

Elsevier required licence: © <2021>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>  
The definitive publisher version is available online at  
[\[https://www.sciencedirect.com/science/article/pii/S0341816221000060?via%3Dihub\]](https://www.sciencedirect.com/science/article/pii/S0341816221000060?via%3Dihub)

Manuscript Number: CATENA11686R2

Title: Usage of antecedent soil moisture for improving the performance of rainfall thresholds for landslide early warning

Article Type: Research Paper

Keywords: landslides; rainfall thresholds; LEWS; soil moisture; Idukki

Corresponding Author: Professor Biswajeet Pradhan, PhD

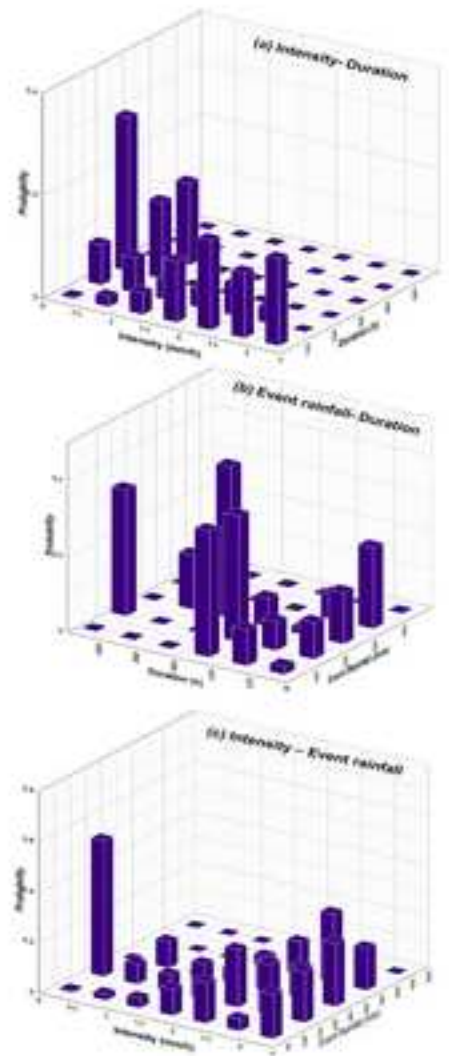
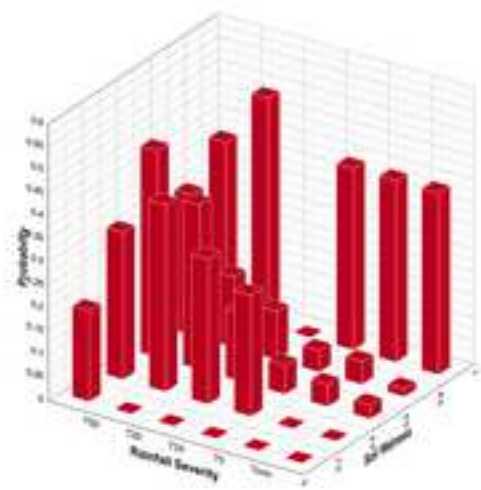
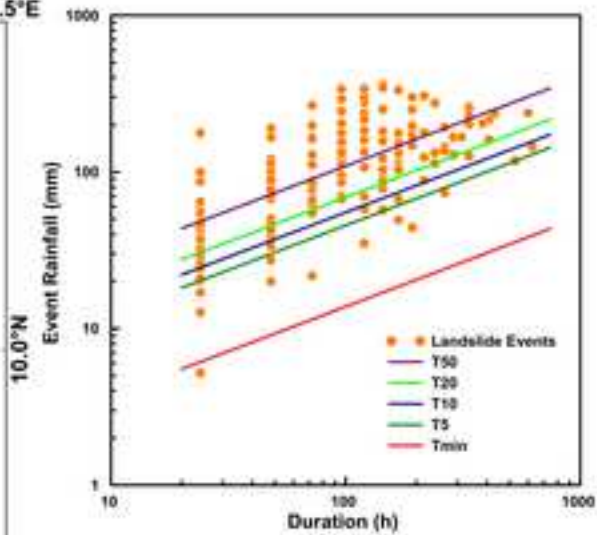
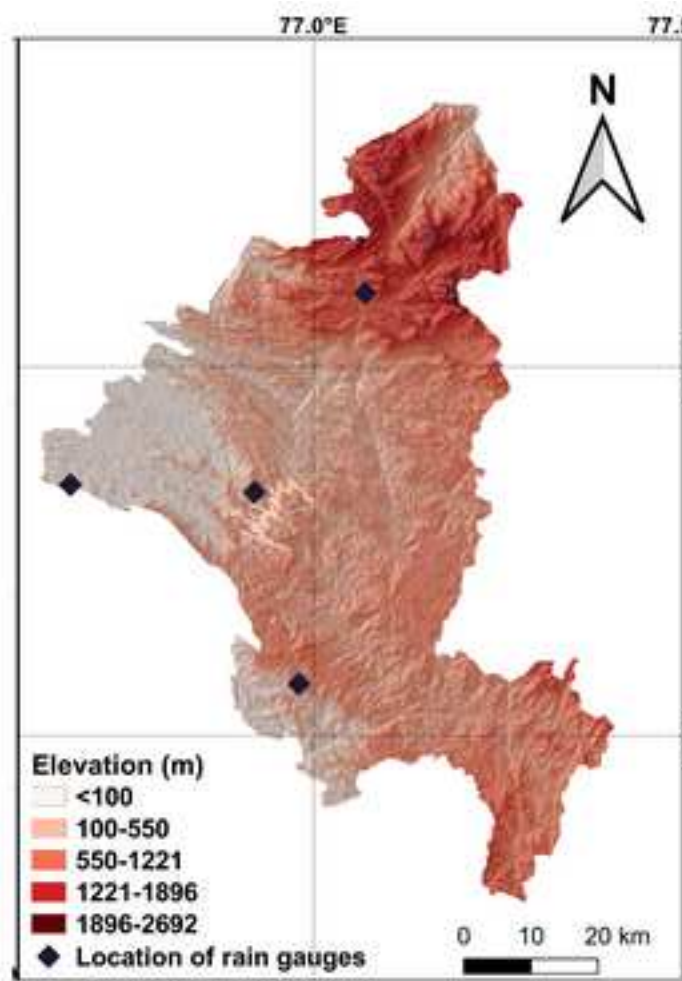
Corresponding Author's Institution: University of Technology Sydney

First Author: Minu Treesa Abraham, Ph.D.,

Order of Authors: Minu Treesa Abraham, Ph.D.,; Neelima Satyam, Ph.D.,; Ascanio Rosi, Ph.D.,; Biswajeet Pradhan, PhD; Samuele Segoni, Ph.D.,

Abstract: Landslides triggered by heavy rains are increasing in number and creating severe losses in hilly regions across the world. Rainfall thresholds on regional and local-scales are being used for forecasting such events, for efficient early warning. Empirical and probabilistic approaches for defining rainfall thresholds are traditional tools which are being used as part of the forecasting system for rainfall induced landslides. Such methods are easy-to-use and are based on statistical analyses. They can be derived without looking into the complex hydro-geological processes involved in slope failures, but are often associated with the disadvantage of higher false alarms, limiting their applications in a regional landslide early warning system (LEWS). This study is an attempt to improve the performance of conventional meteorological thresholds by considering the effect of soil moisture, using a probabilistic approach. Idukki district in southern part of India is highly susceptible to landslides and has witnessed major socio-economical setbacks in the recent disasters happened in 2018 and 2019. This tourist hub is now in need of a landslide forecasting system, which can help in landslide risk reduction. This study attempts to understand the effect of averaged soil moisture estimates derived from passive microwave remote sensing data, for improving the performance of conventional empirical and probabilistic thresholds. For defining empirical thresholds, an algorithm-based approach such as Calculation of Thresholds for Rainfall-induced Landslides Tool (CTRL-T) has been used. Probabilistic thresholds were defined using a Bayesian approach, finding the posterior probability of occurrence using the marginal and conditional probabilities of the control parameters along with the prior probability of occurrence of landslide. The derived rainfall thresholds were quantitatively compared with the Bayesian probabilistic threshold derived using rainfall severity and soil wetness using an area under the curve (AUC) based receiver operating characteristics (ROC) curve method. The results show that when the antecedent moisture content in soil is less, only severe rainfall events can trigger landslides in the study area; while less severe rainfall events can also trigger landslides when the soil is wet. The role of soil wetness in the initiation is used to improve the performance of the conventional methods, and a ROC approach was used for the statistical comparison of different models. Further, the results

indicated that the probabilistic threshold using rainfall severity and soil wetness outperformed the conventional approaches with AUC of 0.96, being the most sensitive and specific among the models considered. This result opens new promising perspectives for the development of an operational LEWS in the Idukki district based on a combination of rainfall and soil moisture data. Moreover, this work contributes to strengthen the advancing trend of hydro-meteorological thresholds based on soil moisture, which is gaining a growing attention in landslide studies and that, to date, was lacking evidences in monsoon regions.



## Highlights

- Landslides can be predicted using empirical and probabilistic rainfall thresholds.
- Soil moisture is critical in slope stability as it affects the infiltration rate.
- Soil moisture can be used with conventional thresholds for better performance.
- Idukki (India) is highly a highly susceptible landslide zone in the Western Ghats.
- Critical rainfall conditions, considering the soil wetness are derived for Idukki.



29 limiting their applications in a regional landslide early warning system (LEWS). This study is an  
30 attempt to improve the performance of conventional meteorological thresholds by considering the  
31 effect of soil moisture, using a probabilistic approach. Idukki district in southern part of India is  
32 highly susceptible to landslides and has witnessed major socio-economical setbacks in the recent  
33 disasters happened in 2018 and 2019. This tourist hub is now in need of a landslide forecasting  
34 system, which can help in landslide risk reduction. This study attempts to understand the effect of  
35 averaged soil moisture estimates derived from passive microwave remote sensing data, for improving  
36 the performance of conventional empirical and probabilistic thresholds. For defining empirical  
37 thresholds, an algorithm-based approach such as Calculation of Thresholds for Rainfall-induced  
38 Landslides Tool (CTRL-T) has been used. Probabilistic thresholds were defined using a Bayesian  
39 approach, finding the posterior probability of occurrence using the marginal and conditional  
40 probabilities of the control parameters along with the prior probability of occurrence of landslide. The  
41 derived rainfall thresholds were quantitatively compared with the Bayesian probabilistic threshold  
42 derived using rainfall severity and soil wetness using an area under the curve (AUC) based receiver  
43 operating characteristics (ROC) curve method. The results show that when the antecedent moisture  
44 content in soil is less, only severe rainfall events can trigger landslides in the study area; while less  
45 severe rainfall events can also trigger landslides when the soil is wet. The role of soil wetness in the  
46 initiation is used to improve the performance of the conventional methods, and a ROC approach was  
47 used for the statistical comparison of different models. Further, the results indicated that the  
48 probabilistic threshold using rainfall severity and soil wetness outperformed the conventional  
49 approaches with AUC of 0.96, being the most sensitive and specific among the models considered.  
50 This result opens new promising perspectives for the development of an operational LEWS in the  
51 Idukki district based on a combination of rainfall and soil moisture data. Moreover, this work  
52 contributes to strengthen the advancing trend of hydro-meteorological thresholds based on soil  
53 moisture, which is gaining a growing attention in landslide studies and that, to date, was lacking  
54 evidences in monsoon regions.

55 **Keywords:** landslides; rainfall thresholds; LEWS; soil moisture; Idukki

## 57 **1. Introduction**

58 Forecasting landslides and evacuating people from hazardous zones is an important risk reduction  
59 strategy (Althuwaynee and Pradhan, 2017). Considering the climate change and associated extreme  
60 rainfall phenomenon, the number of rainfall-induced landslides are expected to rise (Alvioli et al.,  
61 2018; Chen et al., 2019; Gariano and Guzzetti, 2016). Being a geomorphological process in the  
62 landscape evolution (Iida, 1999), the detailed understanding of slope failure mechanisms involves  
63 hydrological studies and forecasting of possible failure planes (Agostini et al., 2014) using relevant  
64 geotechnical and meteorological parameters. However, these parameters are highly site specific and  
65 often difficult to determine with the desired accuracy (Tofani et al., 2017), except that for single  
66 slopes or very small basins (Chae et al., 2017), and sophisticated experimental research is required for  
67 understanding the mechanism in detail (Kim et al., 2018). Hence, a more practiced approach is needed  
68 to forecast the critical conditions which result in the occurrence of landslides using the primary  
69 triggering factor i.e. rainfall – with the aid of rainfall thresholds (Caine, 1980; Keefer et al., 1987;  
70 Piciullo et al., 2018). Rainfall thresholds can be empirical, probabilistic, or algorithm based  
71 (Althuwaynee et al., 2015; Piciullo et al., 2018; Segoni et al., 2018a). All the approaches exploit  
72 historical data to find a mathematical relationship between rainfall and the occurrence of landslides in  
73 a region, to identify critical rainfall conditions which can trigger landslides in the future. A rainfall  
74 event is most commonly characterised in terms of cumulated rainfall event (E), duration (D), and  
75 intensity (I) (which are referred to as “rainfall parameters”). Consequently, the thresholds are often  
76 defined as cumulated event rainfall vs. duration (ED thresholds) (Lainas et al., 2016; Melillo et al.,  
77 2018, 2016; Peruccacci et al., 2017; Teja et al., 2019) or as rainfall intensity vs. duration (ID  
78 thresholds) (Battistini et al., 2017; Brunetti et al., 2010; Guzzetti et al., 2008; Lainas et al., 2016; Wu  
79 et al., 2019).

80 When the definition of thresholds is associated with the generation of many false alarms, their usage  
81 in operational Landslide Early Warning System (LEWS) may be inappropriate (Aleotti, 2004;  
82 Guzzetti et al., 2008; Kirschbaum et al., 2012; Segoni et al., 2018b). Low performances of rainfall



83 thresholds are traditionally related to the uncertainties associated with the definition of rainfall  
84 parameters, the quality and resolution of the historical data and the intrinsic limitations of the  
85 statistical models (Gariano et al., 2020; Marra et al., 2017; Nikolopoulos et al., 2014).

86 Some authors argued that sometimes the statistical correlation between rainfall parameters and  
87 landslide initiation is too weak and that hydro-meteorological thresholds accounting for both rainfall  
88 and hydrological (e.g. soil moisture) parameters could provide a stronger and more accurate  
89 assessment (Bogaard and Greco, 2018; Jakob et al., 2006; Terlien, 1998). Integrating soil moisture  
90 with rainfall thresholds has been proven effective in improving the rainfall thresholds (Abraham et al.,  
91 2020b; Segoni et al., 2018c; Zhao et al., 2019a), as the antecedent moisture content plays a key role in  
92 the shear strength parameters of soil. The soil moisture conditions play a key role in the infiltration  
93 process (Song and Wang, 2019) which significantly influences the initiation of landslides  
94 (Alimohammadlou et al., 2014; Baum et al., 2008; Biccocchi et al., 2019; Iverson, 2000; Wei et al.,  
95 2020; Yang et al., 2019). Weighted indexes (Glade et al., 2000; Ponziani et al., 2012); and satellite  
96 data (Zhao et al., 2019b) can be used for estimating soil moisture values when real-time field  
97 monitoring (Abraham et al., 2020c; Dikshit et al., 2018; Uchimura et al., 2015, 2010) cannot be  
98 conducted. Hydrological models (Abraham et al., 2020b; Zhao et al., 2019a) can also be used for the  
99 estimation of soil moisture content. In the published literature, soil moisture combined with rainfall  
100 thresholds has been tested mainly in Mediterranean, temperate and alpine climatic settings, whereas in  
101 monsoon regions similar types of tests are almost completely missing (Jakob et al., 2006; Mirus et al.,  
102 2018a; Valenzuela et al., 2018; Wicki et al., 2020).

103 The present work attempts to define statistical rainfall thresholds in Idukki district (India) and to  
104 improve their effectiveness by coupling rainfall parameters with soil moisture data. First, ED  
105 thresholds are defined using an automatic algorithm-based approach (Melillo et al., 2014). The  
106 algorithm first identifies the triggering rainfall events using the location of rain gauges and landslides,  
107 the time of occurrence of landslides and the time series rainfall data. It recreates multiple rainfall  
108 conditions which may result in landslides and identifies the maximum probable rainfall condition  
109 based on the location and time. After identifying the triggering rainfall event, the algorithm defines

110 the ED thresholds with multiple exceedance probabilities using frequentist method. Then, by using a  
111 probabilistic approach (Berti et al., 2012), the effect of event rainfall, duration and intensity on the  
112 occurrence of landslides is evaluated (probabilistic rainfall thresholds). Both empirical (Melillo et al.,  
113 2018, 2016; Peruccacci et al., 2017) and probabilistic approaches (Berti et al., 2012; Dikshit and  
114 Satyam, 2019) were considered to establish the relationship between primary triggering factor  
115 (rainfall) and the result (landslide), and these are simple statistical approaches that are easy to derive  
116 by integrating with a rainfall forecasting system. Similar studies have been conducted for Indian  
117 Himalayas (Abraham et al., 2020a; Dikshit and Satyam, 2018, 2019; Teja et al., 2019) and the  
118 Western Ghats (Abraham et al., 2020e, 2019); however, these methods were not always found to be  
119 operational due to a higher number of false alarms or missed alarms, limiting their applications in  
120 LEWS. This study aims to overcome these limitations by integrating soil moisture data along with the  
121 rainfall thresholds. The objective is to find if the addition of soil moisture data can perform better than  
122 the conventional methods based on the rainfall data alone.

## 123 **2. Description of the study area**

124 The Western Ghats of Indian Peninsula is highly susceptible to rainfall-induced landslides. There is a  
125 surge in the number of landslides during monsoon season since 2018, due to very-high intensity  
126 rainfalls. The landslides and floods happened in 2018 severely affected the south Indian states of  
127 Kerala and Karnataka. Among the 14 districts in the state of Kerala, 13 are part of the Western Ghats  
128 and are susceptible to landslides. Nearly 5.3 million people in the state were affected by the disaster in  
129 2018 (United Nations Development Programme, 2018). The Western Ghats scarps, running the whole  
130 extent of the mountain range, are highly prone to landslides. Very-high intensity rainfall, along with  
131 the anthropogenic activities, has accelerated the geological processes leading to landslides, making  
132 the situation alarming (Kuriakose et al., 2009b).

133 Idukki is a hilly district in the Western Ghats and is the second largest district in the state of Kerala, in  
134 terms of area. This district covers an area of 4358 km<sup>2</sup> and derived its name from the word 'Idukku' in  
135 the vernacular dialect meaning *narrow gorge*. This itself indicates the geography of the area. The  
136 district is the major power source of Kerala and houses many hydroelectric projects, including the

137 famous arch dam of Idukki. About 50% of the district is covered by forests and Idukki is drained by  
138 three major rivers, two flowing westward and one eastward. The rainfall across the district is varying  
139 with the least values recorded in the northern side with a long-term average of 1000 mm while the  
140 southern parts record an average rainfall of 5000 mm (Sajeev and Praveen 2014; Department of  
141 Mining and Geology 2016). The southwest monsoon season from June to September contributes 60%  
142 of the annual rainfall and around 24% is contributed by the North-East monsoon from October to  
143 December. Due to varying topography, the climatic conditions in the hill ranges, plateaus and  
144 midlands of the district are different from each other.

145 **Fig. 1.** Location details of study area. (a) India, and (b) Digital Elevation Model of Idukki (modified  
146 using CartoDEM (CartoDEM, 2015)) along with location of rain gauges.

147 Geologically, Idukki can be divided into three different parts from south to north. The charnockite  
148 rocks in the south, migmatitic complex in central portion, and peninsular gneissic complex in the  
149 northern part. Granite gneiss is the oldest and predominant group among the peninsular gneissic  
150 complex while the charnockite group consists of magnetite quartzite, pyroxene granulite and  
151 charnockite (Department of Mining and Geology 2016). Structural cum denudational hills are the  
152 predominant geomorphological feature of Idukki. The hills are generally having a thin soil cover  
153 overlaid on Precambrian basement rocks. The midlands have a rugged topography with small hills and  
154 deep valleys with an average elevation of 50 m. The zone where midlands grades to plateaus are  
155 called the foothills, ranging up to 8 km in width. A major portion of the district belongs to the plateau  
156 region, with a large landmass of moderate slope. The elevation of the plateau region goes up to 1500  
157 m, and the regions at an elevation greater than 1500 m belong to hilly ranges. More than 50% of the  
158 study area is covered by forest loam soils, produced by the weathering of rock under thick forest  
159 cover. The midlands are covered by lateritic soil with high permeability and less organic content. The  
160 valley portion of the terrain are covered with fine particles of sandy loam to clay type, formed by  
161 sedimentation and transportation of hill slopes. The narrow riverbanks consist of fertile alluvial soil  
162 and are more common in the midlands.

