Landslide shape, ellipticity and length-to-width ratios

Faith E. Taylor\textsuperscript{a}, Bruce D. Malamud\textsuperscript{a}, Annette Witt\textsuperscript{b}, Fausto Guzzetti\textsuperscript{c}

\textsuperscript{a}Department of Geography, King’s College London, Strand, London, WC2B 4BG, United Kingdom

\textsuperscript{b}Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany

\textsuperscript{c}Istituto di Ricerca per la Protezione Idrogeologica, Consiglio Nazionale delle Ricerche, 06128 Perugia, Italy

Correspondence to: Faith Taylor, Department of Geography, University of Portsmouth, Portsmouth, PO1 3QL, United Kingdom faith.taylor@port.ac.uk
ABSTRACT: This paper presents a new methodology to systematically quantify the shape of landslides by their ellipticity ($e_L$) and length-to-width ratio ($\Lambda_L$), along with variability in these measures over different geomorphic settings. Two large substantially complete triggered-event landslide inventories (source area and runout) are used: (i) 11,111 earthquake-triggered landslides (1994 Northridge, USA) and (ii) 9,594 rainfall triggered landslides (1998, Guatemala). Three methods are trialled to abstract landslide shapes to ellipses. The best method fits a convex hull to each landslide shape, approximates an ellipse with the equivalent convex hull area and perimeter, and scales this ellipse to match the original landslide area. An ellipticity index ($e_L$) is used based on the intersection of the original landslide shape and the elliptical approximation. We consider an ellipse a reasonable approximation of landslide shape if $e_L \geq 0.5$ (>80% of the two landslide inventories). Landslides with $e_L < 0.5$ reflect processes such as coalescence. We calculate for each landslide an ellipse length-to-width ratio ($\Lambda_E$), finding $1.2 \leq \Lambda_E \leq 15.1$. The statistical distributions of $\Lambda_E$ are examined for ten categories of landslide area ($A_L$). An inverse-Gamma probability density function is found to be a good statistical model for landslide $\Lambda_E$, with model parameters dependent on landslide area category. As landslide area $A_L$ increases, $\Lambda_E$ tends to decrease for the Northridge (earthquake-triggered) inventory and increase for Guatemala (rainfall-triggered). In three additional (rainfall-triggered) landslide inventories, $\Lambda_E$ trends are similar to Guatemala. Our findings show that (i) an ellipse is a reasonable model for >80% of landslide shapes across different geomorphic settings, (ii) those landslides significantly deviating from an ellipse can be related to landscape processes, (iii) the length-to-width ratios of ellipses are non-normally distributed, with implications for modelling landslide hazard and risk. Supplementary material includes code so that the new methodology may be applied more widely.
KEYWORDS: landslide; shape; length-to-width; ellipse; inverse-Gamma; statistics

STATISTICS

Words: 300 (abstract)
11,870 (manuscript main text)
2113 (figure and table captions)

References: 105

Figures: 14
Tables: 4

Supplementary Material A to E: 2475 words, 2 figures
Introduction

This paper addresses the question ‘what shape is a landslide?’ We do this by (i) developing a systematic way to approximate landslide shapes to ellipses, (ii) quantitatively examining the extent landslide shapes vary from an ellipse, and (iii) examining the statistical variation in landslide shape, as measured by the landslide ellipse length-to-width ratio. Understanding the shape and variation of landslide shapes has the potential to give better insights into geomorphological process, spatially explicit simulation modelling, landslide hazard and the intersection of the landslide hazard with infrastructure (exposure). In this introduction we motivate the research in the rest of this paper by giving a background on both general and landslide shape studies.

Shape is the geometrical information that describes an object that remains when location, scale and rotation are filtered out from an object (Kendall, 1984). Across many disciplines, shape and process are often linked (e.g., Hagget and Chorley, 1969; Barrett, 1980; Simon and Darby, 1997; McLellan and Endler, 1998), with shape frequently considered to be the result of a process (e.g., de Boer, 1992). However, it is known that multiple processes can result in the same shape (e.g., Beven, 1996).

Considerable effort across a number of disciplines is devoted to developing systematic ways to quantify the shape of features, with examples including the following: variability in shape of animal skeleton relative to habitat (e.g., Adams and Rohlf, 2000); shape characteristics of different land cover types as a result of faulting and thrusting (e.g., Li et al., 2001); shape of cities and sub-areas within cities as a way to explain urban growth processes (e.g., Batty, 2008).

When mapping landslides, shape refers to the form of the topographic surface modified by the presence of a landslide at a given stage in time. A landslide is “a movement of a mass of rock, earth or debris down a slope” (Cruden, 1991, p.28) and can be conceptualised as a “physical system that develops in time through stages” (Hungr et al., 2014, p.1). The work we develop in this paper primarily focuses on inventories of landslides that have undergone a period of ‘rapid’ or greater movement (i.e., >1.8 m hr⁻¹, Cruden and Varnes, 1996) in a triggered event, although methods developed are also theoretically applicable to slower moving forms of landslide.

The topographic surface is considered by many to be the single most useful
characteristic for the detection and classification of landslides from aerial photographs or other remotely sensed information (Guzzetti et al., 2012). For landslides, different shapes may suggest different types of physical process. For example, high-mobility flows can result in different shapes compared to low-mobility slides (Legros, 2002; Yang and Lee, 2006). However, different processes may result in similar shapes, such as lava flows, rock glaciers, mud flows and debris flows (Beven, 1996).

Here, we focus on the two-dimensional shape of landslides obtained from the landslide planimetric (planar) area, encompassing (i) the landslide source (erosion) area, where the failed material is mainly depleted (eroded), and (ii) the landslide travel and deposition (runout) area, where the landslide material mostly travels and accumulates. For many landslides, the separation between the source, travel, and deposition areas is uncertain, or impossible (Stark and Guzzetti, 2009). Deposition occurs locally in the source and transport areas, and material is eroded in the landslide travel area, particularly for high-mobility landslides such as earth flows, mud flows, and debris flows (Cruden and Varnes, 1986; Hungr et al., 2014).

Although the shape that outlines the combined (source, travel, deposition) landslide area may represent different physical processes (Hovius et al., 2000), we consider in this paper the combined area. We note that some of the previous work on landslide shape (introduced later in Table I) only considers the shape of the source area or does not state what part(s) of the landslide are considered. This means that for some previous studies a direct comparison between descriptions of landslide shape is not possible, highlighting the utility of a more systematic way of describing landslide shape, which we will introduce in this paper.

To illustrate some typical landslide shapes found in nature we present two figures. Figure 1 shows seven landslide shapes across different geological settings and landslide triggering regimes. In Figure 2 we extract seven loosely analogous landslide shapes from an inventory of landslides triggered by Hurricane Mitch in Guatemala (Bucknam et al., 2001). The shape of landslides in Figures 1 and 2 are typically asymmetrical and ‘complex’ in that we could not scale, rotate or translate the shape to fit into a broad category of simple geometric shape (e.g., square, circle). Perhaps because of this variability and irregularity, within the landslide literature landslide ‘shape’ is commonly described qualitatively, based on the type of movement and material involved (Cruden and Varnes, 1986; Hungr et al., 2014), and quantitative
analysis of shape tends to be limited to small numbers of landslides (see Table I). Landslide shapes (of both source area and combined source area and runout) are frequently referred to as ‘elliptical’ (Hovius et al., 1997; Martel, 2004; Marchesini et al., 2009; Martha et al., 2010; Pourghasemi et al., 2014), the validity of which is confronted later in this paper.

The following are some of the physical factors that authors have noted that affect the shape of landslides (including source area, deposition and run out): landslide type (Dikau et al., 1996), topography (Guthrie and Evans, 2004), history of landsliding (McCalpin, 1984), history of landsliding in the immediate vicinity (Samia et al., 2016) and wet or dry triggering mechanisms (Legros, 2002; Yang and Lee, 2006). In addition, soil characteristics have been noted to affect the shape of landslide source areas (Klar et al., 2011; Lehman and Or, 2012; Milledge et al., 2014) and combined source area and run out (Cardinali et al., 2000). Moreover, the methods used to map landslides (Santangelo et al., 2015), and to produce landslide inventories (Guzzetti et al., 2012) can result in differences in the shape of each landslide recorded. For example, the level of detail the landslide perimeter is mapped to (Santangelo et al., 2015), the landslide age at the time of mapping (McCalpin, 1984), and whether or not the source area and runout are separated or combined (Guzzetti et al., 2012).

Some geometrical characteristics of polygons such as area and perimeter have unambiguous, well-established methods (e.g., Meister, 1769; Euler, 1773, described in Michon, 2015) and software tools for quantification (e.g., ESRI, 2016). These methods are commonly applied to landslides, and has led to new insights about the underlying statistical behaviour of landslide areas, and how this might link to process (e.g., Stark and Hovius, 2001; Malamud et al., 2004; Katz et al., 2006; Stark and Guzzetti, 2009). However, systematically describing the shape of asymmetric, irregular forms is not a trivial task, as shape does not have one broadly accepted
method of quantification. We later argue that this is why descriptions of landslide shape are often qualitative and why there is a need for a generally applicable method to systematically quantify landslide shape. Forman (1995, p. 135) notes, “no single measurement or index can unambiguously differentiate all shapes”, and states that optimal methods will satisfy the following four criteria: (i) be easy to calculate, (ii) be applicable to the entire region of interest, (iii) allow the quantitative differentiation of different types of shape, and (iv) allow a shape to be plotted based on the information given in that index.

Across many disciplines, considerable effort has been devoted to the development of methodologies to measure shape (e.g., Barrett, 1980; Li et al., 2001; Benediktsson et al., 2003; Adams et al., 2004; Slice, 2007). Methods to quantify shape include (i) geometric morphometric techniques, based on digitising a set of key points on a shape, and examining the relationships between those points (Adams and Rohlf, 2000), (ii) shape indices, which examine relationships between area and perimeter of shapes (Li et al., 2001), and (iii) shape similarity measures where irregular shapes are compared to regular shapes (Lombardo, 2014).

Within the landslide literature, there is a relatively small body of work quantitatively investigating landslide shape, which we break up into the following three broad categories typified by five exemplar studies: (i) perimeter-area indices (Pourghasemi et al., 2014), (ii) shape similarity to a circle measure (Samia et al., 2016), and (iii) scaling relationships between landslide dimensions and area (Hovius et al., 1997; Guthrie et al., 2008; Milledge et al., 2014).

In the five exemplar studies given above, the most common landslide shape measure was the relationship between landslide length \( (L_L) \) and landslide width \( (W_L) \), typified by the following terms: ‘length-to-width ratio’, ‘length/width’, ‘aspect ratio’, ‘elongation ratio’ and ‘geometrical characteristics’. Using Google Scholar and these five landslide shape terms, we identified 21 key peer-reviewed studies from 1983 to 2017 that characterizes landslide length-to-width ratio (which we refer to as \( \Lambda_L \), where \( \Lambda_L = L_L / W_L \)). In Table I these 21 studies are grouped into three categories (and in brackets, the number of studies that are included in that category, with each study assigned to only one category):

- Category I. Summary values of \( \Lambda_L \) are given, but landslide shape is not the main
focus (6 studies).

- Category II: An in-depth analysis of landslide shape is performed (9 studies).
- Category III: A relationship between landslide area and $\Lambda_L$ is given (6 studies).

The 21 studies use 33 different landslide inventories, representing a range of methods of production and types of inventories (as defined by Guzzetti et al., 2012): triggered event ($n = 16$ inventories), seasonal ($n = 2$), multi-temporal ($n = 4$), historical ($n = 6$) and other/unknown ($n = 5$). As we discuss later in the Data section, the production methods will affect the completeness of the inventory in terms of representing the full ‘population’ of landslide shapes.

The 21 studies in Table I use terms such as ‘lobate’, ‘isosceles triangle’, ‘long and thin’, ‘irregular’, ‘elongate’, ‘rectangular’, ‘spoon-like’, ‘amphitheatre’ and ‘cone-shaped’ to describe landslide forms. Of the 21 studies, landslide shape is defined by the following elements (in brackets percentages of studies): source area (24%), source and runout (transport and depositional) areas combined (62%), and not defined (14%).

Where the landslide length-to-width ratio ($\Lambda_L$) is measured (15 of the 21 studies), this typically ranges $0.3 < \Lambda_L < 10$ with the mean of the stated mean values $\Lambda_L = 1.9$ and standard deviation $= 0.6$ (8 studies) and where stated the median ranges $1.26 < \Lambda_L < 4.2$ (6 studies). The definitions of length and width vary between authors and in 11 of the 21 studies it is not fully clear how the landslide dimensions have been calculated.

As we will discuss later, a ‘complete’ triggered landslide event inventory would contain all landslides associated with a given trigger across the entire region where landslides occurred (Harp et al., 2011). A substantially complete landslide inventory for statistical purposes would contain a representative sample of landslides across all scales and geomorphological regions associated with a given trigger. Of the 33 landslide inventories analysed in Table I, 16 (48%) of these inventories are triggered event inventories, but 7 of these 16 inventories consider either a subset of landslide types (e.g., debris flows) or a subset of the region affected by landslides associated with the
trigger. The remaining 9 of these 16 triggered-event inventories are based on relatively small numbers of landslides (e.g., < 600) within each inventory. A landslide inventory that is not substantially complete may potentially exclude ‘end member’ landslide shapes from analysis.

