Natural Hazards
and Earth System
Sciences

Amalgamation in landslide maps: effects and automatic detection

O. Marc[1,2] **and N. Hovius**[1,2]

[1]Helmholtz Centre Potsdam, German Research Center for Geosciences (GFZ), Telegrafenberg, 14473 Potsdam, Germany
[2]Institute of Earth and Environmental Sciences, University of Potsdam, Potsdam, Germany

Correspondence to: O. Marc (omarc@gfz-potsdam.de)

Abstract. Inventories of individually delineated landslides are a key to understanding landslide physics and mitigating their impact. They permit assessment of area–frequency distributions and landslide volumes, and testing of statistical correlations between landslides and physical parameters such as topographic gradient or seismic strong motion. Amalgamation, i.e. the mapping of several adjacent landslides as a single polygon, can lead to potentially severe distortion of the statistics of these inventories. This problem can be especially severe in data sets produced by automated mapping. We present five inventories of earthquake-induced landslides mapped with different materials and techniques and affected by varying degrees of amalgamation. Errors on the total landslide volume and power-law exponent of the area–frequency distribution, resulting from amalgamation, may be up to 200 and 50 %, respectively. We present an algorithm based on image and digital elevation model (DEM) analysis, for automatic identification of amalgamated polygons. On a set of about 2000 polygons larger than $1000\,\text{m}^2$, tracing landslides triggered by the 1994 Northridge earthquake, the algorithm performs well, with only 2.7–3.6 % incorrectly amalgamated landslides missed and 3.9–4.8 % correct polygons incorrectly identified as amalgams. This algorithm can be used broadly to check landslide inventories and allow faster correction by automating the identification of amalgamation.

1 Introduction

Regional landslide maps are a crucial component of many landslide related studies (Guzzetti et al., 2012): they are necessary to improve our understanding of landslide rupture mechanics and test conceptual models, to produce landslide risk and vulnerability maps, to understand how different climatic and tectonic mechanisms can trigger landslides, and to estimate how mass wasting contributes to sediment production and landscape evolution (Montgomery and Dietrich, 1994; Meunier et al., 2007, 2008; Hovius et al., 1997). Such maps used to be created by manual mapping from remote sensed imagery, often accompanied by partial field checks (e.g. Harp et al., 1981; Harp and Jibson, 1996). Due to the high cost and time associated with manual mapping of thousands or tens of thousands of landslides over large areas, automated mapping techniques are increasingly used (e.g. Martha et al., 2010; Mondini et al., 2011; Parker et al., 2011). These techniques have specific associated errors, amongst which amalgamation, that is the bundling of several adjacent landslides into a single map polygon, is prominent. Amalgamation typically occurs when the spatial density of landslides is high and the resolution of images from which they are mapped relatively low, making it difficult to differentiate multiple landslides in a perturbed area. Automatic mapping algorithms designed to detect change of surface properties, irrespective of the shape of the changed area are especially prone to this effect. If uncorrected, amalgamation can lead to severely erroneous results and interpretations in many domains. For example, studies using landslide maps to estimate the volume of debris produced, whether to understand sediment transfer dynamics (Hovius et al., 2000; Yanites et al., 2010), organic matter mobilization (Hilton et al., 2011), average erosion rates (Hovius et al., 1997) or mountain building (Parker et al., 2011), rely on empirical laws giving landslide volume as a function of landslide area (Guzzetti et al., 2009; Larsen et al., 2010). In this approach, landslide depth is assumed to scale with area, giving rise to strongly non-linear area–volume relations, which assign disproportionate importance to landslides with the largest surface areas. Accurate landslide area mapping, differentiating precisely between indi-

vidual events is therefore of the essence (Li et al., 2014). This also applies to studies considering the area–frequency distribution of landslides, whether to assess landslide hazard and risk associated with extreme events (Malamud et al., 2004b), or to understand the underlying physics of the distribution (Pelletier et al., 1997; Stark and Guzzetti, 2009; Frattini and Crosta, 2013). Finally, any attempt to understand the physics of landslide triggering from mapped landslide patterns could suffer from the effects of incorrectly mapped landslide outlines and the artificial prominence of large disturbed areas (Montgomery and Dietrich, 1994; Meunier et al., 2008).

Here, we survey why and where amalgamation can occur, and determine the minimum error it has introduced to estimates of total landslide volume and the area–frequency distribution of several landslide inventories. Subsequently, we propose an algorithm able to automatically detect amalgamation when provided with a raster file of polygon shapes and a digital elevation model (DEM). Performance of this algorithm is tested on a representative subset of the inventory of landslides triggered by the Northridge earthquake. We finish with a short discussion of the benefits and limitations of this approach and possible alternatives.

## 2 Landslide mapping and amalgamation

Most landslide inventories are derived from analysis of optical or multispectral imagery, exploiting the typical texture, colour and spectral properties of freshly disturbed areas (Guzzetti et al., 2012). Often, landslides are conspicuous because they clear vegetation that has a very different appearance or radiation intensity spectrum. When landslides are mapped as polygons, whether by men or machine, the general assumption is that the polygon represents a single landslide, most often combining a scar area, a deposit area and sometimes a runout area. A mapped polygon is therefore assumed to contain direct or indirect information on the location and size and, implicitly, the volume of one landslide but also potentially about the slope where the landslide initiated and terminated, the runout distance, the drop of potential energy, or the triggering mechanism, such as the local peak ground acceleration or pore pressure at the time of failure.

Amalgamation, the combination of several individual landslides in a single polygon, can be due to the actual coalescence of landslides, or the apparent contiguity of disturbed areas in images with low resolution or poor contrast between affected and unaffected areas (Fig. 1). Indeed, where landsliding is very dense, several adjacent landslides may have joint runout areas or overlapping deposits, or scars separated by a distance too short to be resolved by the available imagery. At a given resolution, multispectral images contain more information than optical images, which may help in delineating individual landslides but this does not always preclude amalgamation in landslide mapping. Even where image resolution would permit accurate mapping of individual

landslides, amalgamation can occur when the primary goal of the mapper is not to map landslide extent precisely, but rather to rapidly evaluate the area affected by slope failure. This seems common for maps predating widespread use of landslide area–volume relationships as well as for more recent inventories, underlining the current lack of care in avoiding or at least flagging amalgamation. In automatic mapping, algorithms that are not object oriented will usually classify single pixels based on their various bulk properties (Guzzetti et al., 2012). If adjacent pixels are classified as disturbed, then the algorithm will combine them in a single polygon, regardless of how many separate landslides are contained within. When image resolution is not very high, then automatic algorithms can bundle hundreds of small landslides, located within a limited area with high propensity to failure, into a single, apparently very large landslide polygon.

