Controls on the grain size distribution of landslides in Taiwan: the influence of drop height, scar depth and bedrock strength.

Odin Marc¹, ², Jens M. Turowski², and Patrick Meunier³

¹G®osciences Environnement Toulouse (GET), UMR 5563, CNRS/IRD/CNES/UPS, Observatoire Midi-Pyrenees, 14 Avenue Edouard Belin, 31400 Toulouse, France.
²German Research Center for Geoscience, GFZ-Potsdam, Section 4.6, Geomorphology, Telegrafenberg, Potsdam, Germany
³Laboratoire de G®ologie, Ecole Normale Sup®rieure, 15 Rue Lhomond, Paris, France

Correspondence: Odin Marc, odin.marc@get.omp.eu

Abstract. The size of grains delivered to river by hillslopes processes is thought to be a key factor controlling sediment transport, long-term erosion as well as the information recorded in sedimentary archives. Recently, models have been developed for estimating the grain size distribution produced in soil, but they may not apply to active orogens where high erosion rates on hillslopes are driven by landsliding. Until now, relatively few studies have focused on landslide grain size distributions. Here we present grain size distribution (GSDs) obtained by grid-by-number sampling on 17 recent landslide deposits in Taiwan, and we compare it to the geometrical and physical properties of the landslides, such as their width, area, rock-type, drop height and estimated scar depth. All slides occurred in slightly metamorphosed sedimentary units, except two, which occurred in younger unmetamorphosed shales, with rock strength expected to be 3 to 10 times weaker from than their metamorphosed counterparts. We found that four landslides displayed a strong grain size segregation on their deposit with downslope toe deposits three to ten times coarser than apex deposits. In three cases, we could also measure the GSD inside the landslides that presented percentiles three to ten times finer than the surface of the deposit. Both observations could be due to either kinetic sieving or deposit reworking after the landslide failure, but we cannot explain why only some deposits had strong segregation. Averaging this spatial variability, we found the median grain size of the deposits to be strongly negatively correlated to drop height, scar width and depth. However, previous work suggests that regolith particles and bedrock blocks should coarsen with increasing depth, opposite to our observation. Accounting for a model of regolith coarsening with depth, we found that the ratio of the estimated original bedrock block size to the D50 was proportional to the potential energy of the landslide normalized to its bedrock strength. Thus the studied landslides agree well with a published, simple fragmentation model, even if that model was calibrated on much larger and much stronger rock avalanches with larger volume and stronger bedrock than those featured in our dataset. This scaling may thus serve for future modeling of grain size transfer from hillslopes to river, aiming to better understand landslide sediment evacuation and coupling to river erosional dynamics.


1 Introduction

Grain size is an essential parameter to understand sediment transport and associated issues in river evolution or hazard related to sediment pulses. For geomorphologists, it is increasingly considered an important parameter for the long-term incision of bedrock streams (Sklar and Dietrich, 2001; Cook et al., 2013, 2014; Turowski, 2018), while it is an essential part of the sedimentological signal which is ultimately archived in sedimentary structure the stratigraphy (e.g., Armitage et al., 2011).

Still, there are many processes that control the grain size distribution (GSD) in rivers, and they are poorly understood (Allen et al., 2015). In recent studies, models have been proposed describing how weathering in the critical zone reduces the original size distribution of bedrock before the grains reach the surface (Marshall and Sklar, 2012; Riebe et al., 2015; Sklar et al., 2017). However, in active orogens with high erosion rates (>0.5 mm/yr), landslides are likely the main providers of sediments to rivers (Hovius et al., 1997; Struck et al., 2015; Marc et al., 2019), and a large fraction of sediment may reach the river incompletely only partially weathered. Indeed, the limits of models predicting soil GSD and the need to account for GSD derived from fractured bedrock was recently shown (Neely and DiBiase, 2020), though the role of mass wasting in delivering and further fragmenting bedrock particles was not explored. In those settings, understanding and modeling the controls on landslide GSD should be an urgent goal, which has been addressed by few studies. Indeed, in contrast to river sediments for which many studies exist (e.g., Ibbeken, 1983; Whittaker et al., 2011; Chung and Chang, 2013; Guerit et al., 2014, 2018) landslide GSDs have rarely been measured, in part because the latter is considerably more difficult, time consuming and potentially dangerous than the former.

A few studies have measured and discussed in detail the GSD of some large historical landslide or rock avalanches, often putting forward to explain their data the various mechanism of rock fragmentation and grain segregation (see Crosta et al., 2007, and references therein). Although interesting for their discussion in terms of rock mechanics, such case studies did not allow to understand the regional variability of landslide GSDs, nor to derive physical scalings that could pave the way to model the GSD of material delivered to river networks by landslides. To our knowledge only seven studies reported detailed GSD measurements from multiple landslide deposits. A pioneering study reported the GSD from 42 landslide dams across the Appenines, with a discussion on the methods to derive the GSD but none on the controls of the GSD variability (Casagli et al., 2003). Locat et al. (2006) presented GSDs from nine large (> 100 M$^3$) rock avalanches from Canada and the Alps, including various rock types, were presented and analyzed in terms of potential energy and fragmentation theories (Locat et al., 2006). They found that the ratio of bedrock initial median block size, $D_i$ (estimated from fracture spacing), to the deposit median grain size, $D_{50}$, was proportional to the change in potential energy per unit of volume, $\rho g H$ (with $H$ the drop height of the center of mass, $\rho$ the rock density and $g$ the gravitational acceleration), normalized by the point-load strength of the bedrock, $\sigma_c$, measured with a point-load test performed on rock sample from the sites. Specifically their nine rock avalanches were best fit by a relation that could be recast as:

$$D_{50} = \frac{D_i}{k_1 \frac{\rho g H}{\sigma_c} - k_2}$$

(1)
With $k_2 = 0.5$ is an empirical threshold for fragmentation, and $k_1 = 83.3$ is an empirical coefficient relating to the conversion of potential energy into fragmentation energy and the effective breaking of particles. Thus, if the scaling is general for landslides, the deposit $D_{50}$ should increase with the source block size and the rock strength but decrease with the drop height of the center of mass, increase with $D_i$ and $\sigma_c$, but decrease $H$. However, subsequent studies often focused on the potential importance of landslide GSD to understand sediment transport dynamics and the expected GSD at the outlet of basins and reported GSDs in Nepal, Japan, California and Southern Italy (Attal and Lavé, 2006; Nishiguchi et al., 2012; Attal et al., 2015; Roda-Boluda et al., 2018). Several of them underlined, qualitatively, the factors influencing the GSDs such as the different lithological units (Attal and Lavé, 2006; Roda-Boluda et al., 2018), or the local hillslope gradient (Attal et al., 2015). Very recently, as a control on the time spent in the weathering engine (Attal et al., 2015), Recently a study presented the GSD of seven medium size rockfalls in Spain, showing that the bedrock block size and the deposit GSD could be related through a fractal fragmentation model (Ruiz-Carulla and Corominas, 2020). They found that potential energy was a main control on the fragmentation, but no clear correlation with rock strength measurements emerged. They did not compare their model and results to the simple scaling proposed by Locat et al. (2006). Thus, none of these more recent works has attempted to frame the landslide GSD in terms of the competition between fragmentation energy and source rock strength, and the scaling for large rock avalanches has not been reproduced on smaller, more common frequent landslides.