163 Because of its topographic variability and heavy rainfall, the district is highly susceptible to rainfall  
164 induced landslides. The typology of landslides in the Western Ghats includes earth and debris slides,  
165 rock falls, creep, slump and debris flows (Abraham et al., 2020d). Due to the thin regolith layer,  
166 shallow landslide (Varnes, 1978) is the most common type during prolonged rainfalls (Kuriakose et  
167 al., 2009a). Idukki district in particular is mostly affected by the cut slope failures along the major  
168 road corridors, disrupting the transportation network in the district. Recent changes in the land use  
169 patterns for infrastructure development and agriculture have affected the stability of slopes of this  
170 ecologically sensitive zone (Gadgil et al., 2011) and has aggravated the number of landslide disasters  
171 (Kuriakose et al., 2009b). Hence the development of an effective regional scale LEWS is highly  
172 needed to forecast the future landslides in the region.

173

### 174 **3. Data and Methodology**

175 The study explores the possibility of using soil moisture data in improving the performance of  
176 statistical thresholds. The overall methodology flow chart adopted in this study is shown in Fig 2. The  
177 methodology involves data collection from multiple sources, the definition of thresholds and their  
178 performance evaluation using different skill scores. For the analysis, historical rainfall, landslide, and  
179 soil moisture data were collected. For developing empirical and probabilistic rainfall thresholds, only  
180 rainfall and landslide data are required, while for developing probabilistic rainfall thresholds based on  
181 rainfall severity and soil wetness (RS threshold), the soil moisture data were integrated with empirical  
182 ED thresholds using a probabilistic approach. While the empirical threshold considers the effect of  
183 rainfall events which resulted in landslides, the probabilistic thresholds consider both triggering and  
184 non-triggering rainfall events for the analysis.

185

186

**Fig. 2.** Methodology of study.

### 187 3.1 Data collection

188 The dataset used for this study spans from 2010 to 2018 and the historical data from this period was  
189 used to derive the empirical and probabilistic thresholds for occurrence of landslides in Idukki district.  
190 The daily rainfall data was collected from the Indian Meteorological Department (India  
191 Meteorological Department 2019) for four rain gauges within the district. The landslide data was  
192 collected from various government agencies and media reports (Abraham et al., 2019) and only  
193 landslides for which the date of occurrence was available were used for the analysis. For each rain  
194 gauge a reference area was defined and multiple landslides triggered in the same day in each area  
195 were considered as one landslide event and rainfall data were collected from the reference rain gauge.  
196 By these criteria, 225 landslide events were identified in the study area which were first used as the  
197 input for empirical thresholds. For probabilistic thresholds, a total of 5028 rainfall events recorded by  
198 the four rain gauges during the study period were considered.

199 The average daily soil moisture data was collected from Giovanni's website by National Aeronautics  
200 and Space Administration Goddard Earth Sciences Data and Information Services Center (NASA  
201 GES DISC) (de Jeu and Owe, 2014, 2012; Giovanni, 2020). The data was derived using land  
202 parameter retrieval model (LPRM), which is a multi-parameter retrieval algorithm focused on  
203 hydrological and climate studies. It retrieves the soil moisture from the microwave observations from  
204 sensors. The observed brightness temperatures were used to derive the soil moisture data, using  
205 LPRM (Owe et al., 2008). LPRM is based on a forward radiative transfer model and the output is the  
206 volumetric soil moisture content in percentage. The soil moisture on the day before the occurrence of  
207 landslide, termed as the 'antecedent soil moisture' was used for the analysis in this research. The  
208 spatial resolution of the data is  $0.25^\circ \times 0.25^\circ$ . The study area (Idukki district) consists of 14 grids of  
209 size  $0.25^\circ \times 0.25^\circ$  (Figure 1). After calculating the area of Idukki within each grid, the weighted  
210 average was calculated for the whole area, for simplified calculation. This value is called the  
211 'averaged moisture content'. Another term, 'soil wetness' is introduced, to represent a range of  
212 antecedent soil moisture, on a scale of 0 to 1. The soil wetness values were divided into five equal  
213 parts, representing different ranges of moisture content. This classification is used to overcome the

214 limitations associated with using averaged data for a larger area. The value of soil wetness is directly  
215 proportional to the moisture content values and indicates the wetness of soil before the landslide.

216 Thus, by using historical rainfall, landslide and soil moisture data, thresholds were defined using  
217 multiple approaches for the study area to find the effect of soil moisture on the forecasting  
218 performance of the thresholds.

219

### 220 3.2 Empirical thresholds

221 The selection of rain gauges and rainfall parameters plays a critical role in the definition of rainfall  
222 thresholds (Abraham et al., 2020e). For the study area, rainfall data from the four available rain  
223 gauges were considered for the analysis. The intensity-duration thresholds for the study area was  
224 earlier derived from using a nearest rain gauge approach (Abraham et al., 2019), considering 225  
225 landslide events occurred from 2010 to 2018. From the pioneering work of Caine (Caine, 1980), ID  
226 thresholds were defined for regions across the globe (Abraham et al., 2020c, 2019; Brunetti et al.,  
227 2010; Dikshit and Satyam, 2018; Guzzetti et al., 2008, 2007; Segoni et al., 2018a). Even though  
228 intensity can easily be converted to event rainfall and vice-versa, recent literature shows a shift  
229 towards defining ED thresholds instead of ID thresholds (Melillo et al., 2018, 2014; Peruccacci et al.,  
230 2012; Teja et al., 2019; Zhao et al., 2019a). The reason is that E and D are two mutually independent  
231 parameters while I is a function of D and E. Hence, for a definition of rainfall thresholds and rainfall  
232 severity, the data points on ED plane was considered in this study. In this study, the reconstruction of  
233 event- duration thresholds was carried out by using Calculation of Thresholds for Rainfall Induced  
234 Landslides - Tool (CTRL-T) (Melillo et al., 2018, 2014). CTRL-T uses an algorithm-based approach,  
235 extracting the rainfall events automatically from the daily precipitation data input. From the extracted  
236 events, rainfall conditions that have triggered landslides were identified; and –used to derive the  
237 rainfall thresholds for the region. The tool considers a buffer zone around each landslide location, to  
238 search for the rain gauge and identify the triggering event. In this study, a search radius of 20 km is  
239 considered, due to the low rain gauge density in the study area. The algorithm also considers a delay

240 time between the end of rainfall and occurrence of landslide. In this study, the delay time is taken as  
 241 48 hours (Melillo et al., 2014). If no rainfall condition is recreated within this delay time before the  
 242 occurrence of landslide, the event will be discarded by the algorithm. The algorithm first determines  
 243 the total event rainfall and duration of rainfall for all identified rainfall events and then to minimise  
 244 the effect of spatial variability of rainfall distribution, single or multiple rainfall conditions (MRC)  
 245 likely to result in failures and a weight is assigned to each of them. Then for each landslide, the  
 246 highest weight was used to identify the reference rain gauge and to choose the maximum probable  
 247 rainfall conditions (MPRC). In this study, five different threshold lines were defined using CTRL-T,  
 248 at different exceedance probabilities of 1%, 5%, 10%, 20% and 50% (termed as  $T_1$ ,  $T_5$ ,  $T_{10}$ ,  $T_{20}$  and  
 249  $T_{50}$ , respectively). Thresholds and related uncertainties were estimated from MPRCs. The defined  
 250 thresholds are in the form of a power law, determined using the frequentist approach (Brunetti et al.,  
 251 2010) and can be expressed as:

252

$$E = (\alpha \pm \Delta\alpha) D^{(\gamma \pm \Delta\gamma)} \quad (1)$$

253

254 where,  $\alpha$  is the scaling parameter or the intercept and  $\gamma$  is the shape parameter which denotes the slope  
 255 of the equation.  $\Delta\alpha$  and  $\Delta\gamma$  represents the uncertainties associated with  $\alpha$  and  $\gamma$ , respectively. The  
 256 uncertainties are determined using a bootstrap approach.

### 257 3.3 Probabilistic approach

258 The empirical thresholds compare an input value with the defined thresholds and will have a single  
 259 output (triggering or non-triggering). It is often difficult to decide the exceedance probability to be  
 260 selected as a threshold beyond which a radical change can be expected in the system (Berti et al.,  
 261 2012). The discretion between triggering and non- triggering rainfall conditions is not trivial in such  
 262 cases. To derive the equation, only the triggering rainfall conditions are considered. This increases the  
 263 chances of false alarms, as numerous rainfall events that cross the threshold line not necessarily  
 264 trigger landslides.

265 By considering both triggering and non- triggering rainfalls for analysis, probability-based models are  
266 more informative and provide a better option to find extreme events. In this study, a Bayesian  
267 approach is used to define probabilistic thresholds (Berti et al., 2012).

### 268 3.3.1 One-dimensional analysis

269 Bayes theorem applies a conditional probability of some event  $L$  (landslide) given the occurrence of  
270 another event  $X$  (rainfall, expressed in terms of E, I or D). This is also called the posterior probability,  
271  $P(X|L)$ . It can be calculated as follows (Berti et al., 2012):

$$P(L|X) = \frac{P(X|L) * P(L)}{P(X)} \quad (2)$$

272 where,  $P(X|L)$  is the conditional probability of occurrence of rainfall of magnitude  $X$ , when a  
273 landslide occurs. This is also called as a likelihood.

274  $P(L)$  is the prior probability of occurrence of landslide regardless of the occurrence rainfall  
275 magnitude.

276  $P(X)$  is the marginal probability of  $X$ , which can be defined as the probability of occurrence of  
277 rainfall regardless of the occurrence of landslides. The terms can be calculated mathematically using  
278 relative frequencies. Let  $N_R$  be the total number of rainfall events during study period,  $N_L$  be the total  
279 number of landslides occurred,  $N_X$  be the number of rainfall events with magnitude  $X$  and  $N_{(X|L)}$  be  
280 the number of rainfall events with magnitude  $X$  that resulted in landslides. The probabilities can be  
281 computed as (Berti et al., 2012):

282

$$P(L) \approx \frac{N_L}{N_R} \quad (3)$$

$$P(X) \approx \frac{N_X}{N_R} \quad (4)$$



$$P(X|L) \approx \frac{N_{(X|L)}}{N_L} \quad (5)$$

283 Considering the rainfalls that resulted in landslides only will give us partial information, the  
 284 likelihood. To understand the influence of rainfall of magnitude  $X$ , it is important to compare the prior  
 285 probability with the posterior probability.

### 286 3.3.2 Two-dimensional analysis

287 Two-dimensional case is the extension of Eq. 2 by considering two conditions  $X, Y$  instead of the  
 288 single condition  $X$  in Eq. 2. In the initial analysis, we consider  $X$  and  $Y$  as magnitude of two rainfall  
 289 parameters (E,D ; I,D; E,I). The calculation of prior, marginal and conditional probabilities are given  
 290 below:

291

$$P(L|X, Y) = \frac{P(X, Y|L) * P(L)}{P(X, Y)} \quad (6)$$

$$P(L) \approx \frac{N_L}{N_R} \quad (7)$$

$$P(X, Y) \approx \frac{N_{X,Y}}{N_R} \quad (8)$$

$$P(X, Y|L) \approx \frac{N_{(X,Y|L)}}{N_L} \quad (9)$$

292 The study explores the effect of antecedent soil moisture content using a two-dimensional  
 293 probabilistic analysis. During the second phase, we considered rainfall severity in ED plane and soil  
 294 wetness as  $X$  and  $Y$ , respectively. Based on the values of soil wetness, five different categories were  
 295 considered for analysis viz, less than 0.2, 0.2 to 0.4, 0.4 to 0.6, 0.6 to 0.8, and 0.8 to 1. The categories

296 based on rainfall severity were less than  $T_1$ ,  $T_1$  to  $T_5$ ,  $T_5$  to  $T_{10}$ ,  $T_{10}$  to  $T_{20}$ ,  $T_{20}$  to  $T_{50}$  and greater than  
297  $T_{50}$ . Thus, the two-dimensional plane was divided into 30 cells as a 6 x 5 matrix as shown in Fig. 6.  
298 These values were used for the definition of RS threshold.

## 299 4. Results

### 300 4.1 Empirical thresholds

301 CTRL-T tool considered 177 landslide events out of the 225 and the rest were discarded to avoid  
302 introduction of relevant spatio-temporal uncertainties in the analysis. The uncertainties are associated  
303 with the less rain gauge density in the study area. As described earlier, the landslides for which  
304 responsible rainfall conditions were not identified were discarded. This can be due to a distance more  
305 than 20 km between the location of rain gauges and landslide or due to a delay time more than 48  
306 hours after the end of any rainfall event. The algorithm forecasted rainfall thresholds with various  
307 exceedance probability both in normal and logarithmic plot (Fig. 3). The threshold lines of 1%, 5%,  
308 10%, 20% and 50% exceedance probabilities were used to classify the events into six categories based  
309 on the severity of rainfall. These lines are named  $T_1$ ,  $T_5$ ,  $T_{10}$ ,  $T_{20}$  and  $T_{50}$ , respectively. The slope of  
310 threshold lines in logarithmic plot was found to be  $0.57 \pm 0.03$ . This value is not in good agreement  
311 with the ID thresholds defined for the area in a previous study (Abraham et al., 2019). Though both  
312 the studies used frequentist approach for the definition of thresholds, the process of identification of  
313 responsible rainfall event was different. In the previous study (Abraham et al., 2019), the responsible  
314 rainfalls were identified using a Thiessen polygon approach manually, while in this study, the  
315 automatic algorithm, CTRL-T is used for identifying the responsible rainfall event. The parameters  
316 of threshold lines and the uncertainties associated are listed in Table 1.

317

318

319 **Fig. 3.** Rainfall event – duration thresholds for Idukki district

320 **Table 1.** Values of  $\alpha$ ,  $\gamma$  and the uncertainties associated with different exceedance probabilities

321

322 The range of duration of rainfalls considered for analysis vary from 1 to 26 days. For the thresholds to  
323 be reliable, the relative uncertainty ( $\Delta x/x$  for any variable  $x$ ) should be less than 10%. Here the  
324 relative uncertainty of  $\gamma$  is 5.2%. But with higher exceedance probabilities, the relative uncertainty of  
325  $\alpha$  is crossing this limiting value.

326 With 5% exceedance probability, 20.19mm rainfall can trigger a landslide in the region for a duration  
327 of 24 hours and when the duration is 624 hours, a rainfall of 129 mm can trigger landslides in the  
328 region. For a better understanding of the effect of each rainfall parameter on the occurrence of  
329 landslides, probabilistic rainfall thresholds were defined for the area.

#### 330 4.2 Probabilistic thresholds

331 The maximum probable rainfall conditions which were used for the definition of ED thresholds were  
332 considered as the triggering rainfall events for the probabilistic analysis. Thus, out of the 5028 rainfall  
333 events considered, 177 events were identified as triggering events by CTRL- T algorithm and the rest  
334 4851 events were considered as non-triggering rainfall events. In the one-dimensional case, six  
335 categories of rainfall duration, five categories of cumulated rainfall event and seven categories of  
336 rainfall intensity were considered. The results are plotted in Fig. 4 (a-f); where Fig. 4a, c and e depict  
337 the prior probability, marginal probability and likelihood, and Fig. 4b, d and f depict the prior and  
338 posterior probabilities. The variable  $X$  in Eq. (2-5) is replaced with  $D$ ,  $E$  and  $I$  in the respective  
339 graphs.  $P(L)$  being a constant parameter (value 0.035 in this study), the ratio of  $P(X|L)$  to  $P(X)$   
340 determines the variation of posterior probability values. Hence when  $P(X|L) > P(L)$ , the posterior  
341 probability is greater than prior probability and vice versa. The more the variation between prior and  
342 posterior probability, the more significant the variable is. It can be seen, that for duration and event  
343 rainfall, for the largest values of variables, the values of  $P(L|X)$  is less than  $P(L)$ , while in the case of  
344 intensity, high intensity rainfalls are more probable to trigger landslides in the region. The plots of  
345  $P(X)$  and  $P(X|L)$  are well above the plot of prior probability in all the cases. Intensity was found to  
346 be the most significant variable, with the maximum ratio between posterior and prior probabilities.

347 The maximum posterior probability when the control parameter is D was found to be 0.053 where the  
348 value is 0.103 and 0.116 in the case of E and I, respectively. Maximum probability occurs when the  
349 duration is between 120 h to 240 h; event rainfall is between 100 mm to 200 mm; and intensity is  
350 greater than 3 mm/h.

351

352 **Fig. 4.** Prior, conditional, marginal and posterior probabilities with respect to rainfall parameters. (a,  
353 b) Duration; (c, d) Event rainfall; and (e, f) Intensity.

354 To evaluate the joint occurrence of two parameters, two-dimensional Bayesian analysis were  
355 conducted with data on three different planes (Fig. 5). The two-dimensional space for each analysis  
356 was divided into small cells based on the categories of parameters used for one-dimensional analysis.  
357 Hence the ID plane is a 7 x 6 matrix, ED plane is a 5 x 6 matrix and the EI plane is a 5 x 7 matrix.  
358 There are several no data points in all three cases, due to the lower number of landslides considered  
359 for the analysis. As identified from the one-dimensional analysis, E and I were found to be more  
360 critical parameters than D. This is the reason why this study has considered all three different  
361 combinations of the control parameters even though the empirical thresholds are defined on ED plane  
362 only. The maximum probability value was obtained on EI plane, when the intensity value is less than  
363 0.5 mm/h and event rainfall is between 100 mm to 200 mm, with a value of 0.54.

364

365 **Fig. 5.** Two-dimensional posterior probabilities of occurrence of landslide on (a) ID plane, (b) ED  
366 plane, and (c) EI plane.

367 It is evident from Fig. 6 that even less severe rainfall events when falling on a moist soil can trigger  
368 landslides in the region. Most of the landslides for which rainfall events were less severe happened on  
369 days with higher soil wetness. Also, when the rainfall event is severe, even dry soil can be susceptible  
370 to landslides. The maximum probability of 0.49 was observed when the rainfall severity was between  
371  $T_{20}$  to  $T_{50}$  and the soil wetness was between 0.8 to 1. With the available data, when the antecedent soil  
372 moisture is less, only extremely severe rainfall conditions can trigger landslides in the area. This

373 affects the performance of the ED thresholds considerably. For different antecedent soil moisture  
374 conditions, this makes it easier to decide the threshold line to be used.

375

376 **Fig. 6.** Two-dimensional Bayesian probabilities for occurrence of landslides based on rainfall severity  
377 and soil wetness.

## 378 **5. Discussions**

379 To verify the performance of all models and to understand which model is performing better for the  
380 study area, different thresholds should be compared quantitatively (Lagomarsino et al., 2015). In this  
381 study, empirical thresholds on ED plane, probabilistic thresholds on all three combinations of control  
382 parameters (ED, ID and EI) and also a two-dimensional Bayesian approach by combining empirical  
383 ED thresholds with soil moisture have been derived. The maximum probability value obtained in the  
384 two-dimensional analysis was in the case of EI thresholds, and the value is 0.54. The value was  
385 obtained when the intensity is less than 0.5 mm/h and event rainfall is between 100 to 200 mm. This  
386 implies a prolonged duration of 8 days or more. The intensity value is too low in this case, yet the  
387 probability value is the maximum. The definition of 2-dimensional Bayesian probability majorly  
388 depends upon the relative occurrence of landslides when the rainfall conditions are satisfied and the  
389 occurrence of rainfall events with specified conditions. The number of events with the specified EI  
390 conditions were less, but more than half of them have resulted in landslides based on the historical  
391 data. Thus, the probability of occurrence of landslides is higher in this case. This result points towards  
392 the significance of using a physical parameter such as soil moisture for the definition of threshold.  
393 The top regolith layer throughout the district consists of forest loam, lateritic soil, alluvial soils etc,  
394 with higher fine fraction (Department of Mining and Geology Kerala, 2016). The less permeable soil  
395 has a higher water holding capacity and the moisture content increases when the rainfall is  
396 continuous. The prolonged rainfall has thus reduced the shear strength of soil and the landslide has  
397 happened at a very less intensity value. This complicated process is simplified by using a statistical

398 approach, by considering the effect of soil wetness. To understand the performance of such a model  
399 with respect to the meteorological thresholds, a quantitative comparison is required.