A striking example of amalgamation can be found in the Jou-Jou Mountain area of Taiwan, where pervasive shallow landsliding occurred during the M_w 7.6 ChiChi earthquake in 1999 (Fig. 2) (Liao and Lee, 2000). In available maps, these landslides have been merged into a few complex shaped polygons, blanketing the steep, gullied hills and covering 9.8 km^2. However, a separate, local survey has found more than one thousand individual, shallow failures, many of which adjoined without making larger landslides (Lee et al., 2010). Together, these landslides had a total area of 7.22 km^2, implying a significant area exaggeration by the automated mapping procedure. The implications of this extreme amalgamation are far reaching. For example, using common landslide area–volume relations (Guzzetti et al., 2009; Larsen et al., 2010), the total volume of the six largest, automatically mapped polygons in the area would be estimated at about 0.19 km^3, with the largest polygon (4.13 km^2) alone contributing about 0.11 km^3. If the total area occupied by these six polygons is arbitrarily repartitioned into 1000 landslides of roughly equal size, set by the characteristic local ridge spacing and slope lengths of 100–150 m, then a 17-fold reduction of the estimated landslide volume would result. This estimate could be refined with access to the local landslide data (Lee et al., 2010), which can be seen to have a non-uniform area–frequency distribution with hundreds of landslides with areas of 100 m^2 and one landslide of 0.1 km^2.

In this example, amalgamation of landslides is easily recognizable due to the complex shape of polygons straddling multiple topographic features, with surface areas much larger than permitted by the characteristic length scale of the topography. Formally, the merging of several landslides can result in a range of geometric or topographic inconsistencies, such as multi-branched polygons, or polygons with orientations inconsistent with local topographic slope or transgressing ridgelines or channels (Fig. 2). We consider that these features are unlikely characteristics of individual landslides, even though failure on multiple scarps, divergence in runout, runout crossing rivers and spreading on the opposing valley
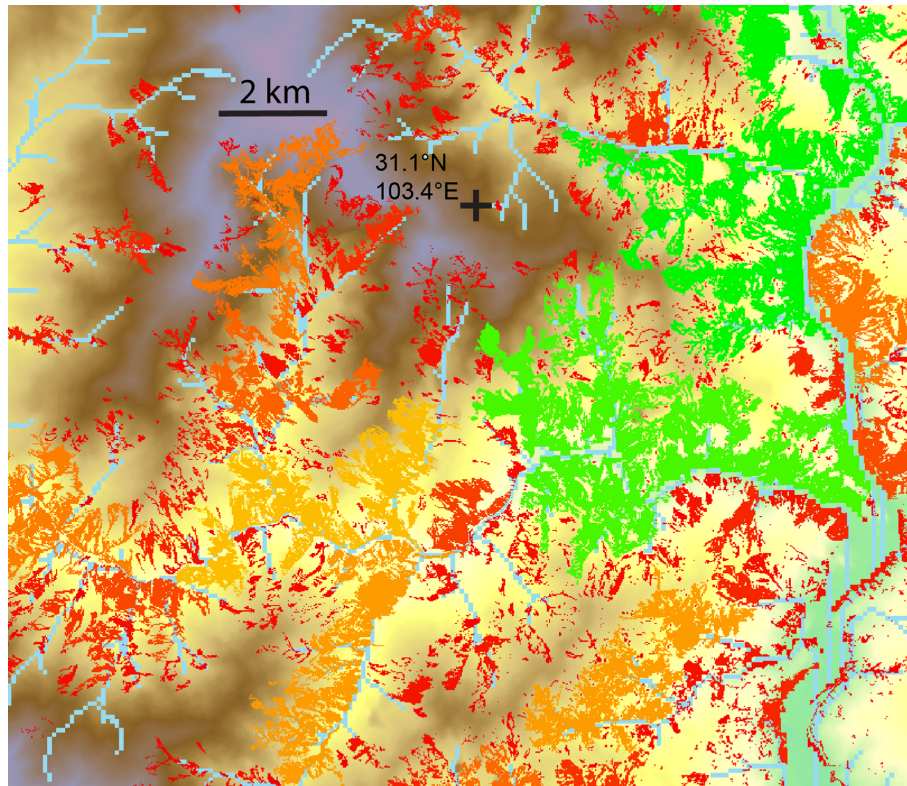
Figure 1. Some polygons from the Parker et al. (2011) data set, representing landsliding caused by the 2008 Wenchuan earthquake. Polygons are colour coded by size (red being the smaller polygons) and overlaid on a DEM and a river network. The density of landsliding is correctly estimated but dozens of small landslides have been connected along slope or even across rivers or ridges. Light orange and green polygons have total area larger than 1 km^2.

side or occasional overtopping of dividing ridges are known to happen.

Some polygons may also appear topographically and geometrically consistent, although they are, in fact, a combination of several adjacent landslides close to or below the resolution of available images, the combined effect of which is to alter the visual or spectral properties of a larger area. This blurring can conjugate amalgamation and an exaggeration of the area affected by landslides, but it cannot be identified without use of very high-resolution images (Fig. 2). It is, therefore, out of the scope of our study and remains a challenge and a caveat for landslide mapping.

3 Data

The recognition of geometric and topographic inconsistencies in landslide inventories is a key to identification of amalgamation of individual landslides and mitigation of its effects. To develop a method for detection of amalgams in large landslide data sets, and to evaluate the effects of amalgamation on scientifically interesting derivatives of these data sets, we have focused on earthquake cases. Large earthquakes can trigger many thousands of landslides in a limited area, re-

ducing the possible effects of geological heterogeneity on landslide populations and their statistics. Moreover, by focusing on landslides with a shared trigger mechanism, we have removed potential complications due to the convolution of trigger-specific effects from our analysis. Finally, earthquake-induced landslide populations tend to span a very large range of landslide sizes, allowing robust computation of area–frequency statistics, one of the key attributes affected by amalgamation.