Based on these studies, we formulate two hypotheses. First, we suggest that Eq (1) could be generalized to landslides of intermediate size and depth, and thus that landslide $D_{50}$ should increase with rock strength, $\sigma_c$, and source materials median size $D_i$, but decrease with drop height, $H$. Second, we hypothesize that materials mobilized by shallow landslides coarsen with the landslide scar thickness, $T$, (i.e., $D_i$ increases with $T$), due to a reduction with depth of the fracture density (Clarke and Burbank, 2011) and/or of the degree of physical and chemical weathering experienced by particles (Cohen et al., 2010; Anderson et al., 2013; Sklar et al., 2017). Testing these hypotheses seems essential to pave the way towards geomorphic models accounting for the GSD of sediments transferred from hillslopes to river and from river and from rivers to sedimentary basins (Allen et al., 2015; Sklar et al., 2017).

With these goals, and given the sparse amount of data on landslide GSDs, we performed detailed measurements on 17 recent landslide deposits in Taiwan. Taiwan is a prime example of an active mountain belt where landslides are the main supplier of sediment to rivers (Hovius et al., 2000) and where reports of river GSDs exist in the literature (Chung and Chang, 2013; Lin et al., 2014). Still, to our knowledge, comprehensive landslide GSD measurements are still lacking in Taiwan. Below we report our measurements and discuss the source of variability of the GSD within given landslides and across the whole dataset. Then, we discuss the validity of the two hypothesis stated above based on the GSD of these landslides. We end by discussing the implications in terms of caveat and opportunity, caveats and opportunities for GSD sampling and implications for fluvial sediment transport.
2 Data and Methods

In this study we report original GSDs for 17 landslide deposits from Taiwan (Fig. 1), as well as basic landslide information that we use to discuss controls on the GSD (Table 1). We detailed below how we constrained landslide characteristics for each deposits and how we measured GSD and measured GSDs for each deposit.

5 2.1 Landslide characteristics

To quantify the variability in landslide GSD and its controls, we have targeted landslides with a known triggering date, and covering a broad range of sizes and areas (40 $m^2$ to 0.1 $km^2$) and lengths (10 to 400 m). Except for four small landslides ($<1000 m^2$), which were opportunistically sampled in parallel to larger neighbouring ones, all landslides were targeted based on satellite imagery and chosen for the accessibility of their deposit. Out of 17, 13 were chosen in the same geographic area and lithological unit, on both side of the southern section of Taiwan Central Range (Fig 1), allowing to have a variability in GSD independent of rock type. The four remaining landslides occurred in other lithological units. Specifically, LS-10 occurred in the emergent topography of Taiwan southern tip, LS-15 and LS-1 on both side of the northern part of the Central Range, and LS-16 in the Northwestern foothills.

Landslide type was difficult to assess, but all landslide most landslides could be called disrupted landslides, involving a mixture debris avalanches (Varnes, 1978), involving variable amount of regolith and bedrock. The only exception is, though LS-13 and LS-14 could also be called rock falls, LS-12, the largest event, may rather be a deep slump, with moderate displacement and probably some rotational components, partly translational, partly rotational. Most landslides correspond to landslide polygons present in the Typhoon Morakot landslide inventory (Marc et al., 2018), and thus occurred in August 2009, about five and a half years before they were surveyed in March 2015. Other more recent landslides were dated based on the time-series of images available in Google Earth (see Table 1). Based on the geological map of Taiwan (available from Taiwan Central Geological Survey), we assign each landslides to a geological formation and the associated rock type (Fig 1).

To assess variability in GSD independent of rock type, 13 out of 17 landslides were chosen in the same geographic area, on both sides of the southern section of Taiwan Central Range, in relatively homogeneous lithological units composed of slate and slightly metamorphosed sandstone (Fig. 1, Table 1). Most landslides (13) are in the Lushan formation, composed of slate and slightly metamorphosed sandstone. Two of the four remaining landslides are also in metasedimentary units, LS-1 LS-1 and LS-15 also occurred on moderately metamorphosed units, on both sides of the northern part of the Central Range, in black schist and LS-15 in metasandstone intercalated with slate. Then, the two last landslides occurred in non metamorphosed units, respectively. The two remaining landslides both occurred in unmetamorphosed units, made of sandstone and shale alternance for LS-10 is in the Nanchuang formation, alternating sandstones and shales, and LS-16 in the Cholan/Chinshui shale which occurred in the emergent topography of Taiwan’s southern tip, and in shales of the Northwestern foothills for LS-16. In LS-16 the high abundance of clay made weak agglomerate that could be broken with hands, many coarse rock fragments ($>10 cm$) were crumbling when touched, highlighting the weakness of this unit rock compared to the other landslide units.

In an effort to associate these formations to quantitative strength estimates, constrain the mechanical strength of these units,
we refer to measurements reported for 128 samples from the Chenyoulan catchment, both for the Nanchuang formation and the metasediments /Nankang formation, which continues until Taiwan southern tip and contains LS-10, and for the metasedimentary units of the Shipachungshi formation, where LS-15 occurred (Lin et al., 2008). The unconfined compressive strength of Nanchuang sandstone ranged from 29 to 117 MPa (mean of 70 MPa), while the metasedimentary units Shipachungshi metasedimentary sandstone ranged from 45 to 179 MPa (mean of 100 MPa) (Lin et al., 2008). However, shales are about equally represented than sandstone in the Nanchuang formation, with sandstones are alternating with weak shale (strength below 10 MPa, while the slates intercalated with the metasedimentary are more irregular and presumably stronger), with an equal proportion of each. In contrast, the Shipachungshi metasedimentary sandstone are intercalated with less frequent slates, often stronger than the Nanchuang shales (Lin et al., 2008). These measurements clearly make the case for highly variable rock strength and are far from encompassing the potential diversity of rock-type sampled by the studied landslides. Still to first order, we expect our goal here is not primarily to constrain the rock strength of individual landslides, but to estimate the relative strength of diverse units. Based on the measurements reported above we make two assumptions. First, that the shales and sediments hosting LS-16 to be very weak, with strength about 10 MPa, LS-10 to have strength around 30 MPa (with large uncertainties depending if strength is rather dominated by the shale or the sandstone), and the rest of the landslide from metasedimentary units, to be stronger than around 100 MPa. The point load strength, \( \sigma_c \), is typically 15 to 25-fold smaller than the unconfined compressive strength (Chau and Wong, 1996) and therefore we consider \( \sigma_c \sim 5 \text{ MPa} \) for the landslides in metasedimentary units and \( \sigma_c \sim 1.5 \text{ MPa} \) and \( \sigma_c \sim 0.5 \text{ MPa} \), for LS-10 and LS16 respectively. LS-10 may be 7 to 13 times and 2 to 4 times weaker, respectively, than the metasediments hosting LS-15. Second, that the slates and metasediments in the Lushan formation have similar strength than the ones in which LS-15, as well as LS-1, occurred and thus that these landslides can be compared without normalizing for strength.

