Forecasting landslides using SIGMA model: a case study from Idukki, India

Minu Treesa Abraham[1], Neelima Satyam[1], Nakshatram Shreyas[1], Biswajeet Pradhan[2,3,4*], Samuele Segoni[5], Khairul Nizam Abdul Maulud[4,6], Abdullah M. Alamri[7]

[1]Department of Civil Engineering, Indian Institute of Technology Indore, Madhya Pradesh, India

[2]Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia

[3]Department of Energy and Mineral Resources Engineering, Sejong University, Choongmu-gwan, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Korea

[4]Earth Observation Centre, Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia

[5]Department of Earth Sciences, University of Florence, Via Giorgio La Pira, 4, 50121 Florence, Italy

[6]Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

[7]Department of Geology and Geophysics, College of Science, King Saud University, Riyadh 11362, Saudi Arabia

*Correspondence: Biswajeet.Pradhan@uts.edu.au or biswajeet24@gmail.com

Abstract

This study proposes a regional landslide early warning system for Idukki (India), using a decisional algorithm. The algorithm forecasts the possibility of occurrence of landslide by comparing the rainfall thresholds with the cumulated rainfall values. The region has suffered severe socio-economical setbacks during the disastrous landslides that happened in 2018 and 2019. Rainfall thresholds are defined for Idukki, using the total amount of precipitation cumulated at different time intervals ranging from 1 to 30 days. The first three-day cumulative values were used for evaluating the effect of short-term rainfall and the remaining days for the effect of long-term rainfall. The derived thresholds were calibrated using historical landslides and rainfall data from 2009-2017, optimised to reduce the false alarms and then validated using the 2018 data. The validation results show that the model is effectively predicting 79% of the landslides that happened in the region during 2018 and can be easily integrated with a rainfall forecasting system for the prediction of landslides. The model can be further improved with the availability of better spatial and temporal resolution of rainfall data and can be used as an effective tool for predicting the occurrence of landslides.

Keywords: LEWS; landslides; rainfall thresholds; SIGMA; Idukki

1. Introduction

Landslides are frequent natural disasters that have severe effects on lives and properties in hilly terrains (Muhammad et al. 2010; Abd Majid and Rainis 2019). Climate change and associated extreme weather conditions result in a surge of natural disasters across the world (Easterling et al. 2000; Morss et al. 2011). In regions where rainfall is the primary triggering mechanism for landslides, prediction of occurrence of landslides is often associated with a rainfall threshold condition beyond which landslides are likely to occur (Guzzetti et al. 2008; Sharir et al. 2017). The threshold defines a critical condition beyond which landslides may occur in the region. The condition can be defined based on physical parameters or statistical analysis, and can be used for providing early warning (Gian et al. 2017; Bordoni et al. 2020). The physically based models make use of rainfall infiltration models and slope stability analysis, to precisely calculate the factor of safety of each cell considered for the analysis (Baum and Godt 2010). Such models are more suitable for site specific or local scale slope stability studies as they require physical parameters as inputs and need detailed field-based data collection process. Even if these approaches are less widespread, recent studies show that they can produce reliable results also at large scale (Fusco et al. 2019; Bordoni et al. 2019). Empirical or statistical models are mostly followed for regional and global scale studies, due to their simplicity and easy exportability. The conventional rainfall thresholds consider the short-term effect of rainfall, or the parameters associated with the immediately preceding rainfall event for identifying the critical conditions. Such thresholds are used for predicting the occurrence of future landslides (Althuwaynee and Pradhan 2017) and can be used as a part of regional Landslide Early Warning System (LEWS) (Ahmed et al. 2020).

LEWS significantly helps in risk reduction by providing more time to the authorities to make decisions and take necessary actions (Piciullo et al. 2018). It is a cost-effective tool to warn the public regarding the imminent danger of landslides (Wicki et al. 2020). LEWS can be considered as a mitigation alternative, subject to upgradation with time, serving the purpose of risk reduction (Piciullo et al. 2018). Forecasting or modelling is a crucial element in a LEWS. Rainfall and landslide inventory database of the study area are analysed statistically to derive threshold models. The most commonly followed thresholds are based on the intensity and duration of the critical rainfall event (Caine 1980; Crosta 1998; Crosta and Frattini 2001; Aleotti 2004; Guzzetti et al. 2008; Brunetti et al. 2010; Abraham et al. 2019, 2020b), but the recent literature shows a shift towards event-duration thresholds (Melillo et al. 2016; Zhao et al. 2019). Intensity, event and duration are the parameters which are used to define a rainfall event; where event is the total amount of rainfall, duration is the total time of continuous rainfall and intensity is the rate of rainfall, calculated as the ratio of event to duration. The parameters of a rainfall event responsible for occurrence of landslides are considered for analysis. This rainfall event is a continuous precipitation, happened immediately before the landslide. It is generally accepted that shallow landslides are triggered by intense rainfalls of short duration (Campbell 1974; Crosta 1998) while slow or deep-seated slides are associated with prolonged rainfall (Bonnard and Noverraz 2001). Hence it is important to consider the effect of long-term rainfall for predicting slow moving landslides. Choosing the extent of antecedent rainfall to be considered is critical, and it has to be decided specifically for each region. In conventional thresholds, a single rainfall event is considered being a triggering factor of landsides and can be used for predicting shallow landslides. It is crucial to consider the effect of both short-term and long-term rainfall for regions, which are affected by both rapid and slow moving landslides. An algorithm-based model, Sistema Integrato Gestione Monitoraggion Allerta (SIGMA) is used for predicting the occurrence of landslides and issuing different warning levels for Idukki district in Kerala, India. The model was first developed for Italy (Martelloni et al. 2012), and has been found effective in predicting landslides in Indian Himalayas (Abraham et al. 2020a). Indian Himalayas contribute to a major share of global landslides (Dikshit et al. 2018; Froude and Petley 2018), is a totally different meteo-geological setting when compared with Italy. The geology of the landslide prone areas in Emilia Romagna region is dominated by highly cemented sandstones and clay beds with complex system of folds, faults and joints. In Darjeeling Himalayas, the study area was a small town, composed of phyllite quartzite and schist. A major portion of the area was formed by schist only. In the case of Idukki, the geology is entirely different, composed of peninsular gneissic complex, charnockite and migmatitic complex. The mean annual precipitation of the study area in Italy was 1072 mm, while in Darjeeling Himalayas, it was 1872 mm and in Idukki it is 3400 mm.