400 An ROC curve approach was used for quantitative comparison. ROC curve is a tool to understand the  
401 performance of a model with a binary outcome. Each threshold value can forecast two outcomes for a  
402 day; 'landslides' or 'no landslides. If the threshold condition is crossed, the model forecasts  
403 'landslides' and otherwise, 'no landslides. When the forecasting is correct and landslide occurs, it is  
404 termed as a true positive (*TP*). Another possibility of correct outcome is the result 'no landslides' on a  
405 day in which landslides do not occur. This can be counted as a true negative (*TN*) result. When the  
406 forecasting goes wrong, it also has two possible outcomes. 'Landslides' forecasted on a non-landslide  
407 day, which is a false positive (*FP*) or simply a false alarm and 'no landslides' forecasted on a day in  
408 which landslides occur, termed as false negatives (*FN*) or missed alarms. A perfect model should only  
409 have true outcomes, without any false alarms or missed alarms.

410 A ROC curve is a plot with a false positive rate of a model on x-axis and a true positive rate on y axis.  
411 It evaluates the overall performance of the model. The true positive rate is also called the sensitivity  
412 of the model. It provides the proportion of landslide occurrences which are correctly identified  
413 ( $TP/(TP + FN)$ ). The specificity of a model is the true negative rate and is the ratio of *TN* to the  
414 sum of *TN* and *FP*. The false positive rate can be calculated by subtracting specificity value from 1.  
415 An ideal model is expected to have both sensitivity and specificity values as 1. Hence the point (0,1)  
416 on ROC curve is called the perfect point. Points which are closer to this perfect point has a better  
417 performance. Also, the model with better performance is the one with a maximum area under the  
418 curve (AUC) among the different models considered. Threat score and True Skill Statistic (TSS) are  
419 two other parameters which were used to understand the performance of a model (Mirus et al.,  
420 2018b). Threat score is defined as the ratio of *TP* to the sum of *TP*, *FP* and *FN*. TSS is the difference  
421 between sensitivity and false positive rate. For an ideal model, the value of both these variables  
422 should be 1.

423 ROC curves for all models considered in the study are plotted in Fig. 7 and the statistical attributes are  
424 listed in Table 2. From Fig. 7, it can be observed that the RS threshold covers the maximum area in

425 the plane with an AUC of 0.96. The empirical ED threshold has the second highest AUC of 0.86. All  
426 the three probabilistic rainfall thresholds have very close AUC values as observed in Fig. 7. EI  
427 threshold covers a larger area than the other two, indicating its better performance in comparison with  
428 the other two probabilistic rainfall thresholds. The distance from perfect point is minimum in the case  
429 of RS thresholds, in the case of critical probabilities 0.1 and 0.15. It can also be observed that the  
430 value of threat score and TSS are maximum in the case of RS thresholds. The maximum value of  
431 threat score is obtained as 0.24 and TSS as 0.90, both in the case of RS thresholds with critical  
432 probability 0.1, which is also the closest one to the perfect point.

433 **Fig. 7.** ROC curves for the derived thresholds. Sensitivity is the ability of a model to correctly identify  
434 the landslide events and Specificity is the ability to correctly identify the non-landslide events

435 Looking into the details in Table 2, it confirms with the literature as the empirical thresholds result in many  
436 false alarms, making it inadequate to use in an LEWS. The number of false alarms can considerably be  
437 reduced by using probabilistic rainfall thresholds, as listed in Table 2. The number of *FP* in the case of  
438 probabilistic ED, ID and EI are much lesser than the other two models considered. But this reduction in  
439 false alarms comes with the cost of a higher number of missed alarms (*FN*). While 171 landslide events  
440 out of the 177 events are correctly forecasted by the empirical ED threshold line  $T_1$ , and 172 are correctly  
441 forecasted by RS threshold when the critical probability is 0.05, the maximum number of correct outcomes  
442 for the other probabilistic models are 106, 105 and 117 on ED, ID and EI planes respectively. For  
443 improving the performance, we need to balance the number of false alarms and missed alarms, which is  
444 achieved by using RS threshold. The RS threshold has *FN* numbers comparable with that of probabilistic  
445 rainfall thresholds, minimising the false alarms and by incorporating an additional filter using soil wetness,  
446 it reduces the number of false alarms when compared to the empirical ED threshold.

447 **Table 2.** Statistical attributes for quantitative comparison.

448

449 The probabilistic rainfall thresholds have high specificity values, pointing to their ability to correctly  
450 forecast the days without landslides, but with very less values of sensitivity. The points on ROC  
451 curves for probabilistic rainfall thresholds are therefore closer to both the axes, reducing the AUC.

452 Even though the points have high specificity values, they are located far from the perfect point, due to  
453 their inefficiency in correctly forecasting the occurrence of landslides. The RS threshold with a  
454 critical probability of 0.1 is the closest one to the perfect point, correctly forecasting 167 landslide  
455 occurrences.

456 From the analysis, the rainfall and soil wetness conditions for which the probability of occurrence is  
457 more than 0.1 should be considered critical. This makes it easier to identify the empirical ED  
458 threshold line for different values of soil wetness. The critical conditions are mentioned in Table 3.

459 **Table 3.** Critical conditions for initiation of landslides in Idukki, based on RS thresholds

460

461 From Table 3, it can be understood that when soil wetness is less than 0.2,  $T_{50}$  line of empirical ED  
462 thresholds should be considered as critical, when the soil wetness is between 0.2 to 0.4,  $T_5$  can be  
463 considered as the critical threshold, ~~for~~ For the next two cases where soil wetness is between 0.4 to  
464 0.8,  $T_{10}$  threshold line can be considered as critical if the critical probability is 0.1. In this case, the  
465 threshold line for 0.2 to 0.4 is  $T_5$ , which is below the threshold line for soil wetness from 0.4 to 0.8.  
466 This variation can be due to the smaller number of data points considered in this study. With the  
467 available data points, very less cases are reported when the soil wetness is between 0.4 to 0.8, and the  
468 rainfall severity is below  $T_{10}$ . To avoid any possible missed alarms due to the limitations of the dataset  
469 considered, the threshold for soil wetness between 0.4 to 0.8 is considered as  $T_1$ , for which the  
470 probability of occurrence of landslides is 0.05 in this study. This variation in the critical probability  
471 ensures the physical validity and easy export of the model. ~~and w~~When the soil wetness is between  
472 0.8 to 1, even rainfall which are is below  $T_1$  can trigger landslides. Hence, for the last condition, we  
473 defined the critical case as  $T_{min}$  where  $T_{min}$  represents the threshold line with minimum exceedance  
474 probability, close to zero. Practically, it represents any possible rainfall condition. ~~The threshold line~~  
475 ~~for 0.2 to 0.4 is  $T_5$ , which is below the threshold line for soil wetness from 0.4 to 0.8. This variation~~  
476 ~~can be due to the less number of data points considered in this study. With the available data points,~~  
477 ~~very less cases are reported when the soil wetness is between 0.4 to 0.8, and the rainfall severity is~~



478 ~~below  $T_{10}$ . If the threshold has to be kept as  $T_5$ , the critical probability should be 0.05, which will~~  
479 ~~increase the number of false alarms as mentioned in Table 2. Hence based on the available data, the~~  
480 ~~threshold is kept as  $T_{10}$ .~~

481 The soil wetness data can be collected from daily satellite observations as taken in this study, or from  
482 real-time field observation using sensors. The severity of rainfall for each day can be estimated from  
483 the rainfall forecasts. Using these two inputs, the possibility of occurrence of landslide can be  
484 estimated using the conditions mentioned in Table 3. With higher exceedance probabilities, the  
485 relative uncertainty of  $\alpha$  of ED threshold is crossing this limiting value. Similar results are observed  
486 when the rainfall data used is of daily temporal resolution (Teja et al., 2019).

487 The type of landslides is also an important factor in identifying the associated rainfall. For example,  
488 rockfalls may be triggered without any rainfall, debris flows are often triggered by short duration  
489 (maybe less than 1 hour) and high intensity (Kean et al., 2011), and shallow landslides are triggered  
490 by short-term rainstorms of high-intensity or long-duration rainfall of low to medium intensity  
491 (Guzzetti et al., 2008). This is the main reason why the models (ED, and RS) often associated with the  
492 disadvantage of higher false alarms. Even though the false alarms are considerably reduced in RS  
493 thresholds, it needs further enhancements to be used in an LEWS. There are chances that the model  
494 may miss alarms for rock-falls, which can be triggered with no rainfall. In the case of flow like  
495 landslides such as debris flows, the failure can be triggered by very short, high intensity rainfalls.  
496 Such rainfall events may trigger landslides in relatively dry soils as well. In this case, even if the  
497 antecedent soil wetness is less than 0.2, if the rainfall severity is greater than  $T_{50}$ , the model will issue  
498 a warning. If an event of severity less than  $T_{50}$  triggers such an event, the model may miss the alarm.

499 With a higher number of data points and better resolution of rainfall data, this can be improved, and  
500 better results can be expected.

## 501 **6. Conclusions**

502 This study has been conducted to evaluate the effect of antecedent soil moisture content to improving  
503 the performance of empirical and probabilistic thresholds for Idukki district in India. The district is

504 suffering from landslides ranging from cut slope failures to debris flows during monsoon seasons. The  
505 recent disasters that happened in 2018 and 2019 in the district emphasises the requirement of a  
506 landslide early warning system for the region.

507 In this study, empirical rainfall thresholds on ED plane was derived for the study area using an  
508 algorithm-based approach. It was found that with 5% exceedance probability, 20.19 mm rainfall can  
509 trigger a landslide in the region for a duration of 24 hours, and when the duration is 624 hours, a  
510 rainfall of 129 mm can trigger landslides in the region.

511 To evaluate the influence of each rainfall parameter on the occurrence of landslides, Bayesian  
512 analyses were conducted for both one-dimensional and two-dimensional cases. It was found that both  
513 intensity and event rainfall have influence on the occurrence of landslides, and most of the events  
514 happened when the rainfall happened in lesser duration. From two-dimensional analysis, the  
515 probabilities on EI plane were found to have the maximum values.

516 To evaluate the effect of soil wetness, another two-dimensional Bayesian approach was conducted,  
517 and it was observed that when the soil is relatively dry, severe rainfall events are required to trigger  
518 landslides and when the soil is wet, also milder rainfall conditions can trigger landslides in the study  
519 area.

520 A statistical comparison between the considered models was used to find out the best performing  
521 model. The comparison was carried out by using a ROC curve approach, where the RS threshold was  
522 found to have the maximum AUC of 0.96, among the models considered in this study. The empirical  
523 ED threshold generated a relevant number of false alarms, resulting in a low specificity value, while  
524 the disadvantage associated with the probabilistic thresholds was the low sensitivity due to a large  
525 number of missed alarms. The proposed method, which combines empirical thresholds with soil  
526 wetness using a probabilistic approach, performs better than both the root models by optimising the  
527 number of false alarms and missed alarms. Based on the comparison, it was found that an RS  
528 threshold of probability 0.1 should be considered critical for the study area and critical rainfall  
529 severity conditions were identified for each soil wetness condition. The performance could be further

530 enhanced in the future by using hourly rainfall data with more dense rain gauge network for the area.  
531 Moreover, real-time monitoring of moisture content data for different locations in the study area can  
532 also contribute to better resolution soil moisture data and thereby improving the performance of the  
533 model.

534 The results of the study therefore open new promising perspectives for the development of an  
535 operational LEWS in the Idukki district, by combining rainfall and soil moisture data. At the same  
536 time, this work provides evidences from a monsoon region about the advances brought by hydro-  
537 meteorological thresholds based on soil moisture, which is gaining a growing attention in landslide  
538 studies all over the world but before today it was relatively unexplored for the setting of LEWS in the  
539 study area. Unfortunately, the use of soil moisture data in operational LEWS with short lead times is  
540 technically difficult, consequently another option to be explored is the use of antecedent rainfall  
541 conditions as a proxy to the soil moisture, which can be a simpler method to express the soil wetness  
542 conditions (Leonarduzzi and Molnar, 2020; Segoni et al., 2018b). More studies must be conducted for  
543 this region, to develop an effective LEWS which could obtain a fair compromise between the  
544 simplicity of the approach and the quality of the forecasting performance.

## 545 **Acknowledgements**

546 Analyses and visualizations used in this study were produced with the Giovanni online data system,  
547 developed and maintained by the NASA GES DISC. The research was supported by Università degli  
548 Studi di Firenze under the Grant 58517\_INTERNAZIONALIZZAZIONE

549

## 550 **References**

- 551 Abraham, M.T., Pothuraju, D., Satyam, N., 2019. Rainfall Thresholds for Prediction of Landslides in  
552 Idukki, India: An Empirical Approach. *Water* 11, 2113. <https://doi.org/10.3390/w11102113>
- 553 Abraham, M.T., Satyam, N., Kushal, S., Rosi, A., Pradhan, B., Segoni, S., 2020a. Rainfall Threshold  
554 Estimation and Landslide Forecasting for Kalimpong, India Using SIGMA Model. *Water* 12,

555 1195. <https://doi.org/10.3390/w12041195>

556 Abraham, M.T., Satyam, N., Pradhan, B., Alamri, A.M., 2020b. Forecasting of landslides using  
557 rainfall severity and soil wetness: A probabilistic approach for Darjeeling Himalayas. *Water*  
558 (Switzerland) 12, 1–19. <https://doi.org/10.3390/w12030804>

559 Abraham, M.T., Satyam, N., Pradhan, B., Alamri, A.M., 2020c. IoT-Based Geotechnical Monitoring  
560 of Unstable Slopes for Landslide Early Warning in the Darjeeling Himalayas. *Sensors* 20, 2611.  
561 <https://doi.org/10.3390/s20092611>

562 Abraham, M.T., Satyam, N., Reddy, S.K.P., Pradhan, B., 2020d. Runout modeling and calibration of  
563 friction parameters of Kurichermala debris flow, India. *Landslides*.  
564 <https://doi.org/10.1007/s10346-020-01540-1>

565 Abraham, M.T., Satyam, N., Rosi, A., Pradhan, B., Segoni, S., 2020e. The Selection of Rain Gauges  
566 and Rainfall Parameters in Estimating Intensity-Duration Thresholds for Landslide Occurrence:  
567 Case Study from Wayanad (India). *Water* 12, 1000. <https://doi.org/10.3390/w12041000>

568 Agostini, A., Tofani, V., Nolesini, T., Gigli, G., Tanteri, L., Rosi, A., Cardellini, S., Casagli, N., 2014.  
569 A new appraisal of the Ancona landslide based on geotechnical investigations and stability  
570 modelling. *Q. J. Eng. Geol. Hydrogeol.* 47, 29–43. <https://doi.org/10.1144/qjegh2013-028>

571 Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. *Eng. Geol.* 73, 247–265.  
572 <https://doi.org/10.1016/j.enggeo.2004.01.007>

573 Alimohammadlou, Y., Najafi, A., Gokceoglu, C., 2014. Estimation of rainfall-induced landslides  
574 using ANN and fuzzy clustering methods: A case study in Saenen Slope, Azerbaijan province,  
575 Iran. *Catena* 120, 149–162. <https://doi.org/10.1016/j.catena.2014.04.009>

576 Althuwaynee, O.F., Pradhan, B., 2017. Semi-quantitative landslide risk assessment using GIS-based  
577 exposure analysis in Kuala Lumpur City. *Geomatics, Nat. Hazards Risk* 8, 706–732.  
578 <https://doi.org/10.1080/19475705.2016.1255670>

579 Althuwaynee, Omar, F., Pradhan, B., Ahmad, N., 2015. Estimation of rainfall threshold and its use in

580 landslide hazard mapping of Kuala Lumpur metropolitan and surrounding areas. *Landslides* 12,  
581 861–875.

582 Alvioli, M., Melillo, M., Guzzetti, F., Rossi, M., Palazzi, E., von Hardenberg, J., Brunetti, M.T.,  
583 Peruccacci, S., 2018. Implications of climate change on landslide hazard in Central Italy. *Sci.*  
584 *Total Environ.* 630, 1528–1543. <https://doi.org/10.1016/j.scitotenv.2018.02.315>

585 Battistini, A., Rosi, A., Segoni, S., Lagomarsino, D., Catani, F., Casagli, N., 2017. Validation of  
586 landslide hazard models using a semantic engine on online news. *Appl. Geogr.* 82, 59–65.  
587 <https://doi.org/10.1016/j.apgeog.2017.03.003>

588 Baum, R.L., Savage, W.Z., Godt, J.W., 2008. TRIGRS — A Fortran Program for Transient Rainfall  
589 Infiltration and Grid-Based Regional Slope Stability Analysis.

590 Berti, M., Martina, M.L.V., Franceschini, S., Pignone, S., Simoni, A., Pizziolo, M., 2012.  
591 Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *J. Geophys.*  
592 *Res. Earth Surf.* 117, 1–20. <https://doi.org/10.1029/2012JF002367>

593 Bicocchi, G., Tofani, V., D’Ambrosio, M., Tacconi-Stefanelli, C., Vannocci, P., Casagli, N., Lavorini,  
594 G., Trevisani, M., Catani, F., 2019. Geotechnical and hydrological characterization of hillslope  
595 deposits for regional landslide prediction modeling. *Bull. Eng. Geol. Environ.* 78, 4875–4891.  
596 <https://doi.org/10.1007/s10064-018-01449-z>

597 Bogaard, T., Greco, R., 2018. Invited perspectives: Hydrological perspectives on precipitation  
598 intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds.  
599 *Nat. Hazards Earth Syst. Sci.* 18, 31–39. <https://doi.org/10.5194/nhess-18-31-2018>

600 Brunetti, M.T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., Guzzetti, F., 2010. Rainfall  
601 thresholds for the possible occurrence of landslides in Italy. *Nat. Hazards Earth Syst. Sci.* 10,  
602 447–458. <https://doi.org/10.5194/nhess-10-447-2010>

603 Caine, N., 1980. The rainfall intensity-duration control of shallow landslides and debris flows: An  
604 update. *Geogr. Ann. Ser. A, Phys. Geogr.* 62, 1–2, 23–27.

605 CartoDEM, 2015. CartoDEM : a national digital elevation model from Cartosat-1 stereo data. Natl.  
606 Remote Sens. Centre, Hyderabad, Dep. Space, Gov. India.

607 Chae, B.G., Park, H.J., Catani, F., Simoni, A., Berti, M., 2017. Landslide prediction, monitoring and  
608 early warning: a concise review of state-of-the-art. *Geosci. J.* 21, 1033–1070.  
609 <https://doi.org/10.1007/s12303-017-0034-4>

610 Chen, C.-W., Tung, Y.-S., Liou, J.-J., Li, H.-C., Cheng, C.-T., Chen, Y.-M., Oguchi, T., 2019.  
611 Assessing landslide characteristics in a changing climate in northern Taiwan. *CATENA* 175,  
612 263–277. <https://doi.org/10.1016/j.catena.2018.12.023>

613 de Jeu, R. (Vrije U.A., Owe, M. (NASA G., 2014. AMSR2/GCOM-W1 surface soil moisture (LPRM)  
614 L3 1 day 25 km x 25 km descending V001, Edited by Goddard Earth Sciences Data and  
615 Information Services Center (GES DISC) (Bill Teng), Greenbelt, MD, USA, Goddard Earth  
616 Sciences Data and Information Services Cente. <https://doi.org/10.5067/CGDEOBASZ178>

617 de Jeu, R. (Vrije U.A., Owe, M. (NASA G., 2012. TMI/TRMM surface soil moisture (LPRM) L3 1  
618 day 25 km x 25 km nighttime V001, Edited by Goddard Earth Sciences Data and Information  
619 Services Center (GES DISC) (Bill Teng), Greenbelt, MD, USA, Goddard Earth Sciences Data  
620 and Information Services Center (GES. <https://doi.org/10.5067/GWHRZEL8SA21>

621 Department of Mining and Geology Kerala, 2016. District Survey Report of Minor Minerals.  
622 Thiruvananthapuram.

623 Dikshit, A., Satyam, D.N., 2018. Estimation of rainfall thresholds for landslide occurrences in  
624 Kalimpong, India. *Innov. Infrastruct. Solut.* 3. <https://doi.org/10.1007/s41062-018-0132-9>

625 Dikshit, A., Satyam, D.N., Towhata, I., 2018. Early warning system using tilt sensors in Chibo,  
626 Kalimpong, Darjeeling Himalayas, India. *Nat. Hazards* 94, 727–741.  
627 <https://doi.org/10.1007/s11069-018-3417-6>

628 Dikshit, A., Satyam, N., 2019. Probabilistic rainfall thresholds in Chibo, India: estimation and  
629 validation using monitoring system. *J. Mt. Sci.* 16, 870–883. <https://doi.org/10.1007/s11629->

630 018-5189-6

631 Gadgil, M., Krishnan, B.J., Ganeshaiyah, K.N., Vijayan., V.S., Borges, R., Sukumar, R., Noronha, L.,  
632 Nayak, V.S., Subramaniam, D.K., Varma, R.V., Gautam, S.P., Navalgund, R.R.,  
633 Subrahmanyam, G.V., 2011. Report of the Western Ghats Ecology Expert Panel (WGEEP).