We have used five published inventories of earthquake-induced landslides, mapped over areas of 10^3–10^4 km^2. Together, these inventories cover a range of mapping approaches from manual mapping with extensive field checking, to fast automated mapping with limited supervision and verification. The 1994 M_w 6.6 Northridge earthquake in California triggered more than 10 000 landslides, which were mapped manually from air photos, with field checks at selected sites (Harp and Jibson, 1996). The same approach was used to map more than 6000 landslides triggered by a M_w 7.6 earthquake in 1976 in Guatemala (Harp et al., 1981). The 1999, M_w 7.6 ChiChi earthquake in west Taiwan also caused severe landsliding, with more than 9000 landslides larger than 625 m^2 (25 m × 25 m) mapped manually from SPOT
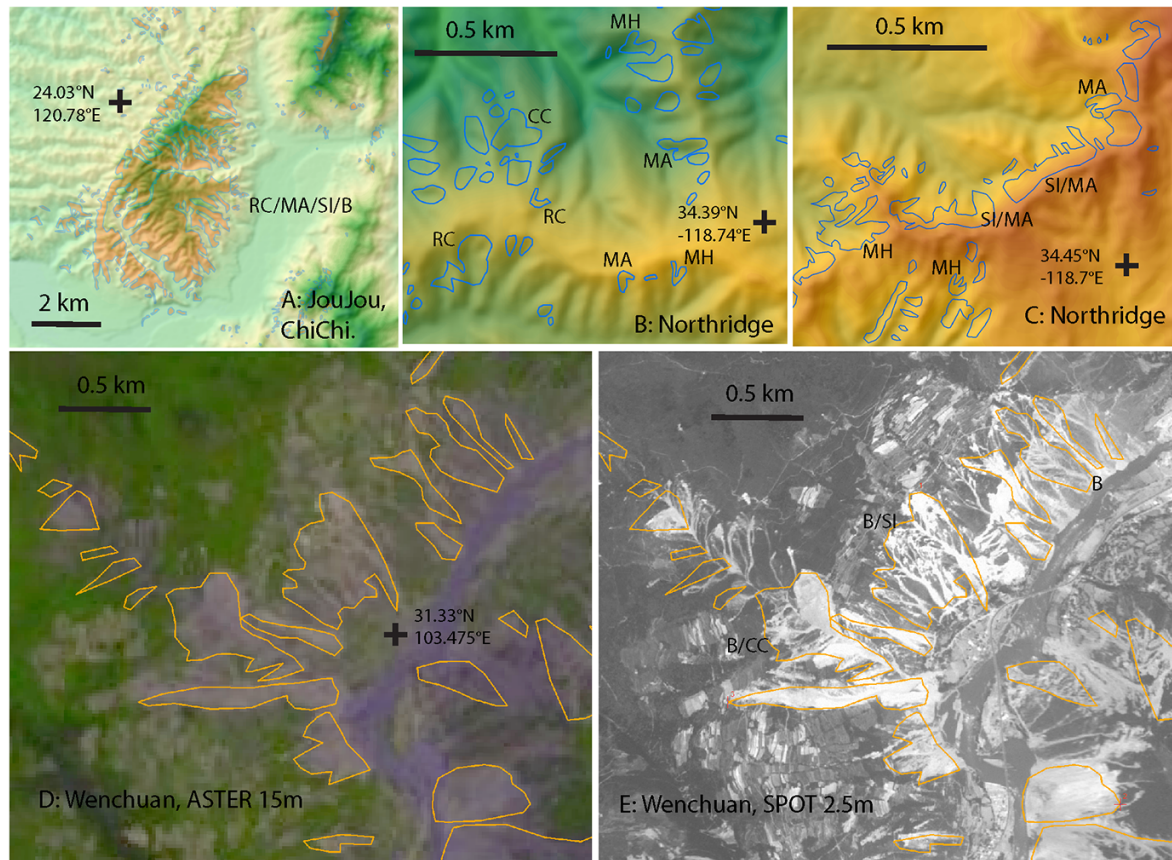
Figure 2. (a), **(b)**, **(c)**: landslide polygons on a DEM topography showing examples of amalgamation in the ChiChi and Northridge inventories. Geometric and topographic inconsistencies that signal amalgamation are specified as follows: RC for ridge crossing, CC for channel crossing, MH for multi-headed, MA for multi-armed and SI for slope inconsistencies. **(d)**, **(e)**: some polygons mapped by Gorum et al. (2011) after the Wenchuan earthquake overlaid on a 15 m resolution ASTER image **(d)** and on a 2.5 m resolution SPOT 5 image **(e)** of the same area. Note the presence of amalgamation but also the significant mapping extent exaggeration when mapping on low resolution relative to the landslide density.

satellite imagery (Liao and Lee, 2000). Finally, for the 2008, M_w 7.9 Wenchuan earthquake in China, many different maps of coseismic landslides exist (Ouimet, 2010; Qi et al., 2010; Dai et al., 2011; Gorum et al., 2011; Parker et al., 2011; Xu et al., 2014), allowing comparison of independent and broadly equivalent data sets. We have used two catalogues containing 50 000 polygons apiece. One was mapped with a semi-automatic algorithm using 2.5 to 10 m resolution SPOT 5 and EO-1 satellite imagery (Parker et al., 2011). The other was mapped by hand, mainly from 15 m resolution ASTER imagery and locally higher resolution imagery (Gorum et al., 2011). In all these inventories, the entire area perturbed by a landslide, including scar, runout and deposit, is delineated by a single polygon.

In addition to these five inventories, we have used Aster GDEM-30m data to evaluate the topographic context of mapped landslide polygons and as an input of our algorithm for detection of amalgams. In the case of the Wenchuan earthquake, we have also used 15 m resolution ASTER images and 2.5 m resolution SPOT 5 images from the epicentral area, taken shortly after the earthquakes, to verify the different landslide maps.

4 Quantifying the effects of amalgamation

The earthquake-induced landslide inventories summarized above are too large for comprehensive manual verification. To assess the possible effects of amalgamation in these data sets, we have focused on the largest polygons in each inventory. These polygons dominate landslide volume estimates and can strongly influence the best fits to area–frequency distributions. Thus, by checking and correcting a limited number of large polygons, the quality of derivatives of landslide inventories can be substantially improved. In checking individual polygons, we considered as anomalous any polygon displaying a geometrical or topographical inconsistency such as branching, traversing of ridges or rivers or orientation inconsistent with the local topographic slope. These polygons

were compared with local topographic data and, when appropriate split to make residual polygons more consistent with the general topography. Nevertheless it is clear that without high-resolution imagery, many landslide polygons were redefined in a relatively crude way.

We have used published area–volume relationships to estimate the volume of landslides from the mapped disturbed areas (Larsen et al., 2010). It was assumed that landslides with area $> 100\,000\,\mathrm{m}^2$ involved bedrock, and that smaller landslides were mixed bedrock and soil failures. Landslide maps typically do not distinguish between scar and deposit, lumping the two into one area measure. According to Larsen et al. (2010), scars and deposits have area–volume relations with the same power-law exponent, implying constant size ratios between scar and deposit areas of 1.1 and 1.9 for mixed and bedrock landslides, respectively. Hence, we have estimated the scar area by dividing the mapped landslide area by 2.1 and 2.9 for mixed and bedrock landslides, respectively, assuming that runout was equal to the scar length. Then we converted scar area A, into volume V, for bedrock and soil landslides with $V = aA^b$ with $a = 0.146$ and 0.234 and $b = 1.33$ and 1.41 for mixed and bedrock landslides, respectively. Computed landslide individual and total volumes appear to be consistent with field estimates for cases where the whole perturbed area is mapped.

Comprehensive landslide inventories have a typical area–frequency distribution with a roll-over and a power-law decay with an exponent, ρ, commonly within a narrow range of values (Malamud et al., 2004a). The roll-over can be caused by censoring of the small landslides due to the mapping resolution (Stark and Hovius, 2001), but can also be related to the physics of landsliding and the transition from cohesion-controlled to friction-controlled hillslope stability with increasing landslide area and depth (Katz and Aharonov, 2006; Stark and Guzzetti, 2009). The roll-over and power-law decay have also been attributed to a combination of the size distribution of continuous local topographic slopes and the distribution of moisture or increasing cohesion with depth (Pelletier et al., 1997; Frattini and Crosta, 2013). We have assessed the impact of amalgamation by comparing the area–frequency distribution of the original data sets with that of our partially corrected data sets. Because the frequency decay with increasing landslide size is usually modelled as a power-law, a specific functional form does not have to be prescribed if we only consider the distribution at areas ten times larger than the roll-over. For these large areas we have obtained ρ with a linear least-square regression of the log-transformed data (Fig. 3).