In this study, SIGMA model, which is found to have a satisfactory performance for Italy and Darjeeling Himalayas, is applied to a different location in the Western Ghats of India. Though the region suffers from a large number of landslides every year, no LEWS is available for Idukki. During 2018 monsoon, thousands of landslides have happened in the Western Ghats, which is being investigated (Vishnu et al. 2019; Kanungo et al. 2020; Meena et al. 2021). Idukki was the worst hit district in the disaster and suffered major social and economic setbacks due to the devastating landslides. The district needs an efficient LEWS to reduce the risk due to landslides. Collecting precise data for physically based models and installation of field monitoring systems are not feasible options, considering the vastness and variations in topography and climatic conditions of the region. The development of statistical rainfall thresholds is the best suited option in such cases, an economical and viable solution for developing an LEWS. Some attempts have been made for forecasting landslides in parts of Western Ghats using rainfall thresholds (Abraham et al. 2019, 2020b; Thennavan et al. 2020) and antecedent soil wetness (Abraham et al. 2021). However, these models are not ready to be used in an operational LEWS due to the higher number of false alarms or the complexities associated with the model. The region is in need for an LEWS model which can balance between the forecasting performance and ease of use. This study is an attempt to develop a regional scale LEWS to reduce the risk due to landslides in the region, using SIGMA model, which has more than 20 years of operational experience.

2. Details of study area

Idukki is a hilly district in the state of Kerala (India), covering an area of 4358 km^2. The district is the major power source of the state and is well known for Idukki dam, one of the highest arch dams in Asia. More than half of the district is covered by forest and the transportation facilities are limited. Idukki belongs to the Western Ghats region and several peaks have an elevation greater than 2000 m (Fig. 1).

Fig. 1 around here

The topography consists of mid lands, plateau regions and hill ranges. The eastern part of Idukki lies within the rain shadow region of Western Ghats and receives less rainfall when compared to the rest of the district. The daily rainfall data for this study has been collected from the Indian Meteorological Department (IMD)(India Meteorological Department 2019) from four rain gauge stations (Fig. 2) in Idukki.

The total area of the district has been divided into four, considering the location of rain gauges, and each unit is called one reference area. This approach has been adopted to account for the spatial and climatic variability across the district (Lu et al. 2014; Pasculli et al. 2014). The demarcation has been done using a multi-step procedure. First, the area is divided by straight lines based on the location of rain gauges, using the concept of Thiessen polygons (Abraham et al. 2019) (this approach considers the nearest rain gauge for each point to be analysed). However, from a practical point of view, division of a region into Thiessen polygons is difficult to execute in an operational LEWS, because local authorities act within their administrative boundaries. Hence, the polygons were modified according to the nearest administrative boundaries (towns or grama panchayats – the administrative divisions). This can help in issuing alarms in a more organised way. Moreover, the new boundaries are more in correspondence with physical elements (e.g. ridges, rivers) than the straight lines of the Thiessen polygons. Since the rainfall data collected is of daily resolution, the model issues a warning which predict the possibility of at least one landslide within the reference area. During calibration and validation, when multiple landslides have occurred in a reference area on a single day, it is considered as a single landslide event.

In the north-south direction, Idukki can be geologically divided into three parts with migmatitic complex lying in between peninsular gneissic complex in the north and charnockite group in the south (Department of Mining and Geology Kerala 2016). The peninsular gneissic complex rocks are well foliated and granite gneiss forms the oldest rock of the region, found in reference area R4. Among the charnockite group, charnockite is widespread in regions R2 and R3 and the presence of magnetite quartzite and pyroxene granulite are also observed in parts of R3(Department of Mining and Geology Kerala 2016). The migmatitic complex comprises of hornblende-biotite gneiss observed in area R4 and biotite gneiss, which covers a major portion of R1.

Structural and denudational hills are the predominant landforms in Idukki. Most of the hills are formed by Precambrian basement rocks with thin regolith thickness. As 60% of the district is covered by forest (major portions of R2 and R4), forest loam is the predominant soil type observed. Forest loam is produced by the weathering of rock under forest cover, characterised by rich organic content. Lateritic soils are found in the midlands of Idukki, formed from laterites with poor fertility. The forest loams consist of silts and clays, rich in organic content with high plasticity, while the grain size of lateritic soil has particles of coarse fraction, with minor fine content and the shear strength is due to the interparticle friction. According to the geotechnical map of India (Geological Survey of India 1995), the rocks of Idukki has low permeability and satisfactory compressive strength, suitable for foundations. But the recent infrastructure developments and the slope cuttings had adverse effects on the stability of slopes in the region. The depth of water level varies from 0 to 8 m (Sindhuraj 2013) throughout the year and during monsoon time, it is close to 0 m for a major share of the district.