634 Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. *Earth-Science Rev.* 162, 227–252.  
635 <https://doi.org/10.1016/j.earscirev.2016.08.011>

636 Gariano, S.L., Melillo, M., Peruccacci, S., Brunetti, M.T., 2020. How much does the rainfall temporal  
637 resolution affect rainfall thresholds for landslide triggering? *Nat. Hazards* 100, 655–670.  
638 <https://doi.org/10.1007/s11069-019-03830-x>

639 Giovanni, 2020. NASA GES DISC [WWW Document].

640 Glade, T., Crozier, M., Smith, P., 2000. Applying probability determination to refine landslide-  
641 triggering rainfall thresholds using an empirical “Antecedent Daily Rainfall Model.” *Pure Appl.*  
642 *Geophys.* 157, 1059–1079. <https://doi.org/10.1007/s000240050017>

643 Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2008. The rainfall intensity-duration control of  
644 shallow landslides and debris flows: An update. *Landslides* 5, 3–17.  
645 <https://doi.org/10.1007/s10346-007-0112-1>

646 Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2007. Rainfall thresholds for the initiation of  
647 landslides in central and southern Europe. *Meteorol. Atmos. Phys.* 98, 239–267.  
648 <https://doi.org/10.1007/s00703-007-0262-7>

649 Iida, T., 1999. A stochastic hydro-geomorphological model for shallow landsliding due to rainstorm.  
650 *Catena* 34, 293–313. [https://doi.org/10.1016/S0341-8162\(98\)00093-9](https://doi.org/10.1016/S0341-8162(98)00093-9)

651 India Meteorological Department, 2019. India Meteorological Department (IMD) Data Supply Portal  
652 [WWW Document].

653 Iverson, R.M., 2000. Landslide triggering by rain infiltration. *Water Resour. Res.* 36, 1897–1910.

654 <https://doi.org/10.1029/2000WR900090>

655 Jakob, M., Holm, K., Lange, O., Schwab, J.W., 2006. Hydrometeorological thresholds for landslide  
656 initiation and forest operation shutdowns on the north coast of British Columbia. *Landslides* 3,  
657 228–238. <https://doi.org/10.1007/s10346-006-0044-1>

658 Kean, J.W., Staley, D.M., Cannon, S.H., 2011. In situ measurements of post - fire debris flows in  
659 southern California : Comparisons of the timing and magnitude of 24 debris - flow events with  
660 rainfall and soil moisture conditions. *J. Geophys. Res.* 116, 1–21.  
661 <https://doi.org/10.1029/2011JF002005>

662 Keefer, D.K., Wilson, R.C., Mark, R.K., Brabb, E.E., Iii, W.M.B., Ellen, S.D., Harp, E.L., Wiczorek,  
663 G.F., Alger, C.S., Zatkint, R.S., 1987. Real-Time Landslide Warning During Heavy Rainfall.  
664 *Science* (80-. ). 238, 921–925.

665 Kim, M.S., Onda, Y., Uchida, T., Kim, J.K., Song, Y.S., 2018. Effect of seepage on shallow  
666 landslides in consideration of changes in topography: Case study including an experimental  
667 sandy slope with artificial rainfall. *Catena* 161, 50–62.  
668 <https://doi.org/10.1016/j.catena.2017.10.004>

669 Kirschbaum, D.B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C., Lerner-Lam, A., 2012.  
670 Advances in landslide nowcasting: Evaluation of a global and regional modeling approach.  
671 *Environ. Earth Sci.* 66, 1683–1696. <https://doi.org/10.1007/s12665-011-0990-3>

672 Kuriakose, S.L., Devkota, S., Rossiter, D.G., Jetten, V.G., 2009a. Prediction of soil depth using  
673 environmental variables in an anthropogenic landscape, a case study in the Western Ghats of  
674 Kerala, India. *Catena* 79, 27–38. <https://doi.org/10.1016/j.catena.2009.05.005>

675 Kuriakose, S.L., Sankar, G., Muraleedharan, C., 2009b. History of landslide susceptibility and a  
676 chorology of landslide-prone areas in the Western Ghats of Kerala, India. *Environ. Geol.* 57,  
677 1553–1568. <https://doi.org/10.1007/s00254-008-1431-9>

678 Lagomarsino, D., Segoni, S., Rosi, A., Rossi, G., Battistini, A., Catani, F., Casagli, N., 2015.



679 Quantitative comparison between two different methodologies to define rainfall thresholds for  
680 landslide forecasting. *Nat. Hazards Earth Syst. Sci.* 15, 2413–2423.  
681 <https://doi.org/10.5194/nhess-15-2413-2015>

682 Lainas, S., Sabatakakis, N., Koukis, G., 2016. Rainfall thresholds for possible landslide initiation in  
683 wildfire-affected areas of western Greece. *Bull. Eng. Geol. Environ.* 75, 883–896.  
684 <https://doi.org/10.1007/s10064-015-0762-5>

685 Leonarduzzi, E., Molnar, P., 2020. Data limitations and potential of hourly and daily rainfall  
686 thresholds for shallow landslides. *Nat. Hazards Earth Syst. Sci. Discuss.* 1–25.  
687 <https://doi.org/10.5194/nhess-2020-125>

688 Marra, F., Destro, E., Nikolopoulos, E.I., Zoccatelli, D., Dominique Creutin, J., Guzzetti, F., Borga,  
689 M., 2017. Impact of rainfall spatial aggregation on the identification of debris flow occurrence  
690 thresholds. *Hydrol. Earth Syst. Sci.* 21, 4525–4532. <https://doi.org/10.5194/hess-21-4525-2017>

691 Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Guzzetti, F., 2016. Rainfall thresholds for  
692 the possible landslide occurrence in Sicily (Southern Italy) based on the automatic  
693 reconstruction of rainfall events. *Landslides* 13, 165–172. [https://doi.org/10.1007/s10346-015-](https://doi.org/10.1007/s10346-015-0630-1)  
694 [0630-1](https://doi.org/10.1007/s10346-015-0630-1)

695 Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Guzzetti, F., 2014. An Algorithm for the  
696 objective reconstruction of rainfall events responsible for landslides. *Landslide Dyn. ISDR-ICL*  
697 *Landslide Interact. Teach. Tools Vol. 1 Fundam. Mapp. Monit.* 12, 311–320.  
698 [https://doi.org/10.1007/978-3-319-57774-6\\_33](https://doi.org/10.1007/978-3-319-57774-6_33)

699 Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Roccati, A., Guzzetti, F., 2018. A tool for  
700 the automatic calculation of rainfall thresholds for landslide occurrence. *Environ. Model. Softw.*  
701 105, 230–243. <https://doi.org/10.1016/j.envsoft.2018.03.024>

702 Mirus, B.B., Becker, R.E., Baum, R.L., Smith, J.B., 2018a. Integrating real-time subsurface  
703 hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning.

704 Landslides 15, 1909–1919. <https://doi.org/10.1007/s10346-018-0995-z>

705 Mirus, B.B., Mopthew, M.D., Smith, J.B., 2018b. Developing hydro-meteorological thresholds for  
706 shallow landslide initiation and early warning. *Water (Switzerland)* 10, 1–19.  
707 <https://doi.org/10.3390/W10091274>

708 Nikolopoulos, E.I., Crema, S., Marchi, L., Marra, F., Guzzetti, F., Borga, M., 2014. Impact of  
709 uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow  
710 occurrence. *Geomorphology* 221, 286–297. <https://doi.org/10.1016/j.geomorph.2014.06.015>

711 Owe, M., de Jeu, R., Holmes, T., 2008. Multisensor historical climatology of satellite-derived global  
712 land surface moisture. *J. Geophys. Res. Earth Surf.* 113, 1–17.  
713 <https://doi.org/10.1029/2007JF000769>

714 Peruccacci, S., Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., Guzzetti, F., 2017. Rainfall  
715 thresholds for possible landslide occurrence in Italy. *Geomorphology* 290, 39–57.  
716 <https://doi.org/10.1016/j.geomorph.2017.03.031>

717 Peruccacci, S., Brunetti, M.T., Luciani, S., Vennari, C., Guzzetti, F., 2012. Lithological and seasonal  
718 control on rainfall thresholds for the possible initiation of landslides in central Italy.  
719 *Geomorphology* 139–140, 79–90. <https://doi.org/10.1016/j.geomorph.2011.10.005>

720 Piciullo, L., Calvello, M., Cepeda, J.M., 2018. Territorial early warning systems for rainfall-induced  
721 landslides. *Earth-Science Rev.* 179, 228–247. <https://doi.org/10.1016/j.earscirev.2018.02.013>

722 Ponziani, F., Pandolfo, C., Stelluti, M., Berni, N., Brocca, L., Moramarco, T., 2012. Assessment of  
723 rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in  
724 the Umbria region (central Italy). *Landslides* 9, 229–237. [https://doi.org/10.1007/s10346-011-](https://doi.org/10.1007/s10346-011-0287-3)  
725 [0287-3](https://doi.org/10.1007/s10346-011-0287-3)

726 Sajeev, R., Praveen, K.R., 2014. Landslide Susceptibility Mapping on Macroscale along the Major  
727 Road Corridors in Idukki District, Kerala. Thiruvananthapuram, India.

728 Segoni, S., Piciullo, L., Gariano, S.L., 2018a. A review of the recent literature on rainfall thresholds

729 for landslide occurrence. *Landslides* 15, 1483–1501. <https://doi.org/10.1007/s10346-018-0966-4>

730 Segoni, S., Rosi, A., Fanti, R., Gallucci, A., Monni, A., Casagli, N., 2018b. A regional-scale landslide  
731 warning system based on 20 years of operational experience. *Water (Switzerland)* 10, 1–17.  
732 <https://doi.org/10.3390/w10101297>

733 Segoni, S., Rosi, A., Lagomarsino, D., Fanti, R., Casagli, N., 2018c. Brief communication: Using  
734 averaged soil moisture estimates to improve the performances of a regional-scale landslide early  
735 warning system. *Nat. Hazards Earth Syst. Sci.* 18, 807–812. [https://doi.org/10.5194/nhess-18-](https://doi.org/10.5194/nhess-18-807-2018)  
736 [807-2018](https://doi.org/10.5194/nhess-18-807-2018)

737 Song, S., Wang, W., 2019. Impacts of antecedent soil moisture on the rainfall- runoff transformation  
738 process based on high- resolution observations in soil tank experiments. *Water (Switzerland)* 11,  
739 15–20. <https://doi.org/10.3390/w11020296>

740 Teja, T.S., Dikshit, A., Satyam, N., 2019. Determination of rainfall thresholds for landslide prediction  
741 using an algorithm-based approach: Case study in the Darjeeling Himalayas, India. *Geosci.* 9.  
742 <https://doi.org/10.3390/geosciences9070302>

743 Terlien, M.T.J., 1998. The determination of statistical and deterministic hydrological landslide-  
744 triggering thresholds. *Environ. Geol.* 35, 124–130. <https://doi.org/10.1007/s002540050299>

745 Tofani, V., Bilocchi, G., Rossi, G., Segoni, S., D'Ambrosio, M., Casagli, N., Catani, F., 2017. Soil  
746 characterization for shallow landslides modeling: a case study in the Northern Apennines  
747 (Central Italy). *Landslides* 14, 755–770. <https://doi.org/10.1007/s10346-017-0809-8>

748 Uchimura, T., Towhata, I., Anh, T.T.L., Fukuda, J., Bautista, C.J.B., Wang, L., Seko, I., Uchida, T.,  
749 Matsuoka, A., Ito, Y., Onda, Y., Iwagami, S., Kim, M.S., Sakai, N., 2010. Simple monitoring  
750 method for precaution of landslides watching tilting and water contents on slopes surface.  
751 *Landslides* 7, 351–357. <https://doi.org/10.1007/s10346-009-0178-z>

752 Uchimura, T., Towhata, I., Wang, L., Nishie, S., Yamaguchi, H., Seko, I., Qiao, J., 2015. Precaution  
753 and early warning of surface failure of slopes using tilt sensors. *Soils Found.* 55, 1086–1099.

754 <https://doi.org/10.1016/j.sandf.2015.09.010>

755 United Nations Development Programme, 2018. Kerala Post Disaster Needs Assessment Floods and  
756 Landslides-August 2018. Thiruvananthapuram, India.

757 Valenzuela, P., Domínguez-Cuesta, M.J., Mora García, M.A., Jiménez-Sánchez, M., 2018. Rainfall  
758 thresholds for the triggering of landslides considering previous soil moisture conditions  
759 (Asturias, NW Spain). *Landslides* 15, 273–282. <https://doi.org/10.1007/s10346-017-0878-8>

760 Varnes, D., 1978. Slope Movement Types and Processes. Transp. Res. Board Spec. Rep.

761 Wei, X., Fan, W., Cao, Y., Chai, X., Bordoni, M., Meisina, C., Li, J., 2020. Integrated experiments on  
762 field monitoring and hydro-mechanical modeling for determination of a triggering threshold of  
763 rainfall-induced shallow landslides. A case study in Ren River catchment, China. *Bull. Eng.  
764 Geol. Environ.* 79, 513–532. <https://doi.org/10.1007/s10064-019-01570-7>

765 Wicki, A., Lehmann, P., Hauck, C., Seneviratne, S.I., Waldner, P., Stähli, M., 2020. Assessing the  
766 potential of soil moisture measurements for regional landslide early warning. *Landslides* 17,  
767 1881–1896. <https://doi.org/10.1007/s10346-020-01400-y>

768 Wu, M.H., Wang, J.P., Chen, I.C., 2019. Optimization approach for determining rainfall duration-  
769 intensity thresholds for debris flow forecasting. *Bull. Eng. Geol. Environ.* 78, 2495–2501.  
770 <https://doi.org/10.1007/s10064-018-1314-6>

771 Yang, Z., Cai, H., Shao, W., Huang, D., Uchimura, T., Lei, X., Tian, H., Qiao, J., 2019. Clarifying the  
772 hydrological mechanisms and thresholds for rainfall-induced landslide: in situ monitoring of big  
773 data to unsaturated slope stability analysis. *Bull. Eng. Geol. Environ.* 78, 2139–2150.  
774 <https://doi.org/10.1007/s10064-018-1295-5>

775 Zhao, B., Dai, Q., Han, D., Dai, H., Mao, J., Zhuo, L., 2019a. Probabilistic thresholds for landslides  
776 warning by integrating soil moisture conditions with rainfall thresholds. *J. Hydrol.* 574, 276–  
777 287. <https://doi.org/10.1016/j.jhydrol.2019.04.062>

778 Zhao, B., Dai, Q., Han, D., Dai, H., Mao, J., Zhuo, L., Rong, G., 2019b. Estimation of soil moisture

779 using modified antecedent precipitation index with application in landslide predictions.

780 Landslides 16, 2381–2393. <https://doi.org/10.1007/s10346-019-01255-y>

781



29 applications in a regional landslide early warning system (LEWS). This study is an attempt to  
30 improve the performance of conventional meteorological thresholds by considering the effect of soil  
31 moisture, using a probabilistic approach. Idukki district in southern part of India is highly susceptible  
32 to landslides and has witnessed major socio-economical setbacks in the recent disasters happened in  
33 2018 and 2019. This tourist hub is now in need of a landslide forecasting system, which can help in  
34 landslide risk reduction. This study attempts to understand the effect of averaged soil moisture  
35 estimates derived from passive microwave remote sensing data, for improving the performance of  
36 conventional empirical and probabilistic thresholds. For defining empirical thresholds, an algorithm-  
37 based approach such as Calculation of Thresholds for Rainfall-induced Landslides Tool (CTRL-T)  
38 has been used. Probabilistic thresholds were defined using a Bayesian approach, finding the posterior  
39 probability of occurrence using the marginal and conditional probabilities of the control parameters  
40 along with the prior probability of occurrence of landslide. The derived rainfall thresholds were  
41 quantitatively compared with the Bayesian probabilistic threshold derived using rainfall severity and  
42 soil wetness using an area under the curve (AUC) based receiver operating characteristics (ROC)  
43 curve method. The results show that when the antecedent moisture content in soil is less, only severe  
44 rainfall events can trigger landslides in the study area; while less severe rainfall events can also trigger  
45 landslides when the soil is wet. The role of soil wetness in the initiation is used to improve the  
46 performance of the conventional methods, and a ROC approach was used for the statistical  
47 comparison of different models. Further, the results indicated that the probabilistic threshold using  
48 rainfall severity and soil wetness outperformed the conventional approaches with AUC of 0.96, being  
49 the most sensitive and specific among the models considered. This result opens new promising  
50 perspectives for the development of an operational LEWS in the Idukki district based on a  
51 combination of rainfall and soil moisture data. Moreover, this work contributes to strengthen the  
52 advancing trend of hydro-meteorological thresholds based on soil moisture, which is gaining a  
53 growing attention in landslide studies and that, to date, was lacking evidences in monsoon regions.

54 **Keywords:** landslides; rainfall thresholds; LEWS; soil moisture; Idukki

55

## 56        **1. Introduction**

57        Forecasting landslides and evacuating people from hazardous zones is an important risk reduction  
58        strategy (Althuwaynee and Pradhan, 2017). Considering the climate change and associated extreme  
59        rainfall phenomenon, the number of rainfall-induced landslides are expected to rise (Alvioli et al.,  
60        2018; Chen et al., 2019; Gariano and Guzzetti, 2016). Being a geomorphological process in the  
61        landscape evolution (Iida, 1999), the detailed understanding of slope failure mechanisms involves  
62        hydrological studies and forecasting of possible failure planes (Agostini et al., 2014) using relevant  
63        geotechnical and meteorological parameters. However, these parameters are highly site specific and  
64        often difficult to determine with the desired accuracy (Tofani et al., 2017), except that for single  
65        slopes or very small basins (Chae et al., 2017), and sophisticated experimental research is required for  
66        understanding the mechanism in detail (Kim et al., 2018). Hence, a more practiced approach is needed  
67        to forecast the critical conditions which result in the occurrence of landslides using the primary  
68        triggering factor i.e. rainfall – with the aid of rainfall thresholds (Caine, 1980; Keefer et al., 1987;  
69        Piciullo et al., 2018). Rainfall thresholds can be empirical, probabilistic, or algorithm based  
70        (Althuwaynee et al., 2015; Piciullo et al., 2018; Segoni et al., 2018a). All the approaches exploit  
71        historical data to find a mathematical relationship between rainfall and the occurrence of landslides in  
72        a region, to identify critical rainfall conditions which can trigger landslides in the future. A rainfall  
73        event is most commonly characterised in terms of cumulated rainfall event (E), duration (D), and  
74        intensity (I) (which are referred to as “rainfall parameters”). Consequently, the thresholds are often  
75        defined as cumulated event rainfall vs. duration (ED thresholds) (Lainas et al., 2016; Melillo et al.,  
76        2018, 2016; Peruccacci et al., 2017; Teja et al., 2019) or as rainfall intensity vs. duration (ID  
77        thresholds) (Battistini et al., 2017; Brunetti et al., 2010; Guzzetti et al., 2008; Lainas et al., 2016; Wu  
78        et al., 2019).

79        When the definition of thresholds is associated with the generation of many false alarms, their usage  
80        in operational Landslide Early Warning System (LEWS) may be inappropriate (Aleotti, 2004;  
81        Guzzetti et al., 2008; Kirschbaum et al., 2012; Segoni et al., 2018b). Low performances of rainfall  
82        thresholds are traditionally related to the uncertainties associated with the definition of rainfall



83 parameters, the quality and resolution of the historical data and the intrinsic limitations of the  
84 statistical models (Gariano et al., 2020; Marra et al., 2017; Nikolopoulos et al., 2014).

85 Some authors argued that sometimes the statistical correlation between rainfall parameters and  
86 landslide initiation is too weak and that hydro-meteorological thresholds accounting for both rainfall  
87 and hydrological (e.g. soil moisture) parameters could provide a stronger and more accurate  
88 assessment (Bogaard and Greco, 2018; Jakob et al., 2006; Terlien, 1998). Integrating soil moisture  
89 with rainfall thresholds has been proven effective in improving the rainfall thresholds (Abraham et al.,  
90 2020b; Segoni et al., 2018c; Zhao et al., 2019a), as the antecedent moisture content plays a key role in  
91 the shear strength parameters of soil. The soil moisture conditions play a key role in the infiltration  
92 process (Song and Wang, 2019) which significantly influences the initiation of landslides  
93 (Alimohammadlou et al., 2014; Baum et al., 2008; Biscopchi et al., 2019; Iverson, 2000; Wei et al.,  
94 2020; Yang et al., 2019). Weighted indexes (Glade et al., 2000; Ponziani et al., 2012); and satellite  
95 data (Zhao et al., 2019b) can be used for estimating soil moisture values when real-time field  
96 monitoring (Abraham et al., 2020c; Dikshit et al., 2018; Uchimura et al., 2015, 2010) cannot be  
97 conducted. Hydrological models (Abraham et al., 2020b; Zhao et al., 2019a) can also be used for the  
98 estimation of soil moisture content. In the published literature, soil moisture combined with rainfall  
99 thresholds has been tested mainly in Mediterranean, temperate and alpine climatic settings, whereas in  
100 monsoon regions similar types of tests are almost completely missing (Jakob et al., 2006; Mirus et al.,  
101 2018a; Valenzuela et al., 2018; Wicki et al., 2020).