In many cases a larger number of smaller polygons were also visibly amalgamated, but we did not correct them, due to the effort and uncertainties involved. Thus, the estimates of errors on total landslide volume and the power law exponent of the landslide area–frequency distribution due to amalgamation, presented below, are likely minimum values. Next, we review the individual landslide inventories and highlight
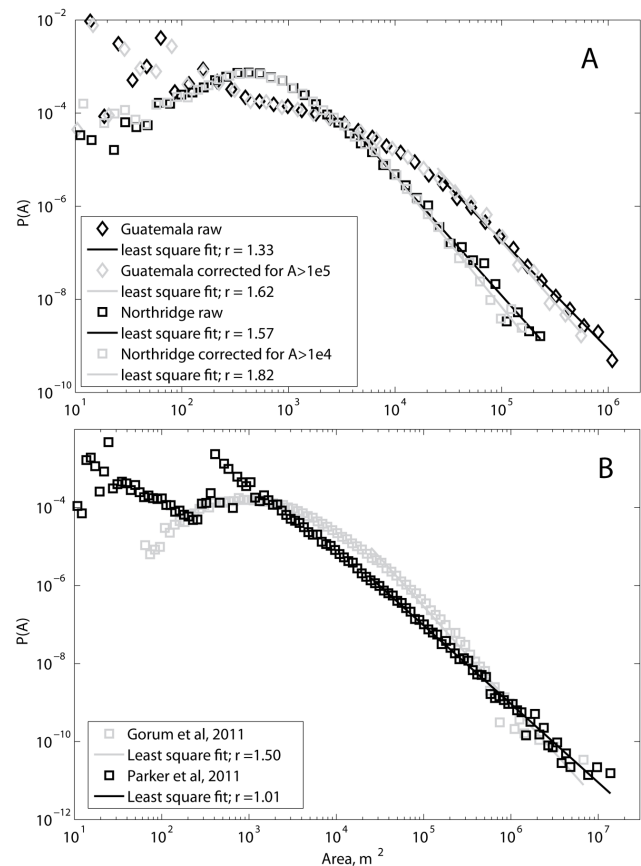
Figure 3. Amalgamation effect on landslide area–frequency distributions. **(a)** Comparison between the raw data from the coseismic landslide maps for the 1976 Guatemala and 1994 Northridge earthquakes and the corrected catalogue where every amalgam larger than $100\,000$ and $10\,000\,\mathrm{m}^2$ was split, respectively. **(b)** For the 2008 Sichuan earthquake, several landslide maps were published. Of these, the Parker et al. (2011) data set is severely affected by amalgamation whereas the Gorum et al. (2011) data set is relatively exempt from amalgams.

the varying degrees to which they are affected by amalgamation and its effects.

Landslides induced by the 1976 Guatemala and 1994 Northridge earthquakes were mapped in detail, apparently to record where landslides had occurred, but not necessarily to distinguish the boundaries of individual landslides. We have inspected all 356 polygons with an area larger $10\,000\,\mathrm{m}^2$ in the Northridge inventory and all 90 polygons exceeding $100\,000\,\mathrm{m}^2$ in the Guatemala data set. Together, these polygons represent 56 and 73 % of the uncorrected volume of the landslide populations of the Northridge and Guatemala earthquakes, respectively. 162 out of 356 and 51 out of 90 of these polygons were found to be amalgams of several landslides. They were split according to their shapes and relation to the local topography. This resulted in a reduction of the total volume of landslides by 16 % in the Northridge case and 35 %

in the Guatemala case, and an increase of the area–frequency scaling exponent, by 16 and 22 %, from 1.57 and 1.33 to 1.82 and 1.62, respectively (Fig. 3). Because polygons smaller than the threshold represent only 44 and 27% of the total volume, respectively, and because they must be less amalgamated and have much smaller individual volumes, their correction would likely add only a minor contribution to the total volume change.

The ChiChi earthquake caused widespread landsliding in the mountains of central west Taiwan. An inventory of these landslides (Liao and Lee, 2000) contains 9272 polygons in an area 150 times larger than the Jou-Jou Mountain, mentioned above, with a total estimated volume of about 0.73 km³. We have inspected all 173 polygons larger 100 000 m², representing 85 % of the total uncorrected volume of the ChiChi inventory. We have found that 100 of them needed corrections ranging from the splitting of minor branches to the artificial fragmentation of the largest polygons in the Jou-Jou Mountain area, where precise correction was impossible. Together, these corrections resulted in a volume reduction of 38 % to 0.45 km³, but an insignificant increase of the area–frequency scaling exponent by 5 %.

The two inventories for the Wenchuan earthquake have similar total landslide areas and similar total numbers of landslides, even though the mapping of Gorum et al. (2011) extended further to the north along the seismogenic fault. We have compared the maps where they overlap, along 150 km of the fault trace, where the majority of landslides occurred (e.g. Fig. 2). There is good overall agreement between the data sets, but the manual mapping of Gorum et al. (2011) has clearly delineated many more individual slides (Fig. 2). Many examples of amalgamation are evident in the Parker et al. (2011) data set (Fig. 1), and although there are some mapping discrepancies between the two inventories, this appears to be the main difference between them. It has resulted in a total landslide volume reduction of 69 %, from 6.30 km³ for the automated-mapping inventory (Parker et al., 2011) to 1.96 km³ for the original manually mapped inventory of Gorum et al. (2011). However, this inventory also contains amalgamation artefacts (Fig. 2). We have visually checked all 152 landslides larger than 300 000 m², representing 51 % of the total volume of the manual inventory (including landslides mapped in areas not surveyed by Parker et al. (2011)). Of these 87 required editing, leading to a final landslide volume estimate of 2.3 km³ instead of 2.45 km, equivalent to a modest reduction of 6 %.

The landslide polygon area–frequency distributions of the Wenchuan inventories also differ significantly (Fig. 3). First, the amalgamated catalogue of Parker et al. (2011) yields a discontinuous distribution, which does not exhibit the roll-over commonly observed in well-mapped data sets (Malamud et al., 2004a; Brardinoni and Church, 2004). Instead the smaller polygons also have a decreasing frequency with increasing size, and they appear to be relatively infrequent compared to medium to large slides. In contrast, the manu-

ally mapped inventory has a area–frequency distribution with a roll-over at 1000 m². The exponent on the best-fit power-law for this data set, after our correction for amalgamation is also much higher than for the Parker et al. (2011) inventory, $\rho = 1.5$ and $\rho = 1.0$, respectively, confirming the relative abundance of large, mostly amalgamated polygons in the latter (Figs. 1, 2). Correction for amalgamation effects results in a slight rise of the scaling exponent of the manually mapped inventory to $\rho = 1.6$.