Geometric landslide metrics were obtained from high resolution satellite imagery available in Google Earth except for four deposits (LS-4, 9n, 13 and 14) deposits, which were too small to be clearly distinguished on the imagery, and which had their dimensions approximated from field observations only, using a laser ranger. Area was obtained by hand mapping the whole disturbed zone on the imagery. Even with high resolution imagery, the transition from scar area to transport channel or from scar area to deposit area could not be clearly assessed for most slides. Length refers to the downslope length between the highest and lowest point of the polygon. The elevation difference between these two points, estimated from the elevation data of Google Earth (in Taiwan mostly from SRTM, predating all of the studied landslides), defined the maximum drop height. The scar width, \( W_{\text{sc}} \), was obtained by measuring the extent of the landslide in the direction orthogonal to flow, in the upper part of the failure only. A smaller drop height and length were obtained when considering the location of the estimated center of mass of the scar and deposits, estimated in Google Earth based on the imagery. An estimate of \( \text{slope-pre-failure scar gradient} \) could be derived from the scar’s approximate length and height difference. For all landslides we estimated the upper, respectively lower, volume of the landslide.

We estimated a maximum and minimum landslide volume using empirical scaling relationships with for the scar area \( A_{\text{sc}} \), of the form \( V = \alpha A_{\text{sc}}^\gamma \), with different parameters values assuming the scar was mobilizing bedrock or soil, respectively, (Larsen et al., 2010) soil or bedrock, respectively. We used \( \gamma = 1.262 \pm 0.009 \) and \( \log 10(\alpha) = 0.649 \pm 0.021 \) for the soil assumption landslides.
in soil, and $\gamma = 1.41 \pm 0.02$ and $\log_{10}(\alpha) = 0.63 \pm 0.06$ for the bedrock assumption in landslides in bedrock (Larsen et al., 2010). Then we derived the upper and lower estimates of landslide mean scar depth as the ratio of volume to scar area for bedrock and soil, respectively, and scar area. For this, we estimated scar areas from scar width, as $A_s = 1.5W_s^2$, consistent with global landslide geometry database (Domej et al., 2017), except for a few landslides where direct observations (Table 1). Using this scaling is necessary because in most cases, field or satellite observations did not allow to determine the ends of the scar dimensions did not match this simple scaling (Table 1) area and the transition to runout or deposition areas. Although approximate, this scaling is still preferable to using total landslide area, as it removes bias in volume estimation associated with variable runout length (e.g., Marc et al., 2019). Last, for a few landslides the deposit volume could be approximated as a fraction (a quarter to a half) of a cone, for which a volume estimate could be obtained as $\pi R^2 h/3$ with $R$ and $h$ the approximate radius and height of the cone which were estimated in the field. This simplified geometry was only suitable for LS-3, 4, 7, 9n, 10, 11, 13 and 15, and yield only a 1st order volume first order "field volume" estimate (Table 1). Nevertheless, these field estimates mostly fall within the bracket of the volumes estimated from global scaling relationships (Fig. S1), lending some support to this approach. Also note that:

Still, for the deposit where we could not obtain a field estimate, better constrained volume estimate could be obtained by choosing one of the two scaling relationships. We note that the field volume estimate of LS-3 and LS-4, in which fresh vegetation debris was still present and a yellowish rock colour is similar to the estimate from the soil scaling relationship (Fig. S1). This is consistent with the observation that they were composed of rock debris with a yellowish color that indicated advanced weathering (See Fig. and contained fresh vegetation debris (see Fig. ). They had their field estimate similar to the soil scaling while for LS-7 and LS-11, which were clearly involving mostly fresh bedrock, the field volume matches estimates match better the bedrock scaling. Thus, to better capture the variability of landslide scar depth for the landslides without field estimate, we either applied the bedrock scaling or the averaged the soil and bedrock scaling, depending on the landslide characteristic. We relationship (Fig. S1). Thus, where field volumes were lacking we used the bedrock estimate for the largest landslides ($W_s > 50$) within which the rock looked mostly fresh (i.e., LS-1, 2, 8, 12). For the other, smaller Some other landslides (LS-5, 6, 9o, ) and for LS 16, we averaged the bedrock and soil thickness estimates. These estimates (together with the ratio of field volume and scar area) are considered-16), featured a mixture of soil and rock material. Consequently, we used the average of the soil and bedrock scaling to estimate their volume. The best estimate for the scar each landslide was divided by its scar area to obtain an estimate of scar thickness (Table 1).

When we give a value with uncertainties on correlation index (Pearson’s $R$ or $R^2$) they always refer to the mean and standard deviation from 10,000 random bootstrapped subsamples of the considered sample.

2.2 Grain size counting

GSDs were obtained using the grid-by-number sampling, following established protocols developed for measuring riverine GSDs (see, Kellerhals and Bray, 1971) and subsequently applied to landslide deposits (Casagli et al., 2003; Attal and Lavé, 2006). We extended survey tapes along an elevation contour over a substantial portion of the deposit width (10 to 50 m)
and sampled grains along the tape at a constant interval, recording the size bin of b-axis measured with rulers. We used bins following a half Phi scale (power of 2 by 0.5 increments) with the smallest bins encompassing all grains finer than 2 mm. When grains could not be moved, we considered the smallest of the two visible axis as the b-axis. The grid step was 0.5 m in most cases but was adjusted to 1 meter for deposits where many boulder scale boulder where present (LS-2s, LS-13, LS-14) to limit the number of time when a grain would be counted. Then we moved the line in parallel, upslope by one to a few meters depending on the deposit’s dimensions and local topography, and repeated the counting. Most slides where sampled with 6-10 survey lines allowing to cover a substantial fraction of the deposit (often 30 to 60%), with total counts often greater than 200–400 typically including several individual grains. This approach also allowed to sample different section of the deposit when a spatial segregation was visible (LS-3, LS-8, LS-9n, LS-10), and to quantitatively assess this spatial variability in grain size (see, Ruiz-Carulla et al., 2015). For two landslides we could separately count grains from the surface and from the interior of the deposit, counting grains on the vertical banks of a deep erosional gully incising the deposit of LS-8, and counting grains on a debris fan next and below the road that had been cleared from the deposit of LS-2. Thus the former case allowed us to survey the internal GSD in place, while the latter more likely represents a remixing from surface and internal part, but which must, by volume and number be more representative of the deposits, and thus must be closer to the inner GSD than what would be derived from surface measurements only. Note that for LS-2 the only undisturbed deposit was the one in the transport channel, where a carapace (a layer of very coarse grains, (Crosta et al., 2007)) seems to have formed (Fig. S1). Finally, on LS-5 we measured separately a debris fan and the terminal section of a channelized deposit which was visibly coarser (Fig. S2 and S3). In this case, given the age of the deposit and its direct contact with the floodplain, it is plausible that the deposit was partly eroded and the fan may be a mixture of internal and superficial material, whereas the higher up channel section may be more representative of the original surface of the deposit. This will be further discussed when talking of segregation. Additionally, on the deposit of LS-7 we distinguished by visual inspection grains made of slate, which were dark, elongated and without visible internal structure, from grains made of metasandstone which were lighter, more cubic and with visible internal grains. We have counted them separately as we found them over the deposit. In many other deposits a large majority of grains either looked fairly homogeneous, or because of lack of time their systematic counting within different rock type could not be done.