102 The present work attempts to define statistical rainfall thresholds in Idukki district (India) and to  
103 improve their effectiveness by coupling rainfall parameters with soil moisture data. First, ED  
104 thresholds are defined using an automatic algorithm-based approach (Melillo et al., 2014). The  
105 algorithm first identifies the triggering rainfall events using the location of rain gauges and landslides,  
106 the time of occurrence of landslides and the time series rainfall data. It recreates multiple rainfall  
107 conditions which may result in landslides and identifies the maximum probable rainfall condition  
108 based on the location and time. After identifying the triggering rainfall event, the algorithm defines  
109 the ED thresholds with multiple exceedance probabilities using frequentist method. Then, by using a

110 probabilistic approach (Berti et al., 2012), the effect of event rainfall, duration and intensity on the  
111 occurrence of landslides is evaluated (probabilistic rainfall thresholds). Both empirical (Melillo et al.,  
112 2018, 2016; Peruccacci et al., 2017) and probabilistic approaches (Berti et al., 2012; Dikshit and  
113 Satyam, 2019) were considered to establish the relationship between primary triggering factor  
114 (rainfall) and the result (landslide), and these are simple statistical approaches that are easy to derive  
115 by integrating with a rainfall forecasting system. Similar studies have been conducted for Indian  
116 Himalayas (Abraham et al., 2020a; Dikshit and Satyam, 2018, 2019; Teja et al., 2019) and the  
117 Western Ghats (Abraham et al., 2020e, 2019); however, these methods were not always found to be  
118 operational due to a higher number of false alarms or missed alarms, limiting their applications in  
119 LEWS. This study aims to overcome these limitations by integrating soil moisture data along with the  
120 rainfall thresholds. The objective is to find if the addition of soil moisture data can perform better than  
121 the conventional methods based on the rainfall data alone.

## 122 **2. Description of the study area**

123 The Western Ghats of Indian Peninsula is highly susceptible to rainfall-induced landslides. There is a  
124 surge in the number of landslides during monsoon season since 2018, due to very-high intensity  
125 rainfalls. The landslides and floods happened in 2018 severely affected the south Indian states of  
126 Kerala and Karnataka. Among the 14 districts in the state of Kerala, 13 are part of the Western Ghats  
127 and are susceptible to landslides. Nearly 5.3 million people in the state were affected by the disaster in  
128 2018 (United Nations Development Programme, 2018). The Western Ghats scarps, running the whole  
129 extent of the mountain range, are highly prone to landslides. Very-high intensity rainfall, along with  
130 the anthropogenic activities, has accelerated the geological processes leading to landslides, making  
131 the situation alarming (Kuriakose et al., 2009b).

132 Idukki is a hilly district in the Western Ghats and is the second largest district in the state of Kerala, in  
133 terms of area. This district covers an area of 4358 km<sup>2</sup> and derived its name from the word 'Idukku' in  
134 the vernacular dialect meaning *narrow gorge*. This itself indicates the geography of the area. The  
135 district is the major power source of Kerala and houses many hydroelectric projects, including the  
136 famous arch dam of Idukki. About 50% of the district is covered by forests and Idukki is drained by

137 three major rivers, two flowing westward and one eastward. The rainfall across the district is varying  
138 with the least values recorded in the northern side with a long-term average of 1000 mm while the  
139 southern parts record an average rainfall of 5000 mm (Sajeev and Praveen 2014; Department of  
140 Mining and Geology 2016). The southwest monsoon season from June to September contributes 60%  
141 of the annual rainfall and around 24% is contributed by the North-East monsoon from October to  
142 December. Due to varying topography, the climatic conditions in the hill ranges, plateaus and  
143 midlands of the district are different from each other.

144 **Fig. 1.** Location details of study area. (a) India, and (b) Digital Elevation Model of Idukki (modified  
145 using CartoDEM (CartoDEM, 2015)) along with location of rain gauges.

146 Geologically, Idukki can be divided into three different parts from south to north. The charnockite  
147 rocks in the south, migmatitic complex in central portion, and peninsular gneissic complex in the  
148 northern part. Granite gneiss is the oldest and predominant group among the peninsular gneissic  
149 complex while the charnockite group consists of magnetite quartzite, pyroxene granulite and  
150 charnockite (Department of Mining and Geology 2016). Structural cum denudational hills are the  
151 predominant geomorphological feature of Idukki. The hills are generally having a thin soil cover  
152 overlaid on Precambrian basement rocks. The midlands have a rugged topography with small hills and  
153 deep valleys with an average elevation of 50 m. The zone where midlands grades to plateaus are  
154 called the foothills, ranging up to 8 km in width. A major portion of the district belongs to the plateau  
155 region, with a large landmass of moderate slope. The elevation of the plateau region goes up to 1500  
156 m, and the regions at an elevation greater than 1500 m belong to hilly ranges. More than 50% of the  
157 study area is covered by forest loam soils, produced by the weathering of rock under thick forest  
158 cover. The midlands are covered by lateritic soil with high permeability and less organic content. The  
159 valley portion of the terrain are covered with fine particles of sandy loam to clay type, formed by  
160 sedimentation and transportation of hill slopes. The narrow riverbanks consist of fertile alluvial soil  
161 and are more common in the midlands.

162 Because of its topographic variability and heavy rainfall, the district is highly susceptible to rainfall  
163 induced landslides. The typology of landslides in the Western Ghats includes earth and debris slides,

164 rock falls, creep, slump and debris flows (Abraham et al., 2020d). Due to the thin regolith layer,  
165 shallow landslide (Varnes, 1978) is the most common type during prolonged rainfalls (Kuriakose et  
166 al., 2009a). Idukki district in particular is mostly affected by the cut slope failures along the major  
167 road corridors, disrupting the transportation network in the district. Recent changes in the land use  
168 patterns for infrastructure development and agriculture have affected the stability of slopes of this  
169 ecologically sensitive zone (Gadgil et al., 2011) and has aggravated the number of landslide disasters  
170 (Kuriakose et al., 2009b). Hence the development of an effective regional scale LEWS is highly  
171 needed to forecast the future landslides in the region.

172

### 173 **3. Data and Methodology**

174 The study explores the possibility of using soil moisture data in improving the performance of  
175 statistical thresholds. The overall methodology flow chart adopted in this study is shown in Fig 2. The  
176 methodology involves data collection from multiple sources, the definition of thresholds and their  
177 performance evaluation using different skill scores. For the analysis, historical rainfall, landslide, and  
178 soil moisture data were collected. For developing empirical and probabilistic rainfall thresholds, only  
179 rainfall and landslide data are required, while for developing probabilistic rainfall thresholds based on  
180 rainfall severity and soil wetness (RS threshold), the soil moisture data were integrated with empirical  
181 ED thresholds using a probabilistic approach. While the empirical threshold considers the effect of  
182 rainfall events which resulted in landslides, the probabilistic thresholds consider both triggering and  
183 non-triggering rainfall events for the analysis.

184

185 **Fig. 2.** Methodology of study.

#### 186 **3.1 Data collection**

187 The dataset used for this study spans from 2010 to 2018 and the historical data from this period was  
188 used to derive the empirical and probabilistic thresholds for occurrence of landslides in Idukki district.

189 The daily rainfall data was collected from the Indian Meteorological Department (India  
190 Meteorological Department 2019) for four rain gauges within the district. The landslide data was  
191 collected from various government agencies and media reports (Abraham et al., 2019) and only  
192 landslides for which the date of occurrence was available were used for the analysis. For each rain  
193 gauge a reference area was defined and multiple landslides triggered in the same day in each area  
194 were considered as one landslide event and rainfall data were collected from the reference rain gauge.  
195 By these criteria, 225 landslide events were identified in the study area which were first used as the  
196 input for empirical thresholds. For probabilistic thresholds, a total of 5028 rainfall events recorded by  
197 the four rain gauges during the study period were considered.

198 The average daily soil moisture data was collected from Giovanni's website by National Aeronautics  
199 and Space Administration Goddard Earth Sciences Data and Information Services Center (NASA  
200 GES DISC) (de Jeu and Owe, 2014, 2012; Giovanni, 2020). The data was derived using land  
201 parameter retrieval model (LPRM), which is a multi-parameter retrieval algorithm focused on  
202 hydrological and climate studies. It retrieves the soil moisture from the microwave observations from  
203 sensors. The observed brightness temperatures were used to derive the soil moisture data, using  
204 LPRM (Owe et al., 2008). LPRM is based on a forward radiative transfer model and the output is the  
205 volumetric soil moisture content in percentage. The soil moisture on the day before the occurrence of  
206 landslide, termed as the 'antecedent soil moisture' was used for the analysis in this research. The  
207 spatial resolution of the data is  $0.25^\circ \times 0.25^\circ$ . The study area (Idukki district) consists of 14 grids of  
208 size  $0.25^\circ \times 0.25^\circ$  (Figure 1). After calculating the area of Idukki within each grid, the weighted  
209 average was calculated for the whole area, for simplified calculation. This value is called the  
210 'averaged moisture content'. Another term, 'soil wetness' is introduced, to represent a range of  
211 antecedent soil moisture, on a scale of 0 to 1. The soil wetness values were divided into five equal  
212 parts, representing different ranges of moisture content. This classification is used to overcome the  
213 limitations associated with using averaged data for a larger area. The value of soil wetness is directly  
214 proportional to the moisture content values and indicates the wetness of soil before the landslide.

215 Thus, by using historical rainfall, landslide and soil moisture data, thresholds were defined using  
216 multiple approaches for the study area to find the effect of soil moisture on the forecasting  
217 performance of the thresholds.

218

### 219 3.2 Empirical thresholds

220 The selection of rain gauges and rainfall parameters plays a critical role in the definition of rainfall  
221 thresholds (Abraham et al., 2020e). For the study area, rainfall data from the four available rain  
222 gauges were considered for the analysis. The intensity-duration thresholds for the study area was  
223 earlier derived from using a nearest rain gauge approach (Abraham et al., 2019), considering 225  
224 landslide events occurred from 2010 to 2018. From the pioneering work of Caine (Caine, 1980), ID  
225 thresholds were defined for regions across the globe (Abraham et al., 2020c, 2019; Brunetti et al.,  
226 2010; Dikshit and Satyam, 2018; Guzzetti et al., 2008, 2007; Segoni et al., 2018a). Even though  
227 intensity can easily be converted to event rainfall and vice-versa, recent literature shows a shift  
228 towards defining ED thresholds instead of ID thresholds (Melillo et al., 2018, 2014; Peruccacci et al.,  
229 2012; Teja et al., 2019; Zhao et al., 2019a). The reason is that E and D are two mutually independent  
230 parameters while I is a function of D and E. Hence, for a definition of rainfall thresholds and rainfall  
231 severity, the data points on ED plane was considered in this study. In this study, the reconstruction of  
232 event- duration thresholds was carried out by using Calculation of Thresholds for Rainfall Induced  
233 Landslides - Tool (CTRL-T) (Melillo et al., 2018, 2014). CTRL-T uses an algorithm-based approach,  
234 extracting the rainfall events automatically from the daily precipitation data input. From the extracted  
235 events, rainfall conditions that have triggered landslides were identified; and used to derive the  
236 rainfall thresholds for the region. The tool considers a buffer zone around each landslide location, to  
237 search for the rain gauge and identify the triggering event. In this study, a search radius of 20 km is  
238 considered, due to the low rain gauge density in the study area. The algorithm also considers a delay  
239 time between the end of rainfall and occurrence of landslide. In this study, the delay time is taken as  
240 48 hours (Melillo et al., 2014). If no rainfall condition is recreated within this delay time before the  
241 occurrence of landslide, the event will be discarded by the algorithm. The algorithm first determines

242 the total event rainfall and duration of rainfall for all identified rainfall events and then to minimise  
243 the effect of spatial variability of rainfall distribution, single or multiple rainfall conditions (MRC)  
244 likely to result in failures and a weight is assigned to each of them. Then for each landslide, the  
245 highest weight was used to identify the reference rain gauge and to choose the maximum probable  
246 rainfall conditions (MPRC). In this study, five different threshold lines were defined using CTRL-T,  
247 at different exceedance probabilities of 1%, 5%, 10%, 20% and 50% (termed as  $T_1$ ,  $T_5$ ,  $T_{10}$ ,  $T_{20}$  and  
248  $T_{50}$ , respectively). Thresholds and related uncertainties were estimated from MPRCs. The defined  
249 thresholds are in the form of a power law, determined using the frequentist approach (Brunetti et al.,  
250 2010) and can be expressed as:

251

$$E = (\alpha \pm \Delta\alpha) D^{(\gamma \pm \Delta\gamma)} \quad (1)$$

252

253 where,  $\alpha$  is the scaling parameter or the intercept and  $\gamma$  is the shape parameter which denotes the slope  
254 of the equation.  $\Delta\alpha$  and  $\Delta\gamma$  represents the uncertainties associated with  $\alpha$  and  $\gamma$ , respectively. The  
255 uncertainties are determined using a bootstrap approach.

### 256 3.3 Probabilistic approach

257 The empirical thresholds compare an input value with the defined thresholds and will have a single  
258 output (triggering or non-triggering). It is often difficult to decide the exceedance probability to be  
259 selected as a threshold beyond which a radical change can be expected in the system (Berti et al.,  
260 2012). The discretion between triggering and non- triggering rainfall conditions is not trivial in such  
261 cases. To derive the equation, only the triggering rainfall conditions are considered. This increases the  
262 chances of false alarms, as numerous rainfall events that cross the threshold line not necessarily  
263 trigger landslides.

264 By considering both triggering and non- triggering rainfalls for analysis, probability-based models are  
265 more informative and provide a better option to find extreme events. In this study, a Bayesian  
266 approach is used to define probabilistic thresholds (Berti et al., 2012).

267 3.3.1 One-dimensional analysis

268 Bayes theorem applies a conditional probability of some event  $L$  (landslide) given the occurrence of  
269 another event  $X$  (rainfall, expressed in terms of E, I or D). This is also called the posterior probability,  
270  $P(X|L)$ . It can be calculated as follows (Berti et al., 2012):

$$P(L|X) = \frac{P(X|L) * P(L)}{P(X)} \quad (2)$$

271 where,  $P(X|L)$  is the conditional probability of occurrence of rainfall of magnitude  $X$ , when a  
272 landslide occurs. This is also called as a likelihood.

273  $P(L)$  is the prior probability of occurrence of landslide regardless of the occurrence rainfall  
274 magnitude.

275  $P(X)$  is the marginal probability of  $X$ , which can be defined as the probability of occurrence of  
276 rainfall regardless of the occurrence of landslides. The terms can be calculated mathematically using  
277 relative frequencies. Let  $N_R$  be the total number of rainfall events during study period,  $N_L$  be the total  
278 number of landslides occurred,  $N_X$  be the number of rainfall events with magnitude  $X$  and  $N_{(X|L)}$  be  
279 the number of rainfall events with magnitude  $X$  that resulted in landslides. The probabilities can be  
280 computed as (Berti et al., 2012):

281

$$P(L) \approx \frac{N_L}{N_R} \quad (3)$$

$$P(X) \approx \frac{N_X}{N_R} \quad (4)$$

$$P(X|L) \approx \frac{N_{(X|L)}}{N_L} \quad (5)$$



282 Considering the rainfalls that resulted in landslides only will give us partial information, the  
 283 likelihood. To understand the influence of rainfall of magnitude  $X$ , it is important to compare the prior  
 284 probability with the posterior probability.

### 285 3.3.2 Two-dimensional analysis

286 Two-dimensional case is the extension of Eq. 2 by considering two conditions  $X, Y$  instead of the  
 287 single condition  $X$  in Eq. 2. In the initial analysis, we consider  $X$  and  $Y$  as magnitude of two rainfall  
 288 parameters (E,D ; I,D; E,I). The calculation of prior, marginal and conditional probabilities are given  
 289 below:

290

$$P(L|X, Y) = \frac{P(X, Y|L) * P(L)}{P(X, Y)} \quad (6)$$

$$P(L) \approx \frac{N_L}{N_R} \quad (7)$$

$$P(X, Y) \approx \frac{N_{X,Y}}{N_R} \quad (8)$$

$$P(X, Y|L) \approx \frac{N_{(X,Y|L)}}{N_L} \quad (9)$$

291 The study explores the effect of antecedent soil moisture content using a two-dimensional  
 292 probabilistic analysis. During the second phase, we considered rainfall severity in ED plane and soil  
 293 wetness as  $X$  and  $Y$ , respectively. Based on the values of soil wetness, five different categories were  
 294 considered for analysis viz, less than 0.2, 0.2 to 0.4, 0.4 to 0.6, 0.6 to 0.8, and 0.8 to 1. The categories  
 295 based on rainfall severity were less than  $T_1$ ,  $T_1$  to  $T_5$ ,  $T_5$  to  $T_{10}$ ,  $T_{10}$  to  $T_{20}$ ,  $T_{20}$  to  $T_{50}$  and greater than  
 296  $T_{50}$ . Thus, the two-dimensional plane was divided into 30 cells as a 6 x 5 matrix as shown in Fig. 6.  
 297 These values were used for the definition of RS threshold.

298 **4. Results**

299 **4.1 Empirical thresholds**

300 CTRL-T tool considered 177 landslide events out of the 225 and the rest were discarded to avoid  
301 introduction of relevant spatio-temporal uncertainties in the analysis. The uncertainties are associated  
302 with the less rain gauge density in the study area. As described earlier, the landslides for which  
303 responsible rainfall conditions were not identified were discarded. This can be due to a distance more  
304 than 20 km between the location of rain gauges and landslide or due to a delay time more than 48  
305 hours after the end of any rainfall event. The algorithm forecasted rainfall thresholds with various  
306 exceedance probability both in normal and logarithmic plot (Fig. 3). The threshold lines of 1%, 5%,  
307 10%, 20% and 50% exceedance probabilities were used to classify the events into six categories based  
308 on the severity of rainfall. These lines are named  $T_1$ ,  $T_5$ ,  $T_{10}$ ,  $T_{20}$  and  $T_{50}$ , respectively. The slope of  
309 threshold lines in logarithmic plot was found to be  $0.57 \pm 0.03$ . This value is not in good agreement  
310 with the ID thresholds defined for the area in a previous study (Abraham et al., 2019). Though both  
311 the studies used frequentist approach for the definition of thresholds, the process of identification of  
312 responsible rainfall event was different. In the previous study (Abraham et al., 2019), the responsible  
313 rainfalls were identified using a Thiessen polygon approach manually, while in this study, the  
314 automatic algorithm, CTRL-T is used for identifying the responsible rainfall event. The parameters  
315 of threshold lines and the uncertainties associated are listed in Table 1.

316

317

318 **Fig. 3.** Rainfall event – duration thresholds for Idukki district

319 **Table 1.** Values of  $\alpha$ ,  $\gamma$  and the uncertainties associated with different exceedance probabilities

320

321 The range of duration of rainfalls considered for analysis vary from 1 to 26 days. For the thresholds to  
322 be reliable, the relative uncertainty ( $\Delta x/x$  for any variable  $x$ ) should be less than 10%. Here the

323 relative uncertainty of  $\gamma$  is 5.2%. But with higher exceedance probabilities, the relative uncertainty of  
324  $\alpha$  is crossing this limiting value.

325 With 5% exceedance probability, 20.19mm rainfall can trigger a landslide in the region for a duration  
326 of 24 hours and when the duration is 624 hours, a rainfall of 129 mm can trigger landslides in the  
327 region. For a better understanding of the effect of each rainfall parameter on the occurrence of  
328 landslides, probabilistic rainfall thresholds were defined for the area.