Fig. 2 around here (Geological Survey of India 2010)

The topography consists of mid lands, plateau regions and hill ranges. The eastern part of Idukki lies within the rain shadow region of Western Ghats and receives less rainfall when compared to the rest of the district. The daily rainfall data for this study has been collected from the Indian Meteorological Department (IMD) (India Meteorological Department 2019) from four rain gauge stations (Figure 2) in Idukki.

The reference area for the first rain gauge, R1 represents the midland region of Idukki with nearly flat terrain, R2 and R3 represents the hilly area in the eastern side centre respectively and R4 consists of the peaks and foothills near the mountains in the northern side. The midland area of Idukki (R1) has a rugged topography, with a slope towards west. R1 is composed of pediment-pediplain complex of denudational origin. The hilly terrains can be divided into high ranges, plateau and foothills. The plateau region (R3 and parts of R2) covers maximum area and is the chief physiographic unit of Idukki. The elevation of this region varies from 500 m to 1500 m above sea level with a slope of around 30 %. A major part of the district is formed by the hill ranges (R2 and R4) of Western Ghats. The slope of this region is between 30 % to 50 % and occasionally goes upto 80 %. The peaks above 1500 m are characterised as high ranges (R4). R4 is the steepest zone with several peaks, composed of low dissected hills and valleys. The region is famous for its tea plantations and the hills have undergone several cutting and filling activities for infrastructure development, in the recent past. R2 region is formed by highly dissected hills and valleys.

The annual and cumulative rainfall from 2009 to 2018 is plotted in Fig. 3. From Fig. 3, it is clear that the rainfall distribution across the district is not uniform. The highest cumulative rainfall is recorded in the southernmost part of the district (R2) and the least value is in the rain shadow region (R4). It should also be noted that during the validation period (2018), the rainfall received is exceptionally high, reaching upto a maximum of 5788 mm in R2. The maximum rainfall was received in the district during the month of August 2018.

Fig. 3 around here

As per the data received from IMD, the monthly rainfall of the region during the study period has crossed 1000 mm once in R1, eight times in R2 and seven times in both R3 and R4. The daily rainfall has crossed 100 mm twenty-four times in R1, with an event in 2010, six in 2011, two each in 2012 and 2013, three each in 2014 and 2017 and seven events in 2018. In R2, the daily rainfall has crossed 100 mm 40 times during the study period and among them five were greater than 200 mm and two were greater than 300 mm. Both the events with daily rainfall greater than 300 mm were recorded in 2018. Similar to R1, the number of severe rainfall events has increased over time. The daily rainfall has crossed 100 mm on 20 days in R3 and 200 mm on 3 days among them. In case of R4, the numbers are 31 and 4, respectively. It can be understood that even if the cumulative rainfall is least recorded in R4, the number of severe rainfall events (greater than 100 mm per day) is the least in R1.

The anthropogenic activities in the recent past have led to cutting of slopes for infrastructure development, which considerably reduced the stability of slopes. The joints and cracks within the rocks are exposed to rain, resulting in slope failures. Earth and debris slides and debris flows have become common landslide types in the region which is mainly affected by shallow landslides (Kuriakose et al. 2008, 2009). Still some earth slides were reported to continue over a long period of time, along the major road corridors which can be attributed as the result of long-term rainfall. The types of landslides vary from translational earth and debris slides along the slope cuts to the long runout debris flows. The region R1 is mostly affected by shallow landslides while most of the debris flows have reported in R3 and R4. Around 65 % of the total landslides considered were shallow landslides, 30 % debris flows, and the remaining were rock falls.

The occurrence of landslides was found to be associated with the occurrence of severe rainfall events. Multiple landslides were recorded on the same day, across the district, following the occurrence of daily rainfall greater than 100 mm. Landslides were recorded on the same day, or within a short span of time after the occurrence of rainfall. Some landslides have occurred on days with very less rainfall recorded in the reference rain gauge. These can either be the effect of prolonged rainfall over the study area, or due to localised heavy rainfalls, which were not recorded in the reference rain gauge. Hence, it is important to study the effect of both long-term and short-term rainfall in the initiation of landslides within the study area. According to the authors who firstly proposed it, SIGMA method is conceived to deal with very different landslide types: shallow landslides (triggered by short and intense rainfalls) and deep-seated landslides (triggered by prolonged rainfalls) (Martelloni et al. 2012; Lagomarsino et al. 2013). This idea is supported by at least 20 years of test and operation use (Lagomarsino et al. 2015; Segoni et al. 2018a). This study is an attempt to explore the use of SIGMA for the study area in Western Ghats.

3. SIGMA model

As the name point out, SIGMA model takes the standard deviation of a statistical distribution as the key parameter for threshold definition. The thresholds are defined as a function of the standard deviation, to predict the possible occurrence of landslides in a region. As the model is purely based on statistical analysis of historical rainfall and landslide data, it can be easily exported to be used in different areas (Martelloni et al. 2012; Segoni et al. 2018b). However, apart from the region for which it was first developed, SIGMA has been applied to very few regions (Abraham et al. 2020a). On account of its predicting capacity and ability to define multiple levels of warning, SIGMA has the potential to be used as an LEWS. This study is an endeavour to evaluate the applicability of SIGMA mode for Idukki district in India. The methodology has been adopted from Martelloni et al. (2012) (Martelloni et al. 2012) and the model has been customised for developing an LEWS for Idukki. The customisations are done according to the statistical distribution of rainfall data of Idukki, to minimise the missed and false alarms generated.