3 Results

3.1 Landslide grain size distributions and their internal variability

The landslide GSDs have the 50th and 84th percentiles ranging from ~15 mm to ~200 mm and ~60 mm to ~600 mm, respectively. This is consistent with the range of observations from previous studies, except the large rock avalanches from Locat et al. (2006) and the volcanic rock avalanches Crosta et al. (2007), which were about 10 times coarser and finer, respectively, than all other studies. LS-2s and LS-16 are much coarser and finer

7
Figure 1. (a) Hillshaded elevation map of Taiwan, with the main lithological units of the central range (based on the geological map of Taiwan available from Taiwan Central Geological Survey) and the locations of the 17 sampled landslides deposits. Note that the limits between young shales (e.g., Chinshui Shale) and quaternary deposits in the west are not color coded. (b–e) Pictures of some sampled landslides, where the yellow line is the approximate contour of the landslide (sometimes going beyond the pictures), and the dashed line indicates the transition from deposit to scar (it is only tentative when associated with “??”). In (b) and (e) the lead author is standing on the deposit for scale.

than the rest of the studied landslides, respectively. Interquartile ratios vary between 3 and 15, but we note that 13 out of 19 GSD have an interquartile ratio of 3 to 6, while only LS-1, 3, 5,–5i, 5s, 8i, 15 and LS-16 have larger spreads (Fig. 2). All distributions seem unimodal, except LS-16 with more than 40% of the grains finer than 2 mm, likely containing a second, sub-millimetric mode that could not be constrained by our methods. Grain size distributions can often be well described by
Table 1. Landslide characteristics for the 17 surveyed deposits. Asterisks indicate landslides for which the geometry was estimated in the field rather than from satellite imagery. Due to its complex displacement LS-12 has large uncertainties on the displacement of its center of mass.

<table>
<thead>
<tr>
<th>Landslide</th>
<th>Rock type</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS-1</td>
<td>Bsc</td>
</tr>
<tr>
<td>LS-2</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-3</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-4</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-5</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-6</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-7</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-8</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-9</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-10</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-11</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-12</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-13</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-14</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-15</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-16</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-17</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-18</td>
<td>Sl/Sd</td>
</tr>
<tr>
<td>LS-19</td>
<td>Sl/Sd</td>
</tr>
</tbody>
</table>

A Weibull or Lognormal distribution (Ibbeken, 1983). For the studied landslides, eight GSDs are better fit (according to both Kolmogorov-Smirnov and Anderson-Darling statistics, (Stephens, 1974)) by a Weibull distribution (LS 2s, 3, 4, 5s, 5i, 6, 9n, 9o, Fig. 2B, S4), while all others are better fit by a log-normal distribution (Fig 2A). Given these two subgroups we refrained from using parameters associated with 2A, S5. Note that LS-16 is poorly fit by both distributions. These two subgroups imply that we cannot prescribe a given distribution form to model landslide GSDs, and future work may aim at understanding why some landslide GSDs obey one or the other distribution. In any case, we refrain from using distribution parameters and will continue to discuss results based on empirical descriptors (i.e., median, interquartile ratio).

GSDs within a single landslide deposit were often heterogeneous, in one case associated with differences between grains of different rock types (slate and sandstone in LS-7), while in seven other cases associated to spatial variability (Fig. 3, S2S6). For LS-7, the slate pieces have grain sizes about three times smaller than the sandstone for a given quantile of the CDFGSD, with a similar distribution shape. The slate grains were typically elongated platelets (i.e., \(a \sim 3b\) and \(b >> c\)), while the sandstone grains were cubic and slightly more abundant than the slate grains (N=196 vs N=167). We observed by with the naked-eye downslope segregation, i.e., an increase in sediment coarseness from the apex to the toe of the deposits (Ruiz-Carulla et al., 2015, for in four cases. The strongest segregation occurred in deposits LS-3 and LS-9n, where the upper part of the deposits have grains 5-10 times finer than the lower part of the deposit, without changing substantially the shape of the distribution. Deposits LS-8 and LS-10 exhibited a more subtle segregation with the upper part of the deposits having distributions finer by a factor 1.5-2 in comparison to the toe of the deposits (FigS2, S6). The toe of LS-10 also displays \(D_{50}\) and \(D_{84}\) twice coarser than at its apex, consistent with other cases, but also has more fine grains, with about 10% of grains finer than 2 mm against less than 5% at the apex. The upper, middle (when differentiated) and lower sections of the deposits...
Figure 2. (a) range of $D_{50}$ and $D_{84}$ for various studies, excluding the coarsest and finest distributions of each study. (b) and Cumulative (c) cumulative distribution function for the 16-17 sampled landslide deposits. For visibility, panels (b) and (c) Note that we show surface and inner distributions best fit by a log-normal distribution for LS-5, LS-8 and a Weibull distribution LS-2, respectively. Note that LS-16 is poorly fit by both and thus 20 distributions in total. For visibility some lines are dashed and the distribution are shown in two panels. Vertical lines are approximate boundary for grain transport by suspension and bedload, for a flood associated with fluid shear stress of 220 Pa (see section Implications for sediment transport).

represented roughly similar proportion of the surface of the deposits, and we obtained count variations below 10% from the different subsections (Fig. 3). Therefore, to study the variability between various landslides we obtained an overall GSD by summing the grain counts from different sub areas of the deposits with spatial segregation.

Then, in two cases we could separately measure the superficial and internal GSD. For LS-8 we observed that the superficial GSD had $D_{16} = 20 \text{ mm}$, $D_{50} = 40 \text{ mm}$ and $D_{84} = 120 \text{ mm}$, $D_{16} = 20 \text{ mm}$, $D_{50} = 40 \text{ mm}$ and $D_{84} = 120 \text{ mm}$ while the internal GSD had $D_{16} = 3 \text{ mm}$, $D_{50} = 10 \text{ mm}$ and $D_{84} = 50 \text{ mm}$, $D_{16} = 3 \text{ mm}$, $D_{50} = 10 \text{ mm}$ and $D_{84} = 50 \text{ mm}$.
Figure 3. Three examples of three types of heterogeneity in grain size distributions within the same landslide deposit. (a) Lithological difference within the deposit of LS-7 with sandstone grains coarser than slate grains. (b) Downslope differences between the upper, middle and lower part of the deposits of LS-3 and LS-9. (c) Difference between the surface and inner part of the deposits of LS-2, LS-5 and LS-8.