The work we present in this paper aims to build upon existing research that examines landslide length-to-width ratio. We do this by developing a systematic measure of $\Lambda_L$ and applying this measure to large numbers of landslides that are digitised as areas represented in planimetric form. We first develop an ellipse length-to-width ratio ($\Lambda_E$) method (and corresponding tool – see Supplementary Material D) that satisfies the four criteria of Forman (1995) given above, such that the method we develop:

- is easy to calculate and apply to large numbers of digitised landslides in a systematic way;
- is applicable to digitised landslide inventories from any region, thus allowing comparison of populations of landslide shapes from different locations;
- allows the quantitative differentiation of different landslide shapes within and between inventories to investigate landslide shape probabilistically; and
- allows landslide shape to be plotted as an ellipse whose dimensions vary for modelling purposes.

We then apply our ellipse length-to-width ratio method to two large triggered landslide event inventories which are considered to be substantially ‘complete’. This allows us to investigate the underlying statistical variability in landslide shape when looking at the complete ‘population’ of potential landslide shapes and areas. By systematically and quantitatively measuring the shape of ‘populations’ of landslides and investigating the relationship between shape and area, we aim to create a link to the growing body of literature focused on the statistical characterisation of landslide areas (e.g., Stark and Hovius, 2000; Malamud et al., 2004; Katz and Aharonov, 2006; Stark and Guzzetti, 2009). The statistical characterisations of landslide shape we establish can be used in spatio-statistical modelling of triggered landslide events. To aid the reader, Table II summarises variables and abbreviations used in this paper.

This paper is organized in the five following major sections: (i) Data describes the data, landslide inventories from the USA and Guatemala, and why these were selected. (ii)
Methods includes an outline of how we approximate landslide shapes to ellipses and measure the length-to-width ratio of those ellipses in a systematic way. We then present (iii) Results from analysing how elliptical the landslide shapes are, and the variability in landslide shape, as measured by ellipse length-to-width ratio. Finally, we give a (iv) Discussion of the results in terms of potential applications and (v) Conclusions. Also included are supplementary material. Supplementary Material A to C give more detail on the methodology and Supplementary Material D and E give links to the model code to apply the methodology to other landslide inventories.

Data

In this section, we discuss the two substantially complete landslide inventories that we use for the majority of this paper to explore landslide shapes. We first describe the criteria by which we select the inventories, and then give details on the inventories themselves.

The criteria we adopt for the creation of landslide inventories used aims to maximise (a) the detail in which landslide shapes are mapped and (b) the statistical representativeness of the sample in terms of landslide areas and geographic settings. The criteria are (with the first two criteria modified from Harp et al., 2011):

(i) Are the landslides resulting from the triggered event present? Imagery used to aid in constructing the landslide inventory (triggered-event, seasonal or multi-temporal) is taken in a short-enough time period after the landslides have occurred to avoid significant under-sampling of smaller landslides. Historical inventories were excluded as smaller landslides are proportionally and systematically under sampled (Guzzetti et al., 2012) due to rapid erosion, and/or revegetation or remediation of smaller landslides (Bell et al., 2012).

(ii) Is the imagery at a high enough resolution to see and record the landslides at all scales systematically? The inventory is created through the use of < 1:60,000 aerial photography over the majority of the affected region, coupled with field investigations. Landslides are systematically mapped down to a small enough
scale so that landslides across all scales are represented. Consideration should also be given to factors such as the flying height, sensor quality and mapping tools (Santangelo et al., 2015) although these details are often not fully described.

(iii) *Does the landslide inventory conform to established statistical behaviour of substantially complete inventories?* Landslides in the inventory follow reasonably well already established landslide frequency-area distributions for substantially complete inventories in the literature (e.g., Stark and Hovius, 2001; Malamud et al., 2004; Van Den Eeckhaut et al., 2007) to ensure representative sampling of different landslide area size categories.

For the primary stages of this research, we require the total number of landslides ($N_L$) in a substantially complete inventory to be large (e.g., $N_L > 7000$). This is required to investigate the probability density distribution of landslide length-to-width ratio ($\lambda_L$) within sub-categories of landslide area ($A_L$), for example, $10 \leq A_L < 100 \text{ m}^2$. For future applications of the methods established in this paper, this criterion may not necessarily be required, as results can be compared to the probability density distributions that we will establish in the Results section.

Of freely available triggered landslide event inventories, we found that the following two inventories matched our criteria well:

(i) *Northridge*: 11,111 landslides triggered by the $M_w = 6.7$, 17 January 1994 Northridge Earthquake in California, USA (Harp and Jibson, 1995). A subset of the Northridge landslide inventory is shown in Figure 3A. According to Harp and Jibson (1995), landslides were systematically mapped immediately after the earthquake using extensive field surveys and 1:60,000 scale aerial photography, with polygons mapped constituting both source and runout areas. The authors state that landslides in the inventory were predominately shallow, highly disrupted falls and slides, occurring in weakly cemented sediments. They estimate that the number of deeper slumps and slides was of the order of hundreds. In an analysis of a high intensity region of landsliding within this 1994 Northridge landslide inventory, Parise and Jibson (1997) found that 4% of landslide polygons were landslide complexes where it had not been possible to delineate individual landslide boundaries due to coalescence. These landslide complexes tend to have a branched, ‘comb-like’ appearance.

(ii) *Guatemala*: 9,594 landslides triggered by heavy rainfall in late October and early
November 1998 from Hurricane Mitch in Guatemala (Bucknam et al., 2001). A subset of the Guatemala inventory is shown in Figure 3B. According to Bucknam et al. (2001), landslides were mapped using 1:40,000 scale aerial photography taken in January – March 2000, with polygons mapped constituting both source and run out areas. They further stated that landslides were primarily debris flows, occurring on moderate to steep hillslopes across diverse geologies, and that the landslides triggered could be broadly split into two types: (a) (the majority) small translational and rotational slides of which a significant proportion mobilised into debris flows; (b) large translational slides of which some mobilised into debris flows that followed river channels, resulting in large, branched shaped polygons.

The authors of the Northridge (Harp and Jibson, 1995) and Guatemala (Bucknam et al., 2001) inventories estimate that the inventories are substantially complete down to landslides with areas \( A_L \approx 25 \text{ m}^2 \) (Northridge) and \( A_L \approx 225 \text{ m}^2 \) (Guatemala). Malamud et al. (2004) discuss the substantial completeness of these two landslide inventories down to small landslide areas, and that both inventories follow the same three-parameter frequency-area distribution (an inverse-Gamma) in terms of landslide area. Malamud et al. (2004) further hypothesize that the inverse-Gamma distribution they have found (along with corresponding parameters) are representative of many landslide-triggered events for low- to medium-mobility landslides. We believe that these two substantially complete landslide event inventories are statistically representative of landslide shapes generally observed in triggered landslide events.

Methods

The Introduction section discussed the assumption of some authors that landslide shapes are approximated by ellipses. Within this context, this section first outlines and compares three methods to abstract landslide shapes to ellipses and second introduces an index to quantify how elliptical landslides are.
Abstracting landslide shape to an ellipse

The bottom box (Step F) in Figure 4 (described later) shows a landslide shape (grey solid line), which is asymmetrical and irregular, and has a given area \( A_L \) and perimeter \( P_L \). Figure 4 also shows an elliptical approximation (black solid line) of the landslide shape, which is defined by its dimensions: area \( A_E \), perimeter \( P_E \), length \( L_E \), width \( W_E \), semi-major axis \( a \), and semi-minor axis \( b \). The ellipse can be characterised by its length-to-width ratio, \( \Lambda_E \). There are many methods that can be used to abstract an irregular shape to an ellipse. Each method will result in ellipses with different areas and length-to-width ratios. Table III outlines three elliptic approximation methods that we have considered. We use the notation \( M \) to indicate the elliptical approximation method with variables (explained in detail further below) inside square brackets:

- (Method 1) Landslide area and perimeter, \( M[A_L, P_L] \)
- (Method 2) Convex hull (CH) fit to landslide shape, \( M[A_{CH}, P_{CH}] \)
- (Method 3) Convex hull (CH) fit to landslide shape and scaled to \( A_L \), \( M[(A_{CH}, P_{CH}) \rightarrow A_L] \)

These methods produce an ellipse based on the original landslide shape from a series of steps, and the landslide \( \Lambda_E \) is approximated from the ellipse’s area and perimeter.

In addition, we experimented with using existing GIS tools to calculate an ellipse from a polygon (using the standard deviational ellipse algorithm described by Yuill, 1971). We found the GIS tool computationally demanding and not producing significantly improved elliptical approximations of landslides in terms of goodness-of-fit, when applied to a small sample of landslide shapes; therefore, this method is not used further in our studies.

[Table III]

In Figure 4 we introduce a flowchart of Steps A to F used for elliptic approximation Methods 1 to 3 as outlined in Table III. Each elliptic approximation method (Methods 1 to 3) uses a different combination of steps A to F shown in Figure 4 which we outline in more detail in Table IV. Steps A to F are justified in further detail in Supplementary
Material A. In the analysis performed here, we define length \( L_E \) as the longest axis of the landslide. This long axis does not necessarily correspond to the direction of the landslide runout. In reality only a small number of landslides are wider than they are long relative to the down-slope direction (Gabet and Dunne, 2002; Rickli et al., 2008; Marchesi et al., 2009).

In Step A of the process (Figure 4, Table IV), the centre of gravity and orientation of the original landslide are calculated for Methods 1 to 3. In Step B (Figure 4, Table IV), a convex hull is fit to the landslide polygon for Methods 2 and 3. For Step C (Figure 4, Table IV) the length-to-width ratio \( \Lambda_E \) for an idealised elliptic shape is calculated using the solution to a quadratic equation, based on the equations for ellipse area (Heath, 1931) and perimeter (Euler, 1773 described in Michon, 2015) outlined in Supplementary Material A. By combining equations for ellipse area and perimeter, we can solve for the length-to-width ratio of an ellipse \( L_E/W_E \), in terms of the area \( A_E \) and perimeter \( P_E \) of the ellipse, by using the solution to a quadratic equation, giving,

\[
\Lambda_E = \frac{L_E}{W_E} = \frac{4\pi A_E}{P_E^2 - \sqrt{P_E^4 - 16\pi^2 A_E^2}} \quad \text{Eq. (1)}
\]

We then substitute into Eq. 1 the area \( A_E \) and perimeter \( P_E \) of the ellipse with the area and perimeter of original landslide shape \( A_L, P_L \), or the convex hull (CH) fit to the landslide shape \( A_{CH}, P_{CH} \), thus approximating the length-to-width ratio \( \Lambda_E \) of an idealised elliptic shape with the same area and perimeter as that of the original landslide or the CH fit.

When substituting the original landslide \( P_L \) for the ellipse \( P_E \), Eq. 1 becomes sensitive to the original landslide’s perimeter ‘sinuosity’ relative to its area \( A_L \). If the original landslide has a ‘sinuous’ perimeter \( P_L \), this has the potential to force the idealised
elliptic shape to longer, thinner forms (shown in Figure 5). For this reason, in Method 3 we fit a CH to the original landslide shape which removes most internal voids from the landslide shape.

In Step D (Table IV) ellipse length \((L_E)\) and width \((W_E)\) can be expressed in terms of the length-to-width ratio of the ellipse \(\Lambda_E\) and the landslide area \((A_L)\) or area of the CH fit \((A_{CH})\) which we denote as \(A:\)

\[
L_E = 2\sqrt{\frac{A_{CH}}{\pi}}
\]

\[
W_E = 2\sqrt{\frac{A}{\pi\Lambda_E}}
\]

Eq. (2)

In Step E (Table IV), for elliptic approximation Method 3 the length \((L_E)\) and width \((W_E)\) of the ellipse approximated from a CH fit to the original landslide shape is scaled so that the area of the CH \((A_{CH})\) matches the area of the original landslide,

\[
L_{Scaled \, CH} = L_{CH} \sqrt{\frac{A}{A_{CH}}}
\]

\[
W_{Scaled \, CH} = W_{CH} \sqrt{\frac{A}{A_{CH}}}
\]

Eq. (3)

This scaling does not affect the ‘shape’ of the ellipse in terms of \(\Lambda_E\), but can result in a better match in terms of area between the idealised elliptic shape \((\Lambda_E)\) and the original landslide shape.

Figure 5 shows examples of elliptical approximations of 12 landslides and their resulting ellipse length-to-width ratio \((\Lambda_E)\) from the Guatemala inventory for Methods 1 to 3.

As shown in Figure 5, many landslides are approximately elliptical in shape, but a few are not, and this varies with method of ellipse fitting. Landslides that are not well modelled by an ellipse tend to be mainly landslide complexes and debris flows, where
complexes are defined as multiple landslides mapped as one polygon due to difficulty delineating the boundaries of individual landslide (Guzzetti et al., 2012). However, as coalescence of individual landslides into a complex may also result in an elliptical shape (in addition to non-elliptical shapes), and vice-versa a single landslide process may result in a non-elliptical shape, one cannot automatically remove landslide complexes in our methodology.

Ellipticity index

To evaluate the degree to which a landslide is elliptical, we use an ellipticity index \( e \) developed by Lombardo (2014) for measuring lake shape and adapt it for landslide shapes:

\[
e_L = 1 - 2 \frac{A_L - A_L \cap E}{A_L}
\]

where \( e_L \) = landslide ellipticity index, \( A_L \) = original landslide area, \( \cap \) represents ‘intersection’, and \( A_L \cap E \) = area of intersection (hashed areas in Figure 5) between original landslide shape and its idealised elliptic shape. The quantity \( (A_L - A_L \cap E) \) is the landslide area not covered by the ellipse.

