Natural Hazards
and Earth System
Sciences

Size scaling of large landslides from incomplete inventories

Oliver Korup[1,2], Lisa V. Luna[1,3], and Joaquin V. Ferrer[1,3]

[1]Institute of Environmental Sciences and Geography, University of Potsdam, Potsdam, Germany
[2]Institute of Geosciences, University of Potsdam, Potsdam, Germany
[3]Potsdam Institute for Climate Impact Research PIK, Potsdam, Germany

Correspondence: Oliver Korup (korup@uni-potsdam.de)

Abstract. Landslide inventories have become cornerstones for estimating the relationship between the frequency and size of slope failures, thus informing appraisals of hillslope stability, erosion, and commensurate hazard. Numerous studies have reported how larger landslides are systematically rarer than smaller ones, drawing on probability distributions fitted to mapped landslide areas or volumes. In these models, much uncertainty concerns the larger landslides (defined here as affecting areas $\geq 0.1\,\mathrm{km}^2$) that are rarely sampled and often projected by extrapolating beyond the observed size range in a given study area. Relying instead on size-scaling estimates from other inventories is problematic because landslide detection and mapping, data quality, resolution, sample size, model choice, and fitting method can vary. To overcome these constraints, we use a Bayesian multi-level model with a generalised Pareto likelihood to provide a single, objective, and consistent comparison grounded in extreme value theory. We explore whether and how scaling parameters vary between 37 inventories that, although incomplete, bring together 8627 large landslides. Despite the broad range of mapping protocols and lengths of record, as well as differing topographic, geological, and climatic settings, the posterior power-law exponents remain indistinguishable between most inventories. Likewise, the size statistics fail to separate known earthquakes from rainfall triggers and event-based triggers from multi-temporal catalogues. Instead, our model identifies several inventories with outlier scaling statistics that reflect intentional censoring during mapping. Our results thus caution against a universal or solely mechanistic interpretation of the scaling parameters, at least in the context of large landslides.

1 Introduction

Keeping records of the size and frequency of landslides is key to estimate rates of erosion, geomorphic work, and hillslope evolution (Dente et al., 2023; Saito et al., 2014; Marc et al., 2019); infer material strength and weathering of hillslopes (Li and Moon, 2021; Alberti et al., 2022); and inform hazard appraisals of slope instability (Guzzetti et al., 2012), particularly in response to contemporary climate change (Smith et al., 2023). A popular way to characterise the relative frequency of landslide-affected areas or volumes is to fit probability distributions to size data compiled in inventories (Malamud et al., 2004). These catalogues contain locations and geometries of individual footprint areas mapped largely from air photos or satellite images. The choice of probability distribution (or "scaling laws") has favoured the inverse power-law or Pareto, the inverse gamma, or the lognormal distributions (Tebbens, 2020) or combinations thereof (Jain et al., 2022). All these distributions are skewed, often heavy-tailed, and capture the widespread observation that larger landslides are systematically rarer than smaller ones.

Reported values of the parameters that define these distributions have seemingly narrow numerical ranges (Tebbens, 2020). This similarity among model fits has led to a lively discussion about whether these parameters reflect generic geometric or mechanistic properties of landslides or the hillslopes that they occur on (Bellugi et al., 2021; Bernard et al., 2021). For example, physical interpretations of the "roll over" that marks the lower bound of inverse power-law distributions include that of a hillslope length scale that is susceptible to failure or the cohesive strength of failure planes (Tebbens, 2020). While landslide size distributions may reflect the nature and spatial intensity of a common trigger such
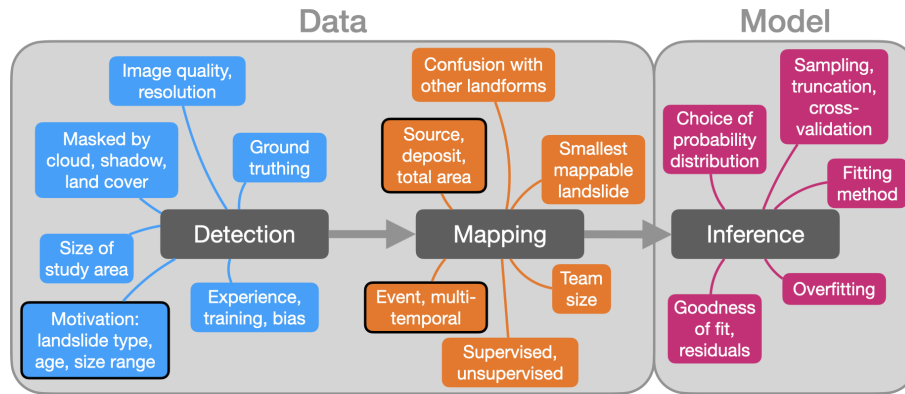
Figure 1. Landslide scaling statistics rely on accurate detection, mapping, and statistical inference. These three (colour-coded) steps are prone to a number of uncertainties at the level of both data (i.e. landslide areas from inventories covering specific single-trigger events or integrating observations over longer time spans) and model; these uncertainties propagate as errors into the derived landslide statistics. Boxes with a black outline are most likely directly tied to physical processes of landsliding.

as a strong earthquake (Valagussa et al., 2019), the average landslide size may contain information about both the cohesive strength of slope material and hillslope relief (Medwedeff et al., 2020). Some of these physical interpretations are backed by, or derived from, numerical simulations of slope instability (Frattini and Crosta, 2013). Yet, a different line of argument proposed that the roll over is a statistical artefact of landslide detection and mapping, approximately marking the smallest discernible landslide in a given study area (Tebbens, 2020). Either way, this discussion has questioned whether these scaling laws are universally applicable to landslides irrespective of environmental setting, mapping methods, and trigger mechanisms (Malamud et al., 2004; Tanyaş et al., 2019). Many of these interpretations have relied on the direct comparison of reported parameter values, and scrutiny concerning possible effects of data sources and quality, mapping method, and statistical errors in the fitted models has increased in more recent work (Bellugi et al., 2021).

Still, most uncertainty remains about the large landslides that are rarely sampled. Hence, the bulk of studies on landslide size has disclosed little about these large landslides, let alone their prediction as first-time failures (Fan et al., 2019). One reason for this knowledge gap is that large landslides are often elusive in catalogues compiled shortly after a landslide-triggering earthquake or rainstorm (Hao et al., 2020; Abancó et al., 2021; Santangelo et al., 2023). Sample sizes often involve only a handful to several dozen large landslides and thus often remain too small for robust statistics in a given study area. Hence, inference is mostly based on the simple extrapolation of model fits beyond the observed size range. Yet, large landslides may often re-shape hillslope geometry and dominate erosion (Korup et al., 2007; Marc et al., 2019) but may involve phases of creep motion and respond differently to triggering conditions than smaller failures because of a longer and more complex slope history of accumulated stress and strain (Lacroix et al., 2020).

In general, the statistics of landslides derive from a sequence of detection, mapping, and statistical inference (Fig. 1). Uncertainties that propagate throughout each step can affect the outcome in terms of landslide scaling statistics.

At the level of the input data, both landslide detection and mapping face several constraints. The mapping objective can dictate whether to focus on landslides attributed to a single trigger, such as a strong earthquake (Meunier et al., 2013; Gorum et al., 2014; Tanyaş et al., 2017) or a rainstorm (Hao et al., 2020; Emberson et al., 2022; Santangelo et al., 2023), or instead to compile landslide traces that have accumulated over years to millennia (LaHusen et al., 2016; Luetzenburg et al., 2022; Fusco et al., 2023). Some of the most comprehensive catalogues today feature thousands to hundreds of thousands of slope failures across entire nations or beyond (Luetzenburg et al., 2022; Fusco et al., 2023). The methods to detect, map, and compile landslide inventories have become more diverse and elaborate beyond traditional mapping from air photos, optical satellite data, or historical records (Xu et al., 2020; Casagli et al., 2023). Newer catalogues are derived from laser scanning (Bernard et al., 2021), radar imagery (Song et al., 2022), object-based image analysis (Milledge et al., 2022), deep neural networks (Schönfeldt et al., 2022), text mining (Franceschini et al., 2022), and seismology (Hibert et al., 2019). Most methods require specific mapping protocols adjusted to the effective resolution of imagery that may be compromised by vegetation, land cover, cloud, and shadow (Brardinoni et al., 2003; Burrows et al., 2022). Debris-flow and snow-avalanche tracks, moraines, and wind-throw gaps in forests can be mistaken for landslide evidence. Overlapping landslide source areas or bodies can obscure the dimensions of slope failure (Marc and Hovius, 2015), and so can subsequent erosion or deposition. Varying image quality, resolution, and coverage all affect landslide size estimates, as does the experience of map-

ping operators (Van Den Eeckhaut et al., 2005). The mapping outcome may depend on whether a single person or a team is at work, as well as on the attention to detail in delineating source, deposit, or total affected areas. Experience and training aid detection and mapping but also introduce bias, for example favouring fresh landslides, certain types of failure, or simply those that are easiest to recognise in the case of limited time or training. Hence, the size range of a given inventory is bounded by the smallest mappable and the largest recognisable landslide (Barlow et al., 2012). Inventories can be incomplete in that they miss out on those large landslides that have indistinct or less obvious geomorphic evidence and thus require experience, skill, and time for accurate detection and mapping. Thus, without any standard mapping protocol in place, landslide researchers have to deal with catalogues of varying extent, detail, and quality for the same task of obtaining traits of landslide size.

At the model level, estimates of size scaling hinge on the choice of probability distribution to characterise the mapped landslides (ten Brink et al., 2009). These estimates also depend on sample size, data pre-processing, fitting method, residuals, and cross-validation. Numerical experiments show that small sample sizes yield volatile estimates of scaling parameters for inverse power-law distributions in particular (Korup et al., 2012). Most estimation methods involve either the regression of log-binned – and thus smoothed – landslide frequencies versus size (Gilham et al., 2018) or maximum likelihood estimates (Clauset et al., 2009); various biases apply to both methods. However, reports of confidence intervals or goodness of fit, and hence ways to assess overfitting, remain rare. Still, the basis for the statistical inference of landslide size distributions varies at the level of the individual inventory, each of which embodies the methods of detection and mapping used and the biases they may induce. In light of these constraints, a direct comparison of landslide scaling estimates between different inventories may be misleading.

We propose a compact solution to compare more fairly the landslide size distributions from diverse inventories by estimating scaling parameters with a single, probabilistically consistent model. We apply this model to large landslides that affect a total area of $> 0.1\,\text{km}^2$ and address the problem of small sample size by using Bayesian inference in a multi-level model that uses data from multiple inventories to estimate the variance of scaling parameters within and across these catalogues (Luna and Korup, 2022). The multi-level approach acknowledges structure in landslide size data in a consistent way. One intuitive grouping of data is by inventory and reflects the diversity in data input quality reviewed above. Our focus on large landslides makes the generalised Pareto distribution (GPD) a natural model choice because extreme value theory predicts that data above a high threshold are approximately generalised Pareto distributed (Castro-Camilo et al., 2022). Another advantage of this distribution is that its parameters can be translated directly into those used most widely in studies of landslide size scaling.

2 Data and methods

We consider data on total landslide-affected areas from several dozen published landslide inventories with open access. We excluded many other detailed catalogues that had no records of landslides meeting our size threshold of $0.1\,\text{km}^2$. Besides information about their size, many databases have landslide types and triggers reported, and many data were recorded following recent (i.e. post-1900) major earthquakes and rainstorms with the intention of characterising the impact of these events. We also included catalogues spanning time intervals of several years to millennia, featuring mostly undated large landslides with unknown triggers to test whether these cumulative inventories have size distributions that differ from those of event-based inventories.

The choice of probability distribution to model landslide area often rests on implicit assumptions. For example, the inverse power law draws on considerations of physical sandpile models and the concept of self-organised criticality (Hergarten and Neugebauer, 1998), whereas the lognormal distribution arises naturally from multiplicative effects of random variables (ten Brink et al., 2009). Here we model reported areas of large landslides with the generalised Pareto distribution (GPD). The GPD is rooted in extreme value theory and approximates the distribution of a continuous random variable x above a high threshold (or location parameter) μ. The GPD thus captures what we would expect theoretically from a sample consisting of observations filtered above a minimum value (Katz et al., 2002). Any physical interpretation of the GPD parameters may need to account for, or correct, this statistical expectation first. The probability density function of the GPD is

$$\text{GPD}(x|\mu,\sigma,k) = \frac{1}{\sigma}\left(1 + \frac{k(x-\mu)}{\sigma}\right)^{-1/k-1}, \qquad (1)$$

where $x \geq \mu$, $\sigma > 0$ is a scale parameter, and $k \geq 0$ is a shape parameter. The scale parameter σ is somewhat comparable to the roll over in studies using an inverse power-law model for estimating landslide size scaling. This roll over approximates the modal landslide size, which is the smallest landslide size above which power-law scaling is assumed. The GPD shape parameter is the inverse of the "scaling exponent" α of the inverse power-law distribution such that $k = 1/\alpha$.

Here, the location parameter μ sets the minimum landslide size for data to be admitted to the GPD and is known as the peak-over-threshold approach in extreme value statistics (Katz et al., 2002). For large landslides, we let $\mu = 0.1\,\text{km}^2$. Empirical relationships between landslide volume and total affected area across a wide range of environmental settings show that an area of $0.1\,\text{km}^2$ corresponds to an average volume of roughly $10^6\,\text{m}^3$ (Larsen et al., 2010), which is the suggested lower threshold for large landslides (McColl and Cook, 2024). This particular choice of μ is a compromise because fewer samples and landslide inventories are available

for higher values of μ, whereas the GPD becomes a less and less valid approximation of the data for lower values of μ.

Fitting the GPD to data can involve maximum likelihood estimates or, in case of few samples ($n < 30$), probability-weighted (L-)moments to avoid volatile parameter estimates (Katz et al., 2002). We use Bayesian inference to learn the GPD parameters from the data, acknowledging explicitly that these come from different inventories that reflect different environmental conditions across study areas and that landslides are likely detected at varying resolution and mapped with different techniques. A Bayesian treatment of this fitting problem seeks a compromise between a likelihood function and a probability distribution of prior knowledge about the model parameters. This approach obviates the need for binned landslide size data to use frequency density (Malamud et al., 2004). Instead, we work with the joint probability that is the numerator of Bayes' rule:

$$p(\boldsymbol{\theta}|\mathcal{D}) = \frac{p(\mathcal{D}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{p(\mathcal{D})}, \tag{2}$$

where $\boldsymbol{\theta}$ is the vector of model parameters that we wish to update from both the landslide size data $\mathcal{D}$ and prior knowledge. We use the GPD as the likelihood function $p(\mathcal{D}|\boldsymbol{\theta})$ and choose (hyper-)prior distributions $p(\boldsymbol{\theta})$ to approximate what we know about landslide size distributions so far and irrespective of the data $\mathcal{D}$ studied here.

Our model uses a multi-level set-up, in which $i \in \{1, \ldots, n\}$ indexes each landslide observation x_i from a sample of size n and $j \in \{1, \ldots, J\}$ indexes each of J different landslide inventories. The idea of the multi-level model is that the size distribution in each landslide inventory j is characterised by an individual set of GPD parameters σ_j and k_j. We further assume that the values of each of these inventory-specific parameter pairs are drawn from the same two probability distributions:

$$x_i \sim \text{GPD}(\mu, \sigma_{j[i]}, k_{j[i]}), \tag{3}$$
$$\sigma_j \sim \text{gamma}(\alpha_\sigma, \beta_\sigma), \tag{4}$$
$$k_j \sim \text{gamma}(\alpha_k, \beta_k). \tag{5}$$

Here we choose independent gamma distributions for both σ_j and k_j to ensure that the parameters are positive and uncorrelated; α_σ and α_k are the corresponding shape parameters, and β_σ and β_k are the inverse-scale (or rate) parameters. The multi-level model thus learns the parameters for each catalogue informed by both its data, the overarching gamma distributions, and prior knowledge. While the model allows σ_j and k_j to vary between landslide inventories, it also draws on information from the full data set via this multi-level structure.

Bayesian reasoning requires that we specify our prior knowledge explicitly. We do this by choosing the hyper-parameters of the two gamma distributions of σ_j and k_j. These hyper-parameters describe the distribution of landslide scaling parameters across all inventories and offer a global summary from all data. Here, we draw on the growing literature of landslide scaling: recent reviews have summarised that the power-law scaling exponent α for landslide inventories is most often reported in the range of $1 < \alpha < 3$ (Tebbens, 2020). Recalling that the GPD shape parameter k equals $1/\alpha$, we can use this reported range to constrain our (hyper-)prior distributions accordingly. We choose hyper-parameter values such that they contain findings from landslide scaling studies based on data other than the ones used here. The exact shape of these distributions may matter little in light of the large sample size that informs our likelihood function. We disregard any correlation between the hyper-parameters and simplistically assume independent distributions:

$$\alpha_\sigma \sim \mathcal{N}(1, 0.25), \tag{6}$$
$$\beta_\sigma \sim \mathcal{N}(5, 5), \tag{7}$$
$$\alpha_k \sim \mathcal{N}(6, 1), \tag{8}$$
$$\beta_k \sim \mathcal{N}(9, 1). \tag{9}$$

We assume independent Gaussian distributions for these hyper-parameters and choose the prior means and standard deviations informed by previous research on landslide scaling properties (Tebbens, 2020).

To avoid having too many inventories with only a handful of large landslides, we consider only those data collections with at least 25 landslides that exceed the threshold size μ. This means that we had to discard many published landslide inventories that only contain smaller slope failures. The data that we need for obtaining the posterior distribution of all GPD parameters consist of the total affected areas by individual landslides and labels of the inventories they belong to. Our data consist of 8627 large landslides filtered from 37 different inventories (Table 1). Together, these large slope failures affected an area of 6407 km^2 or 59 % of the total landslide-affected area recorded in these catalogues. The largest landslide is unnamed and extends over 201 km^2 in the Caspian Sea basin (Pánek et al., 2016). Our data thus span more than 3 orders of magnitude in landslide area; the largest mapped landslide areas per inventory differ by up to 2 orders of magnitude. We note that 19 (or 51 %) of our selected landslide inventories were compiled following an earthquake, including five cases of two inventory versions each for the same event mapped by different research teams. Only 3 catalogues (8 %) are attributed to a rainfall trigger, while 15 catalogues (41 %) are geomorphological inventories that contain information about landslides that accumulated over many years and thus likely reflect various triggers.

The Bayesian implementation of our GPD model requires a numerical approximation of the joint posterior distribution. We use the probabilistic programming language STAN (Carpenter et al., 2017) to code our model and call it via the statistical programming environment R. We ran four independent Hamiltonian Monte Carlo chains to explore the model parameter space with the No U-Turn Sampler (NUTS) coded

Table 1. List of 37 landslide inventories with sample size of landslides $\geq 0.1\,\mathrm{km}^2$, the largest and smallest landslides reported, the count of discarded landslides below μ, and the trigger. Three-letter codes are ISO country identifiers, and numbers refer to different catalogue versions of the same triggering event; catalogues related to earthquakes show estimated magnitude. The * denotes inventories derived from deep learning.

Inventory	Samples	Max area (km^2)	Min area (km^2)	n below μ	Trigger	Reference
Campania ITA	1854	7.3	0.0000	79 362	Various	Fusco et al. (2023)
Caspian Sea KAZ	303	201.1	0.0170	18	Various	Pánek et al. (2016)
Central East NPL	108	0.7	0.0003	12 730	Rainfall	Jones et al. (2021)
Daunia Apennines ITA	712	6.8	0.0001	10 091	Various	Ardizzone et al. (2023)
Far Western NPL	58	0.7	0.0000	26 292	Rainfall	Muñoz-Torrero (2020)
Glacier Bay N.P. CAN 1	52	22.2	0.0972	1	Various	Kim et al. (2022)
High Mountain Asia	96	7.8	0.0155	31	Various	Liu et al. (2021)
Himalayas	35	0.9	0.0219	32	Various	Marc et al. (2019)
Kluane CAN-USA	369	7.0	0.0190	89	Various	Smith et al. (2023)
M6.2 Aisen CHL 1	29	1.0	0.0004	488	Earthquake	Gorum et al. (2014)
M6.2 Aisen CHL 2	43	1.0	0.0002	495	Earthquake	Sepúlveda et al. (2010)
M6.8 Elazig TUR	68	3.0	0.0905	5	Earthquake	Karakas et al. (2021)
M6.9 Nyingchi CHN 1	54	4.9	0.0001	712	Earthquake	Hu et al. (2019)
M6.9 Nyingchi CHN 2	49	3.9	0.0914	10	Earthquake	Zhao et al. (2019)
M7.5 GTM 1	90	1.3	0.0000	6134	Earthquake	Harp et al. (1981)
M7.5 Tari PNG 1	239	5.0	0.0001	11 369	Earthquake	Tanyaş et al. (2022b)
M7.6 PAK 2	84	2.0	0.0000	1369	Earthquake	Basharat et al. (2014)
M7.6 PAK 3	209	1.9	0.0000	2721	Earthquake	Basharat et al. (2016)
M7.8 Gorkha NPL 3	31	1.7	0.0000	24 884	Earthquake	Roback et al. (2018)
M7.8 Gorkha NPL 5	105	2.0	0.0000	21 046	Earthquake	Valagussa et al. (2021)
M7.8 Kaikoura NZL 1	42	1.1	0.0920	5	Earthquake	Tanyaş et al. (2022a)
M7.9 Alaska USA 1	147	9.0	0.0009	1432	Earthquake	Gorum et al. (2014)
M7.9 Sichuan CHN 1	556	7.0	0.0000	196 925	Earthquake	Xu et al. (2014)
M7.9 Sichuan CHN 2	373	7.2	0.0000	69 233	Earthquake	Li et al. (2014)
M8 Haiyuan CHN	519	1.5	0.0902	70	Earthquake	Xu et al. (2020)
M9.1 Tohoku JPN 2	162	62.0	0.0600	4	Earthquake	Tanyaş et al. (2017)
Owyhee USA	412	39.6	0.1253	0	Various	Safran et al. (2011)
Porgera* PNG	41	1.7	0.0002	1516	Earthquake	Bhuyan et al. (2023)
South Central CHL	571	123.9	0.0116	101	Various	Antinao and Gosse (2009)
Southern KGZ 1	56	2.9	0.0905	3	Various	Behling et al. (2016)
St. Elias USA	220	4.5	0.1096	0	Various	Belair et al. (2022)
TC Morakot TWN	340	2.8	0.0000	9896	Rainfall	Emberson et al. (2022)
Tongass USA 2	195	5.3	0.1002	0	Various	Belair et al. (2022)
Utah USA 1	68	2.2	0.0926	9	Various	Belair et al. (2022)
Washington WGS USA	80	58.6	0.1012	0	Various	Belair et al. (2022)
Wenchuan CHN	96	0.6	0.0000	9933	Various	Domènech et al. (2018)
Wenchuan* CHN	161	2.5	0.0000	10 174	Earthquake	Bhuyan et al. (2023)

in STAN and verified that the numerical solutions converged. Unless stated otherwise, we use medians and 95 % highest-density intervals (HDIs) to summarise all posterior distributions. A 95 % HDI means that there is a 95 % probability that a given parameter is in the specified interval (McElreath, 2016).

3 Results

3.1 Model fits and residuals

We express the size distributions of large landslides in cumulative form using the exceedance probability p for a given landslide area (Fig. 2). To measure how well the GPD model fits the data, we compute the residuals in terms of the log-odds ratios between the empirical exceedance probabilities (p) and the predicted averages ($\hat{p}$) for each inventory. The log-odds ratio is $\log\frac{\hat{p}(1-p)}{p(1-\hat{p})}$, conditioned on each observed landslide. A positive (negative) log-odds ratio means
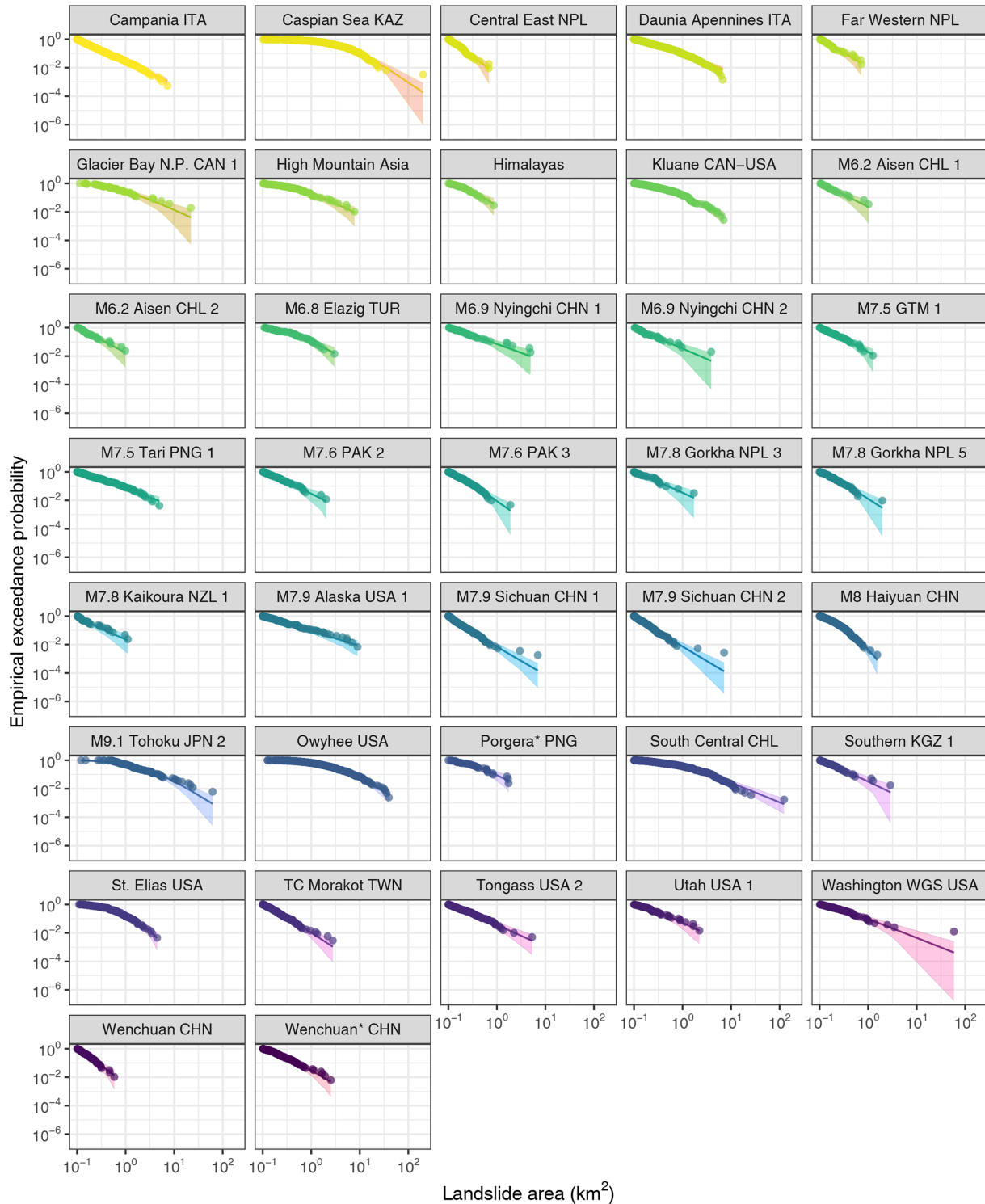
Figure 2. Size and frequency of large landslides from 37 inventories that reported at least 25 slope failures affecting $\geq 0.1\,\mathrm{km^2}$ each (Table 1). Circles are observed data ranked by their empirical exceedance probabilities, and lines are posterior medians of a fitted multi-level generalised Pareto distribution (GPD) with shaded 95 % highest-density intervals (HDIs). Three-letter codes are ISO country identifiers, and numbers refer to different catalogue versions of the same triggering event; catalogues related to earthquakes show estimated magnitude. The * denotes inventories derived from deep learning.

that the model overestimates (underestimates) the empirical exceedance probability of a given landslide area.

We find that the log-odds ratios reveal the most mismatches at either extreme end of the size range, though without any consistency across the inventories (Fig. 3). For example, the model underestimates the exceedance probabilities of landslides $< 0.8\,km^2$ in the M9.1 Tohoku JPN 2 (Tanyaş et al., 2017) and Owyhee USA (Safran et al., 2011) catalogues but overestimates the exceedance probabilities of landslides $> 1.7\,km^2$ in the Daunia Apennines ITA catalogue (Ardizzone et al., 2023).

3.2 Effects of different landslide inventories

Our model estimates that shape parameters k_j vary across the landslide inventories with posterior medians ranging from $\tilde{k}_j = 0.21$ in catalogue M8 Haiyuan CHN (Xu et al., 2020) to $\tilde{k}_j = 0.92$ in catalogue M7.9 Alaska USA 1 (Gorum et al., 2014) (Fig. 4). Narrower posterior distributions mean less uncertainty, mainly owing to more large landslides that inform the model in the relevant catalogue. For example, the Campania ITA inventory (Fusco et al., 2023) contains the most, i.e. 1854, large landslides, and its 95 % HDI is narrowest ($0.57 < k_j < 0.72$). In contrast, the M6.2 Aisen CHL 1 catalogue (Gorum et al., 2014) has the fewest, i.e. 29, large landslides; its broad posterior distribution is thus informed more by the pooled estimate from all inventories together.

The mean of the gamma prior distribution is $\overline{k} = \alpha_k/\beta_k$ by definition, and we derive the power-law exponent α from the identity $\alpha = 1/k$. Similarly, we obtain the mean pooled posterior $\overline{\sigma} = \alpha_\sigma/\beta_\sigma$ from the sampled hyper-parameters. We find that most of the 95 % credible intervals of k_j overlap with that of the mean $\overline{k}$ learned from the pooled model (vertical grey line, flanked by dashed lines marking its 95 % HDI). Only two inventories, i.e. one on historic rock avalanches in the St. Elias Mountains of Alaska, United States (Bessette-Kirton and Coe, 2020), and one on landslides triggered by the 1920 Haiyuan earthquake, China (Xu et al., 2020), stand out with a k_j that is credibly below that of the population average.

Estimates of k_j differ credibly between inventories in the same geographic region, e.g., western Canada and Alaska when comparing St. Elias USA (Bessette-Kirton and Coe, 2020), Kluane CAN-USA (William Smith, personal communication, 2022), and M7.9 Alaska USA 1 (Gorum et al., 2014). In contrast, inventories covering very different geographic regions and time spans have largely overlapping, and thus statistically indifferent, posterior distributions of k_j. This is the case, for example, for a catalogue of landslides triggered by the M7.6 Kashmir earthquake in 2005, M7.6 PAK 3 (Basharat et al., 2016), and one on mostly Quaternary landslides in the Caspian Sea basin, Caspian Sea KAZ (Pánek et al., 2016). Similarly, the inventory of rainfall-triggered landslides in far western Nepal covering 79 time steps between 2002 and 2018, Far Western NPL (Muñoz-

Torrero, 2020), and the one for landslides following the 2018 M7.5 Porgera earthquake in Papua New Guinea, a database fully compiled by a deep learning algorithm, Porgera* PNG (Bhuyan et al., 2023), have indistinguishable posterior distributions of k_j. The same goes for inventories mapped by different teams in response to the same earthquake trigger, such as the 2008 Wenchuan earthquake: M7.9 Sichuan CHN 1 (Xu et al., 2014) and M7.9 Sichuan CHN 2 (Li et al., 2014) (Fig. 4).

The spread of the posterior scale parameter σ_j is more pronounced across the inventories, and the pooled estimate overlaps with those of seven inventories only (Fig. 5). Inventory-specific medians range over 2 orders of magnitude from $\tilde{\sigma}_j = 0.04\,km^2$ in the M7.9 Sichuan CHN 1 catalogue (Xu et al., 2014) to $\tilde{\sigma}_j = 2.56\,km^2$ in the Caspian Sea KAZ catalogue (Pánek et al., 2016). Higher values of $\tilde{\sigma}_j$ mark inventories with the more curved fits in Fig. 2, especially those of Quaternary landslides, such as those of the Caspian Sea, Caspian Sea KAZ (Pánek et al., 2016), or the Columbia River basins, Owyhee USA (Safran et al., 2011), but also the above-mentioned inventories of rock avalanches that happened in the past few decades (St. Elias USA and Kluane CAN-USA). Again, inventories with different environmental settings and landslide triggers have very similar posterior distributions of σ_j, such as the one for landslides triggered during Typhoon Morakot, Taiwan, in 2009, TC Morakot TWN (Emberson et al., 2022), and the one for landslides triggered by the M7.6 Kashmir earthquake in 2005, M7.6 PAK 3 (Basharat et al., 2016). Catalogues addressing the same earthquake trigger have largely overlapping posteriors of σ_j.

3.3 Pooled estimates of landslide scaling

The pooled estimates in our multi-level model express the variance of the learned parameters across all inventories. The sampled hyper-parameters of k_j that describe the shape α_k and rate (or inverse scale) β_k of the gamma-distributed parameter k_j are positively correlated; the same applies for the hyper-parameters α_σ and β_σ (Fig. 6). From these, we find that the numerical approximation of the joint posterior distribution has a distinct maximum. We obtain a mean power-law exponent of $1.37 < \overline{\alpha} < 1.85$ across all inventories with 95 % probability; the posterior median of $\overline{\alpha}$ is 1.6 (Fig. 7). Compared to the prior distribution based on published values of this exponent, our model has gained more certainty from the data considered in this study, yielding a much narrower posterior.

The mean scale parameter is $0.18\,km^2 < \overline{\sigma} < 0.38\,km^2$ across all inventories with 95 % probability; the posterior median of $\overline{\sigma}$ is $0.27\,km^2$. This posterior shifted up from the prior distribution that we centred on our arbitrary size threshold for large landslides. Our model has learned much from the inventory data compared to the priors, especially concerning
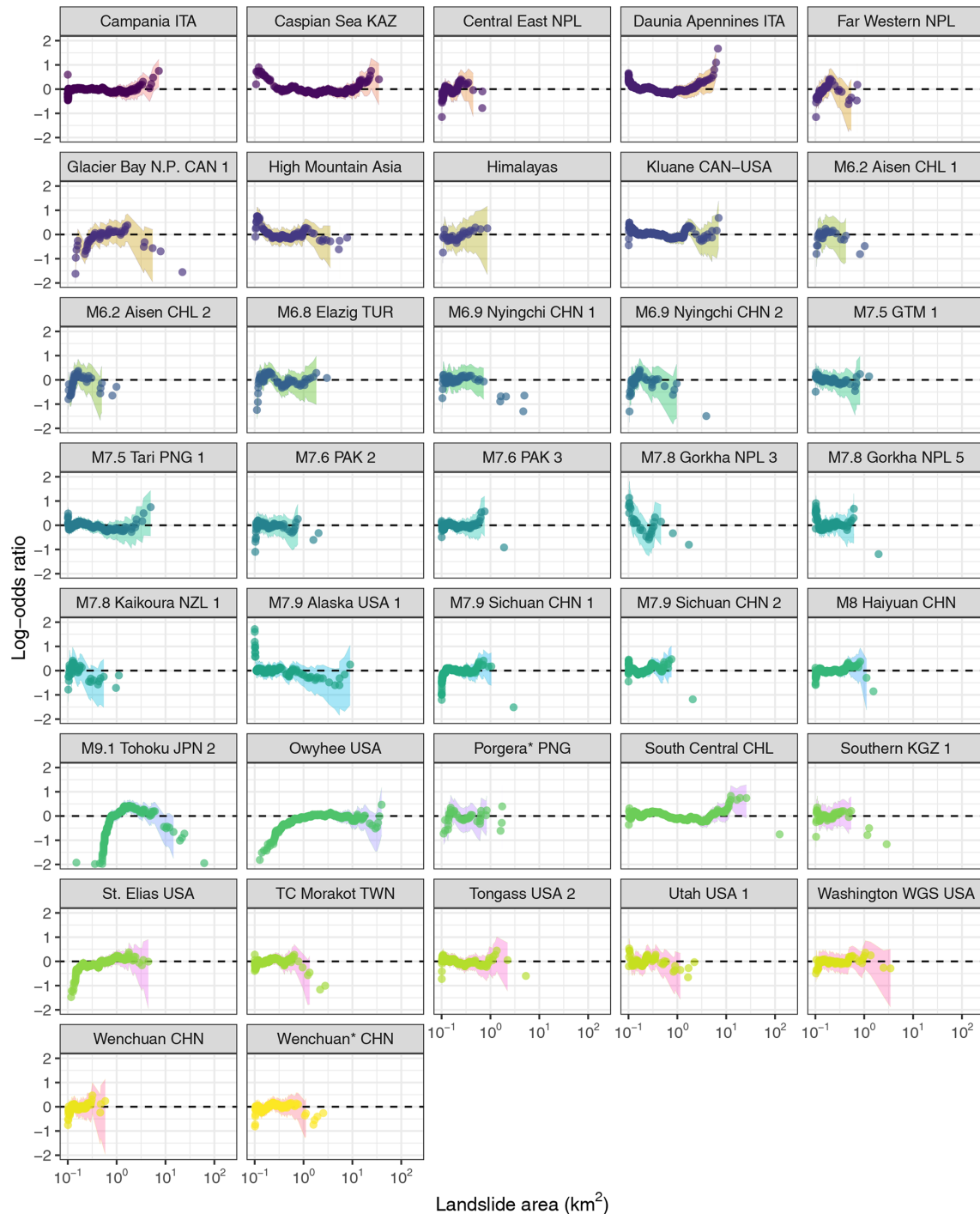
Figure 3. Residuals of multi-level GPD fit to large landslide size distributions (Fig. 2); residuals are expressed as log-odds ratios of observed versus predicted values. Dashed horizontal lines mark perfect fits; positive (negative) ratios indicate over-estimated (under-estimated) exceedance probabilities. Shaded areas are point-wise 95 % HDIs estimated at each landslide observation.

the high variance of σ_j across the individual landslide catalogues.

3.4 Comparison with maximum likelihood estimate

To assess how our choice of Bayesian inference aligns with alternative approaches, we compare our results to maximum likelihood estimates (MLEs) of the exponent α_j of the in-
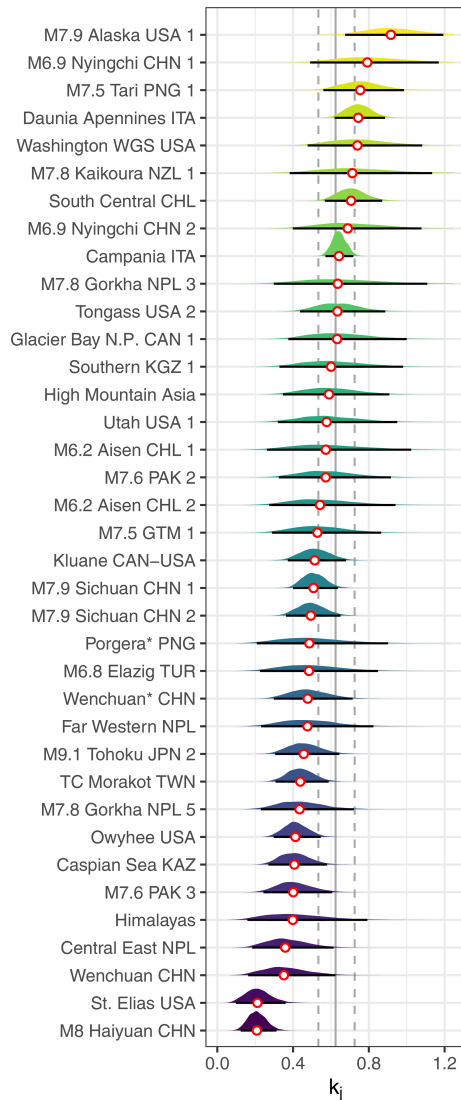
Figure 4. Posterior shape parameters k_j of the multi-level GPD model; this parameter is the inverse of the scaling exponent in power-law distributions. White circles are the medians per inventory, and horizontal black lines are 95 % HDIs; the vertical grey line is the posterior median of the pooled model, and dashed lines delimit its 95 % HDI.

Figure 5. Posterior estimates of the GPD scale parameter σ_j from the multi-level model. White circles are the medians per inventory, and horizontal black lines are 95 % HDIs; the vertical grey line is the posterior median of the pooled model, and dashed lines delimit its 95 % HDI. The dashed vertical red line is the landslide size threshold μ fixed at $0.1\,\text{km}^2$.

verse power-law distribution, based on the Hill estimator (Clauset et al., 2009). By definition, the MLE standard error for each inventory decays with the inverse square root of sample size, whereas the Bayesian estimates are informed by all data via the multi-level model structure. Hence we do not expect a 1 : 1 correspondence from this comparison. Instead, it underlines how variable and uncertain landslide scaling estimates can be for different inventories, regardless of method (Fig. 8).

We obtain inventory-specific MLEs of $0.33 < \hat{\alpha}_j < 2.39$. This spread encompasses most reported values in the literature (Tebbens, 2020). In contrast, the posterior me-

dians of $\alpha_j = 1/k_j$ occupy a seemingly broader range $(1.09 < \tilde{\alpha} < 4.81)$, though it is nominally similar to that of the MLE method for most inventories within the respective errors. However, the coefficient of variation is narrower for the Bayesian median estimates. Two inventories stand out with very high scaling exponents, i.e. M8 Haiyuan CHN and St. Elias USA. Both inventories have only a few landslides that we censored because they were below the size threshold $\mu = 0.1\,\text{km}^2$; in other words, these catalogues contained mostly large landslides originally.
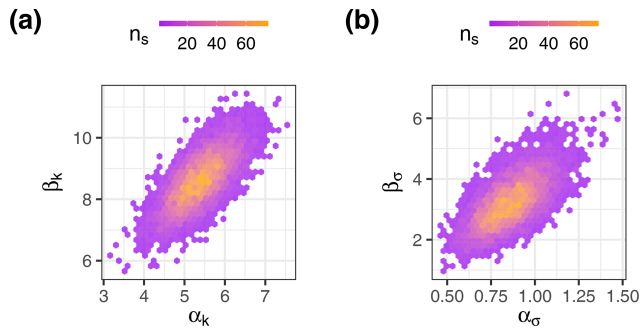
Figure 6. Numerical approximations of the hyper-parameter distributions of the GPD multi-level model; sample densities are used to infer probability densities (n_s is sample size). **(a)** Posterior distribution of the hyper-parameters of the gamma distribution from which the inventory-specific posterior k_j values are drawn (Eq. 5). **(b)** Posterior distribution of the hyper-parameters of the gamma distribution from which the inventory-specific posterior σ_j values are drawn (Eq. 4).

4 Discussion

4.1 Implications

We offer estimates of scaling statistics that characterise the size distribution of rare, large landslides ($\geq 0.1\,\mathrm{km}^2$), informed by thousands of data points from dozens of inventories, as well as findings from previous research. Instead of extrapolating models fit to mainly smaller landslides, we use a dedicated peak-over-threshold approach that uses observations on large landslides exclusively. The narrowest 95 % HDI of α_j, and thus the best we can constrain this parameter, is that of the Campania ITA catalogue with $1.39 < \alpha_j < 1.74$. This numerical range has much overlap with that of previously reported scaling exponents that were obtained for mostly smaller landslides though (Tebbens, 2020). Still, the nearly triangular posterior distribution has much of its probability mass near its peak (Fig. 4), and the same goes for the pooled estimate (Fig. 7a). Except for two cases, the posterior α_j values for landslide inventories remain indistinguishable from the pooled estimate. This low variance of α_j across inventories is striking if we consider the diverse mapping techniques, levels of data quality, coverage, environmental setting, and landslide triggers. The inventories we selected cover several climatic zones with different vegetation and land cover characteristics that likely affect the preservation, detection, and mapping of evidence of large landslides. Moreover, some inventories were generated from deep learning algorithms (Bhuyan et al., 2023), whereas most others were mapped manually. While many posterior estimates of σ_j are close to our arbitrary size threshold for large landslides of $0.1\,\mathrm{km}^2$, the variance in this parameter is high compared to the pooled estimate. Overall, the median estimates of both α_j and σ_j are unaffected by the number of large landslides reported in a given inventory (Fig. 9).

Our model highlights several inventories with scaling statistics that stand out. The M8 Haiyuan CHN and St. Elias USA catalogues have high estimates of α_j, while Caspian Sea KAZ, M9.1 Tohoku JPN 2, and Owyhee USA have high estimates of σ_j well beyond the central tendency of most other catalogues (Fig. 9). These inventories consist almost exclusively of large landslides and have much fewer landslides below the size threshold μ. In contrast, we observe that inventories with the most landslides falling below μ have low values of σ_j consistently. We infer that landslide catalogues that were focused on compiling information about larger landslides tend to have elevated values of σ_j. In this context, our selection of inventories is nearly balanced: 20 of them have less than 10 % large landslides, whereas 17 of them have more than 50 %. One possible explanation for these outlying landslide size statistics is that the GPD is a poor fit to inventories that mainly feature large landslides, at least for the chosen threshold μ. The residuals for most of these inventories show pronounced underestimates for the smallest landslide range, except for the South Central CHL and Caspian Sea KAZ catalogues (Fig. 3). Yet, other inventories with similar residuals (e.g. Glacier Bay N.P. CAN 1) hardly stand out compared to the pooled estimates. Clearly, extrapolating the model across the full size range, for example to infer the number of seemingly missing, underreported, or overlooked landslides (Tanyaş et al., 2019), can be misleading in these cases. Another explanation for the high estimates of k_j and σ_j is that the original mapping was focused on landslides close to, or well above, our choice of $\mu = 0.1\,\mathrm{km}^2$ such that the undersampling of landslides near this size threshold may explain some of the variance of scaling estimates. Reconstructing historic landslide episodes from old air photos and preserved geomorphic evidence (e.g. M8 Haiyuan CHN) may also add variance. Either way, the strategy for keeping such mapping practical is to use a size cutoff. Hence, although smaller landslides may be recognised, they are excluded and thus censored in these inventories. Some of these inventories also contain partly overlapping slope failure deposits of multiple ages, marking several phases of reactivation. Such overlaps may cause more landslides to surpass the size threshold. Hence, the mapping objective would partly bias estimates of σ_j for a size threshold that is too low.

We also find that the 95 % credible intervals of both GPD parameters overlap for landslide inventories regardless of whether they were attributed to recent earthquakes and rainstorms or whether they integrate landslide observations, and thus likely various triggers, over many years (Fig. 10). We infer that the scaling statistics disclose very little about the type of landslide trigger. In this context, our findings caution against a mechanistic interpretation of scaling parameters, at least for large slope failures. These can have longer and more complex histories of precursory slope deformation and failure than smaller landslides (Korup et al., 2007), and they likely respond to stresses that accumulate over repeated episodes of earthquake shaking or rainstorms.
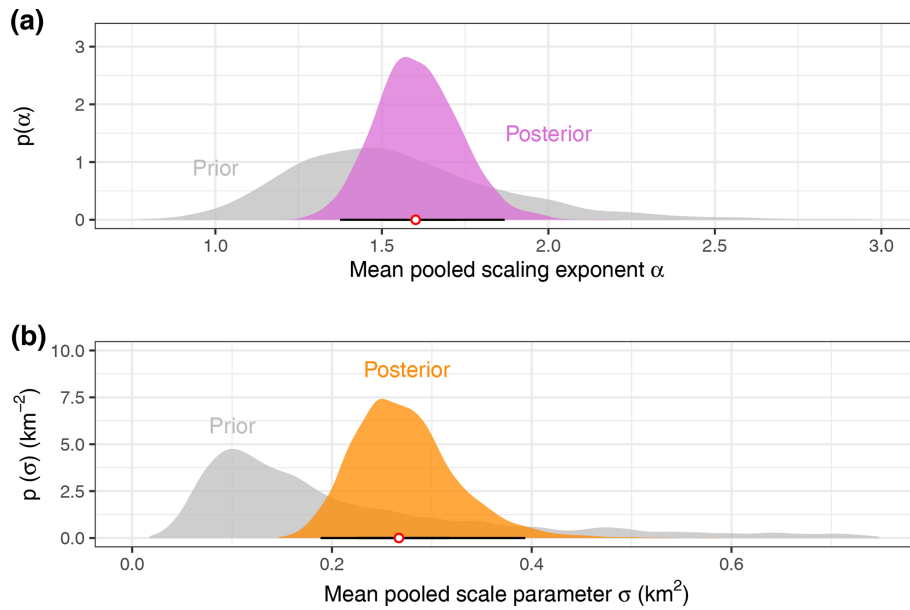
Figure 7. (a) Prior and posterior distributions of inverse power-law scaling exponent α for the pooled model across all landslide inventories. The white circle is the pooled posterior median, and the horizontal black line is 95 % HDI. **(b)** Prior and posterior distributions of GPD scale parameter σ for the pooled model. The white circle is the pooled posterior median, and the horizontal black line is 95 % HDI.
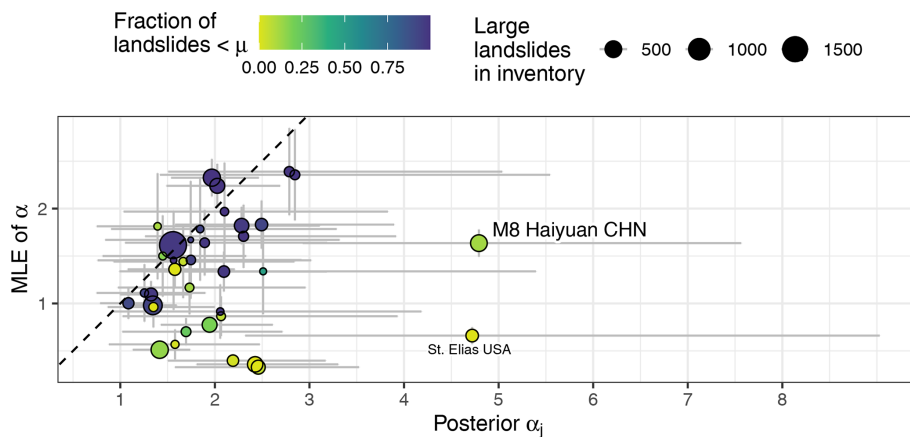

Figure 8. Comparison between posterior GPD estimates and the maximum likelihood estimates (MLEs) of scaling exponents α for each inventory. The dashed $1:1$ line is for visual comparison only. Colour scale shows the fraction of (discarded) landslides below the size threshold μ in each inventory; bubbles are scaled to the sample size of large landslides used for parameter estimation. Vertical grey bars span 2 standard errors around the mean; grey horizontal bars are 95 % HDIs. Axes are scaled equally.

4.2 Role of size threshold and sample size

The heavy-tailed distribution of landslides means that we discarded many samples for small increases in μ. To test the sensitivity of our results to the choice of size threshold μ, we replicated our analyses and recorded the variation in the pooled estimates $\overline{\alpha}$ and $\overline{\sigma}$ as a function of both μ and the minimum number of large landslides that an inventory needs to have to be included in our model. We find that varying the size threshold such that $0.075\,\text{km}^2 < \mu < 0.3\,\text{km}^2$ returns posterior pooled values of α that decrease slightly with in-

creasing μ and the minimum number of samples per inventory, though with much overlap (Fig. 11). Overall, the $\overline{\alpha}$ value is $1.23 < \overline{\alpha} < 1.9$ with 95 % probability regardless of the threshold or sample size that we pick. Estimates of $\overline{\sigma}$ increase slightly with μ but also with some overlap. Preferring larger inventories for a given size threshold reduces the total sample size such that the pooled posterior distributions of σ get broader.

We infer that the choice of μ, and hence the definition of "large" landslides (McColl and Cook, 2024), has limited influence on the scaling statistics and especially $\overline{\alpha}$. Values of μ
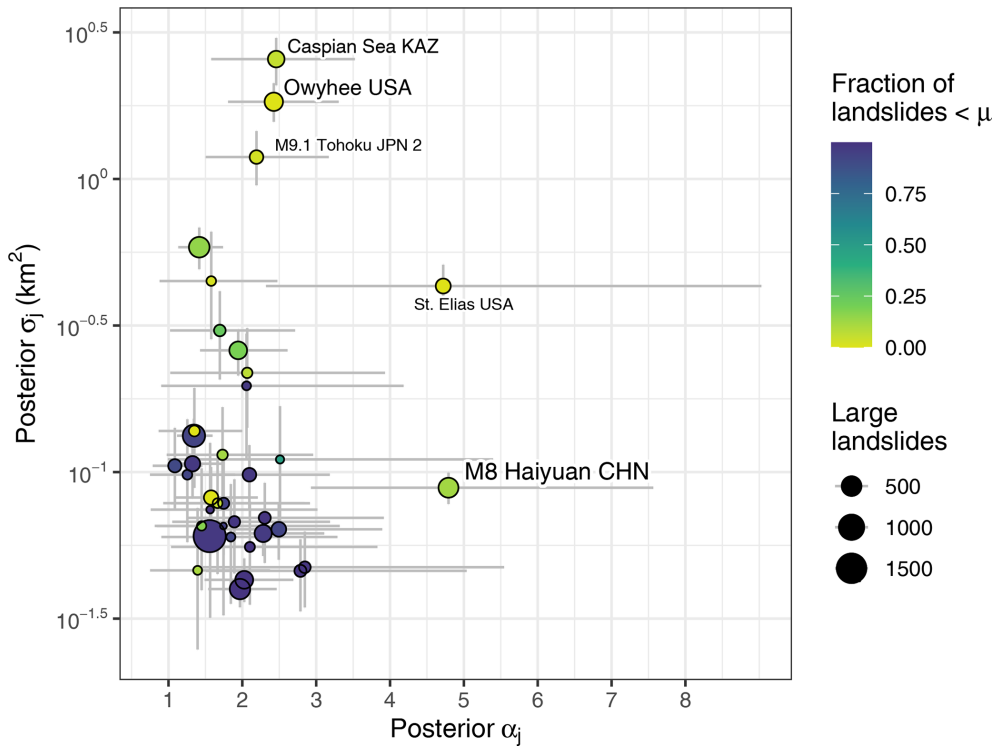
Figure 9. Median posterior parameter estimates as a function of the number of large landslides per inventory (bubble size) and the fraction of discarded landslides below the size threshold (colour scale). Grey bars span the 95 % HDIs.
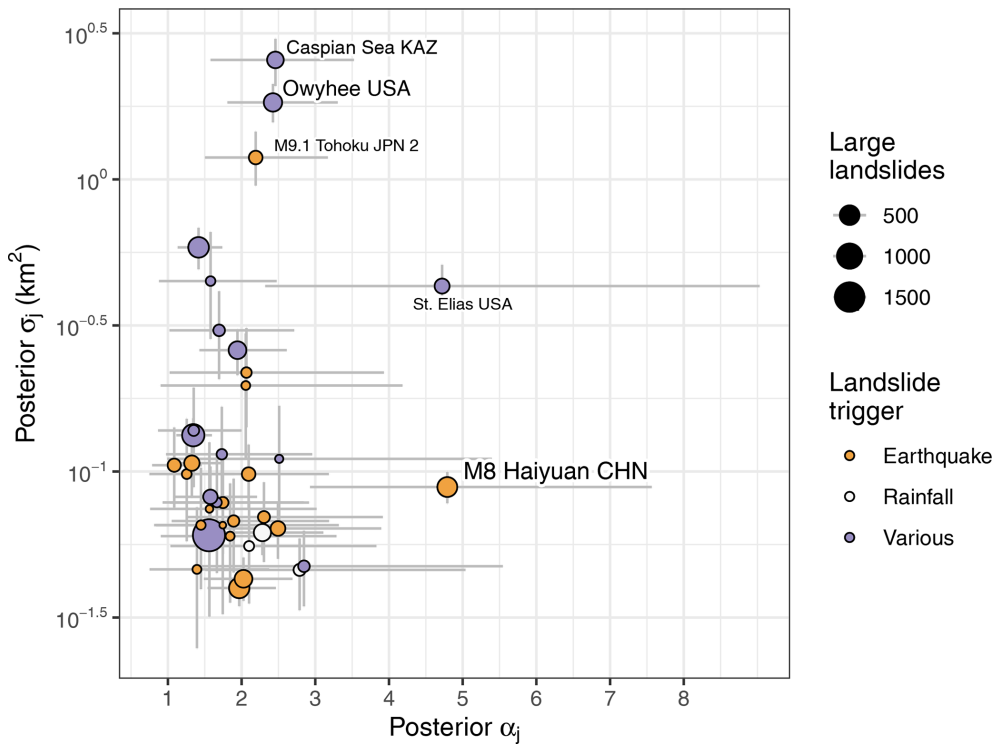

Figure 10. Median posterior parameter estimates as a function of sample size (bubble size) and type of trigger (colour scale).
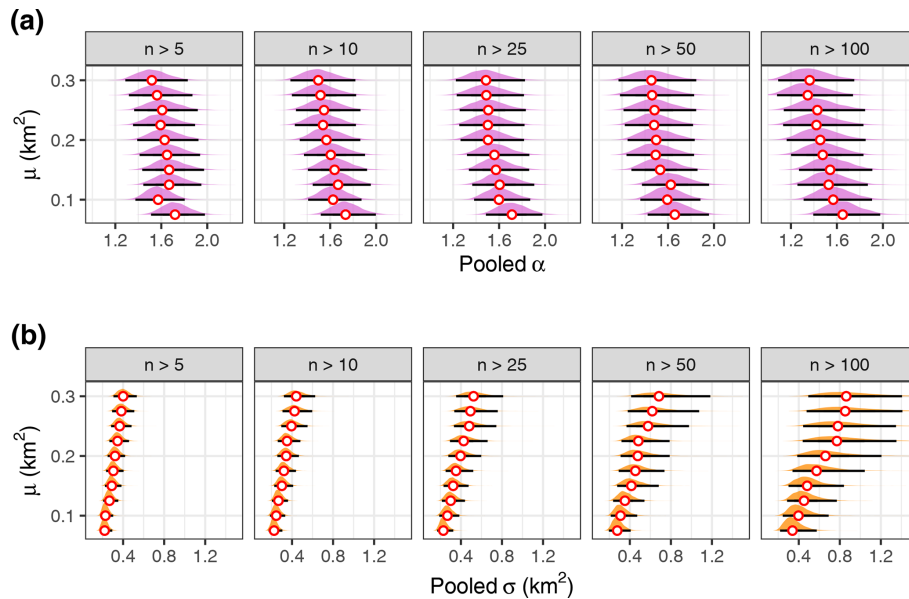
Figure 11. Effect of varying size threshold μ and minimum sample size n in each inventory on posterior estimates of the pooled scaling exponent α and the pooled parameter σ. White bubbles are posterior medians, and horizontal black bars delimit the 95 % HDIs.

below the range that we tested violate the assumption of a high threshold with the result that a GPD would be inappropriate, whereas values above this range suffer from sample sizes that are too small. The pooled scale parameter $\overline{\sigma}$ shows less variation with μ and has largely overlapping posteriors. Hence, our choice of the minimum number of large landslides per inventory ($n = 25$) limits both the number of groups and the overall sample size in our model. Admitting more inventories that contain fewer large landslides changes the posterior $\overline{k}$ and $\overline{\sigma}$ only slightly, but it does narrow the uncertainties, especially for higher thresholds μ.

4.3 Benefits

Our Bayesian multi-level approach expands on previous, though largely separate, efforts of comparing landslide size statistics across different inventories (Tebbens, 2020). We offer here a single, consistent model that has several benefits.

First, the Bayesian set-up can handle the small sample problem of large landslides. Scaling parameters for large landslides from a single landslide inventory are commonly estimated from extrapolating model fits that largely draw on more numerous, smaller landslides. Yet, even simulated power-law distributed data have natural scatter in the largest of observations (Clauset et al., 2009), making it difficult to validate extrapolations and leading to overconfident parameter estimates, especially when ignoring the attached errors (Fig. 8). Using data from other landslide inventories to validate these estimates tacitly assumes that the scaling parameters have similar errors but offers no way of determining whether this assumption is at all valid. The multi-level model instead draws on the larger sample size from all inventories

and explicitly refines this shared knowledge in dedicated posterior distributions for each catalogue. These group-level parameter estimates tend to be closer to the pooled mean than those derived for separate models using fewer data from each group alone. This effect is known as parameter shrinkage (McElreath, 2016) and guards against overfitting, especially for inventories with few data.

Second, the Bayesian treatment quantifies all parameter uncertainties explicitly and especially those that capture previous knowledge about landslide size distributions. We can thus quantify how much we have learned from the data by comparing the posterior and prior distributions (Fig. 7). By design, a Bayesian model seeks a compromise between these previous findings and the data considered here in a probabilistically consistent way. To this end, we made sure to include mostly recently published landslide inventories or those that had not been considered in scaling studies before.

Third, the multi-level model structure enables direct comparison of parameter estimates across and between landslide inventories (Fig. 10). Any differences in the underlying workflows of detecting and mapping landslides and the commensurate sample sizes are being accounted for by separate posterior distributions and their deviation from the pooled estimates (Figs. 4 and 5). Our model measures objectively how similar landslide inventories are in terms of the scaling parameters that jointly, instead of separately, define the size distributions of large landslides.

Fourth, the peak-over-threshold approach that defines the GPD is rooted in extreme value theory and thus expresses what we can expect statistically from the size distribution of large landslides. The parameters of the GPD contain infor-

mation about the power-law scaling and translate readily into parameters of other distributions used to characterise the size scaling of landslides. Given that most inventories focussing on large landslides have to operate on a lower size threshold, we argue that the GPD is a natural choice for characterising the size distributions of more extreme slope failures.

Finally, we can flexibly modify our multi-level model in several ways. One option is to group the data in other ways than by inventories. For example, we can specify the group levels such that they represent dominant landslide, soil, or rock type or any other characteristic that may have been collected during the process of compiling the landslide inventory. We discarded the option of using the type of landslide trigger as a group level because we only have three inventories of rainfall-triggered landslides, so posterior estimates of scaling parameters might rely too much on the more numerous data in earthquake-triggered and multi-event catalogues. Adding inventory type as yet another group level would expand the parameter space and unnecessarily add bias for multi-event catalogues that are likely dominated by an unknown fraction of either rainfall- or earthquake-triggered landslides. Instead, our choice of priors remains impartial to inventory type. We recall that the GPD is by definition "blind" to data below the threshold μ in that it truncates all observations below the threshold. One alternative is to also directly learn μ from the data, either globally or per inventory, and add further covariates that may control the form of the GPD.

5   Conclusions

We propose a multi-level model as common ground for consistently estimating and comparing size distributions of large and rarely observed slope failures across different inventories. In choosing a peak-over-threshold approach, the generalised Pareto distribution (GPD) reflects what we would expect statistically from a given landslide size distribution. The multi-level set-up remediates the problem of low sample size by making use of all available data for estimating scaling parameters while acknowledging inherent differences across inventories. Our model results based on 37 inventories with 8627 large landslides ($\geq 0.1\,\mathrm{km}^2$) show that the power-law exponent for each inventory $\alpha_j = 1/k_j$ discloses little about the different underlying landslide trigger(s), geographic region, or time span concerning a given inventory, i.e. whether it is event-based or compiles landslides of many different ages. Inventories of mostly undated, prehistoric landslides have scaling exponents α_j that hardly differ from those of historic, event-based catalogues. While several studies have attributed a physical meaning to scaling statistics of landslide size, we argue that some of this meaning might get diluted or even lost in empirical data that may combine confounding controls. We surmise that landslide inventories record these physical processes, though in a mixed way that could admit,

for example, different failure types, rock and soil types, and groundwater conditions. We suspect that most landslide inventories have mixed size distributions. For example, mixing data from inventories with differing size thresholds could add variance to σ_j. At least for the large landslides studied here, the scaling statistics likely reflect bulk physical characteristics instead of variables of a single deterministic model of slope stability. Despite thousands of large landslides to learn from, the uncertainty about α_j spans several decimal points. Taking all inventories together, the pooled α captures most of this variance. The GPD scale parameter σ_j has more spread across inventories and is affected especially by those with few or no landslides below our size threshold. We infer that σ_j is sensitive to the desired landslide size range and likely reflects the influence of mapping choices and specifically the compromise of finding a suitable size threshold. Regardless, the choice of probability distribution used to model landslide areas is arbitrary, and parameter estimates disclose nothing about sample size or completeness. We advise against inferring any completeness from the GPD or any other distribution because probability densities are conditional on a model, and models should be fitted to data and not vice versa. Finally, our model measures objectively how much the scaling statistics differ across inventories within estimation error. Such differences can be vital if using scaling statistics for predicting future landslide hazard in terms of size and frequency (Hergarten, 2023).

Author contributions. OK designed the study, carried out the data pre-processing and analyses, and wrote the manuscript. LVL and JVF helped with the collection and compilation of landslide inventories. All authors contributed equally to editing the final version of the manuscript.

lished maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

References

Abancó, C., Bennett, G. L., Matthews, A. J., Matera, M. A. M., and Tan, F. J.: The role of geomorphology, rainfall and soil moisture in the occurrence of landslides triggered by 2018 Typhoon Mangkhut in the Philippines, Nat. Hazards Earth Syst. Sci., 21, 1531–1550, https://doi.org/10.5194/nhess-21-1531-2021, 2021.