There the superficial deposits had almost no fine sediment below 2 mm, whereas the internal body had more than 10% of fine sediments. Thus, the internal GSD had quantiles 10 to 20 times finer than the channel carapace, the largest difference observed in terms of internal variability. Note that the carapace had also a coarser GSD than any other measured landslide deposit in our study. In spite of this massive difference, we note that the internal GSD still had only about 3% grains finer than 2 mm. These two examples clearly show that the superficial GSD can be substantially different from the internal GSD, both in terms of fine grains (<2 mm) but also for coarse to very coarse grains (10 to 100 mm).

Last, in the case of LS-5, it is not entirely clear if the two distributions represent vertical segregation or superficial spatial variability. Given its similar geometry (with a coarser channel than can be seen above, Fig. S4) with LS-2, and given that the fan has a $D_{16} = 4$ mm and a $D_{50} = 20$ mm; $D_{16} = 4$ mm and a $D_{50} = 20$ mm, about three times finer than in the channel, but an almost identical $D_{84}$ of around 200 mm, we consider it to likely be an internal or mixed GSD.

3.2 Relations with landslide properties

The percentiles of the GSDs are highly correlated with linear correlation coefficients $R^2 > 0.9$ between $D_{50}$ and $D_{16}$, $D_{25}$, $D_{75}$, $D_{84}$ and $D_{90}$. However, two other metrics are poorly correlated to the (Fig. S7). We note that more scatter is present for $D_{16}$ and $D_{90}$, suggesting that $D_{50}$, the fraction of fine grains (here considering grains < 8 mm) and the is a good proxy for the bulk of the distribution, but does not completely capture the variability in their tails. The interquartile ratio (here $D_{75}/D_{25}$)
characterizing which characterizes the span of grain size in the distribution, is also independent of the other percentiles (Fig. S8).

For the two landslides for which we have both an internal and superficial grain counting, we only considered one GSD. For LS-8 we considered the superficial one GSD at the surface, to be consistent with all other cases. For LS-5 we considered the coarser distribution from the channel as more representative of the surface deposit. If the fan of LS-5 is representative of the deposit and if the channel is coarser because of some spatial segregation, the percentiles of its GSD would be about two to three times finer than the ones we have selected (Fig. 3). In contrast for LS-2 we considered the internal GSD, because the superficial measurement recorded only what seems to be a carapace over-representing that over-represents coarse grains. For Indeed, with about half of the distribution made of boulders (> 0.5 m) we consider that segregation to be exceptional within our dataset.

Thus, in using LS-2i to study inter-event variability we assume that none of the other landslides had a carapace with similar strong sorting. Nevertheless, LS-2i percentiles may be biased towards finer grains, when compared to surface deposits of the other landslides, and we assume the bias could be up to a factor of two to three, based on LS-5 we considered the coarser distribution from the channel as more representative of the surface deposit. In case this would be due to spatial segregation and the fan distribution would actually represent the average surface the percentiles would be about twice finer, while the opposite is true for LS-2i which may have had a deposit 2-3 times coarser, and LS-8 (Fig. 3).

The drop height, $H$ has the highest, negative, correlation with $D_{50}$ ($R^2 = 0.71 \pm 0.13; N = 17$ with both terms log-transformed). Figure 4: According to Eq (1), with a best fit and assuming similar initial regolith material for all landslides, $D_{50}$ should decrease linearly with the ratio of drop height, $H$, to bedrock strength, $\sigma_b$. Indeed, log-transforming and fitting $D_{50}$ against $H$ for the 15 deposits from the metasedimentary units, which we expect to have relatively similar strength, we obtain $R^2 = 0.71$ and a power-law exponent -0.78. This correlation and exponent are obtained when rescaling the drop height of -0.64 (Fig. 4). Including LS-10 and LS-16 from weaker units yields a substantially poorer fit ($R^2 = 0.31$). Thus, to account for their weaker strength, we rescaled these two landslide drop heights by a factor 3.3 of 3 and 10, given their strength is 3.3 to 10 times smaller than for the metasedimentary units where the other landslides occurred. If instead we only fit the 15 deposit from the metamorphosed units we obtain $R^2 = 0.71$ and a respectively, at the center of values estimated from strength measurements (see methods). Then, we obtained correlation coefficient ($R^2 = 0.71; N = 17$) and a best fit power-law exponent of -0.64. In metamorphosed units, landslide -0. For landslides in metasedimentary units, $D_{50}$ is also negatively correlated (with a larger scatter $R^2 = 0.5 - 0.55$) to landslide size metrics (area, width, volume, depth). However, we note that, for this dataset, these metrics are also strongly correlated with the drop height ($R^2 = 0.56 - 0.66$, Fig. S4). Therefore, these negative correlations between landslide size and GSD may not be causal, especially given that one may expect S9). Given that we would expect larger and deeper landslides to mobilize fresher and coarser grains. We discuss in the next section a simple model for the landslide $D_{50}$ combining a fragmentation scaling (Loo et al., 2006) and a model for regolith coarsening with depth (Cohen et al., 2010), and thus to have positive correlation with $D_{50}$, these negative correlations may simply reflect a sample bias of our dataset in which larger landslides have a larger drop height.

For the spread of the distribution, characterized by the interquartile ratio, we did not find any substantial correlation with any
Figure 4. $D_{50}$ for the 17 landslides (colorcoded by scar thickness) of this study against the drop height of their center of mass. Results are similar for maximal drop height. The best-fit (solid black line) and $R^2$ only consider the drop height rescaled by the landslide rock mass strength (black circles), about three and ten times weaker for LS-10 and LS-16 respectively (see methods). The red circles show the original drop height for these two landslides. Vertical bar show a factor of two uncertainty for LS-5 and LS-2i, for which there may be vertical segregation (see text).

Horizontal error bar represent uncertainties on the drop height of LS-12 and the strength normalization of LS-10 and LS-16 (see methods).

of the landslide variable that we have constrained variables. Even the rock type rock type (or rock strength) does not seem to have an impact on the GSD spread, with several landslide landslides in metasedimentary rocks with very large spreads (LS-3, LS-5, LS-1) and while the two landslides in non-metamorphosed units are on both ends of the spectrum. However, we note that, if the spread is almost uncorrelated to the $D_{50}$ or $D_{84}$, it is substantially negatively correlated to the logarithm of $D_{16}$ ($R^2 = 0.33 \pm 0.17; N = 17$) or in other words to the amount of fine grains (i.e., < 8 mm). Thus, more data is needed to understand the spread of the landslide GSDs.