From these analyses it is clear that amalgamation can significantly distort both landslide population volume estimates and the frequency distribution of mapped landslide areas. However, the frequency distribution itself does not necessarily betray amalgamation, and exhaustive visual screening can be prohibitively time consuming. In the following section, we propose an automatic algorithm, which can be used to differentiate correctly mapped and amalgamated polygons and allow faster and more comprehensive cleaning of affected data sets.

5 Automatic detection of amalgamation

Because amalgamation leads to geometric anomalies and unusual positions of putative landslides in the landscape it is possible to detect amalgams simply by looking at their shape and at the underlying topography. Following the criteria defined in Sect. 2 (Landslide mapping and amalgamation) we have developed an algorithm able to guide a mapper or an end-user towards suspicious polygons, and facilitate a correction or an assessment of the catalogue quality (see Supplement). The algorithm requires a DEM, a raster made from the polygon shapefile and a text file with polygon ID and information. Below, we present the operation of the algorithm and assess its accuracy.

First, the algorithm considers the geometry of a landslide polygon. The branching of polygons is the most common and visible effect of amalgamation. This affects the relation between perimeter, P, and area A, of the polygons, biasing amalgams towards high P. These attributes are easily extracted from a landslide inventory with any GIS. A polygon with given P and A can be compared to an ellipse of equal P and A, and aspect ratio K. Using Fagnano (1750) approximation, ellipse perimeter can be written as:

$$P = \pi \left(\frac{3}{2b}(K+1) - \sqrt{Kb^2} \right), \tag{1}$$

where b is the smallest radius. Since $A = \pi K b^2$, it can be shown that the perimeter of any ellipse varies as

$$P = \left(\frac{3(K+1)}{2\sqrt{K}} - 1 \right) \sqrt{\pi A}. \tag{2}$$

Rearranging Eq. (2), K can be found from P and A as the solution of a second order equation:
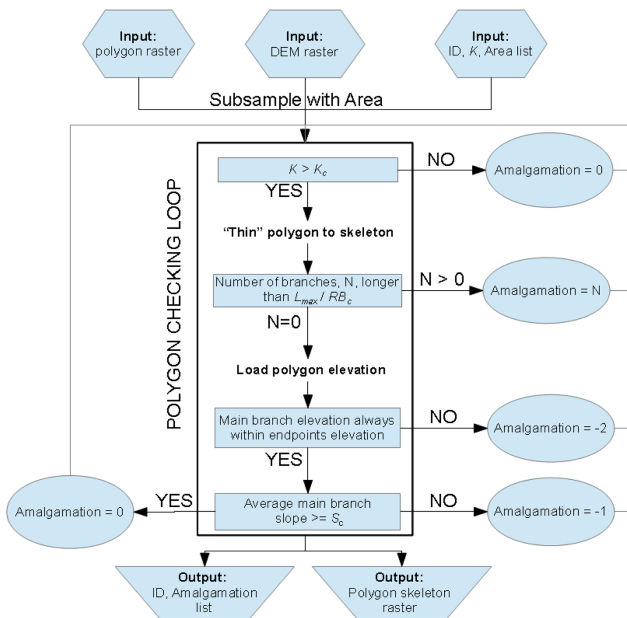
Figure 4. Flowchart of the algorithm for automatic detection of amalgamation. The algorithm is provided in the Supplement. Inputs are used to individually analyse polygons based on geometric and topographic characteristics, following a series of conditional tests that lead to a polygon score. A score of zero means that the polygon is considered clean and any other scores refer to some sort of amalgamation. K is the equivalent ellipse aspect ratio (see Eq. 3), L_{max} is the length of the longest branch of a polygon, RB_c is an arbitrary critical length ratio and S_c is a critical slope angle.

$$K = \frac{1}{2}\left(\frac{4}{9}\left(\frac{P}{\sqrt{\pi A}}+1\right)^2 - 2 + \sqrt{\left(\left(\frac{4}{9}\left(\frac{P}{\sqrt{\pi A}}+1\right)^2 - 2\right)^2 - 4\right)}\right). \quad (3)$$

Thus, any polygon can be described easily and objectively by the aspect ratio, K, of its equivalent ellipse. For reference, a circle would yield $K = 1$, a square $K = 2.3$ and rectangle twice as long as wide $K = 2.7$. A polygon with high K is more likely to be incorrect whereas a polygon with $K < 2$ has a compact shape from which any mapping error cannot easily be recognized. Therefore, to accelerate the algorithm any polygon below a critical aspect ratio, K_c, is assumed to be correct (Fig. 4).

A high K value may signal amalgamation or simply an elongated landslide, for example due to long runout. Therefore, K is a useful input parameter but ultimately it is necessary to explicitly consider the geometry of the polygon. This is achieved by reducing the mapped polygons to their skeleton with a standard image analysis method, which iteratively thins a solid polygon to a branched centre-line (Fig. 5). From this skeleton, branch points and individual branches are easily found. However, even polygons with a relatively simple shape may have skeletons with some branching points and small branches pointing towards a polygon corner or irregular side. To eliminate these spurious branches, we impose

an arbitrary threshold size ratio of branches relative to the longest branch, RB_c. A polygon with a main branch and several smaller branches, all of which are shorter than the main branch by a factor $1/RB_c$ or more is considered to be a correctly mapped, single landslide (Fig. 4). All other polygons receive a score equal to the number of branches, longer than the longest branch divided by RB_c, reflecting qualitatively the degree of amalgamation.

In a second step, the algorithm tests the consistency of a polygon with apparently correct geometry, with the local topography. This is done by extracting the DEM elevation along the longest branch of the polygon, which is assumed to be an adequate representation of the pathway of the landslide. First, the algorithm checks that the highest and lowest elevations along the branch coincide with the top and toe of the mapped landslide. A violation of this condition typically signals that the branch traverses a ridge or valley floor, or that two landslides were merged into a crescent shaped polygon, smooth enough not to be identified as a likely amalgam by the first part of the algorithm. If the polygon passes this second test, then a last check is made to see if the maximum variation of elevation along the main branch is above the minimum slope for landsliding, S_c (Fig. 4). Polygons failing this test are typically oriented perpendicular to the main topographic slope over long distances, as a result of the lateral merging of several small, parallel failures along a ridge or cliff.

Thus, our algorithm is formally based on three adjustable parameters K_c, RB_c, and S_c. Of these, only RB_c may be substantially tuned, depending on the smoothness of the input raster, which in turn depends on the landslide mapping technique and the raster resolution.

S_c is a physical parameter which should normally be close to a 10° threshold for landsliding (e.g. Meunier et al., 2007; Lin et al., 2008), thus requiring minimal tuning, unless the local substrate has exceptional properties. To minimize the number of false negatives (i.e. undetected amalgams), K_c should be set at a low value of about 2, so that only polygons without any geometrical complexity are screened out. Setting K_c at a higher value can be useful to assess the degree of amalgamation and isolate only those polygons that are likely to be composites of many landslides.