The ellipticity index ranges \(-1.0 \leq e_L \leq 1.0\). A completely imperfect fit with 0% overlap between the original landslide shape \( (A_L) \) and the idealised elliptic shape \( (A_E) \) gives \( A_L \cap E = 0.0 \) and \( e_L = -1.0 \). A ‘perfect’ fit with 100% overlap between the original landslide shape \( (A_L) \) and the idealised elliptic shape \( (A_E) \) gives \( A_L \cap E = A_L \) and \( e_L = 1.0 \). Similarly, the relationship of the landslide ellipticity index and overlap (original landslide shape and idealised ellipse) is \( e_L = -0.5 \) (25% overlap), \( e_L = 0.0 \) (50% overlap), and \( e_L = 0.5 \) (75% overlap). When intersecting the original landslide polygon and ellipse in the GIS, a small number of landslides (0.6% Northridge inventory, 3.0% Guatemala inventory) are split by our procedure into non-contiguous sections, generally reflecting long, thin, sinuous landslides (i.e., debris flows that follow the river channel morphology), where an ellipse is a poor approximation of landslide shape. In these cases, the landslides are removed from analysis.

For all landslides in the Northridge and Guatemala inventories, the ellipticity index \( e_L \) of landslide ellipses approximated using Methods 1 to 3 was calculated and is shown
in **Figure 6**. **Figure 6A** and **B** show results from the Northridge inventory separated by elliptical fitting method. **Figure 6C** shows results from Method 3 for both the Northridge and Guatemala inventories.

**Figure 6A** shows the relationship between length-to-width ratio and ellipticity for each method of ellipse fitting (**Table III**). For both Northridge and Guatemala, the method resulting in the highest values of $e_L$ is Method 2 ($M[A_{CH}, P_{CH}]$, convex hull). However, Method 2 may not always be most appropriate as the CH (convex hull) minimum bounding geometry is forced towards larger areas in order to contain the entire landslide shape, and thus larger elliptical approximations of that CH, which can also be seen in **Figure 5**. The larger the area of the elliptical approximation, the greater the probability that there will be a large area of intersection ($A_{L\cap E}$) between the ellipse and the landslide shape, and thus a high value of $e_L$. In addition, Method 1 is particularly sensitive to the level of detail the original landslide perimeter has been mapped to which has potential to vary depending on the geomorphologist(s) preparing the inventory (Santangelo et al., 2016). Method 1 often results in long, thin elliptical approximations when there is a high perimeter to area ratio, hence the greater range of values of $\Lambda_E$ (**Figure 6B**). Consequently, we select elliptic approximation Method 3 (Convex hull fit to landslide shape scaled to $A_L$) as the best approximation. This scaled convex hull (CH) method provides a trade-off between lower values of $e$, but a more realistic approximation of most landslide shapes.

**Figure 6C** illustrates that for both the Northridge and Guatemala inventories, the overall probability density distribution of $e_L$ follows a similar pattern. For Northridge (Guatemala) 97% (95%) of landslide-ellipses have ellipticity index $e_L > 0.0$, corresponding to $\geq 50\%$ overlap between each original landslide shape and the approximated ellipse.

**Landslide ellipticity index and threshold values**

To assess a reasonable ‘threshold’ value of $e_L$ for landslide elliptic approximation, we
used Method 3 (Table III) to calculate ellipticity $e_L$ for landslides in both event inventories. Then a detailed visual inspection was performed, with elliptical approximations ordered smallest to largest, and a sample of landslides and ellipses inspected in 0.1 increments of $e_L$ (Figure 7).

From this inspection (Figure 7), a threshold of $e_L = 0.5$ (75% overlap between the original landslide shape and the idealised ellipse) was assumed to be a reasonable cut-off between landslides that are well approximated ($e_L \geq 0.5$) and those that are not well approximated ($e_L < 0.5$) by an ellipse. The percentage of the total number of landslides with $e_L < 0.5$ (i.e., ‘non-elliptical’) were found to be as follows: 15% (1,670 landslides) of the Northridge inventory and 18% (1,736 landslides) of the Guatemala inventory. In other words, a substantial percentage (82–85%) of the landslides in these two inventories are ‘close to’ elliptical ($e_L \geq 0.5$). This suggests that the assumption within the literature that ‘landslides are elliptical’ (see Introduction section) is reasonable for landslide inventories with geomorphic attributes similar to Northridge and Guatemala.

Across and within both inventories, landslides with $e_L \geq 0.5$ represent a range of landslide types, environments and triggering mechanisms, potentially indicating a convergence of near ellipse form from either (i) different lower-level processes individually resulting in the same form, (ii) the interaction of different lower-level processes resulting in alike shapes or (iii) high-level processes being very similar for many landslide types. Those landslides with $e_L < 0.5$ (e.g., Categories A, B and C in Figure 7) perhaps point to a more interesting distinction in our ability to map landslides, in that they tend to represent polygons where there is a merging of individual low- or high-mobility landslides.

In Northridge, most landslides with $e_L < 0.5$ are landslide complexes, where several smaller landslides have merged, making it difficult to distinguish the individual boundary of each form (Harp and Jibson, 1995, Marc and Hovius, 2015); an example is given in Figure 8A. In Guatemala, many landslides with $e_L < 0.5$ are debris flows that flow to the valley bottom and then follow the river channel morphology, converging
with other debris flows; an example is given in Figure 8B. In both cases, the resultant landslide polygon is the product of multiple source areas and/or run outs. Particularly for flow-type landslides in Guatemala, it may be more appropriate to look to the river morphology literature (e.g., Knighton, 2014) to understand and explain the shape of the landslides.

This section has illustrated that for each landslide, the method of ellipse fitting may result in different values of length-to-width ratio and goodness-of-fit in terms of ellipticity. We find Method 3 (scaled convex hull) to be most appropriate, and this results in >80% of landslides being well represented by an ellipse. With further investigation, the degree to which a landslide is elliptical may give an indication of certain dynamics such as merging and runout.

Results
In this section we examine the distribution of landslide length-to-width ratio for all landslides where an ellipse is a reasonable fit ($e_L \geq 0.5$) for landslide ellipses created using Method 3 (scaled convex hull) described above. The section is split into three sub-sections examining: (i) overall distribution of $\Lambda_E$, (ii) distribution of $\Lambda_E$ by landslide area category and (iii) application of Method 3 to three additional landslide inventories.

Distribution of landslide ellipse length-to-width ratio
In Figure 9, we give the results of length-to-width ratio $\Lambda_E$ for all landslides in the Northridge and Guatemala inventories where an ellipse (using Method 3) was found to be a reasonable approximation of landslide shape ($e_L \geq 0.5$). Figure 9A shows density box plots, which are symmetrical around the vertical axis and indicate what proportion of the data lies at a given $\Lambda_E$. For both Northridge and Guatemala, the density plots are a similar shape, with minimum values of $\Lambda_E \approx 1.3$ and 1.2 respectively, a peak in density (plot width) at $\Lambda_E \approx 2$ and a broad spread of the largest values of $\Lambda_E$.
up to a maximum of 15.1 and 13.9 respectively. In both Northridge and Guatemala, the median ($\Lambda_E = 2.5$ and 2.2) lies below the mean ($\Lambda_E = 2.9$ and 2.4), indicating a positively skewed distribution. The $25^{th}$ percentile is $\Lambda_E = 2.0$ for Northridge and $\Lambda_E = 1.8$ Guatemala. For Northridge (Guatemala) < 1% (<5%) of landslides can be considered very 'compact' where $\Lambda_E < 1.5$ ($\Lambda_E = 1.0$ for a circle).

[Figure 9]

Separation of landslides by area category

To further explore the distribution of $\Lambda_E$, landslide areas for both inventories were split into multiple area categories. As landslide areas are known to span several orders of magnitude (e.g., see Pelletier et al., 1997, for an early review), and medium and large landslide areas are often well described by an inverse power-law distribution (e.g., Van Den Eeckhaut et al., 2007, for a review), landslides here were split into ten approximately logarithmically increasing categories of $A_L$ and $\rho(\Lambda_E)$ calculated for each area category and inventory (Figure 10).

Figure 9B shows that (broadly) for the Northridge landslide event inventory, the $\Lambda_E$ interquartile range (height of the coloured boxes) and median (bold horizontal line) decreases with $A_L$, whereas the opposite is observed in the Guatemala landslide inventory. A Kolmogorov–Smirnov (K–S) two-sample test (Lillefors, 1967) was used to compare the distribution of $\Lambda_E$ for each of the two landslide inventories in the multiple landslide area categories, outlined in detail in Supplementary Material B.

Using a significance level of $p = 0.10$, results in 38 out 45 pairs of $A_L$ categories where the null hypothesis is rejected that the pdf of $\Lambda_E$ is the same in both $A_L$ categories. The $p$ values are generally higher in neighbouring landslide area categories for the ‘tails’ of the categories (i.e., very small and very large landslide areas), which we attribute to small sample sizes in these categories. Broadly speaking, we can say that there is a difference in the distribution of $\Lambda_E$ between pairs of landslide area categories, with some discrepancies.

We now determine the best-fit pdf to the observed $\Lambda_E$ in each landslide area category. TableCurve2D (Sigmaplot, 2015) was used to compare the empirical cumulative
distribution function of $\Lambda_E$ within each landslide area category with 2292 cumulative distribution function fits to the data, and a Lorentzian minimization of errors used for the non-linear functions. A three-parameter inverse-Gamma pdf ranked highly in terms of $r^2$ goodness-of-fit and minimising the number of parameters used to describe the pdf. The inverse-Gamma pdf has been used to model the one-point probability distribution of $A_L$ (Malamud et al., 2004; Guzzetti et al., 2008; Chen, 2009; Ghosh et al., 2012). The inverse-Gamma pdf is given by (Evans et al., 2000):

$$ p(\Lambda_E) = \frac{1}{a\Gamma(\rho)} \left(\frac{a}{\Lambda_E - s}\right)^{\rho+1} \exp \left(-\frac{a}{\Lambda_E - s}\right) $$

Eq. (5)

where $\rho = \text{shape parameter}$, $s = \text{location parameter}$, $a = \text{scale parameter}$ and $\Gamma = \text{Gamma function}$. From Eq. 5, the theoretical mean $\bar{\Lambda}_E$ and mode (position of the maximum probability, the rollover) of the length-to-width-ratio $\Lambda_E$ are given by:

$$ \bar{\Lambda}_E = \left\{\left[\frac{a}{(\rho - 1)}\right] + s\right\}, \rho > 1 $$

Eq. (6)

Mode of $\Lambda_E = \left\{\left[\frac{a}{(\rho + 1)}\right] + s\right\}$

Eq. (7)

Both the mean and mode of the pdf are controlled by all three of the inverse-Gamma parameters. The shape parameter $\rho$ controls the inverse power-law decay (right-hand tail) and the skewness. The location parameter $s$ controls the range over which the inverse-Gamma is defined (i.e., $\Lambda_E > s$). The scale parameter $a$ controls (together with $1/\rho$) the width of the inverse-Gamma distribution.

With regard to describing the probability distribution of $\Lambda_E$, the inverse-Gamma pdf models a relatively low probability of observing small values of $\Lambda_E$, and then as $\Lambda_E$ increases, $p(\Lambda_E)$ increases to a maximum $p(\Lambda_E)$ (producing a characteristic exponential rollover of $p(\Lambda_E)$ for small and medium values of $\Lambda_E$) and then decreases as an inverse power-law decay for medium and large values of $\Lambda_E$. In practical terms, this means there is:

- a relatively *low probability* of observing very ‘compact’, ellipse shapes (i.e., $\Lambda_E$ approaching 1.0, a circle),
- a *medium to low probability* of observing very long, thin landslides (e.g., large $\Lambda_E > 10$)
• a high probability of observing small-medium values of $\Lambda_E$,

Other very similar distributions, such as the double Pareto (e.g., Stark and Hovius, 2001), could be used to model the overall trend in the probability of $\Lambda_E$.

Codes were subsequently developed in R statistical software (R Core Team, 2013) to robustly fit the parameters (including confidence intervals) of the a-priori selected inverse-Gamma pdf to each dataset then test the goodness-of-fit, which are now shared as part of this paper (see Supplementary Material E).

In Figure 10 we show the results of maximum likelihood estimation (MLE) of the three parameter inverse-Gamma pdf parameter values (Eq. 5) to $\Lambda_E$ data (where $e_L \geq 0.5$). This is given for each of the ten landslide area categories for both the Northridge and Guatemala inventories. We also show in Figure 10 for visualization the observed probability density values calculated using approximately logarithmically increasing $\Lambda_E$ bin sizes, where these probability densities are not the basis by which the MLE fitting was done.

[Figure 10]

To obtain the MLE results shown in Figure 10, we used a bootstrapping technique based on Efron and Tibshirani (1993) to obtain an estimate of uncertainty around the range of each of the two inverse-Gamma parameter values fit to the observed distribution. To do this, we created 1000 synthetic samples of the data with the same number of landslide observations ($N_L$) as the original dataset, performed a MLE fit of the pdf to each sample, and investigated the variability in the resulting best-fit parameter values to each synthetic sample. Synthetic samples were created by random sampling of the original dataset’s observed values with replacement, until $N_L$ data were obtained for each sample. This means that some observed values might be sampled multiple times, whereas others might not be sampled at all, resulting in variability in the overall statistical distribution of each sample.

Figure 10 presents the best-fit inverse-Gamma pdf of $\Lambda_E$ in each landslide area category for both Northridge and Guatemala, visually demonstrating a good fit to the probability densities derived from the observed data. For both inventories, the pdfs fit
to the smallest and largest landslide area categories show a wider range of uncertainty, in other words, the 5th and 95th percentile pdfs are quite far apart, which we attribute to small sample sizes in these categories.