The daily precipitation data has been collected for the study area for four different rain gauges (India Meteorological Department 2019) and for each rain gauge, the daily precipitation data were cumulated at 'n' days, with a window which shifts at daily timesteps with 'n' day width. The value of 'n' has been varied from 1 to 365. For each dataset, the cumulative distribution function (F) was calculated with a standard distribution as target function (Martelloni et al. 2012). This target function is used to relate the cumulative rainfall (z) with the distribution $y = a.\sigma$ ('a' is a multiplication constant and 'σ' is the standard deviation of each series). The values of z are sorted in ascending order for each series of n day width.

$$z_1 < z_2 < z_3 < \cdots < z_k < \cdots < z_n \tag{1}$$

The cumulative frequency of sample is defined as

$$P_k = \frac{k}{n} - \frac{0.5}{n} = G(y) \tag{2}$$

for each value of k, varying between 1 to n. The cumulative distribution function of z, $F(z)$ is used to establish the probability that the value of z is less than z_k

By using $P(K)$ and a target function (Goovaerts 1997), the variable z can be transformed to y as:

$$G^{-1}(F(z)) \rightarrow G^{-1}(P_k) = y \quad (3)$$

where G is the target function and P_k is defined as $G(y)$. Once the transformation is complete, for any multiples of standard deviation, the corresponding cumulative frequency of sample can be estimated. For all values of n, the same procedure has been repeated to plot the sigma curves (σ curves or precipitation curves). The algorithm for SIGMA model uses these σ curves as input. The algorithm compares the value of cumulated rainfall recordings for a specific duration with the σ curves. The duration is determined by trial and error, based on the historical rainfall data. SIGMA considers both short term and long-term effect and hence the duration for different levels of warning and different types of slope failures can be different. Using this algorithm, a warning level is issued everyday based on the rainfall. Alerts are issued for every day, based on the rainfall threshold. The cumulated rainfall recordings for daily timesteps were compared to the σ curves to issue an alert (Martelloni et al. 2012). The thresholds take into account the effect of both short term and long-term rainfall. For the short term effect, to issue a warning on highly and moderate critical events which are rapid to very rapid, the effect of cumulative rainfall up to two days were considered. The condition used to check the high and moderate criticality cases are given in equation 4 below.

$$C_{1-3} = \left[\sum_{i=1}^{n} P(t+1-i)\right]_{n=1,2,3} \geq [S_n(\Delta)]_{n=1,2,3} \quad (4)$$

where, $\Delta = a.\sigma,$ the vector C_{1-3} represents the total rainfall cumulated at time t and $S_n(\Delta)$ are the rainfall thresholds for n days and Δ (Martelloni et al. 2012; Segoni et al. 2018b). For slow movements, the algorithm checks for the effect of precipitation from 4 days upto N days, where N is the upper limit of long-term rainfall considered, and is different for the four different rain gauges considered. The condition for issuing an ordinary criticality warning is:

$$C_{4-\mathrm{N}} = [\textstyle\sum_{i=1}^{n+3} P(t-2-i)]_{n=1,2,\ldots 60} \geq [S_{n+3}(\Delta)]_{n=1,2,\ldots \mathrm{N}-3} \quad (5)$$

The definitions of the cumulative rainfall vector C are kept the same as defined by the developers, to derive rainfall thresholds for Idukki.

4. Analysis

The first step of developing SIGMA model is the understanding of distribution of cumulated rainfall data and the selection of target function. The rainfall data from 2009-2018 were used for the analysis, for which the first 9 years (2009-2017) were used for calibration and the last year (2018) for validation. The data of 2009 has been used as a buffer to calculate daily cumulates up to 365 days for the year 2010. From During the calibration period, n day cumulative precipitations were calculated with the value of n ranging from 1 to 365. Then for each value of n, cumulative distribution functions were plotted, after sorting the values in ascending order. It was found that when the number of days is smaller; the distribution is found to be similar to that of log-normal and for higher values of n, the distribution tends towards normal. Hence normal distribution was chosen as the target function and the threshold values for all values of n ($\Delta = a.\sigma$) were calculated using the transformation as mentioned in Eq. 3 (Fig. 4).

The threshold curves were plotted with the values of n on x axis and the threshold values on y axis as shown in Fig. 5. These threshold values were compared with the everyday cumulated values using a decisional algorithm to identify the critical rainfall events.

Fig. 5 around here

For the customised model, a simple algorithm was defined, with four different levels of warning. The alert levels were defined according to the local system, which is in practice for forecasting other disasters. The highest criticality case is considered as a red alert, moderate criticality as orange, ordinary criticality as yellow and no criticality as green. The general public is already aware of these alert levels, hence it is easy to follow the LEWS.

A starting algorithm was used commonly for the whole district after calibration, and was optimised separately for each reference area. The decisional algorithm which was used in the initial stage of calibration is shown in Fig. 6.