To assess the accuracy of the algorithm we have applied it to an inventory of landslides triggered by the 1994 Northridge earthquake in southern California (Harp and Jibson, 1996). Within the bounds of the Santa Susana Mountains, we manually screened all 2083 mapped polygons larger than $1000\,m^2$ for amalgamation. This is close to the roll-over in the landslide area–frequency distribution of the inventory, so that the test set encompasses most of the landslide volume. The Santa Susana subset is representative of the diversity of size and shape that can be found in the Northridge inventory in its entirety. Of all polygons in the subset, the amalgamation state of 136 (6.52 %) could not be ascertained visually. These polygons were removed from the test data set before further analysis. Of the remaining
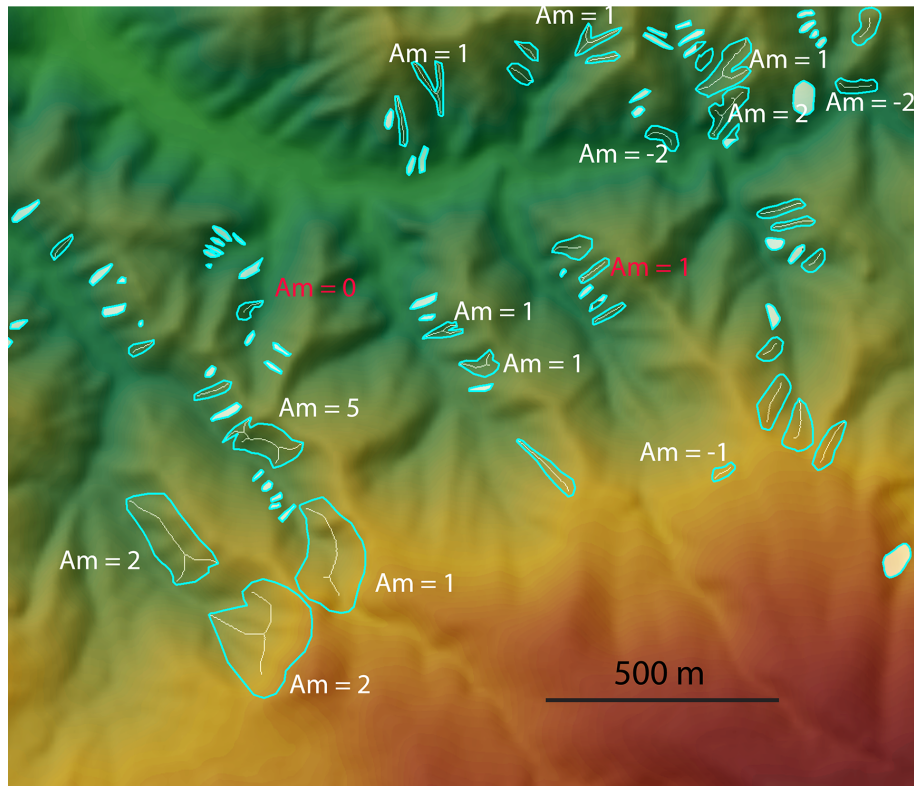
Figure 5. Part of the Northridge landslide polygon inventory overlaid on a hillshaded DEM. The skeleton raster output is shown for all polygons larger than $1000\,\mathrm{m}^2$ and with $K>=2$ (35 polygons). Polygons with $K<2$ are filled in white and considered clean. White labels show erroneous polygons detected by the algorithm, with positive numbers giving the number of secondary branches detected, -2 meaning ridge or river crossing and -1 indicating a slope smaller than $12°$. Red labels show incorrectly diagnosed or dubious results within this sample. Polygons with skeleton but no labels have been correctly classified as unamalgamated.

1950 polygons, 617 amalgams and 1187 single landslides were correctly classified by our algorithm. The algorithm missed 70 amalgams (3.6 % of false negatives, that is undetected amalgams) and incorrectly classified 76 single landslides as amalgams (3.9 % of false positives, that is correct polygons classified as amalgams) (Table 1). About two thirds of all polygons classified as amalgams were detected using the branching criterium, in part because it is the most easily detectable feature but also because it is the first step of the algorithm. One third of amalgamation cases were only diagnosed by the second step of the algorithm, which considers the topographic context of a polygon. Taking results from these two steps together, the overall accuracy of the algorithm was very good, with 1804 of 1950 (92.5 %) polygons in the test set classified correctly (Table 1). Thus, our algorithm provides a relatively rapid and accurate way to assess the quality of a data set and a partial guide to manual correction. It can reduce the workload associated with manual splitting of amalgamated polygons, by foreshortening the amalgam identification phase, and enhancing the detection of smaller amalgamated polygons that may have only subtle distortions. However, the algorithm only yields a minimal

number of branches and the automatic and accurate splitting of complex polygons based on detected branching geometry remains a challenge.

The algorithm can assess the quality of every polygon of an inventory as long as the raster resolution is high enough for a polygon to be made up by at least a few tens of cells, so that a skeleton can be defined. Therefore, at a raster pixel size of $2\,\mathrm{m}$, $100\,\mathrm{m}^2$ polygons would have about 25 pixels and could be analysed by our algorithm. This is lower than the usual roll-over of landslide area–frequency distribution (e.g. Malamud et al., 2004a; Brardinoni and Church, 2004). DEMs with a high spatial resolution will also yield better results and the accuracy of the detection is helped by the fact that the algorithm uses raw elevation data rather than a local derivative such as slope, which is calculated over several adjacent cells.

6 Discussion

We have proposed an algorithm based on polygon geometry and topographic analysis, which allows automatic detection of polygons outlining amalgamated landslides with good but incomplete detection rates and minimal diagnostic error.

Table 1. Confusion matrix of the algorithm tested on the 1950 independently verified polygons larger than $1000\,\mathrm{m}^2$, from an inventory of landslides triggered by the 1994 Northridge earthquake. Positive and negative conditions refers to polygons considered amalgamated and correct, respectively. Therefore, false positives are correctly mapped polygons erroneously identified as amalgams whereas false negatives are amalgams that remain undetected by the algorithm. Values are given as number of landslides and percent of the total population. The algorithm was run with the following parameters: resolution 2 m, $K_c = 2$, $S_c = 12°$ and $RB_c = 5$ for the upper part of the table and $RB_c = 6$ for the lower part.

True positive: 617 (31.6 %)	False positive: 76 (3.9 %)	Positive predictive rate = 89.0 %
False negative: 70 (3.6 %)	True negative: 1187 (60.9 %)	Negative predictive rate = 94.4 %
Sensitivity = 89.8 %	Specificity = 94.0 %	Accuracy = 92.5 %
True positive: 653 (33.5 %)	False positive: 94 (4.8 %)	Positive predictive rate = 87.4 %
False negative: 52 (2.7 %)	True negative: 1151 (59.0 %)	Negative predictive rate = 95.7 %
Sensitivity = 92.6 %	Specificity = 92.5 %	Accuracy = 92.5 %

However, depending on the objective of a study, even a few incorrectly diagnosed polygons may be of concern. Therefore, the algorithm must be tuned towards a reduction of false negative results, by increasing RB_c or S_c, even if the rate of false positive results increases as a consequence. For example, raising RB_c from 5 to 6 in the analysis of landslides in the Santa Susana Mountains results in a useful 24 % reduction of false negative results, from 70 to 52 polygons out of 1950, and a concomitant increase of false positive results by 16 % from 76 to 90 polygons (Table 1). However, increasing S_c to 15° or more may increase significantly the number of false positives but not necessarily the number of true positives as the most common type of amalgamation is related to multiple branches. An increase of false positives is not an issue, if amalgams detected by the algorithm are subsequently split manually. In that case, the operator can decide to leave an incorrectly diagnosed polygon intact. However, false negatives will go unnoticed and could have a large impact. Therefore, it is advisable to perform an additional manual check of the largest polygons in a data set, irrespective of how the classification algorithm has diagnosed them, especially for applications where the importance of polygons is proportional to their size. For example, one false negative within the 10 or 20 largest landslides in an inventory could significantly affect estimated total landslide volume.