We use a modified version of the Kolmogorov–Smirnov (K-S) goodness-of-fit test (Lillefors, 1967) to confront the empirical cumulative distributions for $\Lambda_E$ (i.e., all original landslide $\lambda_E$ in each landslide area category) with the model pdf best-fit MLE parameters (i.e., solid lines in Figure 10). Detailed steps of the modified K-S test are given in Supplementary Material C. The modified K-S testing shows that for all landslide area categories for both the Northridge and Guatemala landslide inventories, the 20 inverse-Gamma pdfs shown in Figure 10 are reasonable models for the distribution of landslide ellipse $\Lambda_E$ at a significance level of $p = 0.05$.

We now explore how the parameters of the inverse-Gamma distribution and thus how the general shape of the length-to-width ratio distribution varies with landslide area category. Figures 11ABC shows boxplots of the three parameter values ($\rho, a, s$) of the inverse-Gamma pdf fit to $\Lambda_E$ in each landslide area category, and in Figure 11D the mode of $\Lambda_E$ (location of the maximum probability, or rollover) based on Eq. (7). The range of parameter values for each of the two locations (Northridge, Guatemala) and ten landslide area categories are derived from the bootstrapping technique described earlier in this section. In the three largest landslide area categories ($A_L \geq 6400 \text{ m}^2$) all three model parameters (Figure 11ABC) extend across a broad range of values, which we attribute primarily to a smaller sample size (compared to the other seven landslide area categories). The smaller number of values results in a greater variability in the distribution parameters when performing the bootstrapping MLE technique, and thus should not necessarily be considered fully indicative of the behaviour of the shape of large landslides.

The shape parameter ($\rho$) primarily controls the inverse power-law decay for the right-hand tail and the skewness of the distribution. For the seven landslide area categories ($A_L < 6400 \text{ m}^2$), the values of the estimated shape parameter within each landslide area category are similar for both inventories (Figure 11A). The shape parameter ($\rho$) first increases with increasing landslide area, then it decreases and finally moderately increases again (Figure 11A). In terms of $\Lambda_E$, larger values of $\rho$ equate to heavier tails (i.e., a steeper gradient) of the right-hand side of the inverse-Gamma pdfs in Figure 10, and thus smaller likelihoods of observing large $\Lambda_E$ values corresponding to longer,
thinner landslides.

[Figure 11]

The scale parameter $a$ controls (together with $1/\rho$) the width of the inverse-Gamma distribution (Figure 10). Higher values of scale parameter in Figure 11B equate to heavier tails (i.e., dying off more slowly) on the left-hand side of the inverse-Gamma pdfs in Figure 10, and thus smaller likelihoods of observing small $\Lambda_E$ values (close to 1.0, i.e., very compact landslide shapes). The scale parameter ($a$) for the two inventories are different: for $A_L < 6400$ m$^2$ the scale parameter for Northridge is significantly larger than for Guatemala (Figure 11B). This is possibly caused by generally larger length-to-width ratios of the Northridge inventory compared to the Guatemala inventory (see above).

The location parameter $s$ controls the range over which the inverse-Gamma is defined (i.e., $\Lambda_E > s$). The $s$ parameter exhibits opposite behaviour to $\rho$ and $a$. For both Northridge and Guatemala inventories, $s$ first decreases with increasing landslide area category, and then increases (Figure 11C). As the length-to-width ratio as introduced above has values of $\Lambda_E > 1.0$, it may be appropriate to truncate the distribution at $\Lambda_E = 1.0$. We do not fix $s = 1.0$ as it is possible to have values of $s > 1.0$ which influences the location of the mean and mode of the distribution, and the potential minimum value of $\Lambda_E$. From trials (not shown here), where $s < 1.0$, this results in less than 1% of the $\Lambda_E$ values having $\Lambda_E < 1.0$.

Figure 11D shows the mode of $\Lambda_E$, in other words, the value of $\Lambda_E$ at which the maximum probability (rollover) occurs in the inverse-Gamma pdf (Figure 10) for each landslide area category in both Northridge and Guatemala inventories. The mode of $\Lambda_E$ (Eq. 7) is governed by all three parameters ($\rho$, $a$, $s$) Figure 11ABC. As the shape parameters are similar for both inventories, it is essentially the different landslide area dependence on the scale ($a$) and location ($s$) parameters that is responsible for the different area dependence of the mode.

For Northridge, the mode of $\Lambda_E$ decreases with increasing $A_L$, whereas for Guatemala it increases with increasing $A_L$. This implies that in Northridge, larger $A_L$ tend towards
more compact shapes (smaller values of $\Lambda_E$) and a slightly narrower range of potential values of $\Lambda_E$, although there is still a spread of $\Lambda_E$ values in each area category. In contrast, for Guatemala, as landslide area increases, there is a greater probability of observing longer and thinner shapes (i.e., larger values of $\Lambda_E$). The behaviour we see in Guatemala (Figure 11D) is more pronounced as many landslides with low values of $e_L$ that were removed from analysis are long, thin debris flows that have become channelized, and would be in the larger landslide area categories and also have large values of $\Lambda_E$.

The behaviour we observe in Figure 11D is in agreement with Figures 9B and 10B. We hypothesize that the difference between the Guatemala and Northridge event inventories may be attributed to differences in the types of landslides in each inventory. The Guatemala inventory is composed of landslides triggered by heavy rainfall that accompanied Hurricane Mitch, with many landslides mobilising into debris flows with long runouts (Bucknam et al., 2001). This indicates that landslides in the Guatemala inventory in the larger area categories are likely to be long-run out debris flows with higher values of $\Lambda_E$. In contrast, the Northridge inventory is composed of landslides triggered by an earthquake, with many of the landslides shallow falls and slides, which generally did not mobilise into long runout slides (Harp and Jibson, 1995).

We also highlight Parise and Jibson (2000) who note that the mean length-to-width ratios for individual landslides in the Northridge inventory is $\Lambda_L = 2.6$; whereas, mean length-to-width ratios for landslide complexes (which tend to be larger in area due to coalescence of smaller landslides) is $\Lambda_L = 1.2$, as many complexes extended along the ridgelines, making them relatively wide in proportion to their length. In an analysis of the largest 356 landslides in the Northridge inventory, Marc and Hovius (2015) found that 44% of these landslides were amalgams of individual landslides. This indicates that landslides in larger landslide area categories in the Northridge inventory are likely to be landslides that have coalesced into approximately elliptical shapes with lower values of $\Lambda_E$.

In summary, we have shown that the inverse-Gamma pdf is a good statistical model for the probability of $\Lambda_E$ within each landslide area category. For Northridge and Guatemala, we observe similar overall trends in the parameter values $\rho$, $a$ and $s$ with increasing landslide area categories (Figure 11), although different absolute parameter values. It is the interaction of the parameters that controls the mode of the
pdf, which characterises the most ‘likely’ landslide shape ($\Lambda E$) for a given landslide area (Eq. 7). The trends in $\rho$ and $a$ mean that for the smallest landslide area categories ($A_L \leq 400 \text{ m}^2$), there is a lower probability of observing compact landslides (low values of $\Lambda E$) and a higher probability of observing elongated landslides. For medium landslide area categories ($400 < A_L \leq 6400 \text{ m}^2$), the trends in $\rho$ and $a$ are opposite of those observed for $A_L \leq 400 \text{ m}^2$. The largest landslide area category ($A_L > 6400 \text{ m}^2$) parameter values show similar trends to the smallest landslide area categories, although the number of observations is relatively low, and thus results are not as robust. The general trend in parameter values with area categories is similar between the two inventories; however, the absolute values and interaction of parameters means the overall probability density functions differ – primarily shifting towards more elongated landslide shapes with increasing area in Guatemala, and more compact landslide shapes with increasing area in Northridge.

Application of shape/ellipticity methodology to three small inventories

Due to the time and effort involved in creating substantially complete triggered landslide event inventories, very few triggered landslide event inventories represent a substantially complete ‘population’ of landslide shapes. Consequently, it is a challenge to draw broader conclusions about ‘generally applicable’ trends in landslide shape within the confines of the data available. Here we use Method 3 (scaled convex hull) and the ellipticity index ($e$) (outlined in the Methods section) to quantify the shape of landslides in three additional triggered event inventories of combined source area and run out landslides:

(i) **Collazzone, Italy.** 422 landslides triggered by rapid snowmelt in 1997 in Collazzone mapped from aerial photo interpretation and extensive field analysis (Cardinali et al. 2000).

(ii) **El Salvador.** 621 landslides triggered by Hurricane Mitch in 1998 in El Salvador mapped from aerial photo interpretation and some field analysis (Crone et al., 2001).

(iii) **Liguria, Italy.** 537 landslides triggered by rainfall in 2011 in Liguria mapped from semi-automatic interpretation of optical satellite imagery (Mondini et al., 2014).

For these three landslide inventories, we have less confidence in the statistical
robustness of our analysis due to smaller sample sizes ($N_L < 700$ landslides) and for
El Salvador and Liguria, Italy as acknowledged by the respective authors, the
production methods result in an under-sampling of some landslides.

We found for Collazzone 96% and for El Salvador 83% of landslides in each inventory
had $e_L \geq 0.5$ (i.e., a 75% overlap between the original landslide and the modelled
ellipse), roughly similar percentages with those observed for Northridge (85%) and
Guatemala (82%). However, we found for Liguria only 66% of landslides had $e_L \geq 0.5$
which we attribute to a high proportion of long-runout debris-flow type landslides, and
the inventory being produced from semi-automated remote sensing techniques,
resulting in relative under-sampling of smaller landslide areas and over sampling of
medium-large landslide areas (Mondini et al., 2014).

Figure 12A shows the range of $\Lambda_E$ for all landslides with $e_L \geq 0.5$ for the three
inventories. The range of $\Lambda_E$ values for El Salvador and Liguria are roughly in line with
those observed for Northridge and Guatemala. The narrower range of $\Lambda_E$ values in
Collazzone may be attributable to the fact that the inventory is a subset of a larger
triggered event inventory ($n = 4222$ landslides) across the broader Umbria region, so
within the smaller Collazzone subregion landslides do not mobilise into debris flows
because of the high clay content of the soil (Cardinali et al., 2000).

Using maximum likelihood estimation (MLE) and Kolmogorov–Smirnov goodness-of-
fit testing, an inverse-Gamma pdf was found to be an appropriate statistical model of
landslide $\Lambda_E$ for each landslide area category in each of the three considered landslide
inventories. Figure 12B indicates a similar trend of increasing mode of $\Lambda_E$ with
increasing landslide area category, as observed in the Guatemala inventory. With
further investigation of other potential controlling factors such as geology or slope unit
size, this may potentially point towards similar behaviour of landslide shape when
moisture (rainfall or snowmelt) is a triggering factor, with larger landslide areas tending
towards larger values of $\Lambda_E$. However, the parameter values fit to each landslide area
category tend to vary between the inventories. It is unclear whether this is simply the
result of the small sample sizes in the inventories used (and thus larger uncertainties),
or whether this reflects different processes controlling landslide shape across different region such as the size and length of slope units (Guthrie and Evans, 2004), initiation locations of landslides (Meunier et al., 2009) or intensity of triggering event (Li et al., 2017). This needs further investigation with additional large, substantially complete triggered landslide event inventories.

Discussion

We have presented methods to fit an ellipse to landslide inventory polygons, measure the goodness-of-fit of that ellipse ($e_L$) and the resultant length-to-width ratio of each ellipse ($\Lambda_E$). The Results section showed application of these methods to two substantially complete inventories and additionally, three ‘lower confidence’ inventories in terms of under-sampling of smaller landslides or total number of landslides. We now discuss these results in terms of how they may contribute to existing and future work on landslide form. This includes: (i) a framework for measuring landslide shape; (ii) use of ‘typical’ values of $\Lambda_E$; (iii) methods of ellipse fitting; (iv) understanding landslide form with regard to process; and (v) modelling of landslides.

A systematic framework for quantifying landslide shape

Our literature review showed that measuring and differentiating the shape of landslides (both in terms of source area and combined source area and run out) has value for (a) characterising landslides (e.g., Parise and Jibson, 2000); (b) understanding physical process (e.g., Guthrie et al., 2008); (c) remote sensing (e.g., Martha et al. 2010); and (d) modelling landslide hazard and risk (e.g., Quinn et al., 2011). In the broader literature, landslide shape is often described qualitatively (e.g., Dikau et al., 1999; Hungr et al. 2014), and of the smaller number of studies quantitatively describing landslide shape, there is a disparity in definition of length and width, and techniques to measure length-to-width ratio (Table I). The method we have developed here attempts to address this lack of consistency in measurement of length-to-width ratio by presenting a framework to systematically quantify landslide shape both in terms of (i) how elliptical a landslide is (for all landslides in an inventory), and (ii) what the length to width ratio of that best fit ellipse is (for landslides considered to be reasonably...
elliptical). The Grass-GIS Python and R codes used to perform this technique are shared with this paper (Supplementary Material D and E) with the hope that this will now be applied to additional inventories to further investigate the behaviour of landslide shape and better answer whether the observations we present in this paper are generally applicable.