Fig. 6 around here

The algorithm is designed very simple, for easy understanding and exportability. The algorithm compares the n day cumulates corresponding to the rainfall forecast, with the threshold curves, to issue an alert. If the threshold is crossed, an alert is issued based on the severity of the possible landslide event. The algorithm first considers the effect of short-term rainfall, to identify the most critical rainfalls, and issue red alert. If the extreme condition does not exist, it searched for the medium criticality case for short-term rainfall and if the threshold value is crossed, an orange alert is issued. For both red and orange alerts, only short-term rainfall is considered as they lead to very fast shallow landslides while long-term rainfall is considered issuing the ordinary criticality level or the yellow alert for slow movements. If both high and moderate levels of criticality conditions are not crossed, the algorithm consider the long-term rainfall and checks if the threshold is crossed within N^{th} day considered, to issue yellow alert. It should also be noted that on days for which red or orange alerts are issued, there are chances that the long-term threshold is also crossed. Hence red and orange alert predicts the possibility of occurrence of both rapid and slow-moving landslides while yellow alert predicts the possibility of occurrence of slow-moving landslides only. The value of N has been selected by trial and error for each reference area separately. For starting the algorithm, it was considered as 63 as in the SIGMA models previously developed (Martelloni et al. 2012; Abraham et al. 2020a).

The threshold is exceeded when any of the elements in the vector C crosses the threshold value. The values used in the starting algorithm were optimised for each reference area separately, using a separate module which uses the threshold criteria with the occurrence of landslides. The thresholds were raised in small increments for each day to verify if false alarms are reduced as shown in Fig. 7. The procedure continues till any true alarm is missed.

Fig. 7 around here

5. Results

The procedure of optimisation is used to reduce the false alarms and fine tuning of the thresholds. After the analysis, 1 σ 1.25σ and 1.5σ considered in the starting algorithm (Fig. 6) were customised for each area. During this process (Figure 7), the threshold values were increased slightly to reduce the number of false alarms. The events which have issued false alarms were considered for this process and threshold value is increased in minor increments, so that the false alarm can be avoided with the condition that no true alarm is missed. The values of N were also customised for each region, to reduce the number of false alarms generated. The process of calibration was a trial-and-error approach. The values of thresholds and N were varied in such a way that the number of false alarms is reduced, at the cost of a minimum number of missed alarms. Several trials were conducted for each reference area, to find a best suited value for N, with a balance between the false and missed alarms. Which means, the value less than N will lead to many missed alarms and any value greater than N will issue more false warnings. Idukki district receives rainfall events of longer duration and the daily resolution of rainfall data makes it extremely difficult to separate events with dry periods less than 24 hours. Hence the long-term rainfall considered for the analysis was customised for each case in order to improve the performance of the model. The values of thresholds modified after optimisation for each reference area are listed in Table 1.

Table 1 around here

The opitimised thresholds were then validated using the rainfall and landslide data of 2018. The region R1 consists of the flat terrains, which is less susceptible to landslides. Most of the cases reported in this area are cut slope failures and other shallow landslides, hence only short-term rainfall is considered for issuing warnings in this region. The threshold values are not too high, implying the possibility of less severe rainfalls triggering landslides in the area.

The optimisation process has effectively reduced the number of false alarms during validation as shown in Fig. 8. It can be observed that the number of false yellow alerts has reduced considerably due to optimisation. The highest number of false alarms generated are yellow, implying ordinary criticality, then red alerts and orange alerts are the least generated. It can also be noted that the optimised values for former 1.25 σ do not vary much and hence the reduction in false orange alarms after optimisation is also the least.

Fig. 8 around here

During the period of validation, the decisional algorithm was used to issue different alert levels for each day, which were compared with the occurrence or non-occurrence of landslides to validate the model. The classical approach of confusion matrix was used for the evaluation, to quantify the performance of SIGMA model for each reference area (Lagomarsino et al. 2015). The number of correct predictions are termed as true positives (TP) and true negatives (TN); where TP is the number of landslides correctly predicted and TN is the number of nonlandslides correctly predicted. Similarly, incorrect predictions are called false positives (FP) and false negatives (FN) where FP indicates the false alarms and FN indicates missed alarms.

The results of validation for each reference are listed in Table 2 below:

Table 2 around here

It can be noted that the algorithm is correctly predicting all the landslides except in the case of R3 and R4, where the topography is highly varying, and the reference area is the largest. The algorithm correctly predicts 79% of the total landslides happened in the region. The performance is the best in the midlands region (R1) where all landslides are correctly predicted, but at the cost of minimum false alarms. Since only short-term rainfall is considered for the analysis of R1, the false alarms generated is very less in this case. The higher number of false alarms are expected as the threshold values are lesser, especially in the case of yellow alert, where there is a high possibility of a threshold being crossed at any of the long-term period considered. The number of false alarms is the maximum in case of yellow alert and the least in case of red alert. This is again due to the less threshold value considered for yellow alert. Another reason for the increase in number of false alarms is the change in rainfall pattern observed during the period of validation, 2018. The rainfall received during 2018 was more than 1.5 times the average annual rainfall during the study period, in all four regions. The rainfall has crossed the derived threshold many times, issuing a number of false alarms in all cases. Hence the model should be improved further to reduce the number of false alarms, to make it operational as a part of LEWS (Segoni et al. 2018b).

6. Discussions

The obtained results show that SIGMA model has a satisfactory performance in three out of four reference areas considered for study. SIGMA model uses a decisional algorithm to predict the landslides based on historical rainfall and landslide data. The model considers the effect of both short-term and long-term rainfall, in order to predict both shallow and deep-seated landslides.

The less rain gauge density and variations in topography of the district have led to some missed alarms in regions R3 and R4 (Fig. 9). When multiple landslides have occurred on the same day, the one closest to rain gauge is considered for representation of TP and FN. The variations in elevation between the location of rain gauge and landslide has resulted in this error in prediction. The variation in topography is a key factor to be considered in identifying the responsible rainfall. The poor rain gauge density in the study area in the major reason of less efficiency in regions R3 and R4. The recorded rain gauge is varying from the actual one, duet to the spatial and topographical variations. This has also resulted in the lesser threshold values, as the thresholds were lowered, to minimise the number of missed alarms. This has resulted in the increased number of false alarms. In the case of R4, the locations near the rain gauge in R4 belongs to the rain shadow region of Idukki which receives very less amount of rainfall. The missed landslides have happened at the southern side of the rain gauge in R4, possibly as a result of a higher amount of rainfall. To identify correctly the responsible rainfall, the district requires a much stronger network of rain gauges.

During the process of optimisation, the threshold values did not vary much, but the false alarms were reduced mainly by varying the number of days considered for the long-term rainfall criteria. The highly varying topography and climate of the region demands for higher rain gauge density, to correctly identify the rainfall events responsible for each landslide. The lesser rain gauge density cannot identify the localised storms or cloud burst that have resulted in slope failures and essentially identifies a less severe rainfall, recorded at the reference gauge as the responsible rainfall event. It can be observed from Fig. 9 that most of the missed landslides happened at locations far from the rain gauges at a different elevation. This leads to the occurrence of landslides at lesser threshold values, which ultimately lead to higher false alarms. If the thresholds are raised, it will result in missed alarms, which is a more critical case. Hence the model can be improved further with the availability of rainfall data with better spatial and temporal resolutions. Even with this limitation, SIGMA model has the advantage of being a simple method which requires only historical landslide and rainfall database as inputs and can be used to predict both rapid and slow failures in the region.

Fig. 9 around here

The procedure of optimisation was adopted to minimise false alarms to the best possible extent, and the procedure involved many trials, in order to finalise the number of days and threshold values considered in the analysis. All four areas differ in their morphology and geology and climatic conditions. Hence the values were customised for each area separately. Optimisation has improved the performance of the model considerably. Increasing any of the threshold values or decreasing the number of days considered for daily cumulates will result in missed alarms and were determined by several trials.

In this study, the cumulated rainfall upto 3 days has been considered for predicting shallow landslides in Idukki district (India). Shallow slides include shallow debris flows, soil slips etc, which are the results of short-term rainfall. The long-term rainfall is used for predicting slow movements and deep-seated landslides in the region. The long-term cumulates are essential for predicting the slow movements observed in the region, but they lead to much more false alarms than the short-term cumulates.

When SIGMA was applied to the study area in Italy, the first prototypal version had a likelihood ratio of 8.38, which was then updated conceptually later and the likelihood ratio was improved to 17.01 (Segoni et al. 2018a). For the second study area in Darjeeling Himalayas, the likelihood ratio of the model was found to be 11.28 (Abraham et al. 2020a). For Idukki, the likelihood ratio is varying from 2.39 to 12.45 which proves the model need further improvements using better resolution rainfall and landslide data.

The rainfall data used from 2009 to 2018 has been used for the analysis, to understand the statistical distribution of cumulated rainfall. The use of a much longer term may result in a lesser mean value and higher threshold limits. Even though the most recent data has been used, the sudden change in rainfall pattern happened during 2018 has issued many false alarms. The model has to be updated continuously with more recent rainfall data, to incorporate the variations in rainfall pattern due to the changing climate.

The base algorithm for SIGMA can easily be exported to other parts of the world also and can be customised using regional specific rainfall and landslide data. Hence the model proves to be a simple tool that can be used as a part of LEWS, with conceptual improvements that can reduce the false alarms in the region.

7. Conclusions

A landslide prediction system for Idukki district (Kerala, India) has been developed using SIGMA model, considering the long-term and short-term effect of rainfall in the initiation of landslides in the region. The model uses statistical distribution of rainfall data and the cumulative distribution function to derive rainfall thresholds which are compared with the daily cumulated rainfall values. A decisional algorithm is used for comparing the rainfall vector with the thresholds, which issues different levels of alert based on the severity of rainfall condition. The has been divided in to four reference areas, considering the topographic variability and location of rain gauges. The database from 2009-2017 were used for calibration of the model. From a common algorithm used for the entire district, the threshold values and number of days considered for the analysis were optimised for each reference area, to reduce the number of false alarms issued. The optimised model was then validated using a completely different dataset of 2018 to evaluate the prediction performance.

SIGMA model for Idukki was found to be effective in predicting all the landslides in three reference areas but with a higher number of false alarms. The best performance of model was found in R1, with an efficiency of 92.05% and likelihood ratio 12.45. If the number of false alarms can be reduced by introducing physical parameters or further constraints in the decisional algorithm, SIGMA can be used as an effective early warning system for the region.

The model has its advantages of being simple and lesser inputs in decision making and can be integrated easily with any rainfall forecasting system to issue the warning. Unlike the conventional empirical approaches, SIGMA can be used to issue multiple levels of warning based on the cumulated rainfall value. The incorporation of multiple warning levels makes the model a better prediction tool for issuing early warning. The use of long term and short term and data helps in predicting both rapid and slow movements within the region, which has helped the algorithm to correctly predict all the landslides in three reference areas. As observed from the study, better spatial and temporal resolutions of rainfall data can considerably reduce the number of false alarms and improve the performance of the model.

The simplified model with good prediction performance is important from the scientific perspective as an important step towards establishing an effective LEWS for the region. If the limitations of poor resolution of data can be improved using a network of rain gauges, the authors believe that the developed tool can help in reducing the risk due to landslides in this hilly district of Kerala, India.