Alberti, S., Leshchinsky, B., Roering, J., Perkins, J., and Olsen, M. J.: Inversions of landslide strength as a proxy for subsurface weathering, Nat. Commun., 13, 6049, https://doi.org/10.1038/s41467-022-33798-5, 2022.

Antinao, J. L. and Gosse, J.: Large rockslides in the Southern Central Andes of Chile (32–34.5° S): Tectonic control and significance for Quaternary landscape evolution, Geomorphology, 104, 117–133, https://doi.org/10.1016/j.geomorph.2008.08.008, 2009.

Ardizzone, F., Bucci, F., Cardinali, M., Fiorucci, F., Pisano, L., Santangelo, M., and Zumpano, V.: Geomorphological landslide inventory map of the Daunia Apennines, southern Italy, Earth Syst. Sci. Data, 15, 753–767, https://doi.org/10.5194/essd-15-753-2023, 2023.

Barlow, J., Lim, M., Rosser, N., Petley, D., Brain, M., Norman, E., and Geer, M.: Modeling cliff erosion using negative power law scaling of rockfalls, Geomorphology, 139–140, 416–424, https://doi.org/10.1016/j.geomorph.2011.11.006, 2012.

Basharat, M., Rohn, J., Baig, M. S., and Khan, M. R.: Spatial distribution analysis of mass movements triggered by the 2005 Kashmir earthquake in the Northeast Himalayas of Pakistan, Geomorphology, 206, 203–214, https://doi.org/10.1016/j.geomorph.2013.09.025, 2014.

Basharat, M., Ali, A., Jadoon, I. A. K., and Rohn, J.: Using PCA in evaluating event-controlling attributes of landsliding in the 2005 Kashmir earthquake region, NW Himalayas, Pakistan, Nat. Hazards, 81, 1999–2017, https://doi.org/10.1007/s11069-016-2172-9, 2016.

Behling, R., Roessner, S., Golovko, D., and Kleinschmit, B.: Derivation of long-term spatiotemporal landslide activity—A multi-sensor time series approach, Remote Sens. Environ., 186, 88–104, https://doi.org/10.1016/j.rse.2016.07.017, 2016.

Belair, G., Jones, E., Slaughter, S., and Mirus, B.: Landslide Inventories across the United States version 2, USGS, https://doi.org/10.5066/P9FZUX6N, 2022.

Bellugi, D. G., Milledge, D. G., Cuffey, K. M., Dietrich, W. E. J., and Larsen, L. G.: Controls on the size distributions of shallow landslides, P. Natl. Acad. Sci. USA, 118, e2021855118, https://doi.org/10.1073/pnas.2021855118, 2021.

Bernard, T. G., Lague, D., and Steer, P.: Beyond 2D landslide inventories and their rollover: synoptic 3D inventories and volume from repeat lidar data, Earth Surf. Dynam., 9, 1013–1044, https://doi.org/10.5194/esurf-9-1013-2021, 2021.

Bessette-Kirton, E. K. and Coe, J. A.: A 36-Year Record of Rock Avalanches in the Saint Elias Mountains of Alaska, With Implications for Future Hazards, Front. Earth Sci., 8, 293, https://doi.org/10.3389/feart.2020.00293, 2020.

Bhuyan, K., Tanyaş, H., Nava, L., Puliero, S., Meena, S. R., Floris, M., van Westen, C., and Catani, F.: Generating multitemporal landslide inventories through a general deep transfer learning strategy using HR EO data, Sci. Rep.-UK, 13, 162, https://doi.org/10.1038/s41598-022-27352-y, 2023.

Brardinoni, F., Slaymaker, O., and Hassan, M. A.: Landslide inventory in a rugged forested watershed: a comparison between air-photo and field survey data, Geomorphology, 54, 179–196, https://doi.org/10.1016/S0169-555X(02)00355-0, 2003.

Burrows, K., Marc, O., and Remy, D.: Using Sentinel-1 radar amplitude time series to constrain the timings of individual landslides: a step towards understanding the controls on monsoon-triggered landsliding, Nat. Hazards Earth Syst. Sci., 22, 2637–2653, https://doi.org/10.5194/nhess-22-2637-2022, 2022.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A.: Stan: A Probabilistic Programming Language, J. Stat. Softw., 76, 1–32, https://doi.org/10.18637/jss.v076.i01, 2017.

Casagli, N., Intrieri, E., Tofani, V., Gigli, G., and Raspini, F.: Landslide detection, monitoring and prediction with remote-sensing techniques, Nature Reviews Earth & Environment, 4, 51–64, https://doi.org/10.1038/s43017-022-00373-x, 2023.

Castro-Camilo, D., Huser, R., and Rue, H.: Practical strategies for generalized extreme value-based regression models for extremes, Environmetrics, 33, e2742, https://doi.org/10.1002/env.2742, 2022.

Clauset, A., Shalizi, C. R., and Newman, M. E. J.: Power-Law Distributions in Empirical Data, SIAM Rev., 51, 661–703, https://doi.org/10.1137/070710111, 2009.

Dente, E., Katz, O., Crouvi, O., and Mushkin, A.: The Geomorphic Effectiveness of Landslides, J. Geophys. Res.-Earth, 128, e2023JF007191, https://doi.org/10.1029/2023JF007191, 2023.

Domènech, G., Yang, F., Guo, X., Fan, X., Scaringi, G., Dai, L., He, C., Xu, Q., and Huang, R.: Two multi-temporal datasets to track the enhanced landsliding after the 2008 Wenchuan earthquake, Zenodo [data set], https://doi.org/10.5281/zenodo.1484667, 2018.

Emberson, R., Kirschbaum, D. B., Amatya, P., Tanyas, H., and Marc, O.: Insights from the topographic characteristics of a large global catalog of rainfall-induced landslide event inventories, Nat. Hazards Earth Syst. Sci., 22, 1129–1149, https://doi.org/10.5194/nhess-22-1129-2022, 2022.

Fan, X., Scaringi, G., Korup, O., West, A. J., van Westen, C. J., Tanyas, H., Hovius, N., Hales, T. C., Jibson, R. W., Allstadt,

K. E., Zhang, L., Evans, S. G., Xu, C., Li, G., Pei, X., Xu, Q., and Huang, R.: Earthquake-Induced Chains of Geologic Hazards: Patterns, Mechanisms, and Impacts, Rev. Geophys., 57, 421–503, https://doi.org/10.1029/2018RG000626, 2019.

Franceschini, R., Rosi, A., Catani, F., and Casagli, N.: Exploring a landslide inventory created by automated web data mining: the case of Italy, Landslides, 19, 841–853, https://doi.org/10.1007/s10346-021-01799-y, 2022.

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

Fusco, F., Tufano, R., De Vita, P., Di Martire, D., Di Napoli, M., Guerriero, L., Mileti, F. A., Terribile, F., and Calcaterra, D.: A revised landslide inventory of the Campania region (Italy), Scientific Data, 10, 355, https://doi.org/10.1038/s41597-023-02155-6, 2023.

Gilham, J., Barlow, J., and Moore, R.: Marine control over negative power law scaling of mass wasting events in chalk sea cliffs with implications for future recession under the UKCP09 medium emission scenario, Earth Surf. Proc. Land., 43, 2136–2146, https://doi.org/10.1002/esp.4379, 2018.

Gorum, T., Korup, O., van Westen, C. J., van der Meijde, M., Xu, C., and van der Meer, F. D.: Why so few? Landslides triggered by the 2002 Denali earthquake, Alaska, Quaternary Sci. Rev., 95, 80–94, https://doi.org/10.1016/j.quascirev.2014.04.032, 2014.

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

Hao, L., Rajaneesh A., van Westen, C., Sajinkumar K. S., Martha, T. R., Jaiswal, P., and McAdoo, B. G.: Constructing a complete landslide inventory dataset for the 2018 monsoon disaster in Kerala, India, for land use change analysis, Earth Syst. Sci. Data, 12, 2899–2918, https://doi.org/10.5194/essd-12-2899-2020, 2020.

Harp, E. L., Wilson, R. C., and Wieczorek, G. F.: Landslides from the February 4, 1976, Guatemala earthquake, Professional Paper 1204-A, USGS, https://doi.org/10.3133/pp1204A, 1981.

Hergarten, S.: The concept of event-size-dependent exhaustion and its application to paraglacial rockslides, Nat. Hazards Earth Syst. Sci., 23, 3051–3063, https://doi.org/10.5194/nhess-23-3051-2023, 2023.

Hergarten, S. and Neugebauer, H. J.: Self-organized criticality in a landslide model, Geophys. Res. Lett., 25, 801–804, https://doi.org/10.1029/98GL50419, 1998.

Hibert, C., Michéa, D., Provost, F., Malet, J.-P., and Geertsema, M.: Exploration of continuous seismic recordings with a machine learning approach to document 20 yr of landslide activity in Alaska, Geophys. J. Int., 219, 1138–1147, https://doi.org/10.1093/gji/ggz354, 2019.

Hu, K., Zhang, X., You, Y., Hu, X., Liu, W., and Li, Y.: Landslides and dammed lakes triggered by the 2017 Ms6.9 Milin earthquake in the Tsangpo gorge, Landslides, 16, 993–1001, https://doi.org/10.1007/s10346-019-01168-w, 2019.

Jain, S., Khosa, R., and Gosain, A. K.: Impact of landslide size and settings on landslide scaling relationship: a study from the Himalayan regions of India, Landslides, 19, 373–385, https://doi.org/10.1007/s10346-021-01794-3, 2022.

Jones, J. N., Boulton, S. J., Bennett, G. L., Stokes, M., and Whitworth, M. R. Z.: Temporal Variations in Landslide Distributions Following Extreme Events: Implications for Landslide Susceptibility Modeling, J. Geophys. Res.-Earth, 126, e2021JF006067, https://doi.org/10.1029/2021JF006067, 2021.

Karakas, G., Nefeslioglu, H. A., Kocaman, S., Buyukdemircioglu, M., Yurur, T., and Gokceoglu, C.: Derivation of earthquake-induced landslide distribution using aerial photogrammetry: the January 24, 2020, Elazig (Turkey) earthquake, Landslides, 18, 2193–2209, https://doi.org/10.1007/s10346-021-01660-2, 2021.

Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Resour., 25, 1287–1304, https://doi.org/10.1016/S0309-1708(02)00056-8, 2002.

Kim, J., Coe, J. A., Lu, Z., Avdievitch, N. N., and Hults, C. P.: Spaceborne InSAR mapping of landslides and subsidence in rapidly deglaciating terrain, Glacier Bay National Park and Preserve and vicinity, Alaska and British Columbia, Remote Sens. Environ., 281, 113231, https://doi.org/10.1016/j.rse.2022.113231, 2022.