Recommendations when using ‘typical’ values of landslide length-to-width ratio

We found for the Northridge earthquake-triggered landslide inventory that the length-to-width ratio for elliptical approximations of landslides with $e_L \geq 0.5$ (75% overlap between the original landslide shape and the best-fit ellipse) was $1.3 \leq \Lambda_E \leq 15.1$, and for the Guatemala rainfall-triggered inventory $1.2 \leq \Lambda_E \leq 13.8$. These can be compared to typical literature values of landslide length-to-width ratios (without the assumption of an ellipse), $\Lambda_L$ from Table I, where $\Lambda_L$ values range $0.1 \leq \Lambda_L \leq 10$ for studies that consider the combined source area and run out of a landslide and define the length the direction of travel (resulting in values of $\Lambda_L < 1$). Our findings differ from the existing literature in two ways: (i) the potential for larger values of $\Lambda_E$ and (ii) we highlight the variation in landslide length-to-width ratio $\Lambda_E$ with landslide area $A_L$. These findings may be attributable to the systematic analysis of large, substantially complete landslide inventories that consider a broad range of landslide types across a number of orders of magnitude of landslide area, rather than the smaller numbers of landslides more typically seen in the literature (Table I).

From the literature reviewed listed in Table I, we identified five studies stating ‘typical’ or mean values of $\Lambda_L$, ranging $2 < \Lambda_L < 8$ (either for only source area, or combined source-area and run out) (Barlow et al., 2006; Martha et al. 2010; Moine et al., 2009; Yang and Lee, 2006; Hovius et al., 1997). We maintain quoting a typical or mean value of $\Lambda_L$ to be a fairly common approximation in the broader literature for identification and modelling of landslides. The power-law distribution that we have shown $\Lambda_E$ to follow highlights that the mean may not be an appropriate way to characterise an inventory’s landslides.

Figure 9A showed that for both Guatemala and Northridge, when all landslides are considered (where $e_L \geq 0.5$), the median and mode value of the distribution $p(\Lambda_E)$ lies well below the mean. The inverse-Gamma pdfs fit to $\Lambda_E$ by $A_L$ category (Figure 10)
indicate that $\Lambda_E$ is non-normally distributed, with a heavy tail for larger values of $\Lambda_E$. The boxplots (Figure 9B) and pdf of $\Lambda_E$ by area categories (Figure 10) show that $\Lambda_E$ varies with landslide area, and that the direction of change varies between inventories. In the rainfall-triggered inventory the median and modal $\Lambda_E$ increases with landslide area, which we attribute potentially to debris flows (discussed later in this section). In the earthquake triggered inventory, the median and mean $\Lambda_E$ decreases with landslide area, which we attribute potentially to large landslide complexes (Parise and Jibson, 2000), although further research of other potential controlling factors (e.g., geology) is required. For these reasons, we recommend that a range of $\Lambda_E$ values should be used when modelling or approximating landslide shapes. When it is not possible to use a range of $\Lambda_E$ values, we suggest using a median value ($\Lambda_E = 2.2$ for Guatemala a $\Lambda_E = 2.5$ for Northridge) or the mode of the inverse-Gamma distribution (Eq. 7) is more appropriate than using the mean ($\Lambda_E = 2.4$ for Guatemala a $\Lambda_E = 2.9$ for Northridge).

Although some of our findings suggest different behaviour between the rainfall inventory and earthquake triggered inventory that we used, there are some similarities that may point towards some general behaviour of landslide shape. These similarities include the following: (i) the inverse-Gamma pdf is a reasonable fit of landslide $\Lambda_E$ when separated by landslide area (Figure 10); (ii) the parameter ($\rho$) that controls the gradient of the power law decay of the distribution for larger values of $\Lambda_E$ is similar for both the Northridge and Guatemala inventories when separated by landslide area category (Figure 11). When applied to the three ‘lower confidence’ inventories, we found slight deviations from these two behaviours, which we attribute to issues of resolution, under sampling, small sample sizes and also possibly that the three lower confidence inventories from Italy and El Salvador have different behaviour than Northridge and Guatemala.

In the introduction, we discussed issues of inventory production that result in an unrepresentative sample of landslides (e.g., Harp et al., 2011; Guzzetti et al., 2012). It has been shown for analysis of landslide inventory areas that considerable deviation from already ‘generally applicable’ established probability density functions (e.g., Stark and Hovius, 2001; Malamud et al., 2004) may indicate that an inventory is incomplete (Turcotte et al., 2006) or there is amalgamation of individual landslide forms (Marc and Hovius, 2015). With further confronting of the probability density functions of $\Lambda_E$ to see how generally applicable these pdfs are, it may be possible to compare distributions.
of $\Lambda_E$ between inventories to indicate the level of detail of mapping. For instance, a high proportion of non-elliptical landslides or deviation of $\Lambda_E$ from expected may indicate that inventory production methods have agglomerated many individual landslide forms into complexes.

Methods of ellipse fitting

The Methodology section presented three methods (and discussed a fourth) for fitting an ellipse to landslide polygons and then measuring the length-to-width ratio of that ellipse. The results confirm that an ellipse is a reasonably good approximation of landslide shape for 82% and 85% of landslides in the Guatemala and Northridge inventories respectively, when using a threshold value of ellipticity index $e_L \geq 0.5$ (Eq. 4). We therefore generalize that for other inventories in similar physiographic settings, it might be appropriate to use an ellipse as a suitable shape to model landslides for a very large percentage of the landslides.

Ellipse fitting is a technique used within the remote sensing literature (e.g., Moine et al., 2009; Martha et al., 2010) to distinguish between potential landslides and tonally similar features such as roads, which would have a considerably longer value of $\Lambda_L$. Because ellipse fitting is typically one of a series of steps in remote sensing workflows, relatively little attention is given to the method of ellipse fitting and how features with long values of $\Lambda_L$ are rejected. Our work highlights that more attention to the method of ellipse fitting may be required. The method can vary considerably in terms of (a) how close the original landslide shape is to an ellipse (Figure 5) and (b) the resultant length-to-width ratio of that ellipse (Figure 7).

Our work has shown that the length-to-width ratio $\Lambda_E$ is heavy tailed and varies with landslide area $A_L$. We believe that these findings will better inform the selection/rejection criteria for landslide inventory production using remote sensing imagery. For example, previous remote sensing classification work has rejected landslides with $\Lambda_L > 3$ (Martha et al., 2010). Our work has shown that for four out of five inventories considered, the 95th percentile is $\Lambda_E \approx 5$. Thus, a remote sensing technique that uses a threshold of $\Lambda_L = 3$ may result in the rejection of long, thin landslides such as debris flows from the inventory. Using both $\Lambda_E$ and $A_L$ of a potential landslide set of pixels identified in remote sensing imagery may assist in more refined
criteria for identifying landslides in remote sensing imagery.

Landslide ellipticity: process and equifinality

The second way we have characterised landslides is by their ellipticity ($e_L$), and in some ways, this may be a more powerful way to characterise landslide shape than by $\Lambda_E$. Although the inventories we have used do not distinguish between different styles of movement and type of material (as per Hungr et al., 2014), our results suggest that the majority of landslide styles have the potential to be elliptical, as shown in Figure 13. In this figure, we have taken ‘typical’ sketches of different landslide types of movement and material, and drawn hypothetical landslide polygons as each landslide might appear in an inventory. Using Method 3 (scaled convex hull), we fit an ellipse to each landslide polygon and calculated both $\Lambda_E$ and $e_L$.

[Figure 13]

Figure 13 highlights that the ‘typical’ landslide forms (in terms of movement style and material) that are well modelled by an ellipse are topple, slide, spread and to a lesser extent, debris/earth fall and some landslide complexes. These landslide forms also tend to be more compact in terms of $\Lambda_E$. Typical landslide forms not well modelled by an ellipse are rock fall, rock flow, debris flow, some earth flows and landslide complexes. Both falls and flows also had higher values of $\Lambda_E$. These falls and flows represent the highest mobility types of landslide movement (Legros, 2002) where the deposited material has the potential to travel long distances from the source area. The idea that most low-mobility landslide forms are reasonably elliptical and have low values of $\Lambda_E$ supports findings from analysis of landslide area (e.g., Stark and Hovius, 2001; Malamud et al., 2004) that regardless of physiographic setting, landslides follow similar underlying statistical behaviour.

Due to the lack of landslide type information within each inventory analysed, we cannot quantitatively summarise the ellipticity $\Lambda_E$ by landslide type. Although, because we make the tools available, this methodology could now be applied to further inventories where landslide type information is available. However, the analysis of ‘typical’
landslide forms (Figure 13) suggests that regardless of movement style or material, most landslide styles result in relatively similar shapes in terms of $\Lambda_E$ and $e_L$. Indeed, further investigation of the landslide ellipticity index as a rough quantitative metric as to landslide shape (and the degree to which it is elliptical or far away from being elliptical) may provide a classification for certain landslide types, and thus give further insights into the link between landslide type and form.

Modelling of landslides

Several studies (e.g., Dietrich et al., 2003; Stark and Guzzetti, 2009) have highlighted the complexity and challenges of developing a physically based, universal model of landslide erosion in the source area and run out in the deposition area, which ultimately controls the resultant form of a landslide. The introduction highlighted the numerous physiographic factors thought to control landslide shape, such as soil depth, vegetation cover, and slope length (e.g., Cardinale et al., 2000; Casadei et al., 2003; Guthrie and Evans, 2004). For many places where landslides occur, it is not possible to collect the amount of data required to physically model landslides (Alcántara-Ayala, 2002), nor is it necessarily appropriate due to the site-specific nature of many process-based models (van Asch et al., 2007).

As an alternative to process-based models, others (e.g., Guzzetti et al., 2005; Guzzetti et al., 2006; Ghosh et al. 2012; Malamud et al., 2015) have used the emergent literature on ‘generally applicable’ nature of landslide statistical behaviour (e.g., Stark and Hovius, 2001; Malamud et al., 2004) to stochastically model landslides. We maintain the results from our work could contribute to this area of research. Four examples of ‘generally applicable’ statistical behaviours are as follows:

- Landslide frequency-size distributions for triggered landslide events have been described by inverse-gamma or double Pareto probability density distributions (i.e., distributions with a large positive skew) across multiple settings (Stark and Hovius, 2001; Malamud et al., 2004; Katz and Ahranov, 2006; Stark and Guzzetti, 2009).
- General relationships between earthquake magnitude and the spatial distribution of landslides have been developed (Keefer et al., 1984; Rodriguez et al., 1999; Marc et al., 2016).
• The relationship between landslide area and volume is constrained to an extent (Hovius et al., 1997; Guzzetti et al., 2009; Larsen et al., 2010).

• There have been some advancements in understanding where within a slope unit landslides are more or less likely to occur depending on the triggering mechanism (Meunier et al., 2009).

These 'generally applicable' statistical behaviours can be used to stochastically model how many landslides might be observed in a particular event, where they might occur and what the size distribution of those landslides might be in terms of area and volume, and what the overall sediment budget from a triggered landslide event might be.

Our work can aid in this stochastic modelling of landslides to better inform the shape of individual landslides. This has implications for risk assessments where the probability of a landslide intersecting with exposure elements such as roads will be greater for a landslide with a higher value of $\Lambda_E$. This may also better inform studies on removal of sediment material from a catchment as landslides with higher values of $\Lambda_E$ and low values of $e_L$ may be more likely to intersect with river channels where material is then removed from the system.

Potential physical insights for $\Lambda_L$

Different physical processes can influence landslide shape, and it is possible for different processes to result in the same general form (Schumm and Lichty, 1965, Hey, 1981, Gerrard, 1984, Haines-Young and Petch, 1983). Indeed, the Northridge and Guatemala inventories are different in terms of landslide types and physiographic setting, but have been shown to follow similar underlying probability density distributions of landslide area (Malamud et al., 2004). Katz and Ahranov (2006) and Stark and Guzzetti (2009) have indicated this may be due to the change from cohesion to friction controlling small versus large landslide areas. In our analysis here, $\Lambda_E$ in both the Northridge and Guatemala inventories follows the inverse-Gamma probability density function (Eq. 6) and has similar values of power-law decay for large values of $\Lambda_E$. However, the parameters controlling the mode of the distribution and the left-hand decay differ ($a$ and $s$, Figure 11), meaning that the inverse-Gamma distribution is either shifted towards smaller or larger values of $\Lambda_E$ depending on the landslide area category. This implies that two different landslides with the same area ($A_L$), one from
the Northridge and the other from the Guatemala inventory, will have different probabilities of being a given shape ($\Lambda_E$). So, although the controls over landslide area may be similar for the two inventories, cohesion and friction may result in slightly different forms between the two inventories.

Broadly, we find that the $\Lambda_E$ at which the maximum probability occurs (the mode for $p(\Lambda_E)$) tends to increase with landslide area for the four inventories we investigated that were triggered by rainfall or snowmelt (Guatemala, and the three smaller inventories, from Collazzone and Liguria, Italy, and El Salvador). For some landslide types, it may be reasonable to assume that the width of the source area remains roughly the same as the width of the final landslide including source area and run out (Stark and Guzzetti, 2009). The length of the landslide is then determined by how far the material is transported from the source area, which could be a result of factors such as slope, geology, and trigger. Where these assumptions hold, our finding of a general increase in landslide $\Lambda_E$ with increasing $A_L$ for rainfall and snowmelt triggered landslides could be supported by Milledge et al. (2014), which builds upon Dietrich et al. (2008) to create a multidimensional slope stability model for shallow landslide source areas. The model is used to investigate the relationship between landslide shape and soil strength. Milledge et al. (2014) found that the minimum area required for a slope to fail increased with landslide length-to-width ratio of the source area $\Lambda_L$.

This supports our findings for Guatemala, where the modal value of $\Lambda_E$ increased with $A_L$ category. Milledge et al. (2014) also found that landslide width is one of the principal controlling factors of whether a landslide of a given area will initiate, as the blocks above and below the landslide source area parcel mainly control resistance to failure. The depth of the block is mainly controlled by soil depth and cohesion (Katz and Aharonov, 2006; Stark and Guzzetti, 2009). Thus, landslide width determines the surface area in contact with these up and down slope land parcels.

Milledge et al. (2014) also found that the least stable landslide source area shape (in terms of $\Lambda_L$) increases with failure plane depth, as a smaller landslide width results in less strength from the contact between the toe end of the landslide and the downslope land parcel. Ultimately, Milledge et al. (2014) state that the shape of an unstable shallow parcel of land will be determined by the spatial pattern of soil strength and water content across a landscape. Pelletier et al. (1997) found that patches of soil with a moisture content greater than a threshold value followed an inverse power-law decay.
which may exert some control over ‘generally observed’ landslide frequency-area statistics. Combined with Milledge et al. (2014) findings about the relationship between soil strength, moisture and landslide shape, this could potentially point towards a relationship between the ‘generally observed’ landslide area frequency size statistics (Stark and Hovius, 2001; Malamud et al., 2004; Katz and Ahranov, 2006; Stark and Guzzetti, 2009) and landslide shape, which would warrant further investigation in terms of process. Other controls over $\Lambda_L$ include slope length (Klar et al., 2011 for landslide source areas), slope angle (Li et al., 2017, for landslide source areas and run out) and the location on the slope at which the landslides initiate (Guthrie and Evans, 2004 for landslide source areas and run out), which could be a result of factors such as geology, local slope, and triggering mechanism (Meunier et al., 2008 for landslide source areas and run out). Although we have not calculated where on the slope the landslides in the inventories we explored were located or how steep the slope is, some accompanying information is included in some of the inventories we analysed, which could support this idea. Landslides in El Salvador typically travelled from top to base of the slope (Crone et al., 2001). This explains large values of $\Lambda_E$ for large landslide areas in this location. Landslides in Liguria, Italy, were located chiefly at the foot of the slopes (Mondini et al., 2014), where long runout landslides with high values of $\Lambda_E$ were either prevented by intersection with drainage channels or were diverted into drainage channels, creating sinuous, asymmetric landslide shapes that were removed from the analysis. Lastly, landslides in Collazzone, Italy, were prevented from mobilising into debris flows because of the high clay content of the soil (Cardinali et al., 2000), explaining the relatively low and narrow range of $\Lambda_E$ values in this inventory. More work is required to better understand the physical controls over landslide shape. This could involve the analysis of parameters such as soil strength and topography for the inventories already investigated here, the analysis of additional large, substantially complete triggered event landslide inventories, the analysis of detailed LiDAR inventories (e.g., Ardizzone et al., 2007) or the comparison of earthquake and rainfall triggered inventories from the same location using the methods and codes we have developed here (Supplementary Material D and E).
Conclusions

This work outlines a methodology to systematically quantify the degree a landslide shape differs from an ellipse using a landslide ellipticity index, e and to quantify the ellipse using a length-to-width ratio, \( \Lambda_E \). These two metrics can be rapidly and systematically applied to all landslides in a triggered event inventory, and allow better quantitative intercomparison of landslide shape across inventories.

This method was applied to two substantially complete, large triggered landslide event inventories, Northridge (Harp and Jibson, 1995) and Guatemala (Bucknam et al., 2001), which we consider to be ‘high confidence’ in terms of data quality and sample size for robust statistical analysis. We find that > 82% of landslides in the two triggered event inventories examined can be reasonably well modelled by an ellipse \( e_L \geq 0.5 \), and that a landslide identified as non-elliptical \( e_L < 0.5 \) may indicate processes of coalescence resulting in branched, irregular shapes.

For Northridge and Guatemala landslides with \( e_L \geq 0.5 \) and considered together, \( \Lambda_E \) ranged \( 1.2 \leq \Lambda_E < 15.1 \), with median \( \Lambda_E = 2.5 \) (Northridge landslides) and \( \Lambda_E = 2.2 \) (Guatemala landslides), indicating the potential for longer, thinner landslides than typically considered in the literature. We found that there was a statistically significant difference in the probability distribution \( p(\Lambda_E) \) when landslides are separated into ten increasing landslide area \( (A_L) \) categories.

Using maximum likelihood estimation and a Monte-Carlo Kolmogorov–Smirnov goodness-of-fit testing, an inverse-Gamma probability density function (pdf) was found to model robustly the range and distribution of \( \Lambda_E \) within each of the ten landslide area categories, and the parameters of this pdf vary with \( A_L \). The inverse-Gamma pdf is an asymmetrical distribution with a power-law decay for larger values of \( \Lambda_E \), indicating that it may not be appropriate to use mean values of \( \Lambda_E \). Although the absolute parameter values differ, the overall trend in variation of parameter values was similar between the Northridge and Guatemala landslide inventories. The overall behaviour of \( \Lambda_E \) with increasing \( A_L \) differs between the inventories; in Northridge (earthquake-triggered), as landslide areas \( A_L \) increase, shapes tend to be more compact; whereas, in Guatemala (rainfall-triggered), as landslide areas increase, the shapes tend towards longer and thinner. Analysis of three additional ‘lower confidence’ (rainfall triggered) inventories points to similar behaviour to the Guatemala inventory of landslide shape.
We expect the methodology presented here, and the associated open source software, to help quantifying landslides in existing landslide inventories, potential landslide shape distributions for landslide simulation modelling, and a metric for better insights as to the link between landslide process and shape.

Acknowledgements

The authors are grateful to the editor Stuart Lane, reviewer David Milledge and Anonymous Reviewer 2 for their helpful comments and feedback. The research of FT was supported by a UK Natural Environment Research Council PhD studentship, and the research of BDM and FG was supported by the EU LAMPRE Project (EC contract no. 312384).

References


Bucknam RC, Coe JA, Chavarria MM, Godt JW, Tarr AC, Bradley L, Rafferty S, Hancock D,


This article is protected by copyright. All rights reserved.


Parise, M, Jibson, R. W. 2000. A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January, 1994 Northridge,


Table I. Summary of 21 studies discussing landslide length-to-width ratio ($\Lambda_L$). (i) Six studies stating values of $\Lambda_L$ where this is not their main focus. (ii) Nine studies of landslide shape in detail. (iii) Six studies stating relationships between landslide $\Lambda_L$ and landslide area. The following notation is also used: ? = unknown, N/A = not applicable, min. = minimum, max. = maximum.

<table>
<thead>
<tr>
<th>Study reference</th>
<th>Location</th>
<th>Landslide type</th>
<th>Inventory type</th>
<th># Landslides</th>
<th>Study theme</th>
<th>Landslide length-to-width ratio, $\Lambda_L$</th>
<th>Notes</th>
<th>How length $L_L$ measured</th>
<th>How width $W_L$ measured</th>
<th>Landslide element(s) considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barlow et al. (2006)</td>
<td>Chilliwack River Basin, Vancouver, Canada</td>
<td>Debris slides; Rock slides; Debris flows</td>
<td>Historical</td>
<td>Aerial photo</td>
<td>Remote sensing; Inventory</td>
<td>Typically $\Lambda_L &gt; 2.5$</td>
<td>From crown to toe</td>
<td>?</td>
<td>Source &amp; runout area combined</td>
<td></td>
</tr>
<tr>
<td>Dikau et al. (1996)</td>
<td>Various European locations</td>
<td>Single rotational slides</td>
<td>?</td>
<td>?</td>
<td>5</td>
<td>Description of landslide types</td>
<td>Min. $\Lambda_L = 0.8$ Median $\Lambda_L = 1.3$ Mean $\Lambda_L = 1.6$ Max. $\Lambda_L = 3.3$</td>
<td>Maximum length perpendicular to slope</td>
<td>Maximum width parallel to slope direction</td>
<td>?</td>
</tr>
<tr>
<td>Gabet and Dunne (2002)</td>
<td>Santa Barbara, California, USA</td>
<td>Soil slips</td>
<td>Triggered event: rainfall</td>
<td>Aerial photo</td>
<td>Soil characteristics; landslide volume</td>
<td>Min. $\Lambda_L = 0.7$ Median $\Lambda_L = 2.4$ Mean $\Lambda_L = 2.5$ Max. $\Lambda_L = 5.6$</td>
<td>Mean</td>
<td>Mean</td>
<td>Source &amp; runout area combined</td>
<td></td>
</tr>
<tr>
<td>Martel (2004)</td>
<td>US Continental Margin (submarine)</td>
<td>Submarine</td>
<td>Historical</td>
<td>11</td>
<td>Model; landslide initiation; submarine landslides</td>
<td>Min. $\Lambda_L = 0.28$ Median $\Lambda_L = 1.26$ Mean $\Lambda_L = 1.47$ Max. $\Lambda_L = 3.90$</td>
<td>From crown to toe</td>
<td>Perpendicular to length</td>
<td>Source &amp; runout area combined</td>
<td></td>
</tr>
<tr>
<td>Martha et al. (2010)</td>
<td>Mandakini River Catchment, Uttarakhand, India</td>
<td>All; Translational rock slides; Rotational rock slides; Debris flows; Shallow translational rock slides</td>
<td>Triggered event: rainfall</td>
<td>Satellite</td>
<td>Remote sensing; inventory</td>
<td>Typically $\Lambda_L &lt; 3.0$</td>
<td>From crown to toe</td>
<td>Width of elliptical approximation (Assumed)</td>
<td>Source &amp; runout area combined</td>
<td></td>
</tr>
<tr>
<td>Study reference</td>
<td>Location</td>
<td>Landslide type</td>
<td>Inventory type</td>
<td>Inventory method</td>
<td># Landslides</td>
<td>Study theme</td>
<td>Landslide length-to-width ratio, $\Lambda_L$</td>
<td>Notes</td>
<td>How length $L_l$ measured</td>
<td>How width $W_l$ measured</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>----------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------------------</td>
<td>-------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Moine et al. (2009)</td>
<td>Barcelonnette basin, Alpes-de-Haute-Provence, France</td>
<td>All; Translational slides; Rotational slides; Earthflows; Block slides</td>
<td>Historical</td>
<td>Aerial photo</td>
<td>156</td>
<td>Remote sensing; inventory</td>
<td>Min. $\Lambda_L = 0.5$ Max. $\Lambda_L = 5.0$</td>
<td>?</td>
<td>?</td>
<td>Source &amp; runout area combined</td>
</tr>
<tr>
<td>(ii) Studies of landslide morphometry in detail (nine papers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barton et al. (1983)</td>
<td>Christchurch Bay Coastal Cliffs, United Kingdom</td>
<td>Cliff top slumps</td>
<td>?</td>
<td>?</td>
<td>42</td>
<td>Morphometry of slumps and spalls</td>
<td>Min. $\Lambda_L = 0.44$ Median $\Lambda_L = 4.2$ Mean $\Lambda_L = 5.2$ Max. $\Lambda_L = 15.2$</td>
<td>Clear difference in dimensions between geological zones. Non-normal distribution.</td>
<td>Effectively the width of the crown in plan view</td>
<td>Effectively the ‘depth’ of the crown in plan view</td>
</tr>
<tr>
<td>Parise and Jibson (2000)</td>
<td>Santa Susana Quadrangle, Northridge, California, USA</td>
<td>All</td>
<td>Triggered event: earthquake</td>
<td>Aerial photo</td>
<td>1562</td>
<td>Morphometry; susceptibility</td>
<td>(Individual slides) Mean $\Lambda_L = 2.6$ (Complexes) Mean $\Lambda_L = 1.2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lewkowicz and Harris (2005)</td>
<td>Mackenzie Valley, Nunavut, Canada</td>
<td>Permafrost active layer detachment failures</td>
<td>Seasonal; multi-temporal</td>
<td>Field survey; aerial photo</td>
<td>50</td>
<td>Morphometry; susceptibility</td>
<td>Median $2.4 \leq \Lambda_L \leq 3.6$</td>
<td>Weak negative correlation between slope, $\Lambda_L$ and $\Lambda_L$. Largest failures: low values of $\Lambda_L$ and occurred on shallow slopes. Mesoscale geomorphic factors control variability in $\Lambda_L$.</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Pourghasemi et al. (2014)</td>
<td>North of Tehran, Iran</td>
<td>All</td>
<td>Historical</td>
<td>Aerial photo; satellite; field survey</td>
<td>528</td>
<td>Morphometry; fractals</td>
<td>Min. $\Lambda_L = 1.001$ Max. $\Lambda_L = 6.084$</td>
<td>Minimum distance from toe to crown in the downslope direction</td>
<td>Maximum breadth perpendicular to length</td>
<td></td>
</tr>
<tr>
<td>Study reference</td>
<td>Location</td>
<td>Landslide type</td>
<td>Inventory type</td>
<td>Inventory method</td>
<td># Landslides</td>
<td>Study theme</td>
<td>Landslide length-to-width ratio, $\Lambda_L$</td>
<td>Notes</td>
<td>How length $L_i$ measured</td>
<td>How width $W_i$ measured</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>----------------</td>
<td>----------------</td>
<td>------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Quinn et al. (2011)</td>
<td>Saint Lawrence Lowlands, East North America</td>
<td>Large landslides in sensitive clay</td>
<td>Historical</td>
<td>Archive; aerial photo; satellite</td>
<td>~62</td>
<td>Morphometry; impact</td>
<td>Min. $\Lambda_L = 0.1$ Max. $\Lambda_L = 10.0$</td>
<td>For spreads: Min. $\approx 0.1$ Max. $\approx 1.4$ For flows: Min. $\approx 0.7$ Max. $\approx 10$</td>
<td>Width at</td>
<td>?</td>
</tr>
<tr>
<td>Rickli and Graf (2009)</td>
<td>Central and East Switzerland</td>
<td>Natural; shallow</td>
<td>6 triggered events (rainfall)</td>
<td>Aerial photo</td>
<td>522 (total)</td>
<td>Morphometry; land use</td>
<td>Min. $\Lambda_L = 0.95$ Median $\Lambda_L = 1.44$ Mean $\Lambda_L = 1.36$ Max. $\Lambda_L = 2.05$</td>
<td>Maximum extent from crown to toe, no run out</td>
<td>Maximum extent perpendicular to length, no run out</td>
<td>Source area only</td>
</tr>
<tr>
<td>Süzen (2002)</td>
<td>Asarsuyu, Turkey</td>
<td>All</td>
<td>Multi-temporal (4 time periods)</td>
<td>Aerial photo</td>
<td>154</td>
<td>Morphometry, temporal change</td>
<td>Min. $\Lambda_L = 1.47$ Mean $\Lambda_L = 1.47$ Max. $\Lambda_L = 3.85$</td>
<td>Longest axis</td>
<td>Unknown</td>
<td>(Assumed) Source &amp; runout area combined</td>
</tr>
<tr>
<td>Tian et al. (2017)</td>
<td>Minxian, China</td>
<td>Large ($A_i &gt; 500m^2$)</td>
<td>Triggered event (earthquake)</td>
<td>Aerial photo; Satellite</td>
<td>635</td>
<td>Morphometry relation to fault and landscape</td>
<td>Min. $\Lambda_L = 0.30$ Mean $\Lambda_L = 2.11$ Max. $\Lambda_L = 8.02$</td>
<td>Minimum length of bounding box perpendicular to sliding direction</td>
<td>Maximum extent of bounding box perpendicular to length</td>
<td>Source area and run out</td>
</tr>
<tr>
<td>Yang and Lee (2006)</td>
<td>Central Taiwan</td>
<td>All</td>
<td>2 triggered events (rainfall &amp; earthquake)</td>
<td>Aerial Photo</td>
<td>468 (Rainfall) 189 (Earthquake)</td>
<td>Morphometry: fractal dimension Rainfall triggered $\Lambda_L &lt; 8.0$ Earthquake triggered $\Lambda_L &lt; 5.0$</td>
<td>Length of equivalent ellipse</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

(iii) Studies of deterministic relationships between landslide area and dimensions (six papers)
<table>
<thead>
<tr>
<th>Study reference</th>
<th>Location</th>
<th>Landslide type</th>
<th>Inventory type</th>
<th>Inventory method</th>
<th># Landslides</th>
<th>Study theme</th>
<th>Notes</th>
<th>How length L measured</th>
<th>How width W measured</th>
<th>Landslide element(s) considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casadei et al. (2003)</td>
<td>Coos Bay, Oregon, USA</td>
<td>Shallow landslides</td>
<td>Historical</td>
<td>?</td>
<td>90</td>
<td>Deterministic model of landslide area</td>
<td></td>
<td>?</td>
<td>?</td>
<td>(Assumed) Source area only</td>
</tr>
<tr>
<td>Frattini and Crosta (2013)</td>
<td>Trento, Italy</td>
<td>All</td>
<td>Multi-temporal</td>
<td>Aerial photo; Lidar</td>
<td>4,175</td>
<td>Physical reasons for observed frequency size statistics</td>
<td></td>
<td>W_L = \frac{Lu}{2}</td>
<td>?</td>
<td>(Assumed) Source &amp; runout area combined</td>
</tr>
<tr>
<td>Guthrie et al. (2008)</td>
<td>British Columbia, Canada</td>
<td>Debris flows and slides</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Cellular automata model of landslide area</td>
<td>Typically 5 &lt; W_L &lt; 50 m</td>
<td>L_L = 0.76A_L^{0.66}</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Hovius et al. (1997)</td>
<td>Western Southern Alps, New Zealand</td>
<td>Falls, slumps, Slides and Debris flows</td>
<td>Multi-temporal</td>
<td>Aerial photo</td>
<td>4,984</td>
<td>Calculation of landslide volume and denudation</td>
<td>Typically \Lambda_L = 2.0</td>
<td>W_L = A_L^{0.5}</td>
<td>?</td>
<td>Source area only</td>
</tr>
<tr>
<td>Milledge et al. (2014)</td>
<td>Multiple: Japan, United Kingdom, USA: Oregon, California, Appalachian Mountains</td>
<td>All</td>
<td>Mixed: multi-temporal, seasonal and Event</td>
<td>Mixed</td>
<td>Unknown</td>
<td>Deterministic model of landslide area and shape</td>
<td>Min. \Lambda_L = 0.6 Max. \Lambda_L = 10.0 Typically \Lambda_L &lt; 3.0</td>
<td>\Lambda_L a function of soil depth, cohesion and friction</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Reneau and Dietrich (1987)</td>
<td>Central California Coast, USA</td>
<td>Shallow debris flows</td>
<td>Triggered event: rainfall</td>
<td>Unknown</td>
<td>61</td>
<td>Deterministic model of landslide area</td>
<td>Min. \Lambda_L = 0.6 Mean \Lambda_L = 1.8 Max. \Lambda_L = 4.2</td>
<td>L_L = \frac{2S_r k W_L}{W_L - 2S_r k} where S_r = root strength, and k is a function of soil and slope</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
### Table II. List of variables and abbreviations used.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Γ</td>
<td>Gamma function ( \Gamma(x) = (x - 1)! ) where ! represents factorial</td>
</tr>
<tr>
<td>Λₗ</td>
<td>Length-to-width ratio of landslide</td>
</tr>
<tr>
<td>Λₑ</td>
<td>Length-to-width ratio of ellipse</td>
</tr>
<tr>
<td>a</td>
<td>Shape parameter for inverse-Gamma pdf</td>
</tr>
<tr>
<td>Aₜₕ</td>
<td>Area of convex hull (CH)</td>
</tr>
<tr>
<td>Aₑ</td>
<td>Area of ellipse (E)</td>
</tr>
<tr>
<td>Aₗ</td>
<td>Area of landslide (L)</td>
</tr>
<tr>
<td>Aₑ∩ₜₕ</td>
<td>Area of intersection (∩) between original landslide shape and elliptical approximation</td>
</tr>
<tr>
<td>CH</td>
<td>Convex Hull</td>
</tr>
<tr>
<td>D</td>
<td>Kolmogorov–Smirnov test vertical distance between two cumulative frequency curves</td>
</tr>
<tr>
<td>D_OBS</td>
<td>Kolmogorov–Smirnov test vertical distance between the two cumulative frequency curves: (i) observed (OBS) and (ii) maximum likelihood estimation fit</td>
</tr>
<tr>
<td>D_SIM</td>
<td>Kolmogorov–Smirnov test vertical distance between two cumulative frequency curves: (i) simulated data (SIM) and (ii) maximum likelihood estimation fit</td>
</tr>
<tr>
<td>e</td>
<td>Ellipticity index</td>
</tr>
<tr>
<td>F(x)</td>
<td>Theoretical cumulative distribution of x</td>
</tr>
<tr>
<td>Fₙ(x)</td>
<td>Empirical cumulative distribution of x</td>
</tr>
<tr>
<td>L</td>
<td>Length</td>
</tr>
<tr>
<td>M[Aₜₕ, Pₜₕ]</td>
<td>Method 2 for elliptic approximation based on convex hull (CH) fit to landslide shape.</td>
</tr>
<tr>
<td>M[(Aₜₕ, Pₜₕ)→Aₗ]</td>
<td>Method 3 for elliptic approximation based on convex hull (CH) fit to landslide shape and scaled to ( Aₗ ).</td>
</tr>
<tr>
<td>M[Aₗ, Pₗ]</td>
<td>Method 1 for elliptic approximation based on ( Aₗ ) and ( Pₗ ).</td>
</tr>
<tr>
<td>n</td>
<td>Number or count (different variable types)</td>
</tr>
<tr>
<td>n ITER</td>
<td>Number of iterations for Monte-Carlo Kolmogorov–Smirnov goodness-of-fit test</td>
</tr>
<tr>
<td>Nₗ</td>
<td>Number of landslides (L) in a substantially complete inventory or landslide area range</td>
</tr>
<tr>
<td>p</td>
<td>Kolmogorov–Smirnov two-sample test significance level</td>
</tr>
<tr>
<td>P</td>
<td>Scale parameter for inverse-Gamma pdf</td>
</tr>
<tr>
<td>Pₜₕ</td>
<td>Perimeter of convex hull</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>$P_E$</td>
<td>Perimeter of ellipse ($E$)</td>
</tr>
<tr>
<td>$P_L$</td>
<td>Perimeter of landslide ($L$)</td>
</tr>
<tr>
<td>$S$</td>
<td>Location parameter for inverse-Gamma pdf</td>
</tr>
<tr>
<td>$Sup$</td>
<td>Supremum (maximum)</td>
</tr>
<tr>
<td>$W$</td>
<td>Width</td>
</tr>
</tbody>
</table>
Table III. Three different methods used in this paper for approximating an ellipse shape from observed landslide shapes. The notation \( M[ ] \) is used to indicate the elliptical approximation method, with variables in the brackets [ ]. The different steps for each method are outlined in Table IV and illustrated in Figure 5, with examples of each method given in Figure 6. See Supplementary Material A for detailed justification of Steps A to F.

