Global detection of rainfall triggered landslide clusters

Susanne A. Benz[1, 2], Philipp Blum[1]

[1] Karlsruhe Institute of Technology (KIT), Institute of Applied Geosciences (AGW), Karlsruhe, Germany

[2] University of California San Diego (UCSD), School of Global Policy and Strategy (GPS), La Jolla, CA, USA

Correspondence to: Philipp Blum (blum@kit.edu) and Susanne A. Benz (sabenz@ucsd.edu)

Abstract

An increasing awareness of the cost of landslides on the global economy and of the associated loss of human life, has led to the development of various global landslide databases. However, these databases typically report landslide events instead of individual landslides, i.e. a group of landslides with a common trigger and reported by media, citizens and/or government officials as a single unit. The latter results in significant cataloging and reporting biases. To counteract this biases, this study aims to identify clusters of landslide events that were triggered by the same rainfall event. An algorithm is developed that finds a series of landslide events that a) is continuous with no more than two days between individual events, and b) precipitation at the location of an individual event correlates with precipitation of at least one other event. The developed algorithm is applied to the Global Landslide Catalog (GLC) maintained by NASA. The results show that more than 40 % of all landslide events are connected to at least one other event, and that 14 % of all studied landslide events are actually part of a landslide cluster consisting of at least 10 events and up to 108 events in one day. Duration of the detected clusters also varies greatly from 1 to 24 days. Our study intends to enhance our understanding of landslide clustering and thus will assist in the development of improved, internationally streamlined mitigation strategies for rainfall related landslide clusters.

22   **Keywords:** Landslide events; Database; Extreme weather; Rainfall induced; Early warning

23   systems;

1.   Introduction

25   The fatal and catastrophic nature of landslides has led to the development and maintenance of

26   various global databases, such as the NASA Global Landslide Catalogue (GLC; e.g. Kirschbaum

27   et al. 2015) and recently the Global Fatal Landslide Database (GFLD) by Froude & Petley (2018).

28   Typically, these databases have a distinct focus. For example, the Global Landslide Catalogue

29   (GLC) operated by NASA focuses on rainfall triggered landslides (Kirschbaum et al., 2010, 2015),

30   whereas the Global Fatal Landslide Database records fatal landslides (Froude and Petley, 2018;

31   Petley, 2012). Through these databases we are able to provide first estimates on the number of

32   recorded fatalities, which were more than 55,000 between 2004 and 2016 (Froude and Petley, 2018)

33   and map near real-time risk for landslides almost on a global scale (Kirschbaum and Stanley, 2018).

34   Still, while they play a key role in understanding the effects of landslides on our society, it is

35   important to note that they are primarily based on news and government reports. These databases

36   therefore do not count landslides, but landslide events, which contain either a single or a multitude

37   of landslides within an area that are assumed to be triggered by the same event (Malamud et al.,

38   2004). The exact number of slope failures in each event is often unknown and depends on the

39   quality of the reporting. For some databases this number is included in a parameter of intensity or

40   size of each event. Typically, for large databases however, this is merely qualitative and describes

41   not only the number of individual landslides, but also an impact such as economic or human losses.

42   This classification is commonly based on press releases and is therefore heavily biased on the news

43   outlet reporting each event (e.g. Carrara et al., 2003).

44 Landslides triggered by catastrophic events, such as earthquakes or major storms, are often counted

45 as one event containing thousands of individual landslides (Kirschbaum et al., 2015). In contrast,

46 landslides caused by non-catastrophic events such as reasonable rainfall, are commonly counted as

47 individual events, disregarding their shared trigger. Consequently, the overall extent of clustering

48 in landslides is often unknown. But only if we better understand the extent of clustering between

49 individual landslide events, will we be able to understand the patterns they occur in and have the

50 chance to utilize these patterns to improve our forecast models (e.g. Martelloni et al., 2012).

51 Until now, few studies have focused on rainfall triggered landslide clusters and rather on temporal

52 clusters over a long time period within a confined region (e.g. Samia et al., 2017; Witt et al., 2010).

53 Biasutti et al. (2016) investigated the spatiotemporal clustering due to rainfall events for three

54 selected urban areas of the US West Coast: Seattle, San Francisco and Los Angeles. Over the nine

55 year study period, they found approximately 20 days within each city with multiple (up to eight)

56 landslide events. Additionally, they could identify close to 40 landslide events that were followed

57 by another event within the next week. However, with a focus on only selected study areas, they

58 did not show the overall extend of these clusters.

59 The objective of this study is therefore to develop an algorithm, which is able to identify such

60 clusters on a global scale. By applying the algorithm to the Global Landslide Catalog (GLC) the

61 overall degree of clustering in the database is shown, and spatial patterns of clusters with at least

62 10 landslide events are described. Additionally, landslide events and rainfall patterns of the most

63 intense and longest clusters are comprehensively discussed. In contrast to previous studies, such as

64 by Biasutti et al. (2016), clusters here are not constricted by a maximum spatial extent, instead they

65 are grouped by analyzing and comparing rainfall prior to the event at the event locations.

2. Material and Method

2.1   Landslide Data

All landslide events within this study are part of the Global Landslide Catalog (GLC) operated by

NASA and introduced in Kirschbaum et al. (2010, 2015). Data within the catalogue is based on

online news articles that are found through search engine options such as Google Alerts. In the

presented study, only events with a location accuracy ≤ 25 km are considered. As the rainfall data

used is only available within ± 50° Latitude, landslide events outside of this range are not

considered. Overall, a total of 9279 landslide events, ranging from 1988 to 2018 are analyzed (Fig.

1). However, only 45 of these events occurred before 2007, when the GLC was established.

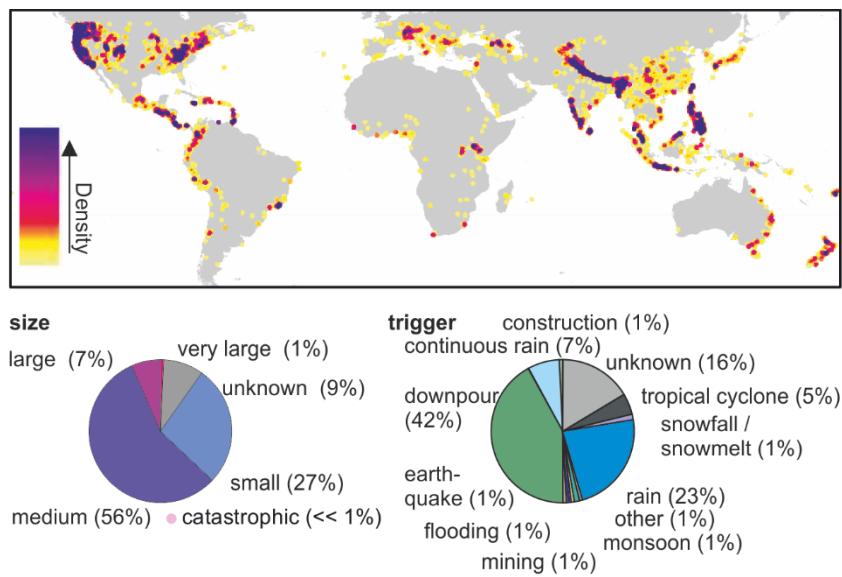



Figure 1. Heat map of all landslide events analyzed in this study and their size and apparent trigger. Overall a total of

9279 events were tested for clustering.

