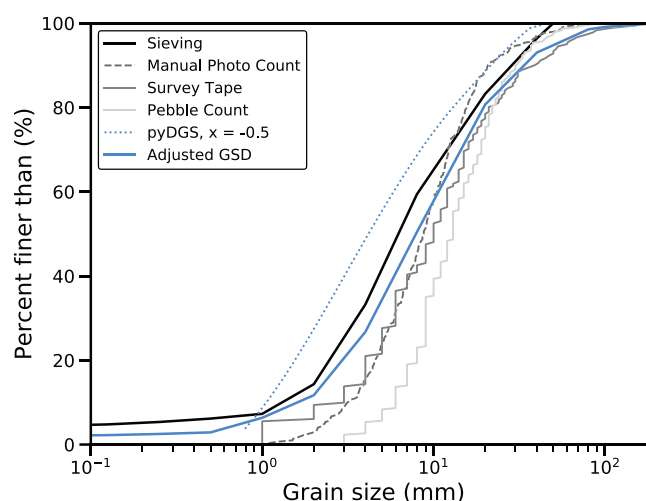
photo counts (Figures S1 and S2). Therefore, for Liusha and Luoquan we combined two pyDGS runs with different shape parameters to obtain a GSD that captured both the finest and coarsest grains

FIGURE 3 The surface GSDs of the Tredegar rockfall based on five sampling methods. The sieving GSD is based on a sample taken within the first 10 cm of the surface near the centre of the deposit. The survey tape count is based on a total of 181 grains across the entire deposit. Both the Wolman pebble count and manual photo count consisted of measuring 300 grains. The manual photo count was based on six photos taken in different parts of the deposit. The pyDGS curve is the average of the 60 GSDs generated using individual photos of the deposit. The adjusted GSD (blue) is calculated by combining the surface sieving GSD (black) and the survey tape GSD (grey) using the method outlined by Fripp and Diplas (1993) and briefly in the main text. We used sieving and survey tape GSDs as these provided the minimum and maximum grain sizes, respectively

measured using sieving (Figures S1 and S2). The *maxscale* parameter defines the maximum grain size that the algorithm searches for in the image as a fraction of the greatest dimension (Buscombe, 2013).

We ran a sensitivity analysis to test how the choices of these three parameters affect the output. For this paper, in Tredegar we use the average GSD obtained by running 60 photos in pyDGS and using an average resolution of 0.12 mm pi^{-1}. We vary the *shape parameter* and *maxscale* throughout.

3.5 | Combining GSDs

We combined distributions to obtain the full GSD for each deposit, following the method of Fripp and Diplas (1993). Each GSD was split into 13 size fractions (0–0.063, 0.063–0.125, 0.125–0.25, 0.25–0.5, 0.5–1, 1–2, 2–4, 4–8, 8–20, 20–40, 40–80, 80–100 and 100–200 mm). The two GSDs were compared and the grain-size fraction with the most similar proportion was chosen to be the match point. The remaining proportions are then rescaled based on the magnitude of the match point. In Tredegar, sieve and survey tape-generated GSDs were combined as these methods covered the largest range of GSD values. In Liusha and Luoquan, pyDGS GSDs with different shape parameters were required to create full GSDs compared to sieving. A *shape parameter* of −1 in Liusha best represents the coarsest percentiles and a *shape parameter* of 0 best fits the finest percentiles. In Luoquan we combine runs with a *shape parameter* of 1 and −1 (Table S5). We tested the sensitivity of the choice of match point, by comparing four possible combined GSDs from our Tredegar data, and found that there was less than a 10% difference in D_{50} values across the combined GSDs.

3.6 | Comparing the different methods

We compared the grain size for the 5th, 10th, 16th, 25th, 50th, 75th, 84th, 90th and 95th percentiles (for all individual methods and combined GSDs) using the normalized root mean square error (NRMSE) (Buscombe, 2013). NRMSE provides a measure of how different two values are; that is, more robust at higher percentiles than standard RMSE. Sieving captures the widest range of grain sizes, so we consider it as the measured value. We calculated NRMSE as outlined in Buscombe (2013):

$$\mathrm{NRMSE}\,(\%) = \frac{\sqrt{\frac{\left(q_{meas}^i - q_{est}^i\right)^2}{n}}}{q_{meas}^i} \times 100 \qquad (2)$$

where n is the number of observations, q_{meas}^i is the percentile grain size from sieving and q_{est}^i is the percentile grain size for the method that we are comparing.

For continuous datasets (manual photo counts, Wolman pebble counts and survey tape counts) we also calculated the percentile uncertainty using the QuantBD function (Eaton et al., 2019; Purinton & Bookhagen, 2021). The output from QuantBD is a minimum and maximum grain-size range for each percentile based on a 95% confidence interval, which we refer to as percentile uncertainty. Finally, two-sample goodness-of-fit chi-squared (χ^2) tests allowed pairwise comparison of different distributions.

4 | RESULTS

For all deposits, the grain-size range varies with measurement method (Figures 3, 5 and 6). In Tredegar, sieving measured the widest grain-size range (from <0.063 mm to 40 mm). Survey tape and pyDGS-derived GSDs spanned two orders of magnitude, from 1 to 170 mm and 0.5 to 45 mm, respectively. Wolman pebble count and manual photo count GSDs recorded an order of magnitude, 3 and 90 mm and 1 and 77 mm, respectively.

Common percentiles used to describe GSDs, D_{16}, D_{50} and D_{84}, all varied by at least an order of magnitude across the different methods (Tables 3 and 4). In Tredegar, the D_{16} values obtained varied by the most compared to D_{50} and D_{84}, as demonstrated by higher NRMSEs for lower percentiles (>50% error for percentiles smaller than D_{50}) (Table 3). The D_{16} value for Wolman pebble counts (7 mm) was five times larger than the D_{16} value obtained using pyDGS (1.4 mm) and sieving (2.2 mm), and exceeded all other D_{16} values and upper limits based on percentile uncertainty (Table 3). D_{16} values for the debris flow deposits also varied by over a factor of two across the different methods (5–13 mm in Liusha and 5.9–56 mm in Luoquan) (Table 4). D_{50} values differed by over a factor of two in Tredegar (4.5–13 mm) and over a factor of three in the Longmen Shan (Liusha: 19–83 mm, Luoquan: 23–150 mm) (Tables 3 and 4). In Tredegar, pyDGS and sieving also obtained the lowest D_{50} and D_{16} values. Survey tape counts and manual photo counts produced similar measurements for both D_{16} (4 and 4.1 mm) and D_{50} (10 and 8.8 mm). Wolman pebble counts obtained the largest D_{50} value. In the Longmen Shan, when only considering the combined pyDGS GSDs, D_{50} values were largest for sieving GSDs (77 and 100 mm). The variation in D_{50} values coincides with the minimum resolutions for each of the respective methods. Larger percentiles, such as D_{84}, and the maximum grain size obtained also differed across methods (Figures 3, 5 and 6; Tables 3 and S1). Photo-based grain-size techniques (both manual and pyDGS) D_{84} values were consistently smaller than the other methods in all locations. In Liusha and Luoquan, manual photo count and pyDGS GSDs under-estimated the upper 20% of the distribution relative to sieving (Figures 5 and 6, Table 4). In some instances, visually calibrated pyDGS runs could be used to produce coarser distributions, however this was at the expense of lower percentile values (Figure S1, Table S1).