Acknowledgments: The authors would like to acknowledge the support from the Geological Survey of India, District Soil Conservation Office, Idukki and Kerala State Disaster Management Authority for the research.

Funding details: This work was supported in part by the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering & IT, University of Technology Sydney (UTS). This APC was funded by Universiti Kebangsan Malaysia, DANA IMPAK PERDANA with grant no: DIP-2018-030. It was also supported by Researchers Supporting Project (number RSP-2020/14), King Saud University, Riyadh, Saudi Arabia.

Conflicts of interest/Competing interests: The authors declare no conflicts of interest

Availability of data and material: The data used for the study has been collected from various government agencies. The satellite data used is available online.

References

Abd Majid N, Rainis R (2019) Application of Geographical Information Systems (GIS) and Discriminant Analysis in Modelling Slope Failure Incidence in Pulau Pinang, Malaysia. Sains Malaysiana 48:1367–1381. https://doi.org/10.17576/jsm-2019-4807-06

Abraham MT, Pothuraju D, Satyam N (2019) Rainfall Thresholds for Prediction of Landslides in

Idukki, India: An Empirical Approach. Water 11:2113. https://doi.org/10.3390/w11102113

Abraham MT, Satyam N, Kushal S, et al (2020a) Rainfall Threshold Estimation and Landslide Forecasting for Kalimpong, India Using SIGMA Model. Water 12:1195. https://doi.org/10.3390/w12041195

Abraham MT, Satyam N, Rosi A, et al (2020b) The Selection of Rain Gauges and Rainfall Parameters in Estimating Intensity-Duration Thresholds for Landslide Occurrence: Case Study from Wayanad (India). Water 12:1000. https://doi.org/10.3390/w12041000

Abraham MT, Satyam N, Rosi A, et al (2021) Usage of antecedent soil moisture for improving the performance of rainfall thresholds for landslide early warning. Catena 200:105147. https://doi.org/10.1016/j.catena.2021.105147

Ahmed B, Rahman MS, Sammonds P, et al (2020) Application of geospatial technologies in developing a dynamic landslide early warning system in a humanitarian context: the Rohingya refugee crisis in Cox's Bazar, Bangladesh. Geomatics, Nat Hazards Risk 11:446–468. https://doi.org/10.1080/19475705.2020.1730988

Aleotti P (2004) A warning system for rainfall-induced shallow failures. Eng Geol 73:247–265. https://doi.org/10.1016/j.enggeo.2004.01.007

Althuwaynee OF, Pradhan B (2017) Semi-quantitative landslide risk assessment using GIS-based exposure analysis in Kuala Lumpur City. Geomatics, Nat Hazards Risk 8:706–732. https://doi.org/10.1080/19475705.2016.1255670

Baum RL, Godt JW (2010) Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides 7:259–272. https://doi.org/10.1007/s10346-009-0177-0

Bonnard C, Noverraz F (2001) Influence of climate change on large landslides: assessment of long term movements and trends. In: International Conference on Landslides causes impact and countermeasures. pp 121–138

Bordoni M, Corradini B, Lucchelli L, et al (2019) Empirical and Physically Based Thresholds for the Occurrence of Shallow Landslides in a Prone Area of Northern Italian Apennines. Water 11:2653. https://doi.org/10.3390/w11122653

Bordoni M, Galanti Y, Bartelletti C, et al (2020) The influence of the inventory on the determination of the rainfall-induced shallow landslides susceptibility using generalized additive models. Catena 193:104630. https://doi.org/10.1016/j.catena.2020.104630

Brunetti MT, Peruccacci S, Rossi M, et al (2010) Rainfall thresholds for the possible occurrence of landslides in Italy. Nat Hazards Earth Syst Sci 10:447–458. https://doi.org/10.5194/nhess-10-

447-2010

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

Campbell RH (1974) Debris flows originating from soil slips during rainstorms in Southern California. Q J Eng Geol 7:339–349. https://doi.org/10.1144/GSL.QJEG.1974.007.04.04

Crosta G (1998) Regionalization of rainfall thresholds: An aid to landslide hazard evaluation. Environ Geol 35:131–145. https://doi.org/10.1007/s002540050300

Crosta GB, Frattini P (2001) Rainfall thresholds for soil slip and debris flow triggering. Proc 2nd EGS Plinius Conf Mediterr Storms 463–487

Department of Mining and Geology Kerala (2016) District Survey Report of Minor Minerals. Thiruvananthapuram

Dikshit A, Satyam DN, Towhata I (2018) Early warning system using tilt sensors in Chibo, Kalimpong, Darjeeling Himalayas, India. Nat Hazards 94:727–741. https://doi.org/10.1007/s11069-018-3417-6

Easterling DR, Meehl GA, Parmesan C, et al (2000) Climate Extremes: Observations, Modeling, and Impacts. 289:2068–2075

Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. Nat Hazards Earth Syst Sci 18:2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Fusco F, De Vita P, Mirus BB, et al (2019) Physically based estimation of rainfall thresholds triggering shallow landslides in volcanic slopes of Southern Italy. Water 11:1–24. https://doi.org/10.3390/w11091915

Geological Survey of India (1995) Geotechnical Map of India

Geological Survey of India (2010) Geology and minerals :District Resource Map, Idukki

Gian QA, Tran DT, Nguyen DC, et al (2017) Design and implementation of site-specific rainfall-induced landslide early warning and monitoring system: a case study at Nam Dan landslide (Vietnam). Geomatics, Nat Hazards Risk 8:1978–1996. https://doi.org/10.1080/19475705.2017.1401561

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, Oxford

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity-duration control of shallow landslides and debris flows: An update. Landslides 5:3–17. https://doi.org/10.1007/s10346-007-0112-1

India Meteorological Department (2019) India Meteorological Department (IMD) Data Supply Portal

Kanungo DP, Singh R, Dash RK (2020) Field observations and lessons learnt from the 2018 landslide disasters in Idukki District, Kerala. Curr Sci 119:1797–1806. https://doi.org/10.18520/cs/v119/i11/1797-1806

Kuriakose SL, Jetten VG, van Westen CJ, et al (2008) Pore water pressure as a trigger of shallow landslides in the Western Ghats of Kerala, India: Some preliminary observations from an experimental catchment. Phys Geogr 29:374–386. https://doi.org/10.2747/0272-3646.29.4.374

Kuriakose SL, Sankar G, Muraleedharan C (2009) History of landslide susceptibility and a chorology of landslide-prone areas in the Western Ghats of Kerala, India. Environ Geol 57:1553–1568. https://doi.org/10.1007/s00254-008-1431-9

Lagomarsino D, Segoni S, Fanti R, Catani F (2013) Updating and tuning a regional-scale landslide early warning system. Landslides 10:91-97.

Lagomarsino D, Segoni S, Rosi A, et al (2015) Quantitative comparison between two different methodologies to define rainfall thresholds for landslide forecasting. Nat Hazards Earth Syst Sci 15:2413–2423. https://doi.org/10.5194/nhess-15-2413-2015

Lu B, Charlton M, Harris P, Fotheringham AS (2014) Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. Int J Geogr Inf Sci 28:660–681. https://doi.org/10.1080/13658816.2013.865739

Martelloni G, Segoni S, Fanti R, Catani F (2012) Rainfall thresholds for the forecasting of landslide occurrence at regional scale. Landslides 9:485–495. https://doi.org/10.1007/s10346-011-0308-2

Meena SR, Ghorbanzadeh O, van Westen CJ, et al (2021) Rapid mapping of landslides in the Western Ghats (India) triggered by 2018 extreme monsoon rainfall using a deep learning approach. Landslides. https://doi.org/10.1007/s10346-020-01602-4

Melillo M, Brunetti MT, Peruccacci S, et al (2016) Rainfall thresholds for the possible landslide occurrence in Sicily (Southern Italy) based on the automatic reconstruction of rainfall events. Landslides 13:165–172. https://doi.org/10.1007/s10346-015-0630-1

Morss RE, Wilhelmi O V., Meehl GA, Dilling L (2011) Improving societal outcomes of extreme weather in a changing climate: An integrated perspective. Annu Rev Environ Resour 36:1–25. https://doi.org/10.1146/annurev-environ-060809-100145

Muhammad M, Idris I, Sharif Salazar A, et al (2010) GIS Based Landslide Hazard Mapping Prediction in Ulu Klang, Malaysia. ITB J Sci 42:163–178. https://doi.org/10.5614/itbj.sci.2010.42.2.7

Pasculli A, Palermi S, Sarra A, et al (2014) A modelling methodology for the analysis of radon potential based on environmental geology and geographically weighted regression. Environ Model Softw 54:165–181. https://doi.org/10.1016/j.envsoft.2014.01.006

Piciullo L, Calvello M, Cepeda JM (2018) Territorial early warning systems for rainfall-induced landslides. Earth-Science Rev 179:228–247. https://doi.org/10.1016/j.earscirev.2018.02.013

Segoni S, Rosi A, Fanti R, et al (2018a) A regional-scale landslide warning system based on 20 years of operational experience. Water 10:1297. https://doi.org/10.3390/w10101297

Segoni S, Rosi A, Lagomarsino D, et al (2018b) Brief communication: Using averaged soil moisture estimates to improve the performances of a regional-scale landslide early warning system. Nat Hazards Earth Syst Sci 18:807–812. https://doi.org/10.5194/nhess-18-807-2018

Sharir K, Roslee R, Lee KE, Simon N (2017) Landslide Factors and Susceptibility Mapping on Natural and Artificial Slopes in Kundasang, Sabah. Sains Malaysiana 46:1531–1540. https://doi.org/10.17576/jsm-2017-4609-23

Sindhuraj S (2013) Ground Water Information Record of Idukki District, Kerala State. Central Ground Water Board Kerala Region, Ministry of Water Resources, Government of India, Thiruvananthapuram

Thennavan E, Ganapathy Pattukandan G, Chandrasekaran SS, Rajawat AS (2020) Probabilistic Rainfall Thresholds for Shallow Landslides Initiation – A Case Study from the Nilgiris District, Western Ghats, India. Int J Disaster Risk Manag 2:1–13. https://doi.org/10.18485/ijdrm.2020.2.1.1

Vishnu CL, Sajinkumar KS, Oommen T, et al (2019) Satellite-based assessment of the August 2018 flood in parts of Kerala, India. Geomatics, Nat Hazards Risk 10:758–767. https://doi.org/10.1080/19475705.2018.1543212

Wicki A, Lehmann P, Hauck C, et al (2020) Assessing the potential of soil moisture measurements for regional landslide early warning. Landslides 17:1881–1896. https://doi.org/10.1007/s10346-020-01400-y

Zhao B, Dai Q, Han D, et al (2019) Estimation of soil moisture using modified antecedent precipitation index with application in landslide predictions. Landslides 16:2381–2393. https://doi.org/10.1007/s10346-019-01255-y