For each event, the GLC provides a landslide type, e.g. land- or mudslide, and a landslide trigger,

e.g. rainfall, downpour, earthquakes or construction work. Detailed descriptions on these

classifications can be found in Kirschbaum et al. (2010, 2015). Furthermore, within the GLC the

intensity, impact, and number of landslides per event is expressed in a variable called "size". While

82   events classified as small in the database are only a single landslide, medium or larger landslide

83   events may consist of multiple landslides within an unspecified range. About 64 % of the studied

84   events are classified as medium or larger in size. However, a precise count of the number of

85   landslides contained within these events does not exist in this database nor in any other of the global

86   scale databases currently available. Within the GLC most of the small events that contain only a

87   single landslide, are located within the United States (Fig. 1).

88   **2.2   Rainfall Data**

89   For the rainfall analysis, the Climate Hazards Group InfraRed Precipitation with Station data

90   (CHIRPS) (Climate Hazards Group, 2015) is used, which has a resolution of $0.05° \times 0.05°$ and

91   daily time steps. For each landslide event location, precipitation data were downloaded for 10 years

92   preceding the event and up to two days after the event using Google Earth Engine (Gorelick et al.,

93   2017).

94   **2.3   Detection of Landslide Clusters**

95   The main objective of this study is to identify clusters of landslide events that occurred during, and

96   are likely triggered by the same rainfall event. To determine if two events, A and B, occurred during

97   the same rainfall event, two conditions have to be fulfilled: (I) A and B occurred within three days

98   of each other, and (II) spearman correlation between daily precipitation at A and at B is greater 0.7

99   and has a p-value less than 0.05 for the 30 days preceding the later of the two events. Other landslide

100 events that fulfill these conditions with either A or B, are considered to be part of the cluster. A

101 schematic drawing of this algorithm is provided in Fig. 2, and a more detailed flowchart in Fig. S1

102 in the supplementary material. The threshold value of three days maximum between two events

103 was used following Biasutti et al. (2016), who found it unlikely that landslide events occurring

104 more than three days apart, occurred during the same rainfall event. However, it is important to

105 note that their study was set in three metropolitan areas on the West Coast of the USA and might

106 not be applicable everywhere.

107 The threshold value of the spearman correlation coefficient was determined by testing the

108 robustness of the identified clusters for different threshold values between zero and one (Fig. S2).

109 Our results indicate that mean duration, area and number of landslide per cluster are comparably

110 robust to changes of the spearman correlation coefficient. In contrast maximum duration, area, and

111 number of landslides per cluster change drastically for different threshold values. From a

112 correlation coefficient threshold of 0.35 to 0.7, maximum number of landslide events per cluster

113 decreases from close to 500 to slightly above 100, maximum duration decreases from more than

114 80 days to approximately 25, and area decreases from 60,000,000 km² (approximately 1/3 of the

115 planet's surface area) to 200,000 km². For threshold values greater 0.7, minor changes are observed.

116 Hence, the latter was set as the correlation threshold value for this study (Fig. S2).

117 Additionally, we tested the robustness of the method to the time period of precipitation for which

118 the correlation coefficient was determined (Fig. S3). It appears that the number of days is much

119 less influential than the set correlation coefficient threshold (Fig. S2). Again, maximum number of

120 landslides, area, and duration are impacted most, however remain stable for time period longer than

121 30 days prior to the second event.

122 It is important to note that the introduced method does not limit the spatial extent of the found

123 landslide clusters. While this ensures that previously undetected, large-scale connections between

124 individual landslide events are found, it is also susceptible to link landslides occurring in different

125 parts of the world, where rainfall coincidentally correlates. Hence, when applying the method to

126 another dataset, the robustness of the threshold values for correlation coefficient and time analyzed

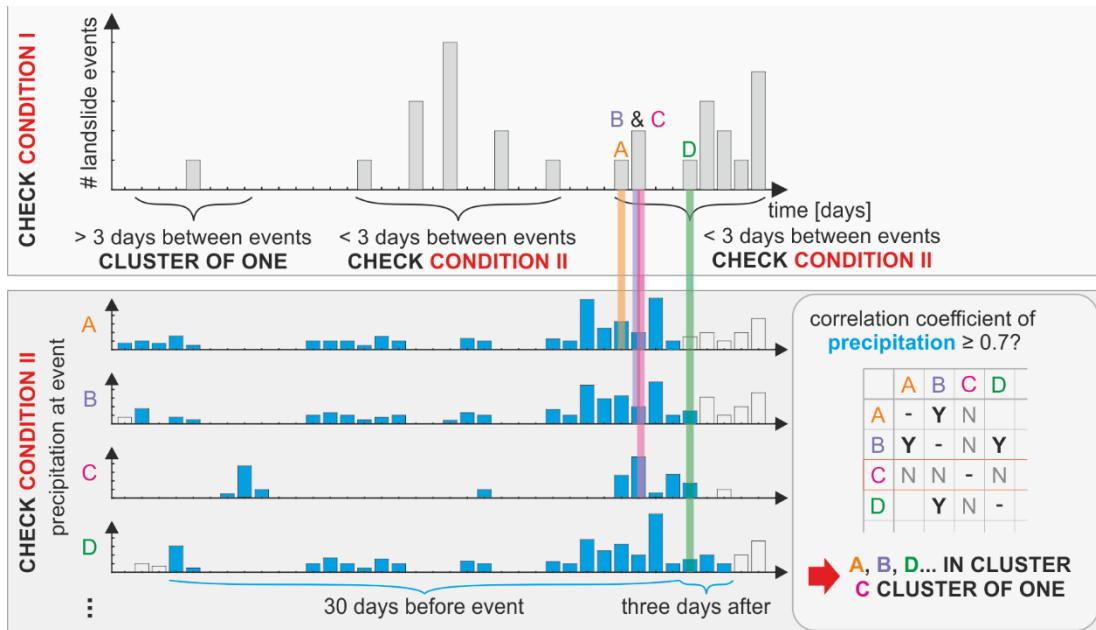
127 needs to be rechecked.

128   The introduced algorithm is independent of subsoil topography and relief parameters. While these

129   impact the precipitation intensity-duration threshold that is commonly expected to trigger

130   landslides, locations with different thresholds might still experience landslides triggered by the

131   same rainfall event.

132

133   **Figure 2.** Schematic drawing of the algorithm used to identify, if two landslide events within the Global Landslide

134   Catalog (GLC) occurred during the same rainfall event and hence belong to the same cluster. For condition II only

135   events occurring within three days of each other are compared.

2.4   Rainfall Analysis

137   In order to compare rainfall during a landslide event to overall rainfall at the location, the 95th

138   percentile of precipitation excluding non-rainy days was determined for 10 years prior to the event.

139   This comparison was also previously used by Kirschbaum et al (2015) to identify rainfall triggered

140   landslide events. However, in their case, rainfall data from the Tropical Rainfall Measuring Mission

141   (TRMM) was used for the time period 2000–2013 independent of the date of the landslide event.