A second, more fundamental issue is that the algorithm considers polygon geometry, in a way which does not allow detection of ellipsoid-shaped amalgams. Examples of this can be found, amongst others, in an inventory of landslides triggered by the 2008 Wenchuan earthquake (Gorum et al., 2011), where several landslides on the same slope were sometimes merged into larger, relatively smooth, polygons with a low K value and without any clear geometric or topographic indication of amalgamation (Fig. 2). In this case, image resolution may have been too low to distinguish the separate landslides, or the mapper may have simplified the geometry for convenience. For such amalgams, even if another criteria, such as alignment of the polygon long axis with the strike of the topographic slope, hints at possible amalgama-

tion, high-resolution imagery would be required to test the diagnosis, as single landslides with similar shape and orientation may exist. Merger of parallel landslide outlines due to image resolution limitations may cause errors of similar magnitude as other types of amalgamation, which are more easily detected, and could critically affect the common argument that at a given pixel resolution small landslides are missed but everything above a cutoff length scale of a few cells is properly mapped. Because the high-resolution imagery required to check visually for the occurrence of low-K amalgamation, or any other type of amalgamation, is rarely available to end-users of landslide inventories, it is important that it is mitigated for by those who develop the mapping techniques and acquire the landslide inventories. This may not always concur with the principal objectives of a particular mapping effort, for example in natural disasters when rapid assessment of the location and total extent of landslides is of the essence. However, if a landslide inventory is to be of general use to the research community, then the risk of amalgamation must be suppressed, both in manual and automatic mapping.

Suppression of landslide polygon amalgamation is hampered by deeper issues, such as image resolution and the uncontrolled subjectivity introduced in binary landslide mapping, where every pixel either is or is not a landslide. We draw into question the general assumption that in a given inventory, every landslide larger than a few image pixels is correctly mapped (e.g. Liao and Lee, 2000). Instead, it is reasonable to expect that many disturbed areas mapped as single medium to large landslides could in fact consist of groups of smaller landslides, giving potentially significantly different erosion volumes and size statistics (Fig. 2). Moreover, satellite imagery does not always yield unambiguous information about the number and shape of landslides, which occurred on a given slope. Where this applies, subjective choices of the mapper are crystallized within the landslide inventory. A Bayesian approach to mapping, aimed at delivering probabilistic instead of binary maps (e.g. Mondini et al., 2013) could be helpful in testing the different possibilities of split-

ting complex disturbed areas (see Fig. 2) and ultimately deliver more accurate, objective and reproducible data sets.

Short of a practicable, comprehensive solution, our method, which has a good reliability, can be used in several ways to mitigate for amalgamation in landslide maps, by helping the mapper to identify mistakes in automatic mapping, and the user to do the same in existing landslide maps. Notably, sorting mapped polygons by K value and size allows rapid, first order vetting of the largest landslides, which, when followed by manual splitting of amalgams, will be enough to yield a reasonable estimate of the total volume of landslides in an inventory. Then, for large populations, one could exclude all polygons with K values above a threshold and consider the correlation between the size or location of remaining landslides and physical parameters such as local topographic slope or triggering effects. Finally, a K value criteria might also be introduced in a semi automatic algorithm detecting landslides, to guide iteratively towards a sound splitting of adjoining landslides.

In the end, we must recall that amalgamation even if it may be a major source of errors, such as in the Wenchuan example, it is not the only one. Firstly, anthropogenic clearance or other disturbance of the landscape may be erroneous for a landslide, especially by automatic algorithms. Secondly, when scar, transport and deposit areas cannot be differentiated the volume of landslides with long runout may be substantially over-estimated. Thirdly, when landslides are reactivated and previously stable parts of the landscape are not implied, then it may be hard for the mapper to delineate the area of the actual failure with accuracy and this new failure may also not yield a volume as large as expected from area–volume relationships. These issues may be difficult to deal with but their effects will be suppressed when high resolution imagery is used by an experienced mapper. Additionally, systematic ways of dealing with these issues, such as the flagging of reactivated landslides, and the differentiation of the transport areas of debris flow or long runout landslides should be practised by mappers and also considered by users analysing old data.

7   Conclusions

We have shown that amalgamation, the bundling of several adjacent landslides into a single map polygon is a common problem in landslide inventories that has inflated estimates of landslide volumes by up to a factor of three, and the power-law exponent of landslide area–frequency distributions by up to 50 %. Even though the design of a comprehensive and fully reliable automatic corrective method remains a challenge, we have presented and tested a practical algorithm for automatic detection of amalgamated polygons based on geometric and topographic considerations. The algorithm performs well with an accuracy of 92.5 % and only 2.7–3.6 % amalgams missed and 3.9–4.8 % correct mapped polygons

incorrectly classified. It can, therefore, be used to automate the identification of landslide amalgams, accelerate the evaluation of data sets, and guide the manual correction of amalgams. Thus, our algorithm is a first step towards setting a quality standard for landslide maps in order to derive scientifically and societally useful variables, such as risk estimates, erosion rates, organic matter fluxes, or correlations between landsliding and physical triggers, as accurately as possible. Further challenges lie in attempting to automatically correct amalgamation and in assessing how mapping errors due to resolution blurring propagate into final products derived from landslide maps.

References

Brardinoni, F. and Church, M.: Representing the landslide magnitude-frequency relation: Capilano River basin, British Columbia, Earth Surf. Process. Landf., 29, 115–124, doi:10.1002/esp.1029, 2004.

Dai, F. C., Xu, C., Yao, X., Xu, L., Tu, X. B., and Gong, Q. M.: Spatial distribution of landslides triggered by the 2008 Ms 8.0 Wenchuan earthquake, China, J. Asian Earth Sci., 40, 883–895, doi:10.1016/j.jseaes.2010.04.010, 2011.

Fagnano, G.: Produzioni matematiche del conte Giulio Carlo di Fagnano, Gavelliana, 1750.

Frattini, P. and Crosta, G. B.: The role of material properties and landscape morphology on landslide size distributions, Earth Planet. Sci. Lett., 361, 310–319, doi:10.1016/j.epsl.2012.10.029, 2013.