The GSDs measured were significantly different from the sieved distributions in Tredegar: Wolman pebble counts ($\chi^2 = 64.14$, d.f. = 3, p-value < 0.05); survey tape counts ($\chi^2 = 13.03$, d.f. = 5, p-value < 0.05); manual photo counts ($\chi^2 = 28.95$, d.f. = 3, p-value < 0.05); and pyDGS ($\chi^2 = 25.95$, d.f. = 5, p-value < 0.05) (Table S6). To infer the magnitude of methodological uncertainty in these values, we truncated the individual distributions in Tredegar to the grain size range covered by all methods (3–34.4 mm, Figure 4; Church et al., 1987). Within this range, survey tape counts, pyDGS and sieving GSDs were similar (sieving and survey tape: $\chi^2 = 6.44$, d.f. = 3, p-value > 0.05; sieving and pyDGS: $\chi^2 = 1.01$, d.f. = 3, p-value > 0.05; survey tape and pyDGS: $\chi^2 = 5.45$, d.f. = 3, p-value > 0.05) (Figure 4, Table S6). Manual photo counts and Wolman pebble counts were significantly different, with 95% confidence to sieving GSDs ($\chi^2 = 12.64$, d.f. = 3, p-value < 0.05; $\chi^2 = 29.91$, d.f. = 3, p-value < 0.05) and survey tape GSDs ($\chi^2 = 8.02$, d.f. = 3, p-value < 0.05; $\chi^2 = 21.07$, d.f. = 3, p-value < 0.05). Wolman pebble count GSDs over-predicted grain sizes in the lower 50% of the distribution when compared with all other methods.

TABLE 3 The D_{16}, D_{50} and D_{84} grain sizes for the five different methods used in Tredegar. The first half of the table gives the percentiles for each method across their entire GSD. The second half of the table shows the percentiles for the single-order grain size covered by all five methods. Values in brackets give the range of grain sizes for each percentile calculated using the three different methods indicated

Percentile	Tredegar entire GSD (mm)				
	Sieving[a]	Wolman pebble count[b]	Survey tape[b]	Manual photo counts[b]	pyDGS $x = -0.5$[c]
16	2.2 (2–4)	7 (6–8)	4 (3–5)	4.1 (3.6–4.6)	1.4 (1–1.8)
50	6.6 (4–8)	13 (12–14)	10 (8.4–12)	8.8 (7.8–9.5)	4.5 (3.4–5.6)
84	21 (20–40)	25 (22–28)	26 (20–33)	17 (15–20)	16 (12–20)

Percentile	Tredegar truncated (3–34.3 mm) (mm)					
	Sieving[a]	Wolman pebble count[b]	Survey tape[b]	Manual photo counts[b]	pyDGS $x = -0.5$[c]	pyDGS $x = 0$[c]
16	4.2 (4–8)	7 (6–7.3)	5 (4–5)	4.9 (4.5–5.4)	4.0 (3–5)	4.8 (3.6–6)
50	7.8 (4–8)	12 (11–13)	10 (8.3–11)	9.1 (8.2–9.9)	8.0 (6–10)	11 (8.3–14)
84	19 (10–20)	21 (20–23)	19 (16–23)	16 (15–18)	20 (15–25)	24 (18–30)

[a]Each sieving percentile was calculated using an assumed linear relationship between the minimum and maximum values of each grain-size bin. Therefore, we have also given the minimum and maximum grain-size bin for each percentile in brackets.
[b]These percentiles were generated using the QuantBD function developed by Eaton et al. (2019) and translated into Python by Purinton and Bookhagen (2021). The percentile uncertainty is quantified using binomial theory for each percentile based on the number of measurements. We provide the minimum and maximum grain-size range for each percentile generated using this technique in brackets for a 95% confidence interval.
[c]These percentiles were generated using pyDGS. The range given in brackets is based on a conservative 25% error estimate based on the errors quantified by Buscombe (2013) for GSDs measured manually and pyDGS GSDs for individual images.

TABLE 4 Common statistical metrics used to describe GSDs. The Liusha pyDGS percentiles are based on a *maxscale* of 6; the Luoquan pyDGS percentiles are based on a *maxscale* of 4

Percentile	Liusha (mm)				
	Sieving[a]	Manual photo counts[b]	pyDGS $x = -0.5$[c]	pyDGS $x = 1$[c]	pyDGS *combined*[c]
16	13 (12–24)	7.4 (5.1–10.4)	5.0 (3.8–6.3)	36 (27–45)	15 (11–19)
50	77 (49–80)	45 (25–61)	19 (14–24)	83 (62–104)	72 (54–90)
84	130[d]	104 (82–130)	65 (49–81)	130 (98–163)	122 (92–153)

Percentile	Luoquan (mm)				
	Sieving[a]	Manual photo counts[b]	pyDGS $x = -0.5$[c]	pyDGS $x = 1$[c]	pyDGS *combined*[c]
16	56 (47–80)	33 (28–49)	5.9 (4.4–7.4)	57 (43–71)	49 (37–61)
50	100[d]	79 (69–89)	23 (17–29)	150 (113–188)	98 (74–123)
84	248 (210–250)	119 (100–190)	89 (67–111)	248 (186–310)	178 (134–223)

[a]Each sieving percentile was calculated using an assumed linear relationship between the minimum and maximum values of each grain-size bin. Therefore, we have also given the minimum and maximum grain-size bin for each percentile in brackets.
[b]These percentiles were generated using the QuantBD function developed by Eaton et al. (2019) and translated into Python by Purinton and Bookhagen (2021). The percentile uncertainty is quantified using binomial theory for each percentile based on the number of measurements. We provide the minimum and maximum grain-size range for each percentile generated using this technique in brackets for a 95% confidence interval.
[c]These percentiles were generated using pyDGS. The range given in brackets is based on a conservative 25% error estimate based on the errors quantified by Buscombe (2013) for manually observed GSDs for individual images and pyDGS-generated GSDs. The pyDGS combined percentiles in Liusha are based on a full GSD generated in pyDGS with a *shape parameter* of 0 and *maxscale* of 8 combined with the GSD for grains >80 mm with a *shape parameter* of −1 and *maxscale* of 6. In Luoquan the GSD are based on a full GSD with a *shape parameter* of 1 and *maxscale* of 8 combined with the GSD for grains >80 mm with a *shape parameter* of −1 and *maxscale* of 4.
[d]No range is given for this percentile as the surrounding grains also have the same *b*-axis.