142   Due to its higher spatial resolution CHIRPS data was used here instead.

143   In addition to the 95th percentile of rainfall, the global rainfall threshold by Guzzetti et al. (2008)

144   was also utilized to determine the likelihood of the individual landslide events being triggered by

145   rainfall. In their study 2626 rainfall events that have resulted in shallow landslides and debris flows

146   were analyzed in order to determine the following global rainfall intensity–duration threshold

147   [http://rainfallthresholds.irpi.cnr.it]:

$$I = 2.2 \cdot D^{-0.44} \tag{1}$$

148   Here the threshold intensity (I) was determined for each 24 hours starting with a duration (D) of

149   12 hours. This results in an average precipitation of 0.73 mm/h for $D = 12$ h, 0.45 mm/h for $D =$

150   36 h, and 0.35 mm/h for $D = 60$ h. The rainfall threshold was then compared to the cumulative

151   mean precipitation of the rainfall event preceding each landslide event.

152   **3.   Results and Discussion**

153   **3.1   Clustering Characteristics**

154   The presented algorithm divided the 9279 landslide events of the Global Landslide Catalog (GLC)

155   into 6474 clusters of events connected through precipitation. However, 85 % of these clusters

156   consist of only a single landslide event, containing in total 59 % of all recorded landslide events.

157   This implies that a large number of landslide events are in fact isolated events with no association

158   to other events. Nevertheless, 67 % of these 'single landslide event'-clusters are categorized as

159   medium or larger and might contain more than one landslide (in comparison 58 % of the landslide

160   events in clusters ≥ one landslide event are categorized as medium or larger). Hence, the number

161   of isolated landslides is likely to be significantly smaller than the number of isolated landslide

162   events.

163   In the Global Landslide Catalog (GLC) only 3 % of the analyzed landslide events are linked to

164   triggers unrelated to rainfall such as construction, volcanos or earthquakes. This number is reduced

165   to 1.5 % for landslides in a cluster of more than one event. Due to the low number of events in this

166   category, future research is necessary to test and thoroughly validate these findings as well as to

167   assess possible reasons and implications of this phenomenon. For now, we assume that this is

168   mainly caused by biased reporting and cataloging of landslide events, where events linked to larger

169   disasters such as earthquakes, might be reported as one large landslide event, whereas landslides

170   linked to rainfall, might be individually reported. Similar observations were previously made by

171   Kirschbaum et al. (2015) for events in the GLC that are linked to major storms. An example of this

172   is the catastrophic magnitude 7.8 Gorkha earthquake in Nepal in 2015. While more than 25,000

173   landslides occurred during the earthquake and its aftershock sequence (e.g. Roback et al., 2018),

174   they are only reported as 13 landslide events in the excerpt from the GLC analyzed here. In it, they

175   are described as ranging in size from small to large and their trigger is given as "unknown",

176   "earthquake" and in one case "snowmelt". Our algorithm sorts these events into eight clusters of

177   up to three events.

178

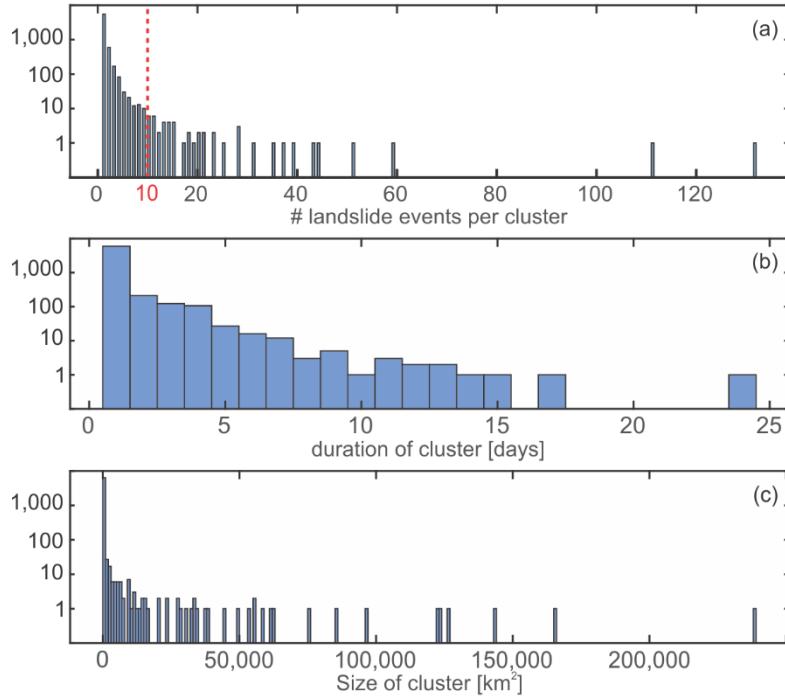



Figure 3. Histogram of the number of events per cluster, duration of clusters and area of the convex hull of each cluster. Clusters with only a single landslide event were appointed an area of zero. Within this study, all clusters with at least 10 landslide events were analyzed more closely.

Figure 3 provides histograms of the landslide events per cluster, duration of clusters and area covered by clusters (convex hull) in a logarithmic scale. As expected, for all three aspects frequency reduces drastically for larger numbers. In the following section all 50 clusters with at least 10 events (marked in red in Figure 3) are evaluated more closely.

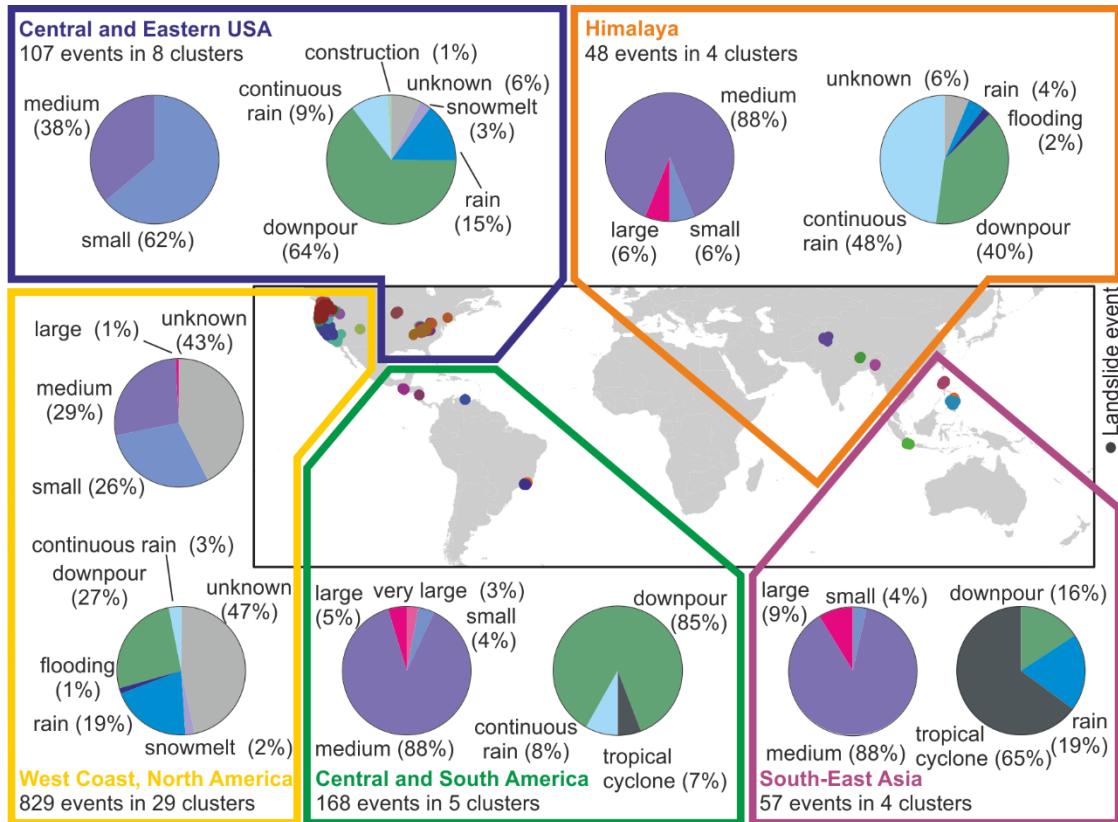



Figure 4. Location of all landslide events within clusters ≥ 10 events (different colors indicate different clusters). Overall, clusters in five distinct regions could be identified in the GLC (see Table S1 for more detail). Size and trigger (GLC categorization) of the associated landslide events are also shown (also see Tables S2 and S3).

3.2 Clusters with more than ten Landslide Events

3.2.1 Global Analysis

Table S1 in the supplementary material gives more detail of the 50 clusters with at least 10 events. In total 13 % of all landslide events are associated with one of these clusters (Table 1). As the database is most likely incomplete, the true number is expected to be higher. Overall the algorithm detects clusters in five distinct regions: (1) West Coast of North America, (2) Central and Eastern USA, (3) Central and Southern America, (4) Himalaya Region and (5) South-East Asia (Fig. 4). However, close to three quarters of all clusters ≥ 10 events are found within the USA mostly due

198   to a bias in the GLC database (Kirschbaum et al., 2015) (Fig. 1). This is also shown in the size of

199   recorded landslide events (Fig. 4 and Table S2).

200   In North America events are often classified as small in size, while clusters in the other regions

201   contain mainly medium events. This might be due to English speaking media, on which the GLC

202   is based, only picking up on large international events that consist of multiple landslides within an

203   area and smaller ones are under or not reported at all.

204   The median clusters with at least 10 events last six days, consist of 15 events, and span over an

205   area of 15,000 km² (Fig. 5). As expected, there is a positive correlation between cluster duration

206   and area (spearman correlation coefficient of 0.70, p-value: 0.001). However, this cannot be

207   observed for cluster duration and number of landslide events within the cluster (spearman

208   correlation coefficient of 0.44, p-value: 0.001). When comparing the different regions, clusters

209   located on the West Coast of North America are on average the longest and cover the largest area.

210   In contrast, events in South America are shortest and smallest, nevertheless they have the highest

211   number of events and clusters per day (Table 1).

212   On a global scale, no significant trend over time can be observed and clusters with ≥ 10 events

213   occur around the year (Fig. S4). Similarly, the total number of reported landslide shows no

214   significant increase in the GLC (Kirschbaum et al., 2015) as well as in other global databases such

215   as the Global Fatal Landslide Database (Froude and Petley, 2018). More regional observations

216   show seasonal variation and are described more closely in the following chapters. However, for

217   three out of the five regions, there are only five clusters or even less.

218

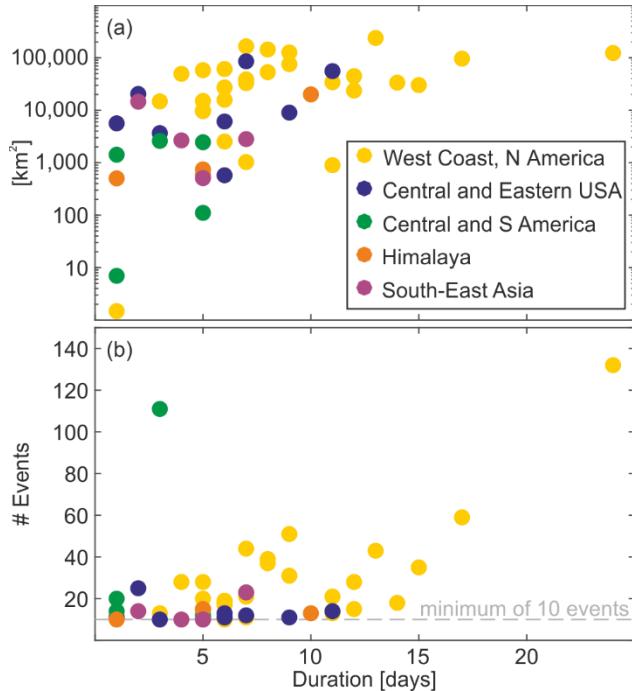



Figure 5. Link between the duration of the individual clusters ≥ 10 events and a) the covered area and b) the number

of landslide events per cluster. The color of the scatter plots indicates the region, in which each cluster occurred.

221   **Table 1.** Regional statistics for all landslide clusters (LC) with at least ten landslide events (LE).

Region	# LC	# LE	LE per LC	Average duration of LCs	LEs per day	Average area of LCs [km²]	Percentage of LE in a LC ≥ 10 LE
Global	50	1,209	24.2	7	3.5	35,441	13
West Coast, North America	29	829	28.6	9	3.3	52,970	31
Central and Eastern USA	8	107	13.4	6	2.4	23,357	12
South and Central America	5	168	33.6	3	11.2	1,320	18
Himalaya	4	48	12.0	5	2.3	5,476	3
South-East Asia	4	57	14.3	5	3.2	5,143	4

222

3.2.2   West Coast, North America

Landslides in the west of North America have been intensively investigated, mainly in the form of

case studies that discuss landslides along the Pacific coast in the states of California (Collins and

Sitar, 2008; Wieczorek, 1988), Oregon (Benda, 1990; Miller and Burnett, 2008) and Washington

227   (LaHusen et al., 2016; Perkins et al., 2017). This region is also one of the few, where the clustering

228   of rainfall triggered landslide events was previously investigated, showing qualitatively that there

229   are many instances in which landslides occur on consecutive days (Biasutti et al., 2016).

230   About 31 % of all landslide events recorded in this area belong to a cluster of at least ten events.

231   This is the highest number compared to the other regions of the world (Table 1). However, this

232   effect might be amplified by the high number of reported landslides. The large number of events

233   and clusters is mainly due to geologic, topographic, climatic conditions and construction practices.

234   For example in Oregon, steep slopes and heavy rainfalls are as well as poor construction practices

235   result in high economic losses (Wang et al., 2002). Burns et al. (2017) estimated an average annual

236   loss of $15.4 million due to landslides in Oregon alone. In years with heavy storms such as 1996,

237   this can accumulate to more than $100 million (Wang et al., 2002).

238   The observed clusters in this area are among the longest and have the largest areas of all regions

239   (Table 1). While the size of landslide events (as given by the GLC) in the west of North America

240   are small compared to most other regions, there is also a considerable amount of events, where the

241   size is unknown (43 %, Fig. 4, Table S2). While about half of the landslide events within clusters

242   ≥ 10 events are classified as "trigger unknown" (47 %), landslide events with a known cause are

243   mainly triggered by downpour (27 %) or rain (19 %) (Fig. 4, Table S3). However, when looking at

244   satellite based rainfall data preceding the clusters, rainfall cannot always be identified as a trigger

245   (Fig. S5). While it generally exceeds the global rainfall threshold (Guzzetti et al., 2008), the 95th

246   percentile of precipitation on rainy days is not reached for the majority of the clusters. Although,

247   several studies linked landslides within California to earthquakes (e.g. Harp and Jibson, 1996;

248   Keefer, 2000), they occurred before 2007 and are not registered in the GLC.

249   While there appears to be no significant change in the number of clusters over time (Fig. S4

250), most clusters occur during the rainy season (November to March), when most landslide events

251 occur. Within the west of North America this time period is therefore often referred to as the

252 "landslide season" (e.g. Mirus et al., 2018). Only one cluster in this region appears in June (Cluster

253 ID 21, Table S1). However, the center of this cluster is located more inland (in San Miguel County,

254 Colorado) and is also the shortest cluster (only one day) within the region as well as the most local

255 of all clusters in this study, covering only 1 km^2. While this cluster is triggered by downpour

256 according to the GLC, this is not apparent from satellite derived precipitation (Fig. S5). The small

257 size of the cluster might be the reason, why low-resolution satellite derived precipitation does not

258 record any anomalies here.

259 **3.2.3   Central and Eastern USA**

260 While most of the clusters with ≥ 10 landslides events of this region, are located in the Appalachian

261 Plateau (Ohio, West Virginia and Kentucky), one cluster can be found in Minnesota (ID 34 in Table

262 S1 and Fig. S6). While it is considerably smaller (580 km^2 compared to more than 9,000 km^2), it

263 is comparable to the Appalachians cluster in its number of landslide events and duration. The

264 Appalachian Plateau is well known for its landslides and the annual direct cost in Kentucky exceeds

265 $10 million (Crawford and Bryson, 2017).

266 Like the landslide clusters observed in the west of North America, clusters here consist mainly of

267 small landslides, which is most likely linked to the news alerts on which the GLC is based.

268 Checking sources in the GLC, they are mainly reported within smaller, more local news outlets

269 compared to landslide events outside of the US. To our knowledge the individual events grouped

270 by our algorithm into clusters have never been linked before. Clusters in this region occur

271 predominantly in spring (February to June), when rainfall is highest, slightly later than events on

272 the West Coast (Fig. S4). According to GLC they are predominantly triggered by downpours (64 %,

273 Fig. 4, Table S3). However, extreme rainfall is not always visible in satellite derived precipitation

274 (Fig. S6). For most clusters, it is below the 95th percentile, but above the global threshold. It is

275 worth noting than one cluster located in West Virginia (Cluster ID 35) shows no rainfall on the

276 satellite before day three of the cluster. Following the GLC, early landslide events within this

277 cluster are linked to snowmelt.

3.2.4 Central and South America

279 In contrast to the clusters in North America, more than 95 % of landslide events within clusters of

280 this region are medium in size or larger and might consists of several landslides themselves (Fig.

281 4). Thus, the number of landslides per cluster and per day is likely to be significantly higher than

282 the number of events per cluster and per day. Still, clusters in this area are on average only two and

283 a half days long, covering an area of slightly over 1,500 km² and they are the smallest and shortest

284 of all regions (Fig. 5, Table 1). It is important to note that this region covers the largest area reaching

285 from Rio de Janeiro in Brazil to Guatemala in Central America. From the few clusters we could

286 identify, it appears that there are dissimilarities between the clusters in Central America and South

287 America. The two clusters in Nicaragua (ID 42) and Guatemala (ID 39) are triggered by continuous

288 rain and a tropical cyclone, respectively. In contrast, all events located in South America (IDs 38,

289 40, and 41) are all triggered by downpour (Table S1 and Fig. S7).

3.2.5 Himalaya

291 Like in South America, most landslide events (94 %) associated with clusters with ≥ 10 events in

292 the Himalaya region are categorized as medium and larger. Thus, the number of landslides per

293 cluster is again expected to be significantly higher than the number of landslide events per cluster.

294 However, there may be differences between regions. Event ID 44, located in India and Pakistan

295 around Jammu and Kashmir, is classified as medium to small, much longer (10 days) and covers

296   an area more than 10 times larger than the other clusters. All of them are classified as medium or

297   large and are located in the East of India with some events in Nepal (Table S1). In both regions,

298   clusters are triggered by continuous rain or downpour. For all clusters satellite based rainfall data

299   exceeds the global threshold, and in most cases the 95th percentile of rainfall on rainy days (Fig.

300   S8). It is important to note that while earthquake triggered landslides are common in the region

301   (e.g. Parkash, 2013; Roback et al., 2018), the presented algorithm is by design only able to pick up

302   clusters that are linked by rainfall.

303   **3.2.6   South-East Asia**

304   As only four clusters are identified in this region, a detailed analysis is impossible. Again, 96 % of

305   the events associated are categorized as medium or larger and the main triggers are tropical

306   cyclones (Cluster IDs 47 and 48), downpour (Cluster ID 49), and rain (ID 50) (Table S1). Here,

307   satellite based rainfall data before clusters is both above the global rainfall threshold and in most

308   cases above the 95th percentile (Fig. S9). While only one of the four clusters (ID 50) is recorded

309   outside of the Philippines (in Indonesia), there is no apparent difference between both countries

310   (Table 1).

311   **3.3   Most Intense Cluster**

312   The cluster with the most events in one day, i.e. most intense cluster, happened in Rio de Janeiro,

313   Brazil, as well as neighboring cities Niteroi and Sao Goncalo in 2010. In an area of approximately

314   2,800 km^2, 111 landslide events were recorded within only three days, however predominantly on

315   6th April 2010 (Table S1, ID 38). This is almost four times as many landslide events in a single day

316   than the second most intense clusters (IDs 1 and 3) located in Washington and Oregon, USA. Both

317   recorded 29 events in one day.

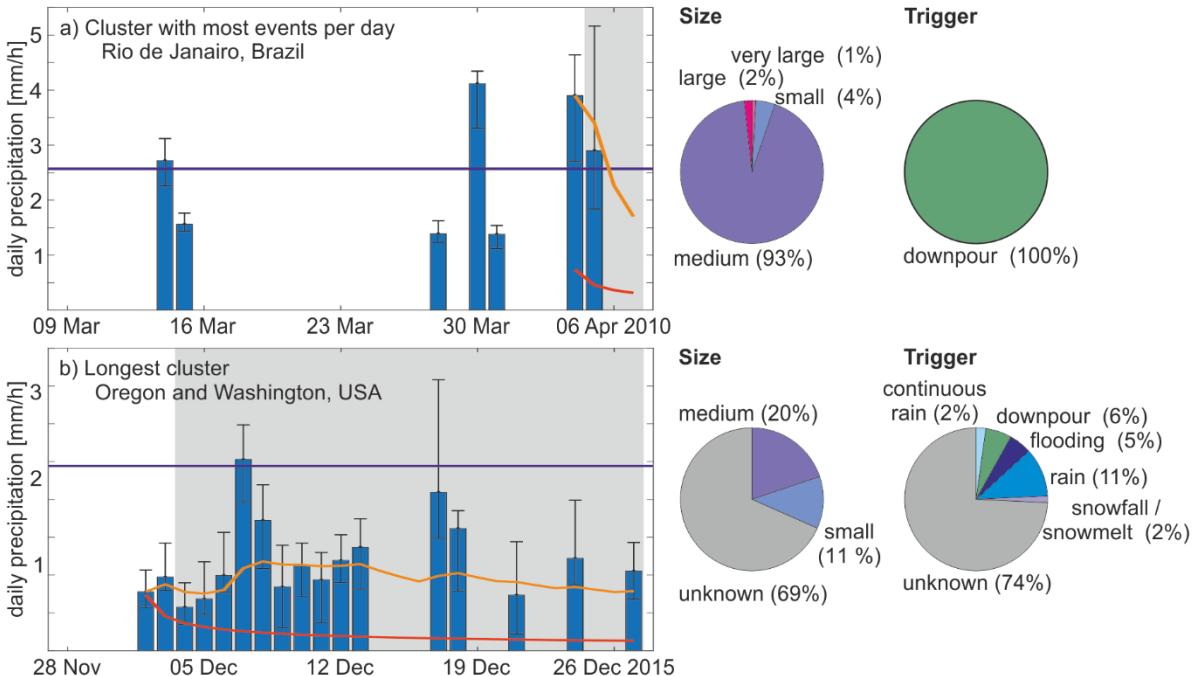



Figure 6. Daily precipitation for 30 days preceding the last landslide event of the cluster with the size of the associated landslide events and their trigger according to the GLC. Shown is the median precipitation for all landslide locations with the inner quartiles as an error bar. The 95th percentile of daily rainfall (rainy days only) in the ten years preceding the event is given in blue, in red the global rainfall threshold ID (Guzzetti et al., 2008) and in orange the cumulative mean for the rainfall event preceding the cluster. a) Cluster with the most events per day (ID 43), and b) longest running cluster (ID 22).

Most of the 111 events associated with the cluster in Rio de Janeiro were recorded as medium in size, all of which were triggered by downpour (Fig. 6a). This is confirmed by satellite derived precipitation. Heavy rainfalls (Figs. 6a, 7) occurred on the 4th and 5th of April of up to 210 mm per day. In comparison, the 95th percentile in the 10 years preceding this cluster is on average only 62 mm per day (rainfall for each individual location shown in Fig. S10). While the rainfall covered a large area, landslide events were primarily reported for steep slopes just outside the densely populated city center. Due to its location close to, and inside the urban area of Rio de Janeiro, the cluster caused approximately 200 fatalities according to CNN news reports (http://www.cnn.com/2010/WORLD/americas/04/12/brazil.flooding.mudslides/).

334   The location in the city might also be the reason for the large number of events being reported, as

335   we can expect more individual landslides being reported here compared to the countryside.

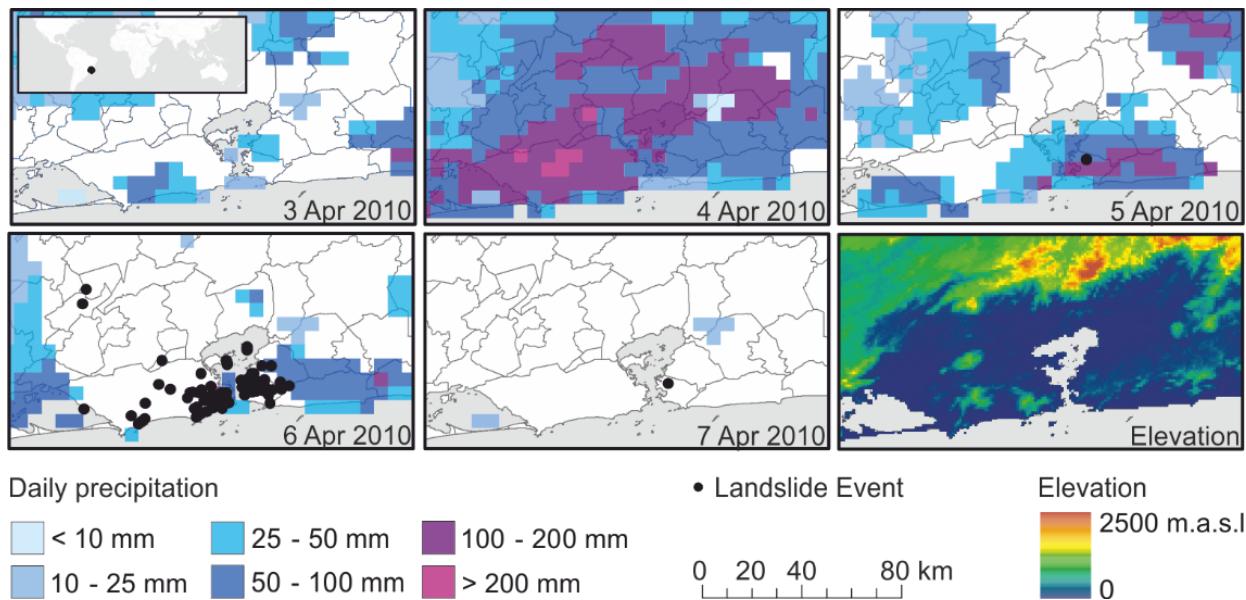
336   While studies not based on English speaking news alerts report a large number of landslides within

337   and around Rio de Janeiro (Calvello et al., 2015; Sandholz et al., 2018), only nine additional

338   landslide events inside the area of this cluster were reported in the GLC between 2009 and 2018.

339   Additionally, just northwest of the cluster another cluster occurred in January 2011 (ID 41 in Table

340   S1, Fig. S7). Although, this cluster only counts 20 individual landslide events within the GLC, it

341   is being reported as thousands individual landslides (Coelho Netto et al., 2013).

343   **Figure 7.** Location of the events in the cluster with the most events per day located in Rio de Janeiro, Brazil. Also

344   shown are daily precipitation and elevation. Elevation data is taken from the US Geological Survey (GTOPO30).

345   **3.4   Longest Cluster**

346   The longest running cluster identified in this study occurred in Oregon and Washington, USA from

347   4th to 27th December 2015 for a total of 24 days with 132 landslide events (Cluster ID 18, Table

348   S1). The second longest cluster lasted 17 days over January and February in 2012 and was also

349 located in Oregon and Washington, USA (Cluster ID 7). Overall, most events within the longest

350 cluster are unknown in size (69 %) and trigger (74 %) (Fig. 6b). However, inspecting satellite based

351 rainfall data, continuous rainfall appears to be the main trigger (Fig. 6b, Fig. 8 and Fig. S11 for

352 rainfall at the individual event locations). While daily rainfall is mainly below the 95th percentile,

353 cumulative mean rainfall is continuously above the global rainfall threshold. Although, heavy

354 rainfall is common in this area during winter times, for this cluster it lasted longer than usual and

355 was followed by shorter rain events in short successions (Fig. 8). Thus, the series of landslides did

356 not halt resulting in the longest cluster in the GLC. Following the information on sources within

357 the GLC, it appears that local media reported about the individual landslide events, but did not

358 detect on the extreme length of the continuous series of landslide events at this point in time (e.g.

359 https://kval.com/news/local/landslide-blocks-i-5-in-sw-washington;

360 https://q13fox.com/2015/12/09/landslide-above-puget-sound-damages-several-homes-at-least-

361 one-vehicle/). As landslide events are such a common occurrence in this region, and due to the

362 large area covered by this cluster, there is currently little to no emphasis on the longevity of this

363 specific series of landslide events in media and scientific studies.