Gorum, T., Fan, X., van Westen, C. J., Huang, R. Q., Xu, Q., Tang, C., and Wang, G.: Distribution pattern of earthquake-induced

landslides triggered by the 12 May 2008 Wenchuan earthquake, Geomorphology, 133, 152–167, 2011.

Guzzetti, F., Ardizzone, F., Cardinali, M., Rossi, M., and Valigi, D.: Landslide volumes and landslide mobilization rates in Umbria, central Italy, Earth Planet. Sci. Lett., 279, 222–229, 2009.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., and Chang, K.-T.: Landslide inventory maps: New tools for an old problem, Earth-Sci. Rev., 112, 42–66, doi:10.1016/j.earscirev.2012.02.001, 2012.

Harp, E. and Jibson, R.: Landslides triggered by the 1994 Northridge, California, earthquake, B. Seismol. Soc. Am., 86, S319–S332, 1996.

Harp, E. L., Wilson, R. C., and Wieczorek, G. F.: Landslides from the February 4, 1976, Guatemala earthquake, U.S.G.S., Prof. Pap., 1204 A, 1981.

Hilton, R. G., Meunier, P., Hovius, N., Bellingham, P. J., and Galy, A.: Landslide impact on organic carbon cycling in a temperate montane forest, Earth Surf. Process. Landf., 36, 1670–1679, doi:10.1002/esp.2191, 2011.

Hovius, N., Stark, C. P., and Allen, P. A.: Sediment flux from a mountain belt derived by landslide mapping, Geology, 25, 231–234, doi:10.1130/0091-7613(1997)025<0231:SFFAMB>2.3.CO;2, 1997.

Hovius, N., Stark, C. P., Hao Tsu, C., and Jiun Chuan, L.: Supply and Removal of Sediment in a Landslide Dominated Mountain Belt: Central Range, Taiwan, The J. Geol., 108, 73–89, doi:10.1086/jg.2000.108.issue-1, 2000.

Katz, O. and Aharonov, E.: Landslides in vibrating sand box: What controls types of slope failure and frequency magnitude relations?, Earth Planet. Sci. Lett., 247, 280–294, doi:10.1016/j.epsl.2006.05.009, 2006.

Larsen, I., Montgomery, D., and Korup, O.: Landslide erosion controlled by hillslope material, Nat. Geosci., 3, 247–251, 2010.

Lee, S. T., Yu, T. T., Peng, W. F., and Wang, C. L.: Incorporating the effects of topographic amplification in the analysis of earthquake-induced landslide hazards using logistic regression, Nat. Hazards Earth Syst. Sci., 10, 2475–2488, doi:10.5194/nhess-10-2475-2010, 2010.

Li, G., West, A. J., Densmore, A. L., Jin, Z., Parker, R. N., and Hilton, R. G.: Seismic mountain building: Landslides associated with the 2008 Wenchuan earthquake in the context of a generalized model for earthquake volume balance, Geochem. Geophys. Geosys, 15, 833–844, doi:10.1002/2013GC005067, 2014.

Liao, H.-W. and Lee, C.: Landslides triggered by Chi-Chi earthquake, in: Proceedings of the 21st Asian conference on remote sensing, Vol. 1, 383–388, 2000.

Lin, G. W., Chen, H., Hovius, N., Horng, M. J., Dadson, S., Meunier, P., and Lines, M.: Effects of earthquake and cyclone sequencing on landsliding and fluvial sediment transfer in a mountain catchment, Earth Surf. Process. Landf., 33, 1354–1373, doi:10.1002/esp.1716, 2008.

Malamud, B., Turcotte, D., Guzzetti, F., and Reichenbach, P.: Landslide inventories and their statistical properties, Earth Surf. Process. Landf., 29, 687–711, 2004a.

Malamud, B., Turcotte, D., Guzzetti, F., and Reichenbach, P.: Landslides, earthquakes, and erosion, Earth Planet. Sci. Lett., 229, 45–59, 2004b.

Martha, T. R., Kerle, N., Jetten, V., van Westen, C. J., and Kumar, K. V.: Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods, Geomorphology, 116, 24–36, doi:10.1016/j.geomorph.2009.10.004, 2010.

Meunier, P., Hovius, N., and Haines, A.: Regional patterns of earthquake-triggered landslides and their relation to ground motion, Geophys. Res. Lett., 34, L20408, doi:10.1029/2007GL031337, 2007.

Meunier, P., Hovius, N., and Haines, J.: Topographic site effects and the location of earthquake induced landslides, Earth Planet. Sci. Lett., 275, 221–232, doi:10.1016/j.epsl.2008.07.020, 2008.

Mondini, A. C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., and Ardizzone, F.: Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images, Remote Sens. Environ., 115, 1743–1757, doi:10.1016/j.rse.2011.03.006, 2011.

Mondini, A. C., Marchesini, I., Rossi, M., Chang, K.-T., Pasquariello, G., and Guzzetti, F.: Bayesian framework for mapping and classifying shallow landslides exploiting remote sensing and topographic data, Geomorphology, 201, 135–147, doi:10.1016/j.geomorph.2013.06.015, 2013.

Montgomery, D. R. and Dietrich, W. E.: A physically based model for the topographic control on shallow landsliding, Water Resour. Res., 30, 1153–1171, doi:10.1029/93WR02979, 1994.

Ouimet, W. B.: Landslides associated with the May 12, 2008 Wenchuan earthquake: Implications for the erosion and tectonic evolution of the Longmen Shan, Tectonophysics, 491, 244–252, doi:10.1016/j.tecto.2009.09.012, 2010.

Parker, R., Densmore, A., Rosser, N., De Michele, M., Li, Y., Huang, R., Whadcoat, S., and Petley, D.: Mass wasting triggered by the 2008 Wenchuan earthquake is greater than orogenic growth, Nat. Geosci., 4, 449–452, 2011.

Pelletier, J. D., Malamud, B. D., Blodgett, T., and Turcotte, D. L.: Scale-invariance of soil moisture variability and its implications for the frequency-size distribution of landslides, Eng. Geol., 48, 255–268, doi:10.1016/S0013-7952(97)00041-0, 1997.

Qi, S., Xu, Q., Lan, H., Zhang, B., and Liu, J.: Spatial distribution analysis of landslides triggered by 2008.5. 12 Wenchuan Earthquake, China, Eng. Geol., 116, 95–108, 2010.

Stark, C. P. and Guzzetti, F.: Landslide rupture and the probability distribution of mobilized debris volumes, J. Geophys. Res.-Earth Surf., 114, F00A02, doi:10.1029/2008JF001008, 2009.

Stark, C. P. and Hovius, N.: The characterization of landslide size distributions, Geophys. Res. Lett., 28, 1091–1094, doi:10.1029/2000GL008527, 2001.

Xu, C., Xu, X., Yao, X., and Dai, F.: Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis, Landslides, 11, 441–461, doi:10.1007/s10346-013-0404-6, 2014.

Yanites, B., Tucker, G., Mueller, K., and Chen, Y.: How rivers react to large earthquakes: Evidence from central Taiwan, Geology, 38, 639–642, 2010.