5 | DISCUSSION

5.1 | Sampling method uncertainty

There are several uncertainties inherent in each sampling method that can lead to systematic bias in the reported results (Table 1). Bias in the sieving GSDs may be introduced as each sample integrates the subsurface and surface grains into a single GSD (Bunte & Abt, 2001b; Dufresne & Dunning, 2017; Johnson et al., 2012). Where deposits are vertically stratified, this will lead to under-estimation of coarse

(or fine) surface fractions. We mitigated this issue by choosing sampling locations that showed no evidence of vertical stratification. We found no systematic change in the grain size with depth, suggesting that differences in GSD are more likely to reflect primary variability in the deposits rather than vertical stratification (Figure S5). For example, while D_{16}, D_{50} and D_{84} values for the uppermost three layers in Luoquan ranged from 30 to 56 mm, 100 to 137 mm and 248 to 306 mm, respectively, there was no evidence of stratification with depth. Additionally, the D_{16} and D_{50} values obtained for sieving were consistently lower than the values obtained using Wolman pebble
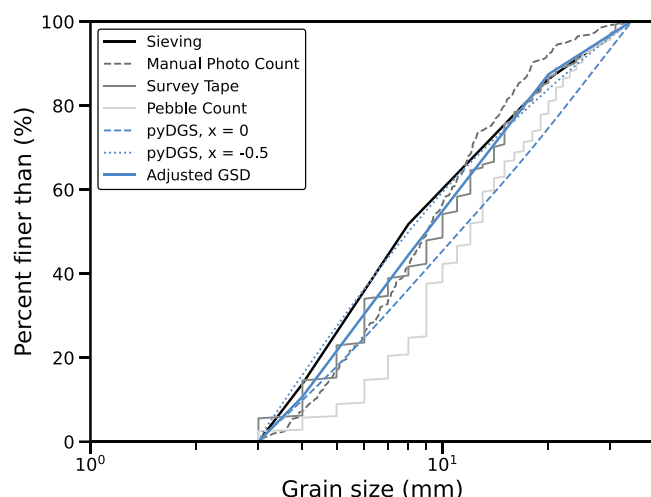
FIGURE 4 Comparison of truncated field-derived GSDs and pyDGS-derived GSDs for the Tredegar rockslide. We found that the range 3–34.4 mm was covered by all five methods and therefore adjusted all curves to fit this range. The sieving GSD is based on a sample taken within the first 10 cm of the surface near the centre of the deposit. Approximately two-thirds of the sample is within the truncated grain size range. The survey tape count is based on a total of 181 grains across the entire deposit. 144 of these grains were within the truncated range. Both the Wolman pebble count and manual photo count consisted of measuring 300 grains. In the Wolman pebble count, 279 grains were within the truncated range. 259 grains from the manual photo count were within the truncated range. The pyDGS curve is the average of the 60 GSDs generated using individual photos of the deposit. Approximately 60% of the full GSD for pyDGS shown in Figure 3 was within the truncated range. The adjusted GSD (blue) is calculated by combining the surface sieving GSD (black) and the survey tape GSD (grey) using the method outlined by Fripp and Diplas (1993) and briefly in the main text. We used sieving and survey tape GSDs as these provided the minimum and maximum grain sizes for the location based on their generated GSDs

counts, survey tape counts and manual photo counts, demonstrating that the changes observed with depth should not result in simply a coarser or finer surface. The D_{84} values in Luoquan and Liusha were also larger than all other surface D_{84} values. Further bias could be introduced if the pit is not constructed correctly, which we avoided by consistently measuring the width and depth of the pit when digging. Whilst sieving presents challenges in terms of efficiency and accessibility, it is the only method able to successfully measure sand grains and finer. Where time or equipment is limited, an alternative method may be chosen, but no other method will be able to sample this fine fraction, which represents up to 20% of the GSD by weight (Casagli et al., 2003). For our statistical comparisons we use sieving as the test statistic due to its larger sample sizes and widest GSDs, making it most likely to be representative of the true distribution across its sampling range.

Photo-based techniques can be limited by photo extent, imbrication, overlap and mistakenly measuring the *c*-axis as opposed to the *b*-axis (Attal & Lavé, 2006; Casagli et al., 2003; Kellerhals & Bray, 1971). These limitations can result in GSDs that under-estimate the coarse end of the distribution (Tables S2 and S3). Furthermore, small sample sizes can also lead to under-estimating the D_{84} values. Despite using the same photos, we found significant differences when

comparing manual photo count and pyDGS GSDs in Tredegar (Table S6; full GSD: $\chi^2 = 62.04$, d.f. = 3, *p*-value < 0.05; truncated GSD: $\chi^2 = 22.21$, d.f. = 3, *p*-value < 0.05). pyDGS with a *shape parameter* of 1 also under-estimated the proportion of grains between 5 and 40 mm relative to manual photo counts in Liusha (Figure 6). The differences in the GSDs obtained using each method may be attributed to the lack of contrast between the fine grains in the image. The lack of contrast results in smaller changes in the texture of the image and therefore reduces the ability of the pyDGS algorithm to register these as grains. Images where the fine grains are all of similar colour are difficult to differentiate, resulting in the individual grains being considered as single larger grains (Buscombe, 2013; Figures 6, S3 and S4). This effect may be enhanced by wet grains in Figures S4A and 6 (Buscombe, 2013).

pyDGS was not well suited to the application of mass movement deposits as it requires calibration and multiple shape factors to capture complex GSDs (Figures 5 and 6; Tables S2, S3 and S4). Even when calibrated, the GSD obtained using pyDGS was statistically different from the full GSDs obtained using sieving in all three locations (Tredegar: $\chi^2 = 25.95$, d.f. = 5, *p*-value < 0.05; Liusha: $\chi^2 = 14.85$, d.f. = 4, *p*-value < 0.05; Luoquan: $\chi^2 = 22.01$, d.f. = 5, *p*-value < 0.05). We acknowledge that pyDGS has the major benefit of automatically generating GSDs from photos, which can enhance our ability to record GSDs over high spatial and temporal resolutions. However, for complex, large mass movement deposits, pyDGS-generated GSDs rely too heavily on the use of another method for calibration that does not increase efficiency.

Wolman pebble counts and survey tape counts cannot measure the finest grains and have minimum grain sizes of 3 and 1 mm for each method, respectively (Figures 3 and 4, Table 3). The statistically different, coarser, GSD for random Wolman pebble counts when compared to survey tape counts ($\chi^2 = 21.07$, d.f. = 3, *p*-value < 0.05) is possibly due to fine pebbles being overlooked by the technique as a result of operator bias (Fripp & Diplas, 1993; Strom et al., 2010). Operator bias may be even more pronounced in heterogeneous, multimodal mass movement deposits towards the extreme small or large grains (Daniels & McCusker, 2010; Strom et al., 2010). A limitation of our approach is the fact that we sampled grains <2 mm using the survey tape method and <4 mm using the Wolman count method (Bunte & Abt, 2001a). Whilst these are below the expected minimum grain sizes in Table 1, we wanted to provide the full GSD of grains visible in the field. The minimum GSDs often used, 4–8 mm, are dictated by work on fluvial GSDs, which are likely to be inundated by shallow water (e.g. Bunte & Abt, 2001a; Kellerhals & Bray, 1971). In a mass movement deposit, the smaller grains on the surface are more likely to be visible, which may allow for the sampling of smaller grains. The higher potential to exclude fine grains when conducting pebble counts, particularly randomly through a Wolman pebble count, will result in mass movement GSDs, which exclude any silt, sand or clay.