4 Discussion

A general limit of our study is that In the following discussion, we propose that the variability of landslide $D_{50}$ can be reconciled with the fragmentation scaling of Locat et al. (2006) when accounting for regolith coarsening with depth (e.g., Cohen et al., 2010). Then we detail processes that may lead to grain size segregation within a given deposit and practical implications for sampling.
Last we explore the implication of the measured GSD on sediment transport and evacuation from river channels. Before turning to these points, we recall that some variable important to model landslide \( D_{50} \) (section 4.1) are only first order estimate, if available at all. First, we rely on rough estimates of the landslide geometry (drop height, volume, depth, ...). Better characterization in the field would have required more field work, and would have remained impossible for many landslides where the scar could not be accessed limited by accessibility, or elaborate construction of DEM based on LIDAR or drone photogrammetry, difficult to perform and limited by the lack of accurate pre-failure DEMs. Thus for the sake of this study we think that having a homogeneous, first-order estimates of these metrics is sufficient to test the dependence of the \( D_{50} / D_{50} \) on landslide geometry. The difficulty to access many scars, as well as the mixed origin (i.e., weathered regolith and bedrock) of several landslide sources also meant that in practice we could not measure the source materials median grain size, \( D_s \).

**4.1 The importance of fragmentation and source material initial grain size**

Here we discuss the hypothesis that Eq (1), proposed and validated by Locat et al. (2006) for large rock avalanches, can also be used for smaller, shallower landslides made of a mixture of regolith and bedrock. For the 17 Taiwanese landslide in our study we found that within a given lithology, drop height seems to be a first order control on the landslide deposit median grain size (Fig. 4). We also found that by rescaling the drop height by their weaker rock strength, LS-10 and LS-16 were consistent with the trend defined by the stronger metamorphosed units. These observations qualitatively agree with Eq (1), but quantitatively, the best fit between \( H \) and \( D_{50} \) was not linear, but a sub-linear power-law. Given that we observe that \( H \) and the landslide scar thickness, \( T \), are correlated in our surveyed landslide (Figs S4, S9) this discrepancy with Eq (1) could be resolved if \( D_s \), which we could not measure, is increasing with \( T \). Models describing the size of particles in a soil or regolith predict upwards fining of grains from the bedrock to the surface due to an increase in the degree of both physical and chemical weathering (Cohen et al., 2010; Anderson et al., 2013; Sklar et al., 2017). In bedrock, fracture density estimated from seismic wave refraction was also found to decrease non-linearly from the surface to a depth of 5-10m (Clarke and Burbank, 2011). Given soils are often thin in Taiwan, and represented a small proportion of the mobilized material, we consider physical weathering is likely dominant and its intensity. Here we consider that the physical weathering rate (i.e., the rate of particle breakdown) can be modelled with an exponential decay from the surface, with a characteristic length scale of \( \lambda = 2 \text{ m} \), consistent with previous modeling works (Cohen et al., 2010; Anderson et al., 2013). Assuming the regolith grain size at a depth \( z \), \( D_r(z) \), to be proportional to the integral of the weathering intensity, we modelled as \( D_r(1 - \exp(-z/\lambda)) \) where \( z \) is the depth and \( D_r \) the rate, we modeled it as \( D_r = D_b(1 - \exp(-z/\lambda)) \) where \( D_b \) is the unweathered bedrock block size, producing a rapid variation near the surface consistent with published models for physical weathering (Cohen et al., 2010; Anderson et al., 2013). Imposing \( D_r = 1000 \text{ mm} \), we obtained \( D_r \) for a given landslide by integrating this profile until. Averaging \( D_r \) from 0 to \( T \), the mean scar thickness of the and choosing a value for \( D_b \) we can obtain \( D_s = D_b(1 - \lambda / T(1 - \exp(-T/\lambda))) \) for each landslide (Fig. 5a).

To compare with the prediction of Locat et al. (2006), we have to assume a value for \( D_b \) and use \( \sigma_d \) values of 0.5 MPa for LS-16, 1.5 MPa for LS-10 and 5 MPa for the other landslides in metasediments. These point load measurements...
are consistent with the typical unconfined compressive strength of intact rock (Lin et al., 2008) after dividing them by 20 (Chau and Wong, 1996). With these strength values and \( D_b = 1900 \text{ mm} \), the ratio of the modeled \( D_i \) and to the measured \( D_{50} \) agree with Eq (1) \((R^2 = 0.9, 0.95, R^2 = 0.9)\), even though its coefficients \((k_1, k_2)\) were calibrated on rock avalanches with rock strength and median grain size order of magnitude larger than the ones from this study (Table 1, Fig. 2a, 5b). All of the surveyed landslides are within a factor of two from the predictions of Eq (1), even LS-12 which likely had a different deformation style than the other landslides. **Note that we chose this good agreement is not very sensitive to how we estimate landslide volume and scar thickness (Fig. S10) Note that by assuming \( D_b = 1900 \text{ mm} \) to match we have matched the \( D_{50} \) measurements, which represent surface deposits. These surface deposits, which may slightly over-estimate the representative grain size relative to the whole landslide deposit, because of kinetic sieving or fine removal by surface runoff (Fig. 3C and discussion below). Based on LS-5 and 8, the inner \( D_{50} \) may be 2-3 times finer than their surface counterpart, and thus field measurement of regolith and bedrock GSD may need to be compared to a model with \( D_b \sim 600-1000 \text{ mm} \) to which uncertainty on rock strength (i.e., probably a factor of two) should also be added. Still, although uncertain, these latter values are intermediate between fracture measurement on surface outcrop in the US (Neely et al., 2019; Verdian et al., 2020) ranging from 10 to 400 mm, and the bedrock block size measured by Locat et al. (2006) on the scar of large rock avalanche, ranging from 600 to more than 10,000 mm. \( D_b \sim 600-1000 \text{ mm} \) also matches quite well the range of \( D_{50} \) to \( D_{75} \) (450 to 900 mm).
found in the carapace of LS-2 which may be a fair first order estimate of the original bedrock block size (Crosta et al., 2007). We also note that LS-10 and 16 which occurred in weaker bedrock may be expected to have a finer $D_b$ than the other slides. We are not able to constrain this but note that even with a $D_b$ three times finer, these two slides would be only a factor of 2 below the prediction ($D_{50}$ predicted) (Fig. 5).

We conclude this section by underlining that more measurements, especially of source rock block size and strength, are needed to fully demonstrate the applicability of the fragmentation theory presented by Locat et al. (2006). Still, we suggest that such fragmentation theory is applicable to understand and predict landslide GSD in a wide range of contexts, at least for rock, soil and mixed avalanches and generally disrupted slides, which are the most commonly triggered (Keefer, 1984). Further, our observations suggest that Eq. (1) can be generalized to account for an exponential reduction of regolith grain size towards the surface (Cohen et al., 2010; Anderson et al., 2013), yielding:

$$D_{50} = \left(1 - \frac{\lambda}{T} (1 - e^{T/\lambda - T/\lambda})\right) \frac{D_b}{k_1 \rho g H k_2}$$