Korup, O., Clague, J. J., Hermanns, R. L., Hewitt, K., Strom, A. L., and Weidinger, J. T.: Giant landslides, topography, and erosion, Earth Planet. Sc. Lett., 261, 578–589, https://doi.org/10.1016/j.epsl.2007.07.025, 2007.

Korup, O., Görüm, T., and Hayakawa, Y.: Without power? Landslide inventories in the face of climate change, Earth Surf. Proc. Land., 37, 92–99, https://doi.org/10.1002/esp.2248, 2012.

Lacroix, P., Handwerger, A. L., and Bièvre, G.: Life and death of slow-moving landslides, Nature Reviews Earth & Environment, 1, 404–419, https://doi.org/10.1038/s43017-020-0072-8, 2020.

LaHusen, S. R., Duvall, A. R., Booth, A. M., and Montgomery, D. R.: Surface roughness dating of long-runout landslides near Oso, Washington (USA), reveals persistent postglacial hillslope instability, Geology, 44, 111–114, https://doi.org/10.1130/G37267.1, 2016.

Larsen, I. J., Montgomery, D. R., and Korup, O.: Landslide erosion controlled by hillslope material, Nat. Geosci., 3, 247–251, https://doi.org/10.1038/ngeo776, 2010.

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

Li, G. K. and Moon, S.: Topographic stress control on bedrock landslide size, Nat. Geosci., 14, 307–313, https://doi.org/10.1038/s41561-021-00739-8, 2021.

Liu, J., Wu, Y., and Gao, X.: Increase in occurrence of large glacier-related landslides in the high mountains of Asia, Sci. Rep.-UK, 11, 1635, https://doi.org/10.1038/s41598-021-81212-9, 2021.

Luetzenburg, G., Svennevig, K., Bjørk, A. A., Keiding, M., and Kroon, A.: A national landslide inventory for Denmark, Earth Syst. Sci. Data, 14, 3157–3165, https://doi.org/10.5194/essd-14-3157-2022, 2022.

Luna, L. V. and Korup, O.: Seasonal Landslide Activity Lags Annual Precipitation Pattern in the Pacific Northwest, Geophys. Res. Lett., 49, e2022GL098506, https://doi.org/10.1029/2022GL098506, e2022GL098506 2022GL098506, 2022.

Nat. Hazards Earth Syst. Sci., 24, 3815–3832, 2024

https://doi.org/10.5194/nhess-24-3815-2024

Malamud, B. D., Turcotte, D. L., Guzzetti, F., and Reichenbach, P.: Landslide inventories and their statistical properties, Earth Surf. Proc. Land., 29, 687–711, https://doi.org/10.1002/esp.1064, 2004.

Marc, O. and Hovius, N.: Amalgamation in landslide maps: effects and automatic detection, Nat. Hazards Earth Syst. Sci., 15, 723–733, https://doi.org/10.5194/nhess-15-723-2015, 2015.

Marc, O., Behling, R., Andermann, C., Turowski, J. M., Illien, L., Roessner, S., and Hovius, N.: Long-term erosion of the Nepal Himalayas by bedrock landsliding: the role of monsoons, earthquakes and giant landslides, Earth Surf. Dynam., 7, 107–128, https://doi.org/10.5194/esurf-7-107-2019, 2019.

McColl, S. T. and Cook, S. J.: A universal size classification system for landslides, Landslides, 21, 111–120, https://doi.org/10.1007/s10346-023-02131-6, 2024.

McElreath, R.: Statistical Rethinking. A Bayesian Course with Examples in R and Stan, 1st Edn., Chapman and Hall/CRC, https://doi.org/10.1201/9781315372495, 2016.

Medwedeff, W. G., Clark, M. K., Zekkos, D., and West, A. J.: Characteristic landslide distributions: An investigation of landscape controls on landslide size, Earth Planet. Sc. Lett., 539, 116203, https://doi.org/10.1016/j.epsl.2020.116203, 2020.

Meunier, P., Uchida, T., and Hovius, N.: Landslide patterns reveal the sources of large earthquakes, Earth Planet. Sc. Lett., 363, 27–33, https://doi.org/10.1016/j.epsl.2012.12.018, 2013.

Milledge, D. G., Bellugi, D. G., Watt, J., and Densmore, A. L.: Automated determination of landslide locations after large trigger events: advantages and disadvantages compared to manual mapping, Nat. Hazards Earth Syst. Sci., 22, 481–508, https://doi.org/10.5194/nhess-22-481-2022, 2022.

Muñoz-Torrero, A.: Multi-temporal Landslide Inventory for the Far-Western region of Nepal, Zenodo [data set], https://doi.org/10.5281/zenodo.4290100, 2020.

Pánek, T., Korup, O., Minár, J., and Hradecký, J.: Giant landslides and highstands of the Caspian Sea, Geology, 44, 939–942, https://doi.org/10.1130/G38259.1, 2016.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org. (last access: 7 November 2024), 2021.

Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., Chamlagain, D., and Godt, J. W.: The size, distribution, and mobility of landslides caused by the 2015 Mw7.8 Gorkha earthquake, Nepal, Geomorphology, 301, 121–138, https://doi.org/10.1016/j.geomorph.2017.01.030, 2018.

RStudio Team: RStudio: Integrated Development for R, RStudio, Inc., Boston, MA, https://posit.co/ (last access: 7 November 2024), 2024.

Safran, E., Anderson, S., Mills-Novoa, M., House, P., and Ely, L.: Controls on large landslide distribution and implications for the geomorphic evolution of the southern interior Columbia River basin, Geol. Soc. Am. Bull., 123, 1851–1862, https://doi.org/10.1130/B30061.1, 2011.

Saito, H., Korup, O., Uchida, T., Hayashi, S., and Oguchi, T.: Rainfall conditions, typhoon frequency, and contemporary landslide erosion in Japan, Geology, 42, 999–1002, https://doi.org/10.1130/G35680.1, 2014.

Santangelo, M., Althuwaynee, O., Alvioli, M., Ardizzone, F., Bianchi, C., Bornaetxea, T., Brunetti, M. T., Bucci, F., Cardinali, M., Donnini, M., Esposito, G., Gariano, S. L., Grita, S., Marchesini, I., Melillo, M., Peruccacci, S., Salvati, P., Yazdani, M., and Fiorucci, F.: Inventory of landslides triggered by an extreme rainfall event in Marche-Umbria, Italy, on 15 September 2022, Scientific Data, 10, 427, https://doi.org/10.1038/s41597-023-02336-3, 2023.

Schönfeldt, E., Winocur, D., Pánek, T., and Korup, O.: Deep learning reveals one of Earth's largest landslide terrain in Patagonia, Earth Planet. Sc. Lett., 593, 117642, https://doi.org/10.1016/j.epsl.2022.117642, 2022.

Sepúlveda, S. A., Serey, A., Lara, M., Pavez, A., and Rebolledo, S.: Landslides induced by the April 2007 Aysén Fjord earthquake, Chilean Patagonia, Landslides, 7, 483–492, https://doi.org/10.1007/s10346-010-0203-2, 2010.

Smith, W. D., Dunning, S. A., Ross, N., Telling, J., Jensen, E. K., Shugar, D. H., Coe, J. A., and Geertsema, M.: Revising supraglacial rock avalanche magnitudes and frequencies in Glacier Bay National Park, Alaska, Geomorphology, 425, 108591, https://doi.org/10.1016/j.geomorph.2023.108591, 2023.

Song, C., Yu, C., Li, Z., Utili, S., Frattini, P., Crosta, G., and Peng, J.: Triggering and recovery of earthquake accelerated landslides in Central Italy revealed by satellite radar observations, Nat. Commun., 13, 7278, https://doi.org/10.1038/s41467-022-35035-5, 2022.

Stan Development Team: Stan Modeling Language Users Guide and Reference Manual, Version 2.35, https://mc-stan.org/ (last access: 7 November 2024), 2024.

Tanyaş, H., van Westen, C. J., Allstadt, K. E., Anna Nowicki Jessee, M., Görüm, T., Jibson, R. W., Godt, J. W., Sato, H. P., Schmitt, R. G., Marc, O., and Hovius, N.: Presentation and Analysis of a Worldwide Database of Earthquake-Induced Landslide Inventories, J. Geophys. Res.-Earth, 122, 1991–2015, https://doi.org/10.1002/2017JF004236, 2017.

Tanyaş, H., van Westen, C. J., Allstadt, K. E., and Jibson, R. W.: Factors controlling landslide frequency–area distributions, Earth Surf. Proc. Land., 44, 900–917, https://doi.org/10.1002/esp.4543, 2019.

Tanyaş, H., Görüm, T., Fadel, I., Yıldırım, C., and Lombardo, L.: An open dataset for landslides triggered by the 2016 Mw 7.8 Kaikōura earthquake, New Zealand, Landslides, 19, 1405–1420, https://doi.org/10.1007/s10346-022-01869-9, 2022a.

Tanyaş, H., Hill, K., Mahoney, L., Fadel, I., and Lombardo, L.: The world's second-largest, recorded landslide event: Lessons learnt from the landslides triggered during and after the 2018 Mw 7.5 Papua New Guinea earthquake, Eng. Geol., 297, 106504, https://doi.org/10.1016/j.enggeo.2021.106504, 2022b.

Tebbens, S. F.: Landslide Scaling: A Review, Earth Space Sci., 7, e2019EA000662, https://doi.org/10.1029/2019EA000662, 2020.

ten Brink, U. S., Barkan, R., Andrews, B. D., and Chaytor, J. D.: Size distributions and failure initiation of submarine and subaerial landslides, Earth Planet. Sc. Lett., 287, 31–42, https://doi.org/10.1016/j.epsl.2009.07.031, 2009.

Valagussa, A., Marc, O., Frattini, P., and Crosta, G.: Seismic and geological controls on earthquake-induced landslide size, Earth Planet. Sc. Lett., 506, 268–281, https://doi.org/10.1016/j.epsl.2018.11.005, 2019.

Valagussa, A., Frattini, P., Valbuzzi, E., and Crosta, G. B.: Role of landslides on the volume balance of the Nepal

2015 earthquake sequence, Sci. Rep.-UK, 11, 3434, https://doi.org/10.1038/s41598-021-83037-y, 2021.

Van Den Eeckhaut, M., Poesen, J., Verstraeten, G., Vanacker, V., Moeyersons, J., Nyssen, J., and van Beek, L. P. H.: The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides, Geomorphology, 67, 351–363, https://doi.org/10.1016/j.geomorph.2004.11.001, 2005.

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

Xu, Y., Allen, M. B., Zhang, W., Li, W., and He, H.: Landslide characteristics in the Loess Plateau, northern China, Geomorphology, 359, 107150, https://doi.org/10.1016/j.geomorph.2020.107150, 2020.

Zhao, B., Li, W., Wang, Y., Lu, J., and Li, X.: Landslides triggered by the Ms 6.9 Nyingchi earthquake, China (18 November 2017): analysis of the spatial distribution and occurrence factors, Landslides, 16, 765–776, https://doi.org/10.1007/s10346-019-01146-2, 2019.