#### 329 4.2 Probabilistic thresholds

330 The maximum probable rainfall conditions which were used for the definition of ED thresholds were  
331 considered as the triggering rainfall events for the probabilistic analysis. Thus, out of the 5028 rainfall  
332 events considered, 177 events were identified as triggering events by CTRL- T algorithm and the rest  
333 4851 events were considered as non-triggering rainfall events. In the one-dimensional case, six  
334 categories of rainfall duration, five categories of cumulated rainfall event and seven categories of  
335 rainfall intensity were considered. The results are plotted in Fig. 4 (a-f); where Fig. 4a, c and e depict  
336 the prior probability, marginal probability and likelihood, and Fig. 4b, d and f depict the prior and  
337 posterior probabilities. The variable  $X$  in Eq. (2-5) is replaced with D, E and I in the respective  
338 graphs.  $P(L)$  being a constant parameter (value 0.035 in this study), the ratio of  $P(X|L)$  to  $P(X)$   
339 determines the variation of posterior probability values. Hence when  $P(X|L) > P(L)$ , the posterior  
340 probability is greater than prior probability and vice versa. The more the variation between prior and  
341 posterior probability, the more significant the variable is. It can be seen, that for duration and event  
342 rainfall, for the largest values of variables, the values of  $P(L|X)$  is less than  $P(L)$ , while in the case of  
343 intensity, high intensity rainfalls are more probable to trigger landslides in the region. The plots of  
344  $P(X)$  and  $P(X|L)$  are well above the plot of prior probability in all the cases. Intensity was found to  
345 be the most significant variable, with the maximum ratio between posterior and prior probabilities.  
346 The maximum posterior probability when the control parameter is D was found to be 0.053 where the  
347 value is 0.103 and 0.116 in the case of E and I, respectively. Maximum probability occurs when the  
348 duration is between 120 h to 240 h; event rainfall is between 100 mm to 200 mm; and intensity is  
349 greater than 3 mm/h.

350

351 **Fig. 4.** Prior, conditional, marginal and posterior probabilities with respect to rainfall parameters. (a,  
352 b) Duration; (c, d) Event rainfall; and (e, f) Intensity.

353 To evaluate the joint occurrence of two parameters, two-dimensional Bayesian analysis were  
354 conducted with data on three different planes (Fig. 5). The two-dimensional space for each analysis  
355 was divided into small cells based on the categories of parameters used for one-dimensional analysis.  
356 Hence the ID plane is a 7 x 6 matrix, ED plane is a 5 x 6 matrix and the EI plane is a 5 x 7 matrix.  
357 There are several no data points in all three cases, due to the lower number of landslides considered  
358 for the analysis. As identified from the one-dimensional analysis, E and I were found to be more  
359 critical parameters than D. This is the reason why this study has considered all three different  
360 combinations of the control parameters even though the empirical thresholds are defined on ED plane  
361 only. The maximum probability value was obtained on EI plane, when the intensity value is less than  
362 0.5 mm/h and event rainfall is between 100 mm to 200 mm, with a value of 0.54.

363

364 **Fig. 5.** Two-dimensional posterior probabilities of occurrence of landslide on (a) ID plane, (b) ED  
365 plane, and (c) EI plane.

366 It is evident from Fig. 6 that even less severe rainfall events when falling on a moist soil can trigger  
367 landslides in the region. Most of the landslides for which rainfall events were less severe happened on  
368 days with higher soil wetness. Also, when the rainfall event is severe, even dry soil can be susceptible  
369 to landslides. The maximum probability of 0.49 was observed when the rainfall severity was between  
370  $T_{20}$  to  $T_{50}$  and the soil wetness was between 0.8 to 1. With the available data, when the antecedent soil  
371 moisture is less, only extremely severe rainfall conditions can trigger landslides in the area. This  
372 affects the performance of the ED thresholds considerably. For different antecedent soil moisture  
373 conditions, this makes it easier to decide the threshold line to be used.

374

375 **Fig. 6.** Two-dimensional Bayesian probabilities for occurrence of landslides based on rainfall severity  
376 and soil wetness.

## 377 **5. Discussions**

378 To verify the performance of all models and to understand which model is performing better for the  
379 study area, different thresholds should be compared quantitatively (Lagomarsino et al., 2015). In this  
380 study, empirical thresholds on ED plane, probabilistic thresholds on all three combinations of control  
381 parameters (ED, ID and EI) and also a two-dimensional Bayesian approach by combining empirical  
382 ED thresholds with soil moisture have been derived. The maximum probability value obtained in the  
383 two-dimensional analysis was in the case of EI thresholds, and the value is 0.54. The value was  
384 obtained when the intensity is less than 0.5 mm/h and event rainfall is between 100 to 200 mm. This  
385 implies a prolonged duration of 8 days or more. The intensity value is too low in this case, yet the  
386 probability value is the maximum. The definition of 2-dimensional Bayesian probability majorly  
387 depends upon the relative occurrence of landslides when the rainfall conditions are satisfied and the  
388 occurrence of rainfall events with specified conditions. The number of events with the specified EI  
389 conditions were less, but more than half of them have resulted in landslides based on the historical  
390 data. Thus, the probability of occurrence of landslides is higher in this case. This result points towards  
391 the significance of using a physical parameter such as soil moisture for the definition of threshold.  
392 The top regolith layer throughout the district consists of forest loam, lateritic soil, alluvial soils etc,  
393 with higher fine fraction (Department of Mining and Geology Kerala, 2016). The less permeable soil  
394 has a higher water holding capacity and the moisture content increases when the rainfall is  
395 continuous. The prolonged rainfall has thus reduced the shear strength of soil and the landslide has  
396 happened at a very less intensity value. This complicated process is simplified by using a statistical  
397 approach, by considering the effect of soil wetness. To understand the performance of such a model  
398 with respect to the meteorological thresholds, a quantitative comparison is required.

399 An ROC curve approach was used for quantitative comparison. ROC curve is a tool to understand the  
400 performance of a model with a binary outcome. Each threshold value can forecast two outcomes for a  
401 day; 'landslides' or 'no landslides. If the threshold condition is crossed, the model forecasts

402 'landslides' and otherwise, 'no landslides. When the forecasting is correct and landslide occurs, it is  
403 termed as a true positive ( $TP$ ). Another possibility of correct outcome is the result 'no landslides' on a  
404 day in which landslides do not occur. This can be counted as a true negative ( $TN$ ) result. When the  
405 forecasting goes wrong, it also has two possible outcomes. 'Landslides' forecasted on a non-landslide  
406 day, which is a false positive ( $FP$ ) or simply a false alarm and 'no landslides' forecasted on a day in  
407 which landslides occur, termed as false negatives ( $FN$ ) or missed alarms. A perfect model should only  
408 have true outcomes, without any false alarms or missed alarms.

409 A ROC curve is a plot with a false positive rate of a model on x-axis and a true positive rate on y axis.  
410 It evaluates the overall performance of the model. The true positive rate is also called the sensitivity  
411 of the model. It provides the proportion of landslide occurrences which are correctly identified  
412 ( $TP/(TP + FN)$ ). The specificity of a model is the true negative rate and is the ratio of  $TN$  to the  
413 sum of  $TN$  and  $FP$ . The false positive rate can be calculated by subtracting specificity value from 1.  
414 An ideal model is expected to have both sensitivity and specificity values as 1. Hence the point (0,1)  
415 on ROC curve is called the perfect point. Points which are closer to this perfect point has a better  
416 performance. Also, the model with better performance is the one with a maximum area under the  
417 curve (AUC) among the different models considered. Threat score and True Skill Statistic (TSS) are  
418 two other parameters which were used to understand the performance of a model (Mirus et al.,  
419 2018b). Threat score is defined as the ratio of  $TP$  to the sum of  $TP$ ,  $FP$  and  $FN$ . TSS is the difference  
420 between sensitivity and false positive rate. For an ideal model, the value of both these variables  
421 should be 1.

422 ROC curves for all models considered in the study are plotted in Fig. 7 and the statistical attributes are  
423 listed in Table 2. From Fig. 7, it can be observed that the RS threshold covers the maximum area in  
424 the plane with an AUC of 0.96. The empirical ED threshold has the second highest AUC of 0.86. All  
425 the three probabilistic rainfall thresholds have very close AUC values as observed in Fig. 7. EI  
426 threshold covers a larger area than the other two, indicating its better performance in comparison with  
427 the other two probabilistic rainfall thresholds. The distance from perfect point is minimum in the case  
428 of RS thresholds, in the case of critical probabilities 0.1 and 0.15. It can also be observed that the

429 value of threat score and TSS are maximum in the case of RS thresholds. The maximum value of  
430 threat score is obtained as 0.24 and TSS as 0.90, both in the case of RS thresholds with critical  
431 probability 0.1, which is also the closest one to the perfect point.

432 **Fig. 7.** ROC curves for the derived thresholds. Sensitivity is the ability of a model to correctly identify  
433 the landslide events and Specificity is the ability to correctly identify the non-landslide events

434 Looking into the details in Table 2, it confirms with the literature as the empirical thresholds result in many  
435 false alarms, making it inadequate to use in an LEWS. The number of false alarms can considerably be  
436 reduced by using probabilistic rainfall thresholds, as listed in Table 2. The number of *FP* in the case of  
437 probabilistic ED, ID and EI are much lesser than the other two models considered. But this reduction in  
438 false alarms comes with the cost of a higher number of missed alarms (*FN*). While 171 landslide events  
439 out of the 177 events are correctly forecasted by the empirical ED threshold line  $T_1$ , and 172 are correctly  
440 forecasted by RS threshold when the critical probability is 0.05, the maximum number of correct outcomes  
441 for the other probabilistic models are 106, 105 and 117 on ED, ID and EI planes respectively. For  
442 improving the performance, we need to balance the number of false alarms and missed alarms, which is  
443 achieved by using RS threshold. The RS threshold has *FN* numbers comparable with that of probabilistic  
444 rainfall thresholds, minimising the false alarms and by incorporating an additional filter using soil wetness,  
445 it reduces the number of false alarms when compared to the empirical ED threshold.

446 **Table 2.** Statistical attributes for quantitative comparison.

447

448 The probabilistic rainfall thresholds have high specificity values, pointing to their ability to correctly  
449 forecast the days without landslides, but with very less values of sensitivity. The points on ROC  
450 curves for probabilistic rainfall thresholds are therefore closer to both the axes, reducing the AUC.  
451 Even though the points have high specificity values, they are located far from the perfect point, due to  
452 their inefficiency in correctly forecasting the occurrence of landslides. The RS threshold with a  
453 critical probability of 0.1 is the closest one to the perfect point, correctly forecasting 167 landslide  
454 occurrences.

455 From the analysis, the rainfall and soil wetness conditions for which the probability of occurrence is  
456 more than 0.1 should be considered critical. This makes it easier to identify the empirical ED  
457 threshold line for different values of soil wetness. The critical conditions are mentioned in Table 3.

458 **Table 3.** Critical conditions for initiation of landslides in Idukki, based on RS thresholds

459

460 From Table 3, it can be understood that when soil wetness is less than 0.2,  $T_{50}$  line of empirical ED  
461 thresholds should be considered as critical, when the soil wetness is between 0.2 to 0.4,  $T_5$  can be  
462 considered as the critical threshold. For the next two cases where soil wetness is between 0.4 to 0.8,  
463  $T_{10}$  threshold line can be considered as critical if the critical probability is 0.1. In this case, the  
464 threshold line for 0.2 to 0.4 is  $T_5$ , which is below the threshold line for soil wetness from 0.4 to 0.8.  
465 This variation can be due to the smaller number of data points considered in this study. With the  
466 available data points, very less cases are reported when the soil wetness is between 0.4 to 0.8, and the  
467 rainfall severity is below  $T_{10}$ . To avoid any possible missed alarms due to the limitations of the dataset  
468 considered, the threshold for soil wetness between 0.4 to 0.8 is considered as  $T_1$ , for which the  
469 probability of occurrence of landslides is 0.05 in this study. This variation in the critical probability  
470 ensures the physical validity and easy export of the model. When the soil wetness is between 0.8 to  
471 1, even rainfall which is below  $T_1$  can trigger landslides. Hence, for the last condition, we defined the  
472 critical case as  $T_{\min}$  where  $T_{\min}$  represents the threshold line with minimum exceedance probability,  
473 close to zero. Practically, it represents any possible rainfall condition.

474 The soil wetness data can be collected from daily satellite observations as taken in this study, or from  
475 real-time field observation using sensors. The severity of rainfall for each day can be estimated from  
476 the rainfall forecasts. Using these two inputs, the possibility of occurrence of landslide can be  
477 estimated using the conditions mentioned in Table 3. With higher exceedance probabilities, the  
478 relative uncertainty of  $\alpha$  of ED threshold is crossing this limiting value. Similar results are observed  
479 when the rainfall data used is of daily temporal resolution (Teja et al., 2019).



480 The type of landslides is also an important factor in identifying the associated rainfall. For example,  
481 rockfalls may be triggered without any rainfall, debris flows are often triggered by short duration  
482 (maybe less than 1 hour) and high intensity (Kean et al., 2011), and shallow landslides are triggered  
483 by short-term rainstorms of high-intensity or long-duration rainfall of low to medium intensity  
484 (Guzzetti et al., 2008). This is the main reason why the models (ED, and RS) often associated with the  
485 disadvantage of higher false alarms. Even though the false alarms are considerably reduced in RS  
486 thresholds, it needs further enhancements to be used in an LEWS. There are chances that the model  
487 may miss alarms for rockfalls, which can be triggered with no rainfall. In the case of flow like  
488 landslides such as debris flows, the failure can be triggered by very short, high intensity rainfalls.  
489 Such rainfall events may trigger landslides in relatively dry soils as well. In this case, even if the  
490 antecedent soil wetness is less than 0.2, if the rainfall severity is greater than  $T_{50}$ , the model will issue  
491 a warning. If an event of severity less than  $T_{50}$  triggers such an event, the model may miss the alarm.  
492 With a higher number of data points and better resolution of rainfall data, this can be improved, and  
493 better results can be expected.

## 494 **6. Conclusions**

495 This study has been conducted to evaluate the effect of antecedent soil moisture content to improving  
496 the performance of empirical and probabilistic thresholds for Idukki district in India. The district is  
497 suffering from landslides ranging from cut slope failures to debris flows during monsoon seasons. The  
498 recent disasters that happened in 2018 and 2019 in the district emphasises the requirement of a  
499 landslide early warning system for the region.

500 In this study, empirical rainfall thresholds on ED plane was derived for the study area using an  
501 algorithm-based approach. It was found that with 5% exceedance probability, 20.19 mm rainfall can  
502 trigger a landslide in the region for a duration of 24 hours, and when the duration is 624 hours, a  
503 rainfall of 129 mm can trigger landslides in the region.

504 To evaluate the influence of each rainfall parameter on the occurrence of landslides, Bayesian  
505 analyses were conducted for both one-dimensional and two-dimensional cases. It was found that both

506 intensity and event rainfall have influence on the occurrence of landslides, and most of the events  
507 happened when the rainfall happened in lesser duration. From two-dimensional analysis, the  
508 probabilities on EI plane were found to have the maximum values.

509 To evaluate the effect of soil wetness, another two-dimensional Bayesian approach was conducted,  
510 and it was observed that when the soil is relatively dry, severe rainfall events are required to trigger  
511 landslides and when the soil is wet, also milder rainfall conditions can trigger landslides in the study  
512 area.

513 A statistical comparison between the considered models was used to find out the best performing  
514 model. The comparison was carried out by using a ROC curve approach, where the RS threshold was  
515 found to have the maximum AUC of 0.96, among the models considered in this study. The empirical  
516 ED threshold generated a relevant number of false alarms, resulting in a low specificity value, while  
517 the disadvantage associated with the probabilistic thresholds was the low sensitivity due to a large  
518 number of missed alarms. The proposed method, which combines empirical thresholds with soil  
519 wetness using a probabilistic approach, performs better than both the root models by optimising the  
520 number of false alarms and missed alarms. Based on the comparison, it was found that an RS  
521 threshold of probability 0.1 should be considered critical for the study area and critical rainfall  
522 severity conditions were identified for each soil wetness condition. The performance could be further  
523 enhanced in the future by using hourly rainfall data with more dense rain gauge network for the area.  
524 Moreover, real-time monitoring of moisture content data for different locations in the study area can  
525 also contribute to better resolution soil moisture data and thereby improving the performance of the  
526 model.

527 The results of the study therefore open new promising perspectives for the development of an  
528 operational LEWS in the Idukki district, by combining rainfall and soil moisture data. At the same  
529 time, this work provides evidences from a monsoon region about the advances brought by hydro-  
530 meteorological thresholds based on soil moisture, which is gaining a growing attention in landslide  
531 studies all over the world but before today it was relatively unexplored for the setting of LEWS in the  
532 study area. Unfortunately, the use of soil moisture data in operational LEWS with short lead times is  
533 technically difficult, consequently another option to be explored is the use of antecedent rainfall  
534 conditions as a proxy to the soil moisture, which can be a simpler method to express the soil wetness  
535 conditions (Leonarduzzi and Molnar, 2020; Segoni et al., 2018b). More studies must be conducted for  
536 this region, to develop an effective LEWS which could obtain a fair compromise between the  
537 simplicity of the approach and the quality of the forecasting performance.

## 538 **Acknowledgements**

539 Analyses and visualizations used in this study were produced with the Giovanni online data system,  
540 developed and maintained by the NASA GES DISC. The research was supported by Università degli  
541 Studi di Firenze under the Grant 58517\_INTERNAZIONALIZZAZIONE

542

## 543 **References**

- 544 Abraham, M.T., Pothuraju, D., Satyam, N., 2019. Rainfall Thresholds for Prediction of Landslides in  
545 Idukki, India: An Empirical Approach. *Water* 11, 2113. <https://doi.org/10.3390/w11102113>
- 546 Abraham, M.T., Satyam, N., Kushal, S., Rosi, A., Pradhan, B., Segoni, S., 2020a. Rainfall Threshold  
547 Estimation and Landslide Forecasting for Kalimpong, India Using SIGMA Model. *Water* 12,  
548 1195. <https://doi.org/10.3390/w12041195>
- 549 Abraham, M.T., Satyam, N., Pradhan, B., Alamri, A.M., 2020b. Forecasting of landslides using  
550 rainfall severity and soil wetness: A probabilistic approach for Darjeeling Himalayas. *Water*  
551 (Switzerland) 12, 1–19. <https://doi.org/10.3390/w12030804>

552 Abraham, M.T., Satyam, N., Pradhan, B., Alamri, A.M., 2020c. IoT-Based Geotechnical Monitoring  
553 of Unstable Slopes for Landslide Early Warning in the Darjeeling Himalayas. *Sensors* 20, 2611.  
554 <https://doi.org/10.3390/s20092611>

555 Abraham, M.T., Satyam, N., Reddy, S.K.P., Pradhan, B., 2020d. Runout modeling and calibration of  
556 friction parameters of Kurichermala debris flow, India. *Landslides*.  
557 <https://doi.org/10.1007/s10346-020-01540-1>

558 Abraham, M.T., Satyam, N., Rosi, A., Pradhan, B., Segoni, S., 2020e. The Selection of Rain Gauges  
559 and Rainfall Parameters in Estimating Intensity-Duration Thresholds for Landslide Occurrence:  
560 Case Study from Wayanad (India). *Water* 12, 1000. <https://doi.org/10.3390/w12041000>

561 Agostini, A., Tofani, V., Nolesini, T., Gigli, G., Tanteri, L., Rosi, A., Cardellini, S., Casagli, N., 2014.  
562 A new appraisal of the Ancona landslide based on geotechnical investigations and stability  
563 modelling. *Q. J. Eng. Geol. Hydrogeol.* 47, 29–43. <https://doi.org/10.1144/qjegh2013-028>

564 Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. *Eng. Geol.* 73, 247–265.  
565 <https://doi.org/10.1016/j.enggeo.2004.01.007>

566 Alimohammadlou, Y., Najafi, A., Gokceoglu, C., 2014. Estimation of rainfall-induced landslides  
567 using ANN and fuzzy clustering methods: A case study in Saenen Slope, Azerbaijan province,  
568 Iran. *Catena* 120, 149–162. <https://doi.org/10.1016/j.catena.2014.04.009>

569 Althuwaynee, O.F., Pradhan, B., 2017. Semi-quantitative landslide risk assessment using GIS-based  
570 exposure analysis in Kuala Lumpur City. *Geomatics, Nat. Hazards Risk* 8, 706–732.  
571 <https://doi.org/10.1080/19475705.2016.1255670>

572 Althuwaynee, Omar, F., Pradhan, B., Ahmad, N., 2015. Estimation of rainfall threshold and its use in  
573 landslide hazard mapping of Kuala Lumpur metropolitan and surrounding areas. *Landslides* 12,  
574 861–875.

575 Alvioli, M., Melillo, M., Guzzetti, F., Rossi, M., Palazzi, E., von Hardenberg, J., Brunetti, M.T.,  
576 Peruccacci, S., 2018. Implications of climate change on landslide hazard in Central Italy. *Sci.*

577 Total Environ. 630, 1528–1543. <https://doi.org/10.1016/j.scitotenv.2018.02.315>

578 Battistini, A., Rosi, A., Segoni, S., Lagomarsino, D., Catani, F., Casagli, N., 2017. Validation of  
579 landslide hazard models using a semantic engine on online news. *Appl. Geogr.* 82, 59–65.  
580 <https://doi.org/10.1016/j.apgeog.2017.03.003>

581 Baum, R.L., Savage, W.Z., Godt, J.W., 2008. TRIGRS — A Fortran Program for Transient Rainfall  
582 Infiltration and Grid-Based Regional Slope Stability Analysis.