where $D_i$ has been replaced by a term depending on the "fresh" bedrock median size, $D_b$, the length scale of weathering decay, $\lambda$, and the landslide thickness, $T$. In a sense, Eq. (2) supports previous qualitative statements on the importance of rock type (Attal and Lavé, 2006; Roda-Boluda et al., 2018) rock type (Attal and Lavé, 2006; Roda-Boluda et al., 2018), which may physically relate to rock strength and regolith block size. Additionally, Eq. (2) combines the concept of physical weathering with the process of fragmentation, controlled by drop height, less often considered in the geomorphological community. Future studies may also try to clarify whether $D_b$ is controlled by $\sigma_c$, which would make the equation more non-linear but reduce the number of parameters to constrain. More complex models of fragmentation have been used to predict landslide GSD (De Blasio and Crosta, 2014; Ruiz-Carulla and Corominas, 2020), and may be better suited to model the full GSD, but Eq. (2), provided it is further validated, opens various interesting perspectives. For example, it suggests that seismically triggered landslides, which occur more often near ridges than rainfall-triggered landslides (see Meunier et al., 2008; Rault et al., 2019), are more likely to deliver finer grains to the river systems, assuming they have a similar size and depth distributions. Eq. More generally it highlights the need for an investigation on how geomorphic factors (e.g., hillslope height, steepness, shape) modulate landslide runout, height drop, and connectivity to channels. Comparing hillslopes in various landscapes could be easily attempted based on comprehensive landslide inventories (Tanyaş et al., 2017; Marc et al., 2018). Eq. (2) would also be well suited for landscape scale modeling of the input of various grain size into rivers (e.g., Benda and Dunne, 1997; Carretier et al., 2016; Neely and DiBiase, 2020), and thus, to better couple landslide and river dynamics in landscape evolution models (Campforts et al., 2020; Egholm et al., 2013).

4.2 Controls on the internal variability of the GSD and implications for future sampling

We found three sources of internal variability of landslide GSD, one associated with the lithology of the individual grains, as reported for Himalayan landslides by Attal and Lavé (2006), and two related to the location of the grains on or in the deposit, as reported for various rock avalanches (Crosta et al., 2007; Ruiz-Carulla et al., 2015). We discuss these observations first in terms of implications for bias and sampling procedure, and second in terms of physical process causing them.
The lithological difference is not likely to be a bias as long as the grains of different lithologies are randomly distributed in the deposit; their sampling frequency should represent their relative abundance in the deposit. Spatial segregation on the surface of the deposit implies that to ensure a representative GSD, the sampling method should be performed ideally across most of the deposit, or at least over the different subunits of the deposit, before doing a weighted average with their relative area of contribution. Measurement based on sieving at a single site or local grain counts along a line or over an area may misrepresent the GSD and should be avoided. For large deposits where access is difficult, the use of pictures from a drone may help to check for segregation and potentially allow to reproduce the grid-by-number counting method using image analysis. However, this requires to scale each drone picture, and thus to deploy reference objects across the deposit which is not always practical, not counting the fact such sampling will be unable to resolve fine grains (< 30 – 100 mm). In contrast, in the presence of a vertical segregation, where superficial and inner GSDs differ, it may be very difficult to estimate a GSD that is representative for the whole deposit. When possible, targeting the banks of incised gullies may offer a good opportunity to characterize the subsurface of the deposit. Some applications mainly require the subsurface GSD, for example modeling the weathering of freshly fragmented bedrock in the landslide deposit and how they can contribute to solute fluxes (Emberson et al., 2016a, b). In contrast, the surface grains matter for sediment transport, and armoring may limit the mobilization of deeper finer grains. Additionally, in the case of a carapace, the question of how to combine the two end-member distribution would require an estimate of the relative thickness of the two end-member GSD, which may be challenging. In the case of a less extreme segregation, as observed for LS-8 and probably LS-5, the proportion of coarse grains (> 200 mm) was similar on the surface and inside the deposit, and only the medium and especially fine grains were more abundant inside the deposit.

The process of kinetic sieving (Savage and Lun, 1988; Gray, 2018) is expected to cause vertical segregation (i.e., a coarser surface and finer subsurface) in granular flows, and a downslope segregation when shear is present, leading to boulder fronts as for LS-3. However, it should be noted that the segregation is favored by transport along moderate slope gradient and tends to disappear for very steep chutes (Vallance and Savage, 2000). Although our gradient estimates are very rough, steep slopes (Table 1, Fig. S5). This excludes LS-12 with likely a complex displacement, and LS-16 for which the weak and clay-rich lithology, prone to form agglomerates, may not have behaved like a typical granular material. Still, it seems hard to explain with kinetic sieving why LS-9n was so clearly segregated downslope, in spite of its very modest size and displacement. Instead, we could hypothesize that on some landslide deposit, episodic reactivation of the scar and channel chute may have sprayed the deposit with finer debris, depositing preferentially the apex of the deposit. Such mechanism might have happened on most of the landslide landslides we have sampled (given their ages), but is hard to for now we cannot constrain its relevance without repeated monitoring of the deposits. Also, it depends heavily on the geometry, and is most likely for compact deposit below a steep, short chute, while on broad deposits below a moderately steep and long transport channel not much material is expected to reach the final deposit. Last, which is left for future studies. Alternatively, for old deposits it is likely that fine materials could have been washed away by repeated storm events. This progressive washing of the fine grains would be consistent with the fact that the superficial deposits are very poor in fine materials, but
have a proportion of coarse blocks fairly similar to the internal part of the deposit (for LS-5 and LS-8, Fig. 3). In these two cases, kinetic sieving may have been limited (although likely present in LS-8 to explain some downslope coarsening) and fines may have been preferentially washed out. On various parts of the deposit of LS-11 we did find finer materials when scraping the top layer of gravels, consistent with this hypothesis. If such a process is expected to happen on all landslide deposits, superficial measurement of very fresh landslide may already represent the bulk of the material (as perhaps LS-15 and LS-3, most recently failed and with high proportion of fine grains), and older deposit may require some correction as medium to fine grains may be underrepresented.

To conclude this discussion, it seems clear that several physical processes can complexify the landslide deposit GSD, and that deconvolving them and applying a process-based correction to the GSD is not straightforward. More datasets are clearly needed to better understand these sources of variability of the GSDs, for example with a more systematic sampling of very fresh landslides where fines should not have been washed out. Thus, we encourage such issue to be anticipated in future studies, and perform field work in a way allowing the spatial variability to be recorded. This would also enable future studies to include various landslide GSDs based on different assumptions or corrections. In this sense, collecting more measurements of landslides where both internal and superficial GSDs can be measured seems essential, especially while trying to have very young landslides with similar characteristics (lithology, height drop).

4.3 Implications for sediment transport in Taiwan

The landslide GSDs we report contain mainly gravel, but also a substantial fraction of boulders, which suggests that, after reaching floodplains and channels, the transport and evacuation of the material will require large floods. To compare these GSDs to typical shear stresses occurring in Taiwanese rivers, we use the shear stress map derived by Yanites et al. (2010b) from a detailed measurement of the width, discharge and slope along the Peikang river. For a 10-year return flood with a discharge of 1000 $m^3.s^{-1}$, they found the shear stress ranged from $\tau \sim 60-400$ Pa. That shear stress, $\tau$, was about 60 to about 380 Pa. For mountain channels with gradient typically about 2% (Yanites et al., 2010b), these shear stresses correspond to flood heights between 0.3 and 2 m. To assess a threshold for bedload transport, we computed for which grain size $D$ the Shields number $\tau/(\rho/\rho_f - 1)gD$, with $\rho$ and $\rho_f$ the grain and fluid density, respectively, was above a transport threshold of 0.045 (Lamb et al., 2008). Similarly, for suspended load transport, we assessed for which $D$ the shear velocity, estimated as $U_s = \sqrt{\tau/\rho_f}$, was larger than the settling velocity of the grain $U_s$ as defined and calibrated by Ferguson and Church (2004) (Fig. 2, 6). Even for an above average 10-year return flood, less than 25% of most landslide deposits could be transported in suspension, except LS-10 and LS-16 with a suspended fraction of up to 50-70%. When accounting for bedload transport, the largest shear stress of $\sim 380$ Pa could not transport 5-25% of the deposits for about half of the landslides, especially LS-6, LS-9n and LS-13. However, the lower range of estimated shear stresses (60-140 Pa) would result in an immobile fraction of 20 to 40% for most landslides, and up to 80% for the three coarsest deposits.

Before discussing the implications of this, we highlight three main limitations which should be addressed by future work.
aiming at constraining the export of landslides deposits. First, the shear stress could not be adjusted to the local channel conditions in which the landslide occurred, neglecting specific width, discharge and gradient, as well as relations between gradient and critical shields (e.g., Lamb et al., 2008) or the influence of landsliding on the channel itself (e.g., Kuo and Brierley, 2014). Second, we ignored armoring effects, in which a superficial layer of coarse grains inhibits the mobility of finer grains (Parker and Sutherland, 1990). In our case, considering armoring could particularly reduce transport for deposits where coarse grains are segregated at the toe or surface of the deposit, such as for LS-2, LS-3 or LS-8. Third, we did not consider debris-flows and hyper-concentrated flows, which are frequent in Taiwan and sometimes reach the Ocean (Dadson et al., 2005; Lin et al., 2005; Hsu et al., 2008) and which would enhance sediment transport given their higher fluid density. Despite these sources of uncertainty, our results suggest that in the relatively strong metasedimentary units, rapid evacuation of the sediment by suspension affects at most 30% of most of the deposits, and most of the transport occurs as bedload over a limited distance. Further, only the largest (10-year return or more) floods will transport substantial parts of the deposit, meaning that large landslide events may load channels with a pulse of coarse sediments requiring several decades to be evacuated. This is much longer than the transient pulse of suspended-sediment enhanced landsliding (Marc et al., 2015) and suspended-sediment transport (Hovius et al., 2011) observed after the Chi-Chi earthquake, which affected river channels for less than 10 years (Hovius et al., 2011), but is consistent with the ~50yr timescales for enhanced lake sediment deposition (including bedload) after earthquakes (Howarth et al., 2012; Wang et al., 2020).

This multi-decadal timescales for sediment export seems consistent with the very large alluviation of the Southern Taiwan river channel, that followed the Morakot intense flood and landsliding channels, after intense flooding and landsliding triggered by the typhoon Morakot (Yanites et al., 2018), and which was still visible in 2015 (e.g., Taimali river), and at the time of writing in satellite imagery. Substantial aggradation, suspected to be long-term, was also observed after the Chi-Chi earthquake (Yanites et al., 2010a; Chen, 2009). More detailed modeling of the evacuation of landslide sediment (e.g., Yanites et al., 2010a; Croissant et al., 2017) could be combined with scenarios based on the detailed GSD-GSDs reported in this study to better quantify the dynamics and timescales of coarse sediment export after large landslides events.

5 Conclusions

We have presented grain size distributions obtained from 17 landslide deposits in Taiwan. They have $D_{50}$ and $D_{84}$ consistent with landslides reported in previous literature, between 45-200 mm an 60-600-15 and 200 mm and between 60 and 600 mm respectively. We found that many deposits had significant spatial segregation in the downslope direction, with the lowest part of the deposits having 2 to 10 times coarser GSDs than the upper part of the deposits. For the three landslides in which we could sample the inner part of deposits, we also found GSDs 3 to 10 times finer than their surface counterparts. We did not find a single process which could explain the presence of these segregation, though this segregation cannot be attributed to a single process, but kinetic sieving and deposit reworking are likely candidates. Still, this internal variability is a strong caveat for very localized sampling (i.e., sieving to play important roles. This internal variability could bias results obtained from local sampling of GSDs, such as sieve samples from a single pit).
Figure 6. Fraction of the landslide GSD that could be transported as both bedload and suspended load (a) and only suspended load (b) as a function of river shear stress during a 10-year return flood. Note that in (b) suspended fraction at 300 and 380 MPa are identical.

Investigating the controls on landslide GSD variability, we observed a strong anticorrelation between the landslide drop height, width and inferred scar depths and the GSD percentiles for all the landslides. Finer GSDs in the two landslides in non-metamorphosed, young sedimentary rocks can be well explained by normalizing the drop height by the literature rock strength. Further, modeling the source material median grain size with an exponential fining towards the surface, consistent with physical weathering models, we found that the reduction ratio from source material to landslide deposits matches the scaling proposed by Locat et al. (2006) and calibrated for much larger and stronger rock avalanches than studied by us, rock avalanches with much larger volume and much higher point-load strength than the ones we studied. Although future measurements on the source rock are needed for a complete demonstration, especially on bedrock strength and fracture spacing, we suggest that simple geomorphic models coupling this fragmentation scaling with a model for regolith grain size could provide a physically-based first order model for the GSD input to rivers by landslides in active orogens. Such an approach could have important implications for landscape evolution models and accounting for sediment transport. Indeed, from our deposits we also noted that even a 10-year flood may not be able to transport the coarsest fraction of many deposits, suggesting that floodplains and channels will likely need several decades to recover after large landslide events.
Data availability. The 28 GSD (for each landslide sub samples) are available in the Hydroshare open repository, together a shapefile with landslide locations and polygons derived from Google Earth. Marc, O., J. Turowski, P. Meunier (2021). Grain Size Distribution of 17 Taiwanese landslide deposits, HydroShare, http://www.hydroshare.org/resource/ade683be61e54fa5b60da97418a5f3df

Author contributions. OM designed the study and the field mission, performed all analyses and wrote the manuscript. JMT and PM provided input for the field methodology, the result interpretations and edited the manuscript. All authors have been collecting the grain size data in the field. We thank Mikael Attal and one anonymous reviewer for their constructive reviews.

Competing interests. The authors declare no competing interests

Acknowledgements. The grain counting during the 2015 field campaign could not have been possible without the additional participation of Antonius Golly, Arnaud Burtin, Anne Schâpa and Niels Hovius, and they are warmly thanked for this. We also thank Kristen Cook and Anne Schâpa for contributing pictures, and Sebastien Carretier for pointing us to the literature on physical weathering models.
References