<table>
<thead>
<tr>
<th>Elliptic Approximation Method Name</th>
<th>Elliptic Approximation Method Abbreviation</th>
<th>Description</th>
<th>Steps in Table IV and Figure 5</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1. Landslide area and perimeter</td>
<td>( M[A_L, P_L] )</td>
<td>Uses the solution of a quadratic equation (Eq. 1) to approximate ellipse ( \Lambda_E ) from perimeter ( P_L ) and area ( A_L ) of the landslide.</td>
<td>A, C, D, F</td>
<td>Figure 6A</td>
</tr>
<tr>
<td>Method 2. Convex hull fit to landslide shape</td>
<td>( M[A_{CH}, P_{CH}] )</td>
<td>Uses a quadratic equation (Eq. 1) to calculate ellipse ( \Lambda_E ) from perimeter ( P_{CH} ) and area ( A_{CH} ) of a convex hull fit to the landslide. A convex hull is the minimum bounding area that encloses a polygon where all internal angles connecting vertices are convex (de Berg et al., 2008).</td>
<td>A, B, C, D, F</td>
<td>Figure 6B</td>
</tr>
<tr>
<td>Method 3. Convex hull fit to landslide shape scaled to ( A_L )</td>
<td>( M[A_{CH}, P_{CH}) \rightarrow A_L )</td>
<td>Same as Method 2, but the area of the ellipse is scaled to match the area of the original landslide ( A_L )</td>
<td>A, B, C, D, E, F</td>
<td>Figure 6C</td>
</tr>
</tbody>
</table>
Table IV. Description of Steps A to F used to calculate the length-to-width ratio of a best fit ellipse from a landslide polygon. **Figure 4** outlines the overall workflow of Steps A to F. **Table III** outlines elliptic approximation methods. See **Supplementary Material A** for detailed justification of each step.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Output</th>
<th>Used in which Elliptic Approximation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step A:</strong> Preliminary landslide spatial attributes</td>
<td>Calculation of landslide polygon centre of gravity based on weighted mean of landslide Northing and Easting Coordinates (Khorshidi, 2009). Calculation of landslide orientation from a best fit bounding box.</td>
<td>Landslide centre of gravity coordinates ((X, Y)). Landslide orientation (^\circ).</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td><strong>Step B:</strong> Fit convex hull to original landslide shape</td>
<td>A convex hull is fit to the original landslide shape. A convex hull is the smallest shape that completely contains the perimeter of the original landslide shape, where all internal angles of the CH are less than 180° (de Berg et al., 2008). Area of convex hull ((A_{CH})) calculated using shoelace algorithm (ESRI, 2016). Perimeter of convex hull ((P_{CH})) calculated using Pythagorean theorem (Prashker, 1999).</td>
<td>Area of convex hull ((A_{CH})). Perimeter of convex hull ((P_{CH})).</td>
<td>2, 3</td>
</tr>
<tr>
<td><strong>Step C:</strong> Calculate ellipse length-to-width ratio</td>
<td>Using Eqs. 1–3 the length-to-width ratio of an idealised ellipse ((\Lambda_E)) is calculated from the area ((A_L)) and perimeter ((P_L)) of the landslide (Method 1) or the area ((A_{CH})) and perimeter ((P_{CH})) of the convex hull from <strong>Step B</strong> (Method 2 and 3).</td>
<td>Length-to-width ratio of an idealised ellipse ((\Lambda_E)).</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td><strong>Step D:</strong> Calculate ellipse length and width</td>
<td>Using Eq. 2 the length ((L_E)) and width ((W_E)) of the idealised ellipse are calculated from the ellipse length-to-width ratio ((\Lambda_E)) and ellipse area ((A_E)).</td>
<td>Ellipse length ((L_E)). Ellipse width ((W_E)).</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td><strong>Step E:</strong> Scale ellipse length and width</td>
<td>Using Eq. 3 the area of the best fit ellipse ((A_E)) is scaled to match the area of the original landslide ((A_L)).</td>
<td>Rescaled ellipse length ((L_E)). Rescaled ellipse width ((W_E)).</td>
<td>3</td>
</tr>
<tr>
<td><strong>Step F:</strong> Plot ellipse</td>
<td>Using GIS buffer tools, the landslide centre of gravity ((X, Y)) (<strong>Step A</strong>) is buffered using values of the ellipse length ((L_E)) and width ((W_E)) from <strong>Step E</strong> and/or <strong>Step D</strong>. The long axis of the ellipse is aligned with the landslide orientation calculated in <strong>Step A</strong>.</td>
<td>Ellipse polygon that approximately aligns in orientation with original landslide polygon.</td>
<td>1, 2, 3</td>
</tr>
</tbody>
</table>
Figure 1. Photographs of seven commonly observed landslide shape types, loosely analogous to the landslide inventory shapes shown in Figure 2, with a rough landslide shape outline (red dashed line) drawn by eye onto each photograph. Landslide information and photograph source are as follows: (A) translational sliding, topples and flows, Mameyes, Puerto Rico complex slide, 1985. Approximate size 350,000m$^2$ (Photograph R.W. Jibson, USGS, 2013). (B) complex slump-earth flow, La Conchita, California, USA, 2005. Approximate size 35,000 m$^2$ (Photograph R.L. Schuster, USGS, 2013). (C) Thistle Utah complex earthflow, 1983. Approximate size 490,000 m$^2$ (Photograph R.L. Schuster, USGS, 2013). (D) South Coast Haiti landslide, 2010 (Photograph E.L. Harp, USGS, 2013), size unknown. (E) Shallow landslide in the San Gabriel Mountains, California, USA, 2005 (USGS, 2013), size unknown. (F) Rockslide in Guerrero, Mexico, 1989 (Photograph, Wikipedia User NotHome, Public domain), size unknown. (G) Debris flow near Glenwood Springs, Colorado, 2002 (Photograph A. Holland-Sears, U.S.D.A., USGS, 2013), size unknown. Photos (A–E) and (G) are 'courtesy of the U.S. Geological Survey' and were all obtained via the USGS archive of landslide photographs (USGS, 2013). This article is protected by copyright. All rights reserved.
Figure 2. Examples of seven commonly observed types of landslide shapes and descriptions based on our visual analysis of a triggered event landslide inventory of 9,594 landslides triggered by intense rain in the 1998 Hurricane Mitch in Guatemala (Bucknam et al., 2001). Approximate downslope is indicated by arrows. Scales are different for each of the seven landslide shape examples. The typical landslide shapes here can be compared with the range given in Figure 1. (A) A landslide complex with a very irregular, branched shape. (B) A roughly symmetrical triangular shaped landslide that narrows in the flow direction. (C) A landslide that is initially approximately symmetrical and regularly shaped but appears to have an additional ‘lobe’ which is attributed to the landslide flowing around a topographical feature. (D) A compact, nearly circular landslide. (E) A relatively long, thin landslide with a fairly regular shape. (F) Similar to (E) but more compact and broadens in the direction of flow. (G) A very long, thin, sinuous shape typical of debris flows which follow channel morphology.
Figure 3. Subsets of the two triggered landslide event inventories used in this study. (A) Subset of 1600 landslides (grey filled polygons) for the Northridge landslide event inventory where 11,111 landslides were triggered by the 17 January 1994 Northridge Earthquake in California, USA (Harp and Jibson, 1995). (B) Subset of 650 landslides (grey filled polygons) for the Guatemala landslide event inventory where 9,594 landslides were triggered by heavy rains in late October and early November 1998 associated with Hurricane Mitch in Guatemala (Bucknam et al., 2001). Digital elevation models in (A) and (B) from USGS (2006). Inset maps for each panel (bottom right) show the extent of the entire inventory (grey filled shapes) and the location of the zoomed inventory subsets within each country.
Figure 4. Flowchart of Steps A to F used for elliptic approximation Methods 1 to 3 (outlined in Table III and IV). Line and arrow colours represent the combination of steps used in a particular elliptic approximation method: Method 1 (M[A_L, P_L], green solid line) is based on the original landslide area A_L and landslide perimeter P_L; Method 2 (M[A_{CH}, P_{CH}], orange dashed line) is based on a convex hull (CH) fit to the original landslide shape; Method 3 (M[(A_{CH}, P_{CH})→A_L], purple dotted line) is based on the convex hull fit from Method 2 (M[A_{CH}, P_{CH}]) but scaled to the landslide area A_L. The final ellipse is defined by its area (A_E), perimeter (P_E), length (L_E), width (W_E), semi-major axis (a), and semi-minor axis (b).
Figure 5. Elliptical approximations of 13 landslide shapes from the Guatemala landslide inventory using three different elliptic approximation methods; also given are the corresponding ellipticity index, $e_L$ (Eq. 4). The three elliptic approximation methods are described in Table III and are based on the following: (A) $M[A_L, P_L]$ landslide area ($A_L$) and landslide perimeter ($P_L$); (B) $M[A_{CH}, P_{CH}]$ convex hull (CH) area ($A_{CH}$) and perimeter ($P_{CH}$); (C) $M[(A_{CH}, P_{CH})\rightarrow A_L]$ convex hull (CH) area ($A_{CH}$) and perimeter ($P_{CH}$) with ellipse area ($A_E$) scaled to match area of original landslide shape ($A_L$). Landslide shapes in this figure are actual ones taken from the inventory of 9,594 landslides triggered by Hurricane Mitch in Guatemala (Bucknam et al., 2001).
Figure 6. (A) Relationship between ellipticity index ($e_L$) (Equation 4) and length-to-width ratio of ellipse ($\lambda_E$) of elliptical approximations of landslides shapes using Methods 1 to 3 described in Table III ($A_L$ and $P_L$ the original landslide area, $A_{CH}$ and $P_{CH}$ the area and perimeter of the convex hull (CH) fit to the original landslide shape) for 11,111 landslides triggered by the 1994 Northridge Earthquake, USA. Horizontal boxplots within each panel show the overall distribution of ellipticity index for each method, where whiskers indicate the 5th and 95th percentile of the distribution, the central bar indicates the 50th percentile and the diamond marker indicates the mean value. (B) Vertical box plots of length-to-width ratios ($\lambda_L$) for Methods 1 to 3, shown side-by-side for comparison. (C) Probability density distributions of ellipticity index, $e_L$ (Equation 4) of elliptical approximations of landslide shapes using Methods 3 for 11,111 landslides triggered by the Northridge Earthquake (red) and 9,594 landslides triggered by the 1998 Hurricane Mitch in Guatemala (blue). Shown also is the cut-off value of $e_L = 0.5$ that we use in this paper to reject landslides from further analysis where the elliptical approximation of the original landslide is not a good fit. Note that 66 [292] landslides were removed from the Northridge [Guatemala] inventories when the intersection of the elliptical approximation with the original landslide shape resulted in more than one (non-contiguous) segment.
Figure 7. Examples of elliptical approximations of real landslides with varying levels of the ellipticity index $e_L$ (Eq. 4) using elliptic approximation Method 3 $M[A_{CH}, P_{CH}] ightarrow (A_L)$ (convex hull fit to landslide shape and scaled to $A_L$). Examples of ellipticity index are shown in $e_L = 0.1$ increments. Real landslide shapes are from the inventory of 9,594 landslides triggered by Hurricane Mitch in Guatemala (Bucknam et al., 2001).
Figure 8. Examples of a landslide complex (Northridge inventory) and some debris flows (Guatemala inventory) with ellipticity index $e_L < 0.5$ (see text), draped over elevation and Google Earth Imagery. (A) A large landslide complex (orange outline) located in the Northridge, California USA region and triggered by the 17 January 1994 Northridge Earthquake. (Landslide shape from Harp and Jibson, 1995). It is possible that several smaller landslides have coalesced to create an irregular, branched form that is not well modelled by an ellipse. Google Earth Imagery from 1 June 1994 (approximately 5 months after triggering event) (Map Data: Google, US Geological Survey, 2015) (B) A large, branched debris flow (purple outline) triggered by heavy rains in late October and early November 1998 from Hurricane Mitch in Guatemala. The debris flow is mapped as one shape (shapes from Bucknam et al., 2001), but is most likely the result of coalescence of several runouts from source areas on different slopes. Google Earth Imagery from April 2003 (approximately 5 years after the triggering event although scars are still visible) (Map Data: Google, US Geological Survey, Digital Globe, 2015). Note, in both (A) and (B) there is a slight offset in alignment between the landslide shapes and the imagery, due to conversion to Google Earth Projection and visualisation at an angle.
Figure 9. Landslide length-to-width ratios for the Northridge and Guatemala landslide inventories. (A) Density box plots (width of the plot denotes the density of observations for a given value of $\Lambda_E$) of landslide ellipse length-to-width ratio ($\Lambda_E$) on a linear y-axis scale for all landslide ellipses where ellipticity index $e_L \geq 0.5$ for 9,441 landslides triggered by the 1994 Northridge Earthquake, USA and 8,031 landslides triggered by the 1998 Hurricane Mitch in Guatemala. (B) Boxplots of landslide ellipse length-to-width ratio ($\Lambda_E$) on a logarithmic y-axis scale with landslides separated into ten ranges of landslide area ($A_L$) which increase approximately logarithmically. Boxplot whiskers indicate the 5th and 95th percentile of the distribution within each landslide area category, the central bar indicates the 50th percentile and the diamond marker indicates the mean value. For both (A) and (B) the number of values ($n$) contributing to the plot or box plot is given.
Figure 10. Inverse-Gamma probability density function (pdf) (Eq. 5) fit to landslide ellipse length-to-width ratios ($\Lambda_E$) using maximum likelihood estimation (MLE) where landslides are split into categories based on the landslide area ($A_L$) with $n$ the number of landslides in that category. (A) 9,441 landslides triggered by the 1994 Northridge Earthquake, California USA (B) 7,858 landslides triggered by 1998 Hurricane Mitch in Guatemala. Shown for each of the 20 panels are small coloured squares that represent the probability densities $p(\Lambda_E)$ of the observed $\Lambda_E$ for the considered landslide area ($A_L$) category; we derive probability densities so they are distributed in approximately logarithmically equal bin sizes. Calculated separately, and from ‘all’ landslide areas (not the probability densities) are solid lines represent the best-fit inverse-Gamma pdfs (with a fixed location parameter of $s = 1$) using MLE. The shaded area on both the upper and lower sides of the pdf line represents 5th/95th percentile confidence intervals around the pdf. These are calculated using a bootstrapping technique where observed data is repeatedly sampled with replacement and an inverse-Gamma pdf fit to each sample. The mode of $\Lambda_E$, i.e. the location of the maximum probability (rollover), is shown as a vertical dashed line and the 5th/95th percentile values of rollover from the bootstrapping technique are shown as the shaded vertical area on either side of the vertical dashed line.
Parameter values and characteristics describing the inverse-Gamma pdf fit to landslide ellipse $\Lambda_E$ values in each landslide area category for 9,441 landslides in the 1994 Northridge inventory (red) and 8,031 landslides in the 1998 Guatemala inventory (blue). (A) Shape parameter ($\rho$) controls the inverse power-law decay (right-hand tail) and the skewness. (B) Scale parameter ($a$) controls (together with $1/\rho$) the width of the inverse-gamma distribution. (C) Location parameter ($s$) primarily controlling the position of the mode, and (D) the Mode of $\Lambda_E$, i.e., the location of the ’rollover’ of the inverse-Gamma pdf, which denotes the value of $\Lambda_E$ at which the maximum probability density occurs. Boxplot whiskers indicate the 5th and 95th percentile of the distribution within each landslide area category, the central bar indicates the 50th percentile and the diamond marker indicates the mean value.
Figure 12. Landslide ellipse length-to-width ratios ($\Lambda_E$) for all landslide ellipses for three additional smaller rainfall/snowmelt triggered inventories from Collazzone (Italy), El Salvador and Liguria (Italy) compared to the Northridge (USA) and Guatemala inventories. (A) Boxplots of ($\Lambda_E$) for all landslide ellipses where ellipticity index ($e$) ≥ 0.5 separated by landslide inventory. (B) Boxplots of ($\Lambda_E$) for all landslide ellipses where ellipticity index ($e$) ≥ 0.5 separated by landslide area category for each smaller inventory. Boxplot whiskers indicate the 5th and 95th percentile of the distribution, the central bar indicates the 50th percentile and the diamond marker indicates the mean value. For landslide area categories where n < 50, individual values of $\Lambda_E$ are plotted rather than a box denoting percentiles of the data.
**Figure 13.** Comparison of ‘typical’ landslide forms with their hypothetical landslide polygons and best-fit ellipses. Left: drawings of typical landslide forms based on their material and style of movement. Created by the British Geological Survey (2015) after Hungr et al. (2014). Right: hypothetical landslide polygons (grey) of the equivalent landslide style of movement and material. For each hypothetical landslide, the best-fit ellipse is overlaid (using Method 3 outlined in Table III) and the ellipticity $e_L$ (Eq. 4). The length-to-width ratio $\lambda_L$ of each ellipse is also shown. Rock and debris spreads are not shown as these forms do not produce contiguous polygonal shapes in the same way as other landslide forms.
What shape is a landslide? Landslide ellipticity and length-to-width ratios

Faith E. Taylor*, Bruce D. Malamud, Annette Witt, Fausto Guzzetti

We present a methodology to systematically quantify the shape of landslides by their ellipticity ($e_L$) and length-to-width ratio ($\Lambda_L$). An ellipse is a reasonable model for >80% of landslides. Landslides that deviate from an ellipse are related to landscape processes. $\Lambda_E$ ranges 1.2 – 15.1 and is non-normally distributed.

$e_L = 0.34$  \hspace{1cm}  $e_L = 0.82$

$\Lambda_E = 1.90$  \hspace{1cm}  $\Lambda_E = 1.90$

Original Landslides
Convex Hull
Elliptical Approximation