5.2 | Methodological uncertainty, sample size and sample type

No single method accurately measured the full GSD in any of the mass movement deposits studied (Figures 3, 5 and 6). In Tredegar, we combined two GSDs collected using different methods to obtain a full
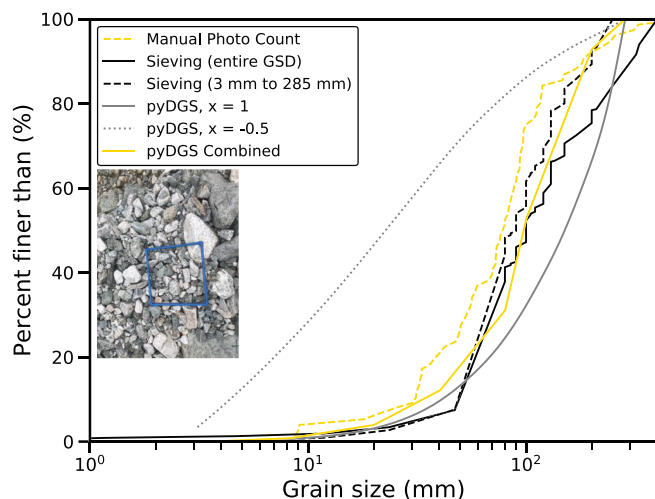
FIGURE 5 Sieving, manual and automated photo analysis-based surface GSDs for the Luoquan (Figure 1c) debris flow deposit. The solid gold line shows the GSD derived by combining two pyDGS runs. The inset photo shows the pit image used to estimate surface GSDs from manual photo counts and pyDGS. In total, 76 grains were measured using a manual photo grid sampling technique
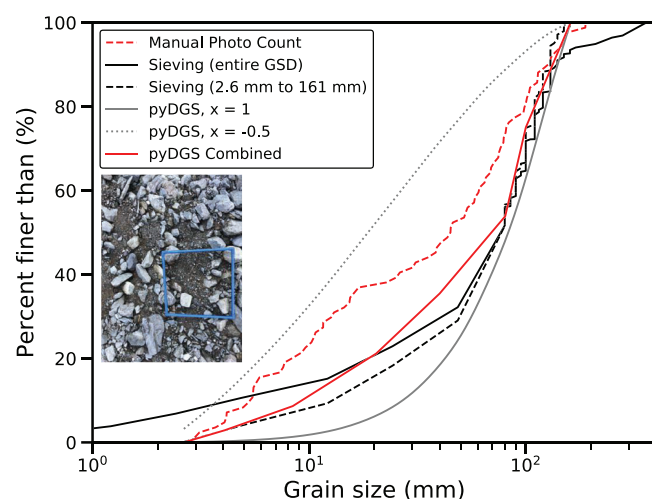

FIGURE 6 Sieving, manual and automated photo analysis-based surface GSDs for the Liusha (Figure 1b) debris flow deposit. The solid red line shows the GSD derived using two pyDGS runs. The inset photo shows the pit image used to estimate surface GSDs from manual photo counts and pyDGS. In total, 84 grains were measured using a manual photo grid sampling technique

GSD. When selecting which methods to combine, it is important to consider the differences in the distribution produced based on methodological uncertainty, sample size and sampling method. To identify differences in the GSDs associated with methodological uncertainties, we compared the GSDs measured using different methods across a restricted set of grain sizes, where issues of resolution are likely to be minimal. Wolman pebble counts and manual photo counts had significantly different distributions across this restricted grain size range, suggesting that these methods are the least comparable to sieving and survey tape counts, and therefore the least reliable. Survey tape counts and pyDGS GSDs were consistent with sieving GSDs over a restricted grain-size range (Table S6), implying that they are strong

candidates for combining to create a full GSD. The statistically similar relationship between survey tape counts and sieving over a single order of magnitude suggests that a more systematic approach to pebble counts can be used to represent the fraction of grains larger than fine gravel in a mass movement deposit better than a random pebble count (Kellerhals & Bray, 1971). Consequently, any statistical differences across the full GSD measured by these three methods are likely to be a result of the sample size and sample range of each method. The importance of method choice, and grain-size range, was further reflected in the percentile values for each method. For full GSDs, survey tape counts, manual photo counts and Wolman pebble counts all over-estimated D_{16} relative to the sieving D_{16} value, due to their inability to sample grains smaller than gravel (Table 1; Casagli et al., 2003; Wolman, 1954). Thus, the use of Wolman pebble counts or manual photo counts introduces methodological uncertainties to the sampling of mass movement deposits and results in statistically different, unreliable, GSDs.

Methodological differences in sample size may also affect the measured GSDs and explain the differences in manual photo count and Wolman pebble count GSDs (Church et al., 1987; Purinton & Bookhagen, 2021; Storz-Peretz & Laronne, 2013). Primarily, the issue of sample size relates to the ability to accurately capture the coarse end of the distribution (Church et al., 1987). Previous studies have suggested recommended sample sizes for different methods based on coarse fluvial deposits (e.g. Eaton et al., 2019; Fripp & Diplas, 1993; Graham et al., 2010; Kellerhals & Bray, 1971; Purinton & Bookhagen, 2021). When we applied these methods to large, complex mass movement deposits, such as Liusha and Luoquan where the coarse grains were much larger than is typical in fluvial settings, it was challenging to strictly apply these sample sizes. Recommended sample sizes vary for survey tape and Wolman pebble count methods based on the range of grain sizes found in the deposit. Measuring enough grains for at least the 84th percentile to converge provides a helpful criterion (Purinton & Bookhagen, 2021). For the finer mass movement deposit in Tredegar, the 84th percentile converged after 300 measurements, which took approximately 1–2 h of sample time. This timescale is not significantly different from that required to construct and sieve a pit in a fine deposit.

Photo counts have a recommended areal coverage of 100–200 times the D_{max} to obtain <10% errors (Eaton et al., 2019; Graham et al., 2010; Purinton & Bookhagen, 2021; Storz-Peretz & Laronne, 2013). In the Longmen Shan, where the D_{max} from the images used were 189 and 552 mm, respectively, this would require a photo with a width of 11 m (an area of >100 m^2). Such a photo could only be taken with an unmanned aerial vehicle (UAV) and would subsequently compromise the resolution of the finest grains (unless combined with a higher-resolution photo) (Graham et al., 2010; Storz-Peretz & Laronne, 2013). This example highlights the primary challenges of sample size, as it is common to find grains >500 mm in mass movement deposits that may be smaller than 100 m^2 or where larger areas are not spatially uniform, for example due to segregation.

In fluvial environments, there is a volumetric sieving target where a maximum of 5% of the total weight limit can be made up of the largest grain (Church et al., 1987). Occasionally, boulders >50 kg were still recorded in the debris flow deposit pits, which meant this criterion was not always achievable. Where deposits are small or only a fraction of the deposit needs to be sampled, sieving may be a more

appropriate technique for obtaining a GSD, though it is difficult to achieve the recommended sample sizes in mass movement deposits for any individual sampling method. As such, accurate GSDs, which meet the recommended sample sizes, are more likely to be achieved by combining multiple methods that are optimized to sample certain grain-size ranges (Attal & Lavé, 2006; Casagli et al., 2003; Fripp & Diplas, 1993).

Whilst we chose areas of the deposits that were spatially uniform, we note that each method has slightly different sampling frequencies and depths. We refer to the uncertainty associated with differences in the location of the sample as the sample type. Our field sites were chosen to minimize differences in sample type. Within the three pits sampled, there was no vertical stratification by grain size across the top 50 cm (which we sampled at 10 cm intervals) (Figure S5). Thus, there is no evidence that a surface sample would be significantly different from a sieved sample (Attal & Lavé, 2006). We combined the sieving and survey tape GSDs to produce a full distribution. We assumed that sieving and survey tape GSDs could be merged due to the fact that they do not produce statistically different distributions over a truncated range of grain sizes (Figure 4, Table S6). Whilst the combined GSD will produce the widest GSD, we note that the uncertainties associated with combined methods are likely to be propagated in the adjusted distribution, for example the effect of sample size. In Luoquan and Liusha it was not necessary to combine multiple GSDs because sieving recorded the minimum and maximum grain size. Thus, combining complementary methods that sample different grain size ranges, but without significant methodological uncertainty (e.g. sieving and survey tape), may provide the best opportunity to accurately report the full GSD of mass movement deposits.

5.3 | Applying these methods to different types of mass movement

A solution to the challenges associated with developing accurate GSDs across the wide range of mass movement grain sizes is to vary the method based on the research question being asked. In many cases, only a portion of the entire GSD is required to identify the transport and depositional mechanisms occurring within a deposit and subsequently interpret the types of mass movement (Blair, 1999; Cruden & Varnes, 1996; Kaitna et al., 2016; McKenna et al., 2012; Wang & Sassa, 2003). For example, flow-like failures are commonly associated with processes such as inverse grading and kinetic sieving, which result in a coarse surface layer, front and levees (Johnson et al., 2012). The GSD of levees may require characterization of grain size across a wider spatial scale, using survey tape counts or manual point counts. In contrast, sieving may be better suited when deposits have a high proportion of fine material, such as for viscous flows (Kaitna et al., 2016; Wang & Sassa, 2003).

Measurements of deposit GSDs have been used to infer the source of the material mobilized from the relationship between bedrock strength and the GSD of rock avalanche, rockfall and landslide deposits (Dunning, 2006; Marc et al., 2021). GSDs can also help to identify the source of the mobilized material. For example, in California, finer, sandier debris flows were hillslope triggered, whereas the coarser debris flows mobilized material from within the channel (Kean et al., 2011). These findings may also be supported in the Longmen

Shan, where rock type variability may explain the higher proportion of grains <10 mm in Liusha in comparison to Luoquan (Figures 5 and 6). As the fracture spacing of metasediments is smaller than the granitoids found in Luoquan (Figure 1D), this difference may have been overlooked using a method that is biased towards coarser grain sizes.

Mass movement GSDs are more commonly obtained for rock avalanches, debris flows and landslides, where grain size plays a role in controlling mobility through processes such as comminution, fragmentation and segregation (Crosta et al., 2007; Dufresne & Dunning, 2017; Dunning, 2006; Locat et al., 2006). These processes produce GSDs with potentially large spatial variability, a wide range of grain sizes and bimodal or multimodal distributions (Crosta et al., 2007; Dufresne & Dunning, 2017; Makris et al., 2020). An understanding of the entire GSD of rock avalanche deposits can also help to understand what controls the rate of different transport and depositional processes. All grain sizes were found to control segregation in an experimental setting for dry granular flows, which includes rock avalanches (Gray & Ancey, 2011). Here, a higher proportion of fine grains resulted in a longer distance being required for medium and large particles to segregate (Gray & Ancey, 2011). The efficiency of fragmentation in deposits is also thought to relate to GSDs. For example, there is a decrease in the efficiency of fragmentation when the number of fines increases as the fines act to buffer interactions between larger grains (Locat et al., 2006). Whilst Locat et al. (2006) obtained this conclusion using photographs of grains, they did note that their proportions of fines were likely to be an under-estimate. Hence, whilst broad patterns can be well captured using more accessible, common methods (Marc et al., 2021), it is important to capture full GSDs for deposits, using multiple methods, when identifying depositional and transport processes.

Examples of where a restricted sampling of the GSD of mass movement deposits might be useful is when considering their contribution to rockfall hazard and fluvial bedload transport. In rockfalls, deposited grain volume can predict runout hazards better than the initial volume, which tends to over-estimate kinetic energy and runout (Ruiz-Carulla et al., 2015). Subsequently, only the coarse fraction (rocks >0.01 m^3) is required, as this can provide an indication of the furthest point to which the runout will travel, which is most important for hazard models. Depending on the nature of the hazard, using a single method to rapidly constrain the GSD of coarser boulders may therefore outweigh the importance of spending considerable time extracting the entire GSD of the deposit using sieving. In terms of fluvial bedload, the GSD > 1 mm of landslides has been successfully compared directly to the GSD of weathering products to understand the importance of landslides in hillslope and fluvial sediment budgets (Roda-Boluda et al., 2018). This was achievable because the study only focused on the surface material, where most fines have been washed away. However, the appropriate method will vary depending on the mass movement deposit sampled and the GSD of the other processes acting within the catchment. The importance of the entire GSD has been shown for the Marsyandi River, where the pebble and suspended/bedload ratio were both affected by hillslope processes, including landslides (Attal & Lavé, 2006).

The methodological uncertainties associated with comparing GSDs and percentiles obtained using different methods can have consequences for accurate process interpretation. For example, the factor of two difference in grain-size percentile estimates from survey tape

counts relative to sieving for a fine deposit could shift the D_{50} value from suspended load to bedload, which would have implications for estimates of sediment export rates and onward transport (Croissant et al., 2021; Marc et al., 2021). Similarly, by excluding up to 20% by weight of the finest grains, all non-sieving methods are unable to find evidence for processes where the proportion of sand and silt is influential (de Haas et al., 2015; Kaitna et al., 2016; Makris et al., 2020). The rates and calibre of hillslope sediment supply to channels have also been used increasingly to drive landscape evolution and fluvial modelling (Attal et al., 2015; Croissant et al., 2021; Egholm et al., 2013; Roda-Boluda et al., 2018). Given that mass movement-derived sediment is an essential component in these problems (Sklar & Dietrich, 2006), improvements are needed in our ability to characterize this material to provide robust conclusions about the timescales and rates of bedrock incision and sediment transport.

6 | CONCLUSION

Measurements of mass movement GSDs present concerns over accuracy, precision and pragmatism. Each study is required to make choices about methodology, sampling locations and size that suit both the research question being asked and the practical challenges of field sites. Here we show that these choices about methodology can introduce up to a factor of five difference in simple metrics like D_{16} and D_{50}. This results in GSDs and grain-size percentiles that are not directly comparable to GSDs measured using different methods, especially when the same grain-size range is not considered. We demonstrate that for smaller, finer mass movement deposits, survey tape counts and pyDGS are a suitable alternative to sieving for measuring the GSD over a single order of magnitude. Whilst pyDGS could be used to obtain a representative GSD over a single order of magnitude for the smaller landslide deposit, once trained, we were unable to obtain a representative GSD using a single curve for the larger debris flow deposits. In the larger, coarser debris flow deposits in the Longmen Shan, manual photo counts were unable to obtain the maximum resolution measured using sieving. We were also unable to reach the desired sample size for manual photo counts for coarse deposits. In all cases clear, detailed descriptions of the protocol are essential so that uncertainties introduced by different methods can be quantified and the implications for process interpretation can be better understood.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

AUTHOR CONTRIBUTIONS

a. Conceptualization: ELH, TCH, DEJH, XF

b. Funding acquisition: ELH, TCH, XF

c. Methodology: ELH, TCH, DEJH, JL

d. Investigation (data collection): ELH, TCH, JL

e. Resources (provision of data): ELH, TCH

f. Software: ELH, TCH, DEJH

g. Supervision: TCH, DEJH, XF

h. Writing (initial draft): ELH, TCH

i. Writing (reviewing and editing): ELH, TCH, DEJH, JL, XF

DATA AVAILABILITY STATEMENT

All data used in this paper is available upon request. The code for pyDGS was written by Daniel Buscombe and can be freely accessed at https://github.com/dbuscombe-usgs/pyDGS.

ORCID

Erin L. Harvey https://orcid.org/0000-0002-6167-9438
Tristram C. Hales https://orcid.org/0000-0002-3330-3302
Daniel E. J. Hobley https://orcid.org/0000-0003-2371-0534

REFERENCES

Attal, M. & Lavé, J. (2006) Changes of bedload characteristics along the Marsyandi River (central Nepal): Implications for understanding hillslope sediment supply, sediment load evolution along fluvial networks, and denudation in active orogenic belts. *Tectonics, Climate, and Landscape Evolution*, 398, 143–171. Available from: https://doi.org/10.1130/2006.2398(09)

Attal, M., Mudd, S.M., Hurst, M.D., Weinman, B., Yoo, K. & Naylor, M. (2015) Impact of change in erosion rate and landscape steepness on hillslope and fluvial sediments grain size in the Feather River basin (Sierra Nevada, California). *Earth Surface Dynamics*, 3(1), 201–222. Available from: https://doi.org/10.5194/esurf-3-201-2015

Barclay, W.J., Jackson, D.I., Mitchell, M., Owens, B., Riley, N.J., White, D.E., Strong, G.E. & Monkhouse, R.A. (1989) *Geology of the South Wales Coalfield: Memoir for 1:50,000 Geological Sheet 232 (England and Wales)*. London: HMSO.

Bitelli, G., Dubbini, M. & Zanutta, A. (2004) Terrestrial laser scanning and digital photogrammetry techniques to monitor landslide bodies. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35, 246–251.

Blair, T.C. (1999) Cause of dominance by sheetflood vs. debris-flow processes on two adjoining alluvial fans, Death Valley, California. *Sedimentology*, 46(6), 1015–1028. Available from: https://doi.org/10.1046/j.1365-3091.1999.00261.x

Brasington, J., Vericat, D. & Rychkov, I. (2012) Modeling river bed morphology, roughness, and surface sedimentology using high resolution terrestrial laser scanning. *Water Resources Research*, 48(11), 1–18. Available from: https://doi.org/10.1029/2012WR012223

Bunte, K. & Abt, S.R. (2001a) Sampling frame for improving pebble count accuracy in coarse gravel-bed streams. *Journal of the American Water Resources Association*, 37(4), 1001–1014. Available from: https://doi.org/10.1111/j.1752-1688.2001.tb05528.x

Bunte, K. & Abt, S.R. (2001b) *Sampling Surface and Sub-surface Particle Size Distributions in Wadeable Gravel- and Cobble-Bed Streams for Analysis in Sediment Transport, Hydraulics, and Streambed Monitoring*. Report No. GTR-74. USDA Forest Service, Rocky Mountain Research Station: Fort Collins, CO.

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

Carbonneau, P.E., Bergeron, N. & Lane, S.N. (2005) Automated grain size measurements from airborne remote sensing for long profile measurements of fluvial grain sizes. *Water Resources Research*, 41(11), 1–9. Available from: https://doi.org/10.1029/2005WR003994

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), 1–11. Available from: https://doi.org/10.1029/2003WR002759

Casagli, N., Ermini, L. & Rosati, G. (2003) Determining grain size distribution of the material composing landslide dams in the Northern Apennines: Sampling and processing methods. *Engineering Geology*, 69(1–2), 83–97. Available from: https://doi.org/10.1016/S0013-7952(02)00249-1

Chen, H., Su, D. & Chen, K. (2001) Some case studies on the engineering geological characteristics of debris flows in Taiwan. *Western Pacific Earth Sciences*, 1, 265–296.

Church, M.A., McLean, D.G. & Wolcott, J.F. (1987) River bed gravels: Sampling and analysis. In: Thorne, C.R., Bathurst, J.C. & Hey, R.D. (Eds.) *Gravel Bed Rivers*. Chichester: Wiley, pp. 43–87.

Croissant, T., Hilton, R.G., Li, G.K., Howarth, J., Wang, J., Harvey, E.L., Steer, P. & Densmore, A.L. (2021) Pulsed carbon export from mountains by earthquake-triggered landslides explored in a reduced-complexity model. *Earth Surface Dynamics Discussions*, 9(4), 823–844. Available from: https://doi.org/10.5194/esurf-9-823-2021

Crosta, G.B., Frattini, P. & Fusi, N. (2007) Fragmentation in the Val Pola rock avalanche, Italian Alps. *Journal of Geophysical Research – Earth Surface*, 112(F1), 1–23. Available from: https://doi.org/10.1029/2005JF000455

Cruden, D.M. & Varnes, D.J. (1996) Landslide types and processes. In: Turner, A.K. & Schuster, R.L. (Eds.) *Landslides Investigation and Mitigation. Special Report 247 of the Transportation Research Board*. Washington, DC: National Academy Press, pp. 36–75.

Cucchiaro, S., Cavalli, M., Vericat, D., Crema, S., Llena, M., Beinat, A., Marchi, L. & Cazorzi, F. (2018) Monitoring topographic changes through 4D-structure-from-motion photogrammetry: Application to a debris-flow channel. *Environmental Earth Sciences*, 77(18), 1–21. Available from: https://doi.org/10.1007/s12665-018-7817-4

Cuttler, M.V.W., Lowe, R.J., Falter, J.L. & Buscombe, D. (2017) Estimating the settling velocity of bioclastic sediment using common grain-size analysis techniques. *Sedimentology*, 64(4), 987–1004. Available from: https://doi.org/10.1111/sed.12338

Daniels, M.D. & McCusker, M.H. (2010) Operator bias characterizing stream substrates using Wolman pebble counts with a standard measurement template. *Geomorphology*, 115(1–2), 194–198. Available from: https://doi.org/10.1016/j.geomorph.2009.09.038

de Haas, T., Braat, L., Leuven, J.R.F.W., Lokhorst, I.R. & Kleinhans, M.G. (2015) Effects of debris flow composition on runout, depositional mechanisms, and deposit morphology in laboratory experiments. *Journal of Geophysical Research – Earth Surface*, 120(9), 1949–1972. Available from: https://doi.org/10.1002/2015JF003525

de Scally, F.A. & Owens, I.F. (2005) Depositional processes and particle characteristics on fans in the Southern Alps, New Zealand. *Geomorphology*, 69(1–4), 46–56. Available from: https://doi.org/10.1016/j.geomorph.2004.11.021

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.

Diplas, P. & Fripp, J.B. (1992) Properties of various sediment sampling procedures. *Journal of Hydraulic Engineering*, 118(7), 955–970. Available from: https://doi.org/10.1061/(asce)0733-9429(1992)118:7(955)

Diplas, P. & Sutherland, A.J. (1988) Sampling techniques for gravel-sized sediments. *Journal of Hydraulic Engineering*, 114(5), 484–501. Available from: https://doi.org/10.1061/(asce)0733-9429(1988)114:5(484)

Dufresne, A. & Dunning, S.A. (2017) Process dependence of grain size distributions in rock avalanche deposits. *Landslides*, 14(5), 1555–1563. Available from: https://doi.org/10.1007/s10346-017-0806-y

Dunning, S.A. (2006) The grain-size distribution of rock avalanche deposits in valley-confined settings. *Italian Journal of Engineering Geology and Environment*, 1, 117–121. Available from: https://doi.org/10.4408/IJEGE.2006-01.S-15

Dunning, S.A., Massey, C.I. & Rosser, N.J. (2009) Structural and geomorphological features of landslides in the Bhutan Himalaya derived from Terrestrial Laser Scanning. *Geomorphology*, 103(1), 17–29. Available from: https://doi.org/10.1016/j.geomorph.2008.04.013

Eaton, B.C., Dan Moore, R. & Mackenzie, L.G. (2019) Percentile-based grain size distribution analysis tools (GSDtools) – estimating confidence limits and hypothesis tests for comparing two samples. *Earth Surface Dynamics*, 7(3), 789–806. Available from: https://doi.org/10.5194/esurf-7-789-2019

Egholm, D.L., Knudsen, M.F. & Sandiford, M. (2013) Lifespan of mountain ranges scaled by feedbacks between landsliding and erosion by rivers. *Nature*, 498(7455), 475–478. Available from: https://doi.org/10.1038/nature12218

Folk, R.L. & Ward, W.C. (1957) Brazos River bar: A study in the significance of grain size parameters. *Journal of Sedimentary Petrology*, 27(1), 3–26. Available from: https://doi.org/10.1306/74D70646-2B21-11D7-8648000102C1865D

Fripp, J.B. & Diplas, P. (1993) Surface sampling in gravel streams. *Journal of Hydraulic Engineering*, 119(4), 473–490. Available from: https://doi.org/10.1061/(asce)0733-9429(1993)119:4(473)

Genevois, R., Galgaro, A. & Tecca, P.R. (2001) Image analysis for debris flow properties estimation. *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science*, 26(9), 623–631. Available from: https://doi.org/10.1016/S1464-1917(01)00059-9

George, G.T. (2015) *The Geology of South Wales*, 2nd edition. Maidstone: Geoserv Publishing.

Graham, D.J., Reid, I. & Rice, S.P. (2005) Automated sizing of coarse-grained sediments: Image-processing procedures. *Mathematical Geology*, 37(1), 1–28. Available from: https://doi.org/10.1007/s11004-005-8745-x

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), 1–15. Available from: https://doi.org/10.1029/2008WR006940

Gray, J.M.N.T. & Ancey, C. (2011) Multi-component particle-size segregation in shallow granular avalanches. *Journal of Fluid Mechanics*, 678, 535–588. Available from: https://doi.org/10.1017/jfm.2011.138

Guerit, L., Barrier, L., Liu, Y., Narteau, C., Lajeunesse, E., Gayer, E. & Métivier, F. (2018) Uniform grain-size distribution in the active layer of a shallow, gravel-bedded, braided river (the Urumqi River, China) and implications for paleo-hydrology. *Earth Surface Dynamics*, 6(4), 1011–1021. Available from: https://doi.org/10.5194/esurf-6-1011-2018

Gupta, S.K. & Shukla, D.P. (2018) Application of drone for landslide mapping, dimension estimation and its 3D reconstruction. *Journal of the Indian Society of Remote Sensing*, 46(6), 903–914. Available from: https://doi.org/10.1007/s12524-017-0727-1

Hubert, J.F. & Filipov, A.J. (1989) Debris-flow deposits in alluvial fans on the west flank of the White Mountains, Owens Valley, California, U.S.A. *Sedimentary Geology*, 61(3–4), 177–205. Available from: https://doi.org/10.1016/0037-0738(89)90057-2

Ibbeken, H., Warnke, D.A. & Diepenbroek, M. (1998) Granulometric study of the Hanaupah Fan, Death Valley, California. *Earth Surface Processes and Landforms*, 23(6), 481–492. Available from: https://doi.org/10.1002/(SICI)1096-9837(199806)23:6<481::AID-ESP906>3.0.CO;2-T

Johnson, C.G., Kokelaar, B.P., Iverson, R.M., Logan, M., LaHusen, R.G. & Gray, J.M.N.T. (2012) Grain-size segregation and levee formation in geophysical mass flows. *Journal of Geophysical Research – Earth Surface*, 117(F1), 1–23. Available from: https://doi.org/10.1029/2011jf002185

Kaitna, R., Palucis, M.C., Yohannes, B., Hill, K.M. & Dietrich, W.E. (2016) Effects of coarse grain size distribution and fine particle content on pore fluid pressure and shear behavior in experimental debris flows. *Journal of Geophysical Research – Earth Surface*, 121(2), 415–441. Available from: https://doi.org/10.1002/2015JF003725

Kean, J.W., Staley, D.M. & Cannon, S.H. (2011) In situ measurements of post-fire debris flows in southern California: Comparisons of the

timing and magnitude of 24 debris-flow events with rainfall and soil moisture conditions. *Journal of Geophysical Research – Earth Surface*, 116(F4), 1–21. Available from: https://doi.org/10.1029/2011JF002005

Kellerhals, R. & Bray, D. (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

Kim, B.C. & Lowe, D.R. (2004) Depositional processes of the gravelly debris flow deposits, South Dolomite alluvial fan, Owens Valley, California. *Geosciences Journal*, 8(2), 153–170. Available from: https://doi.org/10.1007/BF02910191

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

Locat, P., Couture, R., Leroueil, S., Locat, J. & Jaboyedoff, M. (2006) Fragmentation energy in rock avalanches. *Canadian Geotechnical Journal*, 43(8), 830–851. Available from: https://doi.org/10.1139/T06-045

Ma, L. (2002) *Geological Atlas of China*. Beijing: Geological Publishing House.

Major, J.J. & Voight, B. (1986) Sedimentology and clast orientations of the 18 May 1980 southwest-flank lahars, Mount St. Helens, Washington. *Journal of Sedimentary Petrology*, 56, 691–705. Available from: https://doi.org/10.1306/212F8A1C-2B24-11D7-8648000102C1865D

Makris, S., Manzella, I., Cole, P. & Roverato, M. (2020) Grain size distribution and sedimentation in volcanic mass-wasting flows: Implications for propagation and mobility. *International Journal of Earth Sciences*, 109(8), 2679–2695. Available from: https://doi.org/10.1007/s00531-020-01907-8

Marc, O., Turowski, J.M. & Meunier, P. (2021) Controls on the grain size distribution of landslides in Taiwan: The influence of drop height, scar depth and bedrock strength. *Earth Surface Dynamics*, 9, 995–1011. Available from: https://doi.org/10.5194/esurf-9-995-2021

McKenna, J.P., Santi, P.M., Amblard, X. & Negri, J. (2012) Effects of soil-engineering properties on the failure mode of shallow landslides. *Landslides*, 9(2), 215–228. Available from: https://doi.org/10.1007/s10346-011-0295-3

Michaelides, K., Hollings, R., Singer, M.B., Nichols, M.H. & Nearing, M.A. (2018) Spatial and temporal analysis of hillslope–channel coupling and implications for the longitudinal profile in a dryland basin. *Earth Surface Processes and Landforms*, 43(8), 1608–1621. Available from: https://doi.org/10.1002/esp.4340

Neely, A.B. & DiBiase, R.A. (2020) Drainage area, bedrock fracture spacing, and weathering controls on landscape-scale patterns in surface sediment grain size. *Journal of Geophysical Research – Earth Surface*, 125(10), 1–22. Available from: https://doi.org/10.1029/2020JF005560

Neverman, A.J., Fuller, I.C., Procter, J.N. & Death, R.G. (2019) Terrestrial laser scanning and structure-from-motion photogrammetry concordance analysis for describing the surface layer of gravel beds. *Progress in Physical Geography*, 43(2), 1–22. Available from: https://doi.org/10.1177/0309133318822966

Prodger, S., Russell, P. & Davidson, M. (2017) Grain-size distributions on high-energy sandy beaches and their relation to wave dissipation. *Sedimentology*, 64(5), 1289–1302. Available from: https://doi.org/10.1111/sed.12353

Purinton, B. & Bookhagen, B. (2019) Introducing PebbleCounts: A grain-sizing tool for photo surveys of dynamic gravel-bed rivers. *Earth Surface Dynamics*, 7(3), 859–877. Available from: https://doi.org/10.5194/esurf-2019-20

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), 1–29. Available from: https://doi.org/10.1029/2021jf006260

Roda-Boluda, D.C., D'Arcy, M., McDonald, J. & Whittaker, A.C. (2018) Lithological controls on hillslope sediment supply: Insights from landslide activity and grain size distributions. *Earth Surface Processes and Landforms*, 43(5), 956–977. Available from: https://doi.org/10.1002/esp.4281

Rubin, D.M. (2004) A simple autocorrelation algorithm for determining grain size from digital images of sediment. *Journal of Sedimentary Research*, 74(1), 160–165. Available from: https://doi.org/10.1306/052203740160

Ruiz-Carulla, R., Corominas, J. & Mavrouli, O. (2015) A methodology to obtain the block size distribution of fragmental rockfall deposits. *Landslides*, 12(4), 815–825. Available from: https://doi.org/10.1007/s10346-015-0600-7

Saunders, G. (2014) Development of photogrammetric methods for landslide analysis. Available at https://www.duo.uio.no/bitstream/handle/10852/42338/saunders_thesis.pdf?sequence=1

Sklar, L.S. & Dietrich, W.E. (2006) The role of sediment in controlling steady-state bedrock channel slope: Implications of the saltation–abrasion incision model. *Geomorphology*, 82(1–2), 58–83. Available from: https://doi.org/10.1016/j.geomorph.2005.08.019

Sklar, L.S., Riebe, C.S., Genetti, J., Leclere, S. & Lukens, C.E. (2020) Downvalley fining of hillslope sediment in an alpine catchment: Implications for downstream fining of sediment flux in mountain rivers. *Earth Surface Processes and Landforms*, 45(8), 1828–1845. Available from: https://doi.org/10.1002/esp.4849

Sklar, L.S., Riebe, C.S., Marshall, J.A., Genetti, J., Leclere, S., Lukens, C.L. & Merces, V. (2017) The problem of predicting the size distribution of sediment supplied by hillslopes to rivers. *Geomorphology*, 277, 31–49. Available from: https://doi.org/10.1016/j.geomorph.2016.05.005

Sosio, R., Crosta, G.B. & Frattini, P. (2007) Field observations, rheological testing and numerical modelling of a debris-flow event. *Earth Surface Processes and Landforms*, 32(2), 290–306. Available from: https://doi.org/10.1002/esp

Storz-Peretz, Y. & Laronne, J.B. (2013) Automatic grain sizing of vertical exposures of gravelly deposits. *Sedimentary Geology*, 294, 13–26. Available from: https://doi.org/10.1016/j.sedgeo.2013.05.004

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

Vallance, J.W. & Savage, S.B. (2000) Particle segregation in granular flows down chutes. In: Rosato, A. & Blackmore, D. (Eds.) *International Union of Theoretical and Applied Mechanics Symposium on Segregation in Granular Flows*. Dordrecht: Springer, pp. 31–51.

Vallance, J.W. & Scott, K.M. (1997) The Osceola mudflow from Mount Rainier: Sedimentology and hazard implications of a huge clay-rich debris flow. *Bulletin of the Geological Society of America*, 109(2), 143–163. Available from: https://doi.org/10.1130/0016-7606(1997)109<0143:TOMFMR>2.3.CO;2

Vázquez-Tarrío, D., Borgniet, L., Liébault, F. & Recking, A. (2017) Using UAS optical imagery and SfM photogrammetry to characterize the surface grain size of gravel bars in a braided river (Vénéon River, French Alps). *Geomorphology*, 285, 94–105. Available from: https://doi.org/10.1016/j.geomorph.2017.01.039

Wang, G. & Sassa, K. (2003) Pore-pressure generation and movement of rainfall-induced landslides: Effects of grain size and fine-particle content. *Engineering Geology*, 69(1–2), 109–125. Available from: https://doi.org/10.1016/S0013-7952(02)00268-5

Westoby, M.J., Dunning, S.A., Woodward, J., Hein, A., Marrero, S.M., Winter, K. & Sugden, D.E. (2015) Sedimentological characterization of Antarctic moraines using UAVs and Structure-from-Motion photogrammetry. *Journal of Glaciology*, 61(230), 1088–1102. Available from: https://doi.org/10.3189/2015jog15j086

Whipple, K.X. & Dunne, T. (1992) The influence of debris-flow rheology on fan morphology, Owens Valley, California. *Geological Society of America Bulletin*, 104(7), 887–900. Available from: https://doi.org/10.1130/0016-7606(1992)104<0887:TIODFR>2.3.CO;2

Wohl, E.E., Anthony, D.J., Madsen, S.W. & Thompson, D.M. (1996) A comparison of surface sampling methods for coarse fluvial sediments. *Water Resources Research*, 32(10), 3219–3226. Available from: https://doi.org/10.1029/96WR01527

Wolman, M.G. (1954) A method of sampling coarse river-bed material. *Eos, Transactions of the American Geophysical Union*, 35(6), 951–956. Available from: https://doi.org/10.1029/TR035i006p00951

Zhang, L.M., Xu, Y., Huang, R.Q. & Chang, D.S. (2011) Particle flow and segregation in a giant landslide event triggered by the 2008 Wenchuan earthquake, Sichuan, China. *Natural Hazards and Earth System Sciences*, 11(4), 1153–1162. Available from: https://doi.org/10.5194/nhess-11-1153-2011

Zhang, S., Zhang, L.M. & Chen, H.X. (2014) Relationships among three repeated large-scale debris flows at Pubugou Ravine in the Wenchuan earthquake zone. *Canadian Geotechnical Journal*, 51(9), 951–965. Available from: https://doi.org/10.1139/cgj-2013-0368

Zhang, W., Wang, Q., Chen, J., Li, H., Que, J. & Kong, Y. (2015) Grain-size analysis of debris flow alluvial fans in Panxi Area along Jinsha River, China. *Sustainability (Switzerland)*, 7(11), 15219–15242. Available from: https://doi.org/10.3390/su71115219

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Harvey, E.L., Hales, T.C., Hobley, D.E.J., Liu, J. & Fan, X. (2022) Measuring the grain-size distributions of mass movement deposits. *Earth Surface Processes and Landforms*, 1–16. Available from: https://doi.org/10.1002/esp.5337