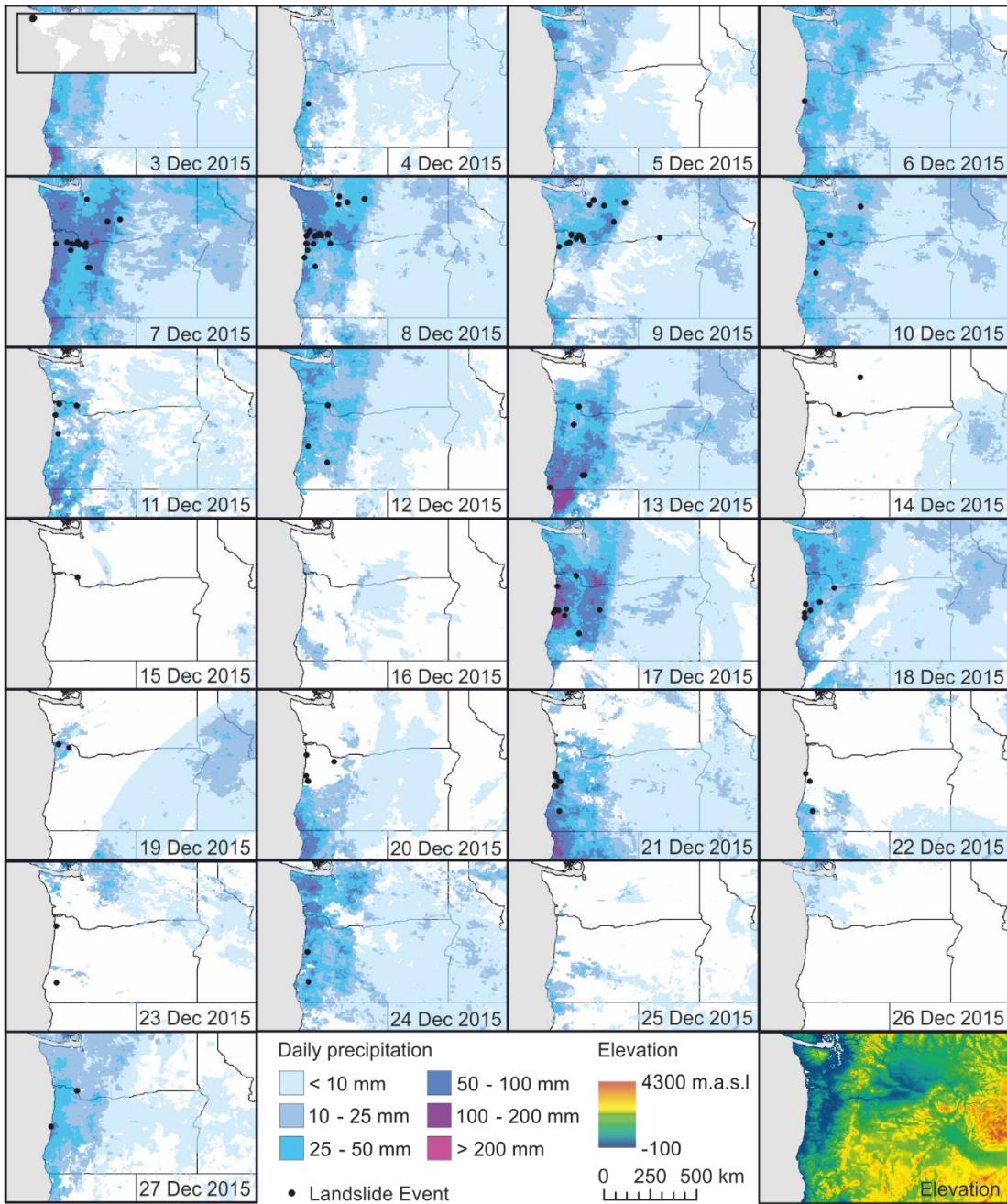



Figure 8. Location and time series of the longest cluster, located mainly in Oregon, USA. Also shown are daily rainfall
and elevation. Elevation data is available from the US Geological Survey (GTOPO30).

4.   Conclusion

In this study an algorithm is presented that detects clusters of landslide events that occur during,
and are likely triggered by the same rainfall events. Here this algorithm is applied to the Global

370   Landslide Catalog (GLC), where it detects that more than 40 % of all recorded events can be linked

371   to at least one other event. The global analysis shows that 14 % of all landslide events are part of a

372   cluster $\geq$ 10 events. However, this percentage varies dramatically by the region, ranging from 30 %

373   on the West Coast of North America to 3 % in the Himalayas. Part of this is caused by sampling

374   and reporting bias. As the GLC is based on English speaking media, events in the USA are reported

375   and cataloged in much greater detail than events abroad. Nevertheless, within the GLC we could

376   detect clusters $\geq$ 10 landslide events in five distinct regions: (1) West Coast of North America, (2)

377   Central and Eastern USA, (3) Central and Southern America, (4) Himalaya Region, and (5) South-

378   East Asia. In South America, the studied clusters are the shortest, but contain the most events per

379   day. However, this is mainly due to a cluster in Rio de Janeiro, where 108 of events were recorded

380   on 6th April 2010. As most of these events are classified as medium or larger, the absolute number

381   of landslides is expected to be significantly higher. In contrast, the longest and largest clusters are

382   observed on the West Coast of North America. On average clusters here last nine days and cover

383   an area of more than 50,000 km^2. The steep slopes and continuous rainfalls present in the area

384   combined with the above average reporting of landslide events, makes a more detailed analysis of

385   rainfall related landslide clusters possible. The longest of all detected clusters $\geq$ 10 landslide events

386   is also located in this region: In December 2015, 132 landslide events were recorded over a time

387   period of 24 days spanning more than 120 thousand km^2, which were all triggered by the same

388   rainfall event. Detection of large scale clusters such as this one can not only help to improve our

389   understanding of the link between individual events, but also be used in our mitigation strategies.

390   Only once we improve our understanding of the relation between individual landslide events, we

391   will be able to predict their behavior and forecast their economic losses and fatalities. While our

392   study does not replace case specific and small scale studies, as well as the identification of threshold

393   values, it can provide an improved understanding for managing landslide mitigations on a larger

394 scale. Within the area covered by individual clusters the same mitigation strategies, including early

395 warning systems (EWS) based on weather forecast simulations, can be developed and validated.

396 For future research we recommend to use the presented algorithm not only for the correlation with

397 precipitation data, but also to include the geometry of atmospheric rivers during cluster detection.

398 Finally, the algorithm could be applied to more regional and other global landslide databases

399 thereby improving our understanding on the spatial and temporal occurrence of landslide clusters.

400 **References**

401 Benda, L.: The influence of debris flows on channels and valley floors in the Oregon Coast Range,

402 U.S.A., Earth Surface Processes and Landforms, 15(5), 457–466, doi:10.1002/esp.3290150508,

403 1990.

404 Biasutti, M., Seager, R. and Kirschbaum, D. B.: Landslides in West Coast metropolitan areas: The

405 role of extreme weather events, Weather and Climate Extremes, 14, 67–79,

406 doi:10.1016/j.wace.2016.11.004, 2016.

407 Burns, W. J., Calhoun, N. C., Franczyk, J. J., Koss, E. J. and Bordal, M. G.: Estimating losses from

408 landslides in Oregon, Roanoke, VA., 2017.

409 Calvello, M., d'Orsi, R. N., Piciullo, L., Paes, N., Magalhaes, M. and Lacerda, W. A.: The Rio de

410 Janeiro early warning system for rainfall-induced landslides: Analysis of performance for the years

411 2010–2013, International Journal of Disaster Risk Reduction, 12, 3–15,

412 doi:10.1016/j.ijdrr.2014.10.005, 2015.

413 Carrara, A., Crosta, G., and Frattini, P.: Geomorphological and historical data in assessing landslide

414 hazard, Earth Surf. Proc. Land., 28, 1125–1142, https://doi.org/10.1002/esp.545, 2003.

415 Climate Hazards Group: CHIRPSv2.0, , doi:10.15780/G2RP4Q, 2015.

416 Coelho Netto, A. L., Sato, A. M., de Souza Avelar, A., Vianna, L. G. G., Araújo, I. S., Ferreira, D.

417 L. C., Lima, P. H., Silva, A. P. A. and Silva, R. P.: January 2011: The Extreme Landslide Disaster

418 in Brazil, in Landslide Science and Practice, edited by C. Margottini, P. Canuti, and K. Sassa, pp.

419 377–384, Springer Berlin Heidelberg, Berlin, Heidelberg., 2013.

420 Collins, B. D. and Sitar, N.: Processes of coastal bluff erosion in weakly lithified sands, Pacifica,

421 California, USA, Geomorphology, 97(3), 483–501, doi:10.1016/j.geomorph.2007.09.004, 2008.

422 Crawford, M. and Bryson, L.: Field Investigation of an Active Landslide In Kentucky: A

423 Framework to Correlate Electrical Data and Shear Strength, Kentucky Geological Survey Report

424 of Investigations, doi:https://doi.org/10.13023/kgs.ri01.13, 2017.

425 Froude, M. J. and Petley, D. N.: Global fatal landslide occurrence from 2004 to 2016, Natural

426 Hazards and Earth System Sciences, 18(8), 2161–2181, doi:10.5194/nhess-18-2161-2018, 2018.

427 GLC: Global Landslide Catalog, NASA's Open Data Portal [online] Available from:

428 https://data.nasa.gov/d/h9d8-neg4?category=Earth-Science&view_name=Global-Landslide-

429 Catalog (Accessed 3 May 2018), n.d.

430 Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R.: Google Earth

431 Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, 202,

432 18–27, doi:10.1016/j.rse.2017.06.031, 2017.

433 Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: The rainfall intensity–duration control of

434 shallow landslides and debris flows: an update, Landslides, 5(1), 3–17, doi:10.1007/s10346-007-

435 0112-1, 2008.

436 Guha S, Below R, Ph. Hoyois Ph (2015) EM-DAT: International Disaster Database –

437 www.emdat.be – Université Catholique de Louvain – Brussels – Belgium

438 Harp, E. L. and Jibson, R. W.: Landslides triggered by the 1994 Northridge, California, earthquake,

439 Bulletin of the Seismological Society of America, 86(1B), S319–S332, 1996.

440 Keefer, D. K.: Statistical analysis of an earthquake-induced landslide distribution — the 1989

441 Loma Prieta, California event, Engineering Geology, 58(3), 231–249, doi:10.1016/S0013-

442 7952(00)00037-5, 2000.

443 Kirschbaum, D. and Stanley, T.: Satellite-Based Assessment of Rainfall-Triggered Landslide

444 Hazard for Situational Awareness, Earth's Future, 6(3), 505–523, doi:10.1002/2017EF000715,

445 2018.

446 Kirschbaum, D., Stanley, T. and Zhou, Y.: Spatial and temporal analysis of a global landslide

447 catalog, Geomorphology, 249, 4–15, doi:10.1016/j.geomorph.2015.03.016, 2015.

448 Kirschbaum, D. B., Adler, R., Hong, Y., Hill, S. and Lerner-Lam, A.: A global landslide catalog

449 for hazard applications: method, results, and limitations, Natural Hazards, 52(3), 561–575,

450 doi:10.1007/s11069-009-9401-4, 2010.

451 LaHusen, S. R., Duvall, A. R., Booth, A. M. and Montgomery, D. R.: Surface roughness dating of

452 long-runout landslides near Oso, Washington (USA), reveals persistent postglacial hillslope

453 instability, Geology, 44(2), 111–114, doi:10.1130/G37267.1, 2016.

454 Malamud, B. D., Turcotte, D. L., Guzzetti, F. and Reichenbach, P.: Landslide inventories and their

455 statistical properties, Earth Surface Processes and Landforms, 29(6), 687–711,

456 doi:10.1002/esp.1064, 2004.

457 Martelloni, G, Segoni, S., Fanti, R. and Catani, F.: Rainfall thresholds for the forecasting of

458 landslide occurrence at regional scale. Landslides, 9(4), 485-495, doi:10.1007/s10346-011-0308-

459 2, 2012.

460   Miller, D. J. and Burnett, K. M.: A probabilistic model of debris-flow delivery to stream channels,

461   demonstrated for the Coast Range of Oregon, USA, Geomorphology, 94(1), 184–205,

462   doi:10.1016/j.geomorph.2007.05.009, 2008.

463   Mirus, B., Morphew, M. and Smith, J.: Developing Hydro-Meteorological Thresholds for Shallow

464   Landslide Initiation and Early Warning, Water, 10(9), 1274, doi:10.3390/w10091274, 2018.

465   Parkash, S.: Earthquake Related Landslides in the Indian Himalaya: Experiences from the Past and

466   Implications for the Future, in Landslide Science and Practice: Volume 5: Complex Environment,

467   edited by C. Margottini, P. Canuti, and K. Sassa, pp. 327–334, Springer Berlin Heidelberg, Berlin,

468   Heidelberg., 2013.

469   Perkins, J. P., Reid, M. E. and Schmidt, K. M.: Control of landslide volume and hazard by glacial

470   stratigraphic architecture, northwest Washington State, USA, Geology, 45(12), 1139–1142,

471   doi:10.1130/G39691.1, 2017.

472   Petley, D.: Global patterns of loss of life from landslides, Geology, 40(10), 927–930,

473   doi:10.1130/G33217.1, 2012.

474   Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., Chamlagain, D. and Godt,

475   J. W.: The size, distribution, and mobility of landslides caused by the 2015 M w 7.8 Gorkha

476   earthquake, Nepal, Geomorphology, 301, 121–138, doi:10.1016/j.geomorph.2017.01.030, 2018.

477   Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F. and Rossi, M.: Do

478   landslides follow landslides? Insights in path dependency from a multi-temporal landslide

479   inventory, Landslides, 14(2), 547–558, doi:10.1007/s10346-016-0739-x, 2017.

480 Sandholz, S., Lange, W. and Nehren, U.: Governing green change: Ecosystem-based measures for

481 reducing landslide risk in Rio de Janeiro, International Journal of Disaster Risk Reduction, 32, 75–

482 86, doi:10.1016/j.ijdrr.2018.01.020, 2018.

483 Wang, Y., Summers, R. D. and Hofmeister, R. J.: Landslide loss estimation - Pilot project in

484 Oregon, Oregon Department of Geology and Mineral Industries (DOGAMI)., 2002.

485 Wieczorek, G. F.: Landslides, Floods, and Marine Effects of the Storm of January 3-5, 1982, in the

486 San Francisco Bay Region, California, USGS Numbered Series, Geological Survey (U.S.). [online]

487 Available from: http://pubs.er.usgs.gov/publication/pp1434 (Accessed 6 November 2018), 1988.

488 Witt, A., Malamud, B. D., Rossi, M., Guzzetti, F. and Peruccacci, S.: Temporal correlations and

489 clustering of landslides, Earth Surface Processes and Landforms, 35(10), 1138–1156,

490 doi:10.1002/esp.1998, 2010.

491