583 Berti, M., Martina, M.L.V., Franceschini, S., Pignone, S., Simoni, A., Pizziolo, M., 2012.  
584 Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *J. Geophys.*  
585 *Res. Earth Surf.* 117, 1–20. <https://doi.org/10.1029/2012JF002367>

586 Bicocchi, G., Tofani, V., D'Ambrosio, M., Tacconi-Stefanelli, C., Vannocci, P., Casagli, N., Lavorini,  
587 G., Trevisani, M., Catani, F., 2019. Geotechnical and hydrological characterization of hillslope  
588 deposits for regional landslide prediction modeling. *Bull. Eng. Geol. Environ.* 78, 4875–4891.  
589 <https://doi.org/10.1007/s10064-018-01449-z>

590 Bogaard, T., Greco, R., 2018. Invited perspectives: Hydrological perspectives on precipitation  
591 intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds.  
592 *Nat. Hazards Earth Syst. Sci.* 18, 31–39. <https://doi.org/10.5194/nhess-18-31-2018>

593 Brunetti, M.T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., Guzzetti, F., 2010. Rainfall  
594 thresholds for the possible occurrence of landslides in Italy. *Nat. Hazards Earth Syst. Sci.* 10,  
595 447–458. <https://doi.org/10.5194/nhess-10-447-2010>

596 Caine, N., 1980. The rainfall intensity-duration control of shallow landslides and debris flows: An  
597 update. *Geogr. Ann. Ser. A, Phys. Geogr.* 62, 1–2, 23–27.

598 CartoDEM, 2015. CartoDEM : a national digital elevation model from Cartosat-1 stereo data. *Natl.*  
599 *Remote Sens. Centre, Hyderabad, Dep. Space, Gov. India.*

600 Chae, B.G., Park, H.J., Catani, F., Simoni, A., Berti, M., 2017. Landslide prediction, monitoring and  
601 early warning: a concise review of state-of-the-art. *Geosci. J.* 21, 1033–1070.

602 <https://doi.org/10.1007/s12303-017-0034-4>

603 Chen, C.-W., Tung, Y.-S., Liou, J.-J., Li, H.-C., Cheng, C.-T., Chen, Y.-M., Oguchi, T., 2019.

604 Assessing landslide characteristics in a changing climate in northern Taiwan. *CATENA* 175,

605 263–277. <https://doi.org/10.1016/j.catena.2018.12.023>

606 de Jeu, R. (Vrije U.A., Owe, M. (NASA G., 2014. AMSR2/GCOM-W1 surface soil moisture (LPRM)

607 L3 1 day 25 km x 25 km descending V001, Edited by Goddard Earth Sciences Data and

608 Information Services Center (GES DISC) (Bill Teng), Greenbelt, MD, USA, Goddard Earth

609 Sciences Data and Information Services Cente. <https://doi.org/10.5067/CGDEOBASZ178>

610 de Jeu, R. (Vrije U.A., Owe, M. (NASA G., 2012. TMI/TRMM surface soil moisture (LPRM) L3 1

611 day 25 km x 25 km nighttime V001, Edited by Goddard Earth Sciences Data and Information

612 Services Center (GES DISC) (Bill Teng), Greenbelt, MD, USA, Goddard Earth Sciences Data

613 and Information Services Center (GES. <https://doi.org/10.5067/GWHRZEL8SA21>

614 Department of Mining and Geology Kerala, 2016. District Survey Report of Minor Minerals.

615 Thiruvananthapuram.

616 Dikshit, A., Satyam, D.N., 2018. Estimation of rainfall thresholds for landslide occurrences in

617 Kalimpong, India. *Innov. Infrastruct. Solut.* 3. <https://doi.org/10.1007/s41062-018-0132-9>

618 Dikshit, A., Satyam, D.N., Towhata, I., 2018. Early warning system using tilt sensors in Chibo,

619 Kalimpong, Darjeeling Himalayas, India. *Nat. Hazards* 94, 727–741.

620 <https://doi.org/10.1007/s11069-018-3417-6>

621 Dikshit, A., Satyam, N., 2019. Probabilistic rainfall thresholds in Chibo, India: estimation and

622 validation using monitoring system. *J. Mt. Sci.* 16, 870–883. [https://doi.org/10.1007/s11629-](https://doi.org/10.1007/s11629-018-5189-6)

623 [018-5189-6](https://doi.org/10.1007/s11629-018-5189-6)

624 Gadgil, M., Krishnan, B.J., Ganeshiah, K.N., Vijayan., V.S., Borges, R., Sukumar, R., Noronha, L.,

625 Nayak, V.S., Subramaniam, D.K., Varma, R.V., Gautam, S.P., Navalgund, R.R.,

626 Subrahmanyam, G.V., 2011. Report of the Western Ghats Ecology Expert Panel (WGEEP).

627 Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. *Earth-Science Rev.* 162, 227–252.  
628 <https://doi.org/10.1016/j.earscirev.2016.08.011>

629 Gariano, S.L., Melillo, M., Peruccacci, S., Brunetti, M.T., 2020. How much does the rainfall temporal  
630 resolution affect rainfall thresholds for landslide triggering? *Nat. Hazards* 100, 655–670.  
631 <https://doi.org/10.1007/s11069-019-03830-x>

632 Giovanni, 2020. NASA GES DISC [WWW Document].

633 Glade, T., Crozier, M., Smith, P., 2000. Applying probability determination to refine landslide-  
634 triggering rainfall thresholds using an empirical “Antecedent Daily Rainfall Model.” *Pure Appl.*  
635 *Geophys.* 157, 1059–1079. <https://doi.org/10.1007/s000240050017>

636 Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2008. The rainfall intensity-duration control of  
637 shallow landslides and debris flows: An update. *Landslides* 5, 3–17.  
638 <https://doi.org/10.1007/s10346-007-0112-1>

639 Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2007. Rainfall thresholds for the initiation of  
640 landslides in central and southern Europe. *Meteorol. Atmos. Phys.* 98, 239–267.  
641 <https://doi.org/10.1007/s00703-007-0262-7>

642 Iida, T., 1999. A stochastic hydro-geomorphological model for shallow landsliding due to rainstorm.  
643 *Catena* 34, 293–313. [https://doi.org/10.1016/S0341-8162\(98\)00093-9](https://doi.org/10.1016/S0341-8162(98)00093-9)

644 India Meteorological Department, 2019. India Meteorological Department (IMD) Data Supply Portal  
645 [WWW Document].

646 Iverson, R.M., 2000. Landslide triggering by rain infiltration. *Water Resour. Res.* 36, 1897–1910.  
647 <https://doi.org/10.1029/2000WR900090>

648 Jakob, M., Holm, K., Lange, O., Schwab, J.W., 2006. Hydrometeorological thresholds for landslide  
649 initiation and forest operation shutdowns on the north coast of British Columbia. *Landslides* 3,  
650 228–238. <https://doi.org/10.1007/s10346-006-0044-1>

651 Kean, J.W., Staley, D.M., Cannon, S.H., 2011. In situ measurements of post - fire debris flows in  
652 southern California : Comparisons of the timing and magnitude of 24 debris - flow events with  
653 rainfall and soil moisture conditions. *J. Geophys. Res.* 116, 1–21.  
654 <https://doi.org/10.1029/2011JF002005>

655 Keefer, D.K., Wilson, R.C., Mark, R.K., Brabb, E.E., Iii, W.M.B., Ellen, S.D., Harp, E.L., Wieczorek,  
656 G.F., Alger, C.S., Zatkint, R.S., 1987. Real-Time Landslide Warning During Heavy Rainfall.  
657 *Science* (80-. ). 238, 921–925.

658 Kim, M.S., Onda, Y., Uchida, T., Kim, J.K., Song, Y.S., 2018. Effect of seepage on shallow  
659 landslides in consideration of changes in topography: Case study including an experimental  
660 sandy slope with artificial rainfall. *Catena* 161, 50–62.  
661 <https://doi.org/10.1016/j.catena.2017.10.004>

662 Kirschbaum, D.B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C., Lerner-Lam, A., 2012.  
663 Advances in landslide nowcasting: Evaluation of a global and regional modeling approach.  
664 *Environ. Earth Sci.* 66, 1683–1696. <https://doi.org/10.1007/s12665-011-0990-3>

665 Kuriakose, S.L., Devkota, S., Rossiter, D.G., Jetten, V.G., 2009a. Prediction of soil depth using  
666 environmental variables in an anthropogenic landscape, a case study in the Western Ghats of  
667 Kerala, India. *Catena* 79, 27–38. <https://doi.org/10.1016/j.catena.2009.05.005>

668 Kuriakose, S.L., Sankar, G., Muraleedharan, C., 2009b. History of landslide susceptibility and a  
669 chorology of landslide-prone areas in the Western Ghats of Kerala, India. *Environ. Geol.* 57,  
670 1553–1568. <https://doi.org/10.1007/s00254-008-1431-9>

671 Lagomarsino, D., Segoni, S., Rosi, A., Rossi, G., Battistini, A., Catani, F., Casagli, N., 2015.  
672 Quantitative comparison between two different methodologies to define rainfall thresholds for  
673 landslide forecasting. *Nat. Hazards Earth Syst. Sci.* 15, 2413–2423.  
674 <https://doi.org/10.5194/nhess-15-2413-2015>

675 Lainas, S., Sabatakakis, N., Koukis, G., 2016. Rainfall thresholds for possible landslide initiation in



676 wildfire-affected areas of western Greece. *Bull. Eng. Geol. Environ.* 75, 883–896.  
677 <https://doi.org/10.1007/s10064-015-0762-5>

678 Leonarduzzi, E., Molnar, P., 2020. Data limitations and potential of hourly and daily rainfall  
679 thresholds for shallow landslides. *Nat. Hazards Earth Syst. Sci. Discuss.* 1–25.  
680 <https://doi.org/10.5194/nhess-2020-125>

681 Marra, F., Destro, E., Nikolopoulos, E.I., Zoccatelli, D., Dominique Creutin, J., Guzzetti, F., Borga,  
682 M., 2017. Impact of rainfall spatial aggregation on the identification of debris flow occurrence  
683 thresholds. *Hydrol. Earth Syst. Sci.* 21, 4525–4532. <https://doi.org/10.5194/hess-21-4525-2017>

684 Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Guzzetti, F., 2016. Rainfall thresholds for  
685 the possible landslide occurrence in Sicily (Southern Italy) based on the automatic  
686 reconstruction of rainfall events. *Landslides* 13, 165–172. [https://doi.org/10.1007/s10346-015-](https://doi.org/10.1007/s10346-015-0630-1)  
687 [0630-1](https://doi.org/10.1007/s10346-015-0630-1)

688 Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Guzzetti, F., 2014. An Algorithm for the  
689 objective reconstruction of rainfall events responsible for landslides. *Landslide Dyn. ISDR-ICL*  
690 *Landslide Interact. Teach. Tools Vol. 1 Fundam. Mapp. Monit.* 12, 311–320.  
691 [https://doi.org/10.1007/978-3-319-57774-6\\_33](https://doi.org/10.1007/978-3-319-57774-6_33)

692 Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Roccati, A., Guzzetti, F., 2018. A tool for  
693 the automatic calculation of rainfall thresholds for landslide occurrence. *Environ. Model. Softw.*  
694 105, 230–243. <https://doi.org/10.1016/j.envsoft.2018.03.024>

695 Mirus, B.B., Becker, R.E., Baum, R.L., Smith, J.B., 2018a. Integrating real-time subsurface  
696 hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning.  
697 *Landslides* 15, 1909–1919. <https://doi.org/10.1007/s10346-018-0995-z>

698 Mirus, B.B., Mopthew, M.D., Smith, J.B., 2018b. Developing hydro-meteorological thresholds for  
699 shallow landslide initiation and early warning. *Water (Switzerland)* 10, 1–19.  
700 <https://doi.org/10.3390/W10091274>

701 Nikolopoulos, E.I., Crema, S., Marchi, L., Marra, F., Guzzetti, F., Borga, M., 2014. Impact of  
702 uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow  
703 occurrence. *Geomorphology* 221, 286–297. <https://doi.org/10.1016/j.geomorph.2014.06.015>

704 Owe, M., de Jeu, R., Holmes, T., 2008. Multisensor historical climatology of satellite-derived global  
705 land surface moisture. *J. Geophys. Res. Earth Surf.* 113, 1–17.  
706 <https://doi.org/10.1029/2007JF000769>

707 Peruccacci, S., Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., Guzzetti, F., 2017. Rainfall  
708 thresholds for possible landslide occurrence in Italy. *Geomorphology* 290, 39–57.  
709 <https://doi.org/10.1016/j.geomorph.2017.03.031>

710 Peruccacci, S., Brunetti, M.T., Luciani, S., Vennari, C., Guzzetti, F., 2012. Lithological and seasonal  
711 control on rainfall thresholds for the possible initiation of landslides in central Italy.  
712 *Geomorphology* 139–140, 79–90. <https://doi.org/10.1016/j.geomorph.2011.10.005>

713 Piciullo, L., Calvello, M., Cepeda, J.M., 2018. Territorial early warning systems for rainfall-induced  
714 landslides. *Earth-Science Rev.* 179, 228–247. <https://doi.org/10.1016/j.earscirev.2018.02.013>

715 Ponziani, F., Pandolfo, C., Stelluti, M., Berni, N., Brocca, L., Moramarco, T., 2012. Assessment of  
716 rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in  
717 the Umbria region (central Italy). *Landslides* 9, 229–237. [https://doi.org/10.1007/s10346-011-](https://doi.org/10.1007/s10346-011-0287-3)  
718 [0287-3](https://doi.org/10.1007/s10346-011-0287-3)

719 Sajeev, R., Praveen, K.R., 2014. Landslide Susceptibility Mapping on Macroscale along the Major  
720 Road Corridors in Idukki District, Kerala. Thiruvananthapuram, India.

721 Segoni, S., Piciullo, L., Gariano, S.L., 2018a. A review of the recent literature on rainfall thresholds  
722 for landslide occurrence. *Landslides* 15, 1483–1501. <https://doi.org/10.1007/s10346-018-0966-4>

723 Segoni, S., Rosi, A., Fanti, R., Gallucci, A., Monni, A., Casagli, N., 2018b. A regional-scale landslide  
724 warning system based on 20 years of operational experience. *Water (Switzerland)* 10, 1–17.  
725 <https://doi.org/10.3390/w10101297>

726 Segoni, S., Rosi, A., Lagomarsino, D., Fanti, R., Casagli, N., 2018c. Brief communication: Using  
727 averaged soil moisture estimates to improve the performances of a regional-scale landslide early  
728 warning system. *Nat. Hazards Earth Syst. Sci.* 18, 807–812. [https://doi.org/10.5194/nhess-18-](https://doi.org/10.5194/nhess-18-807-2018)  
729 807-2018

730 Song, S., Wang, W., 2019. Impacts of antecedent soil moisture on the rainfall- runoff transformation  
731 process based on high- resolution observations in soil tank experiments. *Water (Switzerland)* 11,  
732 15–20. <https://doi.org/10.3390/w11020296>

733 Teja, T.S., Dikshit, A., Satyam, N., 2019. Determination of rainfall thresholds for landslide prediction  
734 using an algorithm-based approach: Case study in the Darjeeling Himalayas, India. *Geosci.* 9,  
735 <https://doi.org/10.3390/geosciences9070302>

736 Terlien, M.T.J., 1998. The determination of statistical and deterministic hydrological landslide-  
737 triggering thresholds. *Environ. Geol.* 35, 124–130. <https://doi.org/10.1007/s002540050299>

738 Tofani, V., Bicocchi, G., Rossi, G., Segoni, S., D'Ambrosio, M., Casagli, N., Catani, F., 2017. Soil  
739 characterization for shallow landslides modeling: a case study in the Northern Apennines  
740 (Central Italy). *Landslides* 14, 755–770. <https://doi.org/10.1007/s10346-017-0809-8>

741 Uchimura, T., Towhata, I., Anh, T.T.L., Fukuda, J., Bautista, C.J.B., Wang, L., Seko, I., Uchida, T.,  
742 Matsuoka, A., Ito, Y., Onda, Y., Iwagami, S., Kim, M.S., Sakai, N., 2010. Simple monitoring  
743 method for precaution of landslides watching tilting and water contents on slopes surface.  
744 *Landslides* 7, 351–357. <https://doi.org/10.1007/s10346-009-0178-z>

745 Uchimura, T., Towhata, I., Wang, L., Nishie, S., Yamaguchi, H., Seko, I., Qiao, J., 2015. Precaution  
746 and early warning of surface failure of slopes using tilt sensors. *Soils Found.* 55, 1086–1099.  
747 <https://doi.org/10.1016/j.sandf.2015.09.010>

748 United Nations Development Programme, 2018. Kerala Post Disaster Needs Assessment Floods and  
749 Landslides-August 2018. Thiruvananthapuram, India.

750 Valenzuela, P., Domínguez-Cuesta, M.J., Mora García, M.A., Jiménez-Sánchez, M., 2018. Rainfall

751 thresholds for the triggering of landslides considering previous soil moisture conditions  
752 (Asturias, NW Spain). *Landslides* 15, 273–282. <https://doi.org/10.1007/s10346-017-0878-8>

753 Varnes, D., 1978. Slope Movement Types and Processes. Transp. Res. Board Spec. Rep.

754 Wei, X., Fan, W., Cao, Y., Chai, X., Bordoni, M., Meisina, C., Li, J., 2020. Integrated experiments on  
755 field monitoring and hydro-mechanical modeling for determination of a triggering threshold of  
756 rainfall-induced shallow landslides. A case study in Ren River catchment, China. *Bull. Eng.  
757 Geol. Environ.* 79, 513–532. <https://doi.org/10.1007/s10064-019-01570-7>

758 Wicki, A., Lehmann, P., Hauck, C., Seneviratne, S.I., Waldner, P., Stähli, M., 2020. Assessing the  
759 potential of soil moisture measurements for regional landslide early warning. *Landslides* 17,  
760 1881–1896. <https://doi.org/10.1007/s10346-020-01400-y>

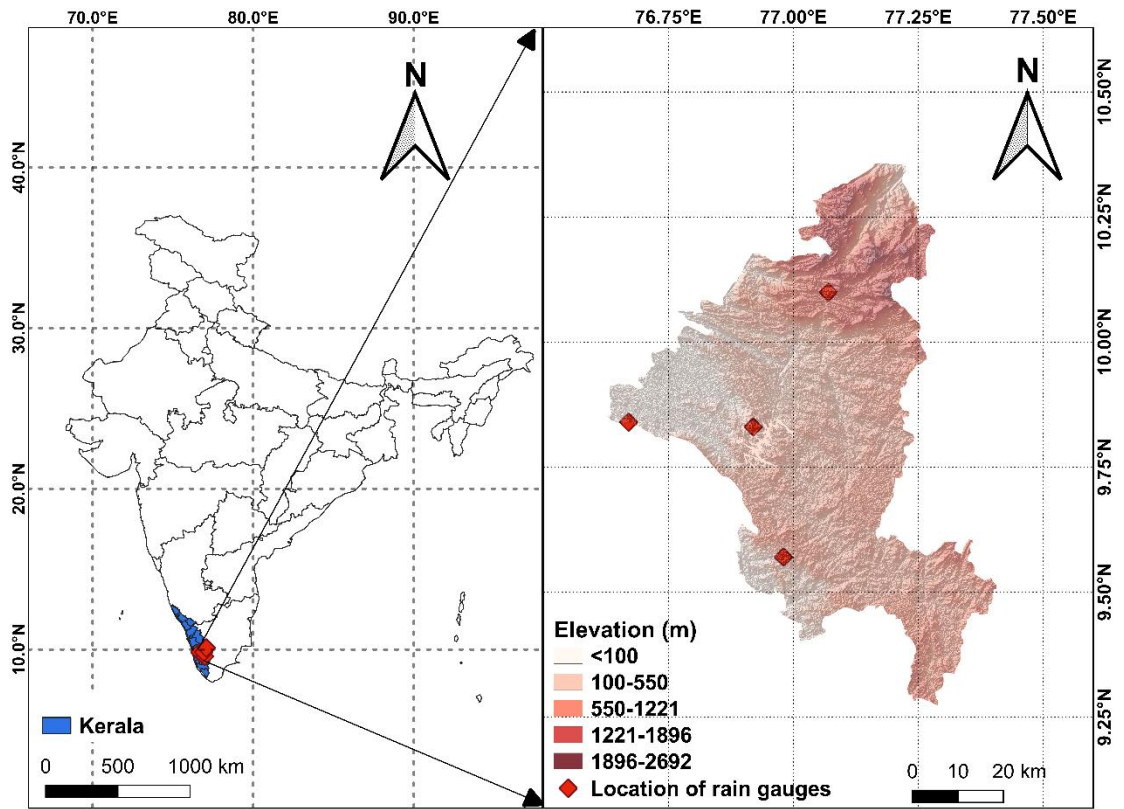
761 Wu, M.H., Wang, J.P., Chen, I.C., 2019. Optimization approach for determining rainfall duration-  
762 intensity thresholds for debris flow forecasting. *Bull. Eng. Geol. Environ.* 78, 2495–2501.  
763 <https://doi.org/10.1007/s10064-018-1314-6>

764 Yang, Z., Cai, H., Shao, W., Huang, D., Uchimura, T., Lei, X., Tian, H., Qiao, J., 2019. Clarifying the  
765 hydrological mechanisms and thresholds for rainfall-induced landslide: in situ monitoring of big  
766 data to unsaturated slope stability analysis. *Bull. Eng. Geol. Environ.* 78, 2139–2150.  
767 <https://doi.org/10.1007/s10064-018-1295-5>

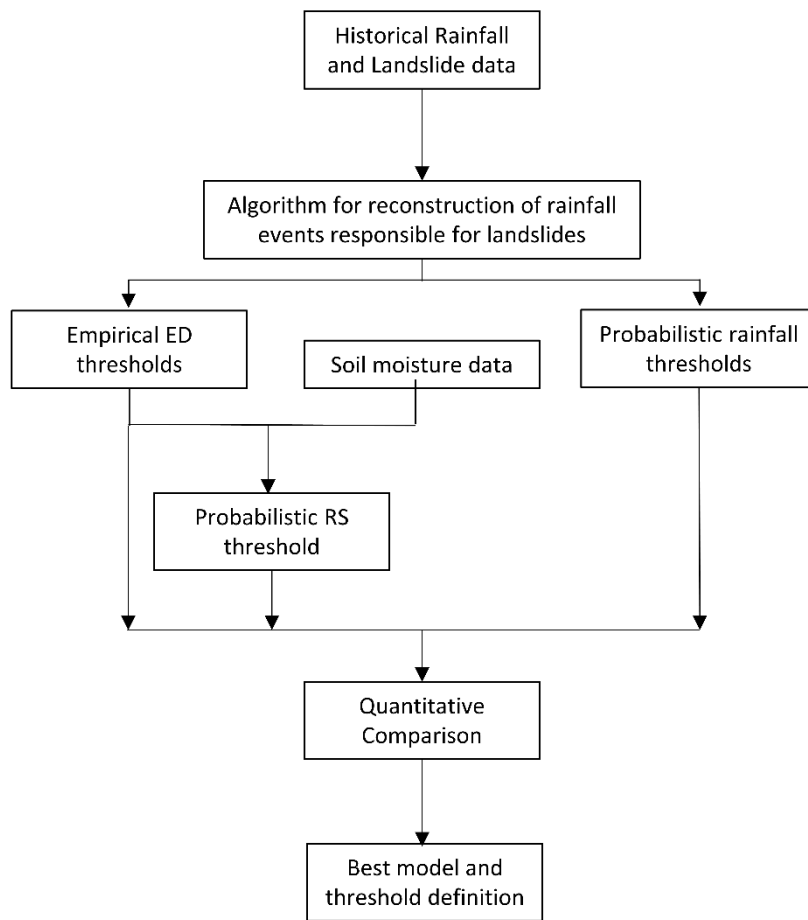
768 Zhao, B., Dai, Q., Han, D., Dai, H., Mao, J., Zhuo, L., 2019a. Probabilistic thresholds for landslides  
769 warning by integrating soil moisture conditions with rainfall thresholds. *J. Hydrol.* 574, 276–  
770 287. <https://doi.org/10.1016/j.jhydrol.2019.04.062>

771 Zhao, B., Dai, Q., Han, D., Dai, H., Mao, J., Zhuo, L., Rong, G., 2019b. Estimation of soil moisture  
772 using modified antecedent precipitation index with application in landslide predictions.  
773 *Landslides* 16, 2381–2393. <https://doi.org/10.1007/s10346-019-01255-y>

774



**Fig. 1.** Location details of study area. (a) India, and (b) Digital Elevation Model of Idukki (modified using CartoDEM (CartoDEM, 2015)) along with location of rain gauges.



**Fig. 2.** Methodology of study.

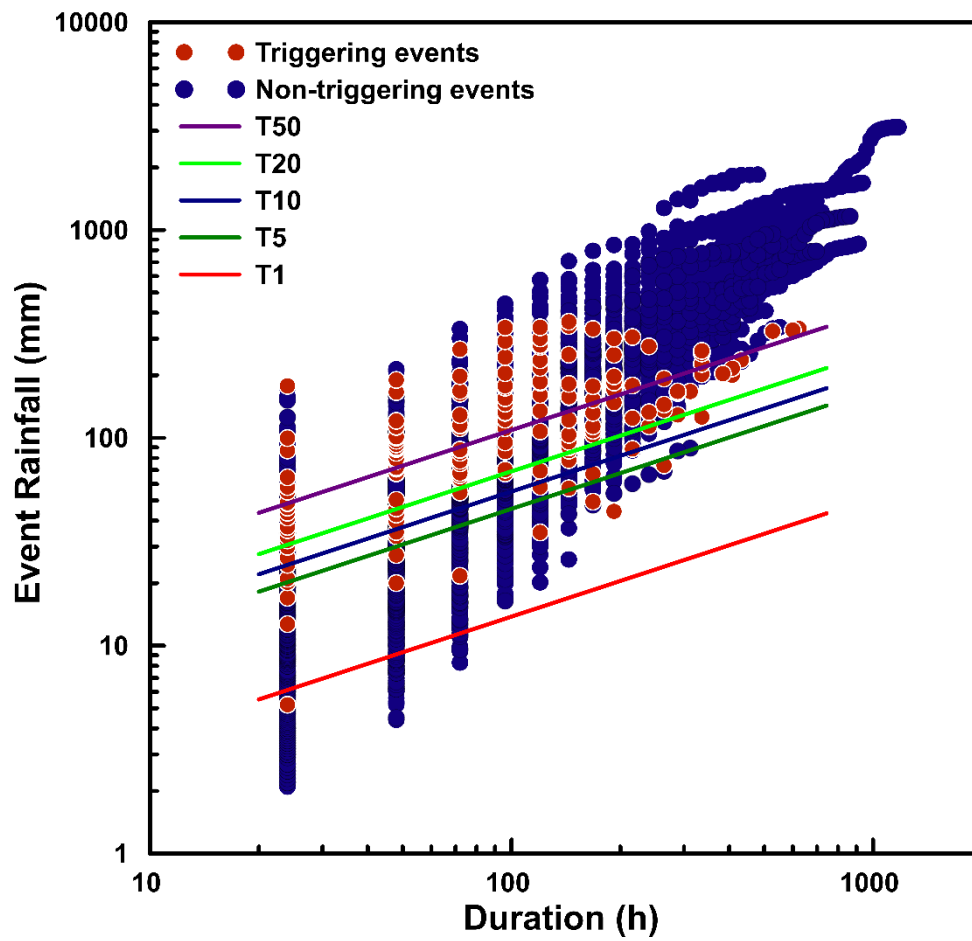
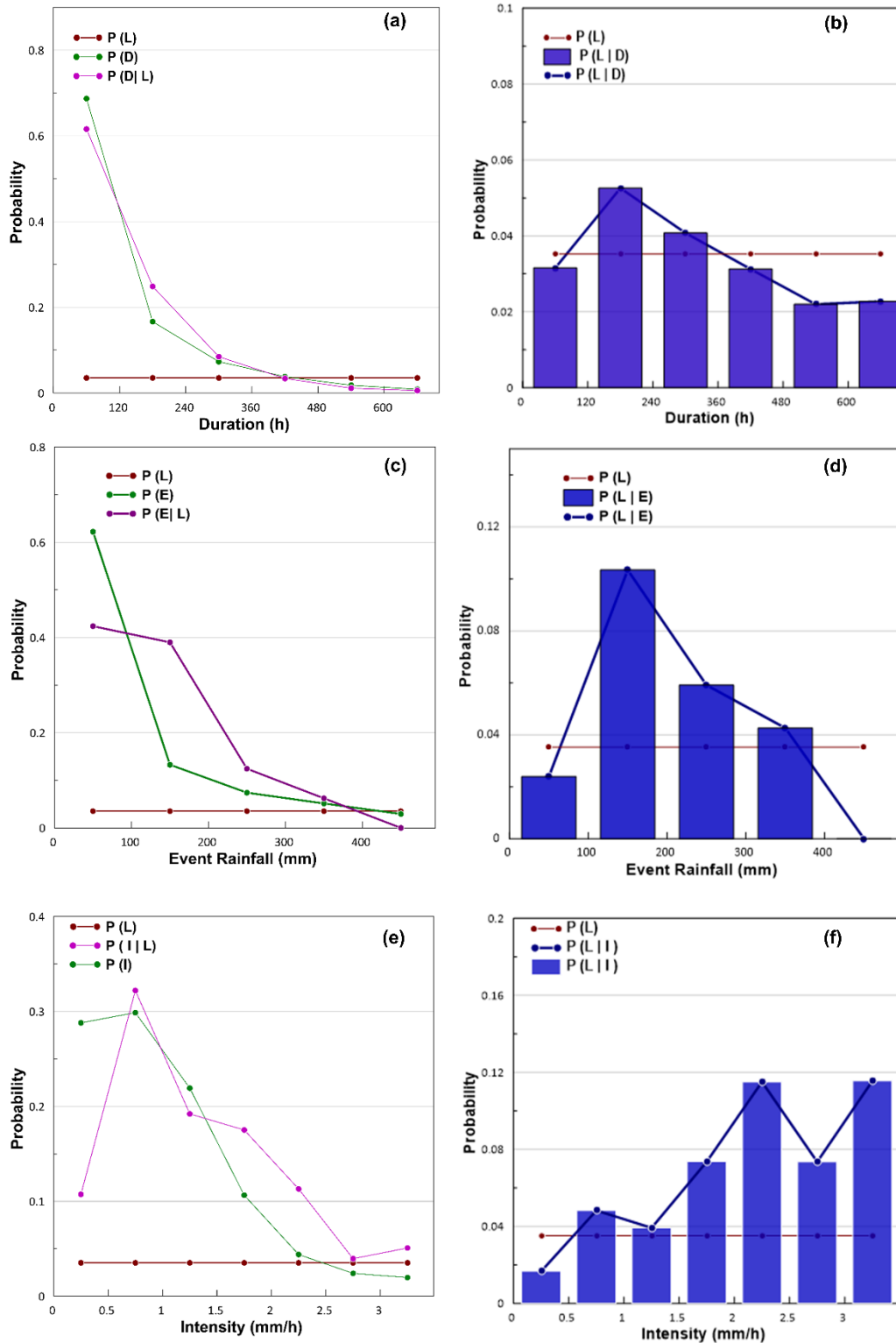
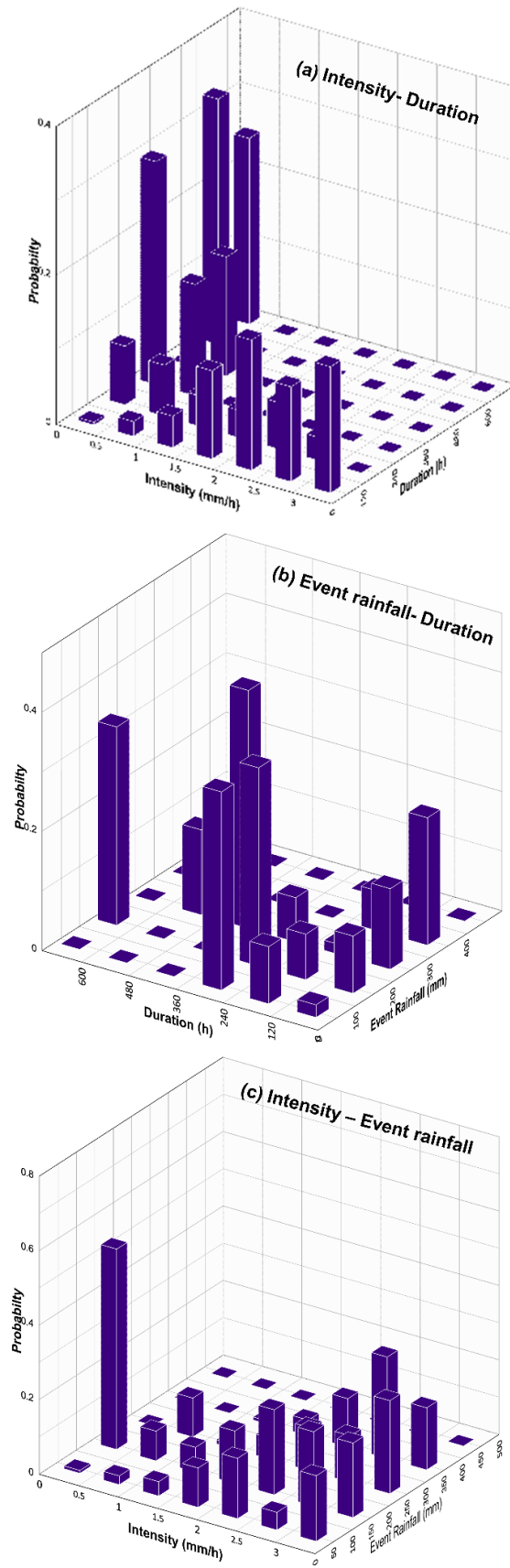


Fig. 3. Rainfall event – duration thresholds for Idukki district.



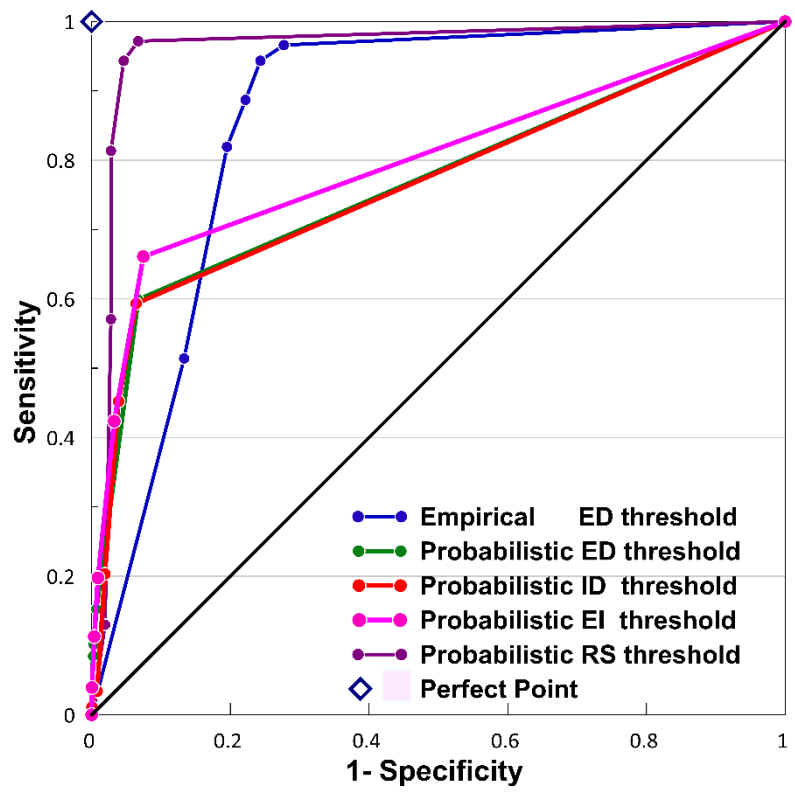
**Fig. 4.** Prior, conditional, marginal and posterior probabilities with respect to rainfall parameters. (a, b) Duration; (c, d) Event rainfall; and (e, f) Intensity.





**Fig. 5.** Two-dimensional posterior probabilities of occurrence of landslide on (a) ID plane, (b) ED plane, and (c) EI plane.





**Fig. 7.** ROC curves for the derived thresholds. Sensitivity is the ability of a model to correctly identify the landslide events and Specificity is the ability to correctly identify the non-landslide events

### Figure captions

**Fig. 1.** Location details of study area. (a) India, and (b) Digital Elevation Model of Idukki (modified using CartoDEM (National Remote Sensing Centre 2015)) along with location of rain gauges.

**Fig. 2.** Methodology of study.

**Fig. 3.** Rainfall event – duration thresholds for Idukki district.

**Fig. 4.** Prior, conditional, marginal and posterior probabilities with respect to rainfall parameters. (a, b) Duration; (c, d) Event rainfall; and (e, f) Intensity.

**Fig. 5.** Two-dimensional posterior probabilities of occurrence of landslide on (a) ID plane, (b) ED plane, and (c) EI plane.

**Fig. 6.** Two dimensional Bayesian probabilities for occurrence of landslides based on rainfall severity and soil wetness.

**Fig. 7.** ROC curves for the derived thresholds. Sensitivity is the ability of a model to correctly identify the landslide events and Specificity is the ability to correctly identify the non-landslide events

**Table 1.** Values of  $\alpha$ ,  $\gamma$  and the uncertainties associated with different exceedance probabilities

<b>Threshold</b>	<b><math>\alpha</math></b>	<b><math>\Delta\alpha</math></b>	<b><math>\gamma</math></b>	<b><math>\Delta\gamma</math></b>
T <sub>1</sub>	2.3	0.8	0.57	0.03
T <sub>5</sub>	3.3	1.1	0.57	0.03
T <sub>10</sub>	4.0	1.3	0.57	0.03
T <sub>20</sub>	5.0	1.6	0.57	0.03
T <sub>50</sub>	7.9	2.4	0.57	0.03

**Table 2.** Statistical attributes for quantitative comparison.

Thresh old model	Threshold value	TP	FP	FN	TN	Sensi tivity	Speci ficity	Distance			AUC
								from perfect point	Thre at score	True skill statistic	
Empiric al ED	T1	171	3594	6	9377	0.97	0.72	0.28	0.05	0.69	0.86
	T5	167	3156	10	9815	0.94	0.76	0.25	0.05	0.70	
	T10	157	2878	20	10093	0.89	0.78	0.25	0.05	0.67	
	T20	145	2531	32	10440	0.82	0.80	0.27	0.05	0.62	
	T50	91	1729	86	11242	0.51	0.87	0.50	0.05	0.38	
Probabil istic ED	0.05	106	870	71	12101	0.60	0.93	0.41	0.10	0.53	0.77
	0.1	27	86	150	12885	0.15	0.99	0.85	0.10	0.15	
	0.15	18	36	159	12935	0.10	1.00	0.90	0.08	0.10	
	0.2	18	36	159	12935	0.10	1.00	0.90	0.08	0.10	
	0.3	15	25	162	12946	0.08	1.00	0.92	0.07	0.08	
Probabil istic ID	0.05	105	830	72	12141	0.59	0.94	0.41	0.10	0.53	0.77
	0.1	80	502	97	12469	0.45	0.96	0.55	0.12	0.41	
	0.15	36	247	141	12724	0.20	0.98	0.80	0.08	0.18	
	0.2	6	102	171	12869	0.03	0.99	0.97	0.02	0.03	
	0.3	2	6	175	12965	0.01	1.00	0.99	0.01	0.01	
Probabil istic EI	0.05	117	966	60	12005	0.66	0.93	0.35	0.10	0.59	0.79
	0.1	75	421	102	12550	0.42	0.97	0.58	0.13	0.39	
	0.15	35	117	142	12854	0.20	0.99	0.80	0.12	0.19	
	0.2	20	47	157	12924	0.11	1.00	0.89	0.09	0.11	
	0.3	7	6	170	12965	0.04	1.00	0.96	0.04	0.04	
Probabil istic RS	0.05	172	3133	5	9838	0.97	0.76	0.24	0.05	0.73	0.96
	0.1	167	527	10	12444	0.94	0.96	0.07	0.24	0.90	
	0.15	144	477	33	12494	0.81	0.96	0.19	0.22	0.78	
	0.2	101	470	76	12501	0.57	0.96	0.43	0.16	0.53	
	0.3	23	98	154	12873	0.13	0.99	0.87	0.08	0.12	

**Table 3.** Critical conditions for initiation of landslides in Idukki, based on RS thresholds.

Soil Wetness	Critical ED threshold line
0.0 – 0.2	$T_{50}$
0.2 – 0.4	$T_5$
0.4 – 0.6	$\underline{T_{40}T_1}$
0.6 – 0.8	$\underline{T_{40}T_1}$
0.8 – 1.0	$T_{\min}$

**Declaration of interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: