Engineering Geology xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities

Saibal Ghosh [a,b,*], Cees J. van Westen [b], Emmanuel John M. Carranza [b], Victor G. Jetten [b], Mauro Cardinali [c], Mauro Rossi [c], Fausto Guzzetti [c]

[a] Engineering Geology Division, Geological Survey of India, Eastern Region, Kolkata, India
[b] Department of Earth System Analysis, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands
[c] CNR-IRPI, via Delta Madonna Alta 126, 06128 Perugia, Italy

ARTICLE INFO

Article history:
Received 30 September 2009
Received in revised form 20 June 2010
Accepted 16 March 2011
Available online xxxx

Keywords:
Event-based landslide inventory
Frequency–magnitude relation
Temporal probability
Magnitude probability
Darjeeling Himalaya

ABSTRACT

Event-based landslide maps display landslide events of different magnitudes that are associated with a single major trigger. Event-based landslide maps allow the estimation of spatio-temporal probabilities of landslide events that are required for the assessment of landslide hazard. In this paper, we discuss the generation of event-based landslide inventory maps for a region in Darjeeling Himalayas (India) with limited historical landslide information. We used eight different sets of source data, spanning the period from 1968 to 2007, which include topographic maps, satellite images and post-event field maps. For each period we mapped the various landslide types, and calculated landslide density and activity in individual terrain units to analyze changes in the spatial behavior of landslides through time. For the assessment of the temporal landslide probability, we used landslide event-days and associated daily and antecedent rainfall amounts in a multivariate statistical analysis to model the temporal relationship between landslide events and the amount of triggering rainfall. For estimating the magnitude probability, we applied magnitude–frequency analysis for the different event-based landslide maps using landslide area (m^2) as a proxy. Despite the incompleteness of the source data sets used, the resulting landslide inventory maps can be used successfully in determining temporal and magnitude probabilities of future landslides.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Incorporation of landslide hazard analysis into land use planning is important in mitigating the impact of landslides. Whereas this is slowly becoming a standard practice in developed countries, many developing countries still lack proper land use planning, let alone the inclusion of landslide risk due to lack of detailed and reliable landslide hazard maps. In many cases, available landslide hazard maps are actually landslide susceptibility maps with qualitative legends (Kanungo et al., 2006; Ghoshal et al., 2008; Gupta et al., 2008; Mathew et al., 2009), that are difficult to use in the quantification of the possible impact of landslides. Conversion of landslide susceptibility maps into landslide hazard maps requires estimates of landslide spatial, temporal and magnitude probabilities (Guzzetti et al., 1999; Glade et al., 2005; Fell et al., 2008; van Westen et al., 2008). These can only be derived accurately through the generation and analysis of landslide inventory maps.

Landslide inventory maps can be prepared through various methods, such as historical archive studies, interviews, detailed geomorphologic fieldwork, image interpretation and remote sensing techniques (Hansen, 1984; Wieczorek, 1984; Guzzetti et al., 2000; Cardinali et al., 2002; Galli et al., 2008; van Westen et al., 2008; Martha et al., 2010). An event-based inventory is prepared soon after a prominent triggering event (e.g., rainfall and earthquake) to depict all slope failures that are due to that triggering event (Harp and Jibson, 1996; Guzzetti et al., 2004; Sorriso-Valvo et al., 2004). Due to lack of sufficient historical information on landslides or post-event field maps, air photos or satellite images from the past are often used as the main remote sensing source data for the generation of event-based landslide maps (Rib and Liang, 1978; Guzzetti et al., 2008).

The success of generating an event-based landslide inventory via stereoscopic interpretation of images from different periods depends on several factors. Most importantly the availability of images of suitable quality and resolution, taken soon after a major landslide triggering event, determines whether it is possible to extract individual event-based landslide maps. The characteristics of landslides may not be recognizable anymore in remote sensing imagery at later dates, as they may be obscured by vegetation re-growth or even by quick reactivation. If no images are available directly after a triggering event it becomes difficult to make separate event-based inventories for those triggering events that occurred between the dates of two successive images. Also skills in image interpretation play an important role (Carrara, 1993), as

* Corresponding author at: Engineering Geology Division, Geological Survey of India, Eastern Region, Kolkata, India. Fax: +91 33 2359 6393.
E-mail addresses: ghosh@itc.nl, saibal_ghosh@rediffmail.com (S. Ghosh).

0013-7952/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.enggeo.2011.03.016

ARTICLE IN PRESS

well as the accuracy with which landslides are located on base maps (Malamud et al., 2004). Through time, multiple small slope failures may merge into larger landslides, which render detection of smaller landslides problematic and thereby introducing bias in frequency–size statistics of landslides. Accordingly, lack of spatial and temporal accuracy in landslide inventories, incompleteness of individual inventories and gaps in time between inventories are serious bottlenecks in determining temporal and magnitude probabilities of landslides and, therefore, pose difficulties in quantitative hazard assessment (van Westen et al., 2006).

In this paper, we discuss how to generate event-based landslide inventory maps for triggering events over a 40 year period (1968 to 2007) in a highly landslide-prone area in Darjeeling Himalayas (India) based on available source data. In a data-scarce environment like India, the source data for landslide inventory mapping are incomplete, there are data gaps for certain periods, and the source data has different scales and resolution. Notwithstanding such constraints, we generated an event-based landslide database for utilization in quantitative hazard assessment.

2. Study area

The study area is located around the town of Kurseong in the Darjeeling district of West Bengal province in India (Figure 1a). The area is characterized by a prominent NE–SW trending ridge in the central part dividing the study area into sub-catchments of two major streams (Figure 1a and d): Balason toward the west and Tista toward the east. Elevations in the study area vary from 236 m to 2189 m and slopes vary from 0° to 84°. Climate is humid with a long period (June–October) of monsoon-controlled heavy precipitation. Annual precipitation varies between 2000 mm and 5000 mm (Soja and Starkel, 2007).

The study area is part of a tectono-stratigraphic sequence of metamorphic rocks of the Himalayan Fold-Thrust-Belt (FTB) in the north and the foreland molasse basin in the south (Figure 1b and c). Toward the north, high-grade metamorphic rocks (migmatites) are present, whereas its southern boundary is marked by a high-strain ductile shear zone called the Main Central Thrust (MCT) that coincides with high-to-low grade metasediments (cf. Hubbard, 1996; Searle and Szule, 2005). The study area is situated in the southern part of the Darjeeling klippe (Figure 1b and c), where high-grade metamorphic rocks of the Central Crystalline Gneissic Complex (CCGC) are thrusted over low-grade metasediments of the Daling Group along the MCT (Mallet, 1875; Sinha-Roy, 1982). Toward the south, foreland molasse sediments of the Siwalik Group are underlain by an intra-thrusted slice of Gondwana sediments. Toward the north, Gondwana sediments are overthrusted by Daling Group metasediments along the abrupt southern-most front of Himalayan FTB known as the Main Boundary Thrust (MBT). The MCT and MBT represent the main regional tectonic discontinuity surfaces in this part of the Himalayas (Figure 1b–d).

Denudational geomorphic processes are predominant due to the active tectonic processes in Darjeeling Himalayas. It has been demonstrated by several workers in the adjacent Nepalese Himalayas (Caine and Mool, 1982; Selby, 1988; Petley et al., 2007) that the high frequency of slope failures in the Eastern Himalayas can be attributed to a very high rate of denudation (i.e., 5–14 mm year^{-1}) associated with active tectonic processes of Himalayan orogeny. Slopes are gentler on ridge tops but become increasingly steeper downward to the streams (Dutta, 1966; Burbank et al., 1996; Binnie et al., 2007).

In general, three types of landslides are observed in the study area (Figure 2). Most frequent are shallow translational rockslides, followed by shallow translational debris slides and some deep-seated rockslides that are much larger than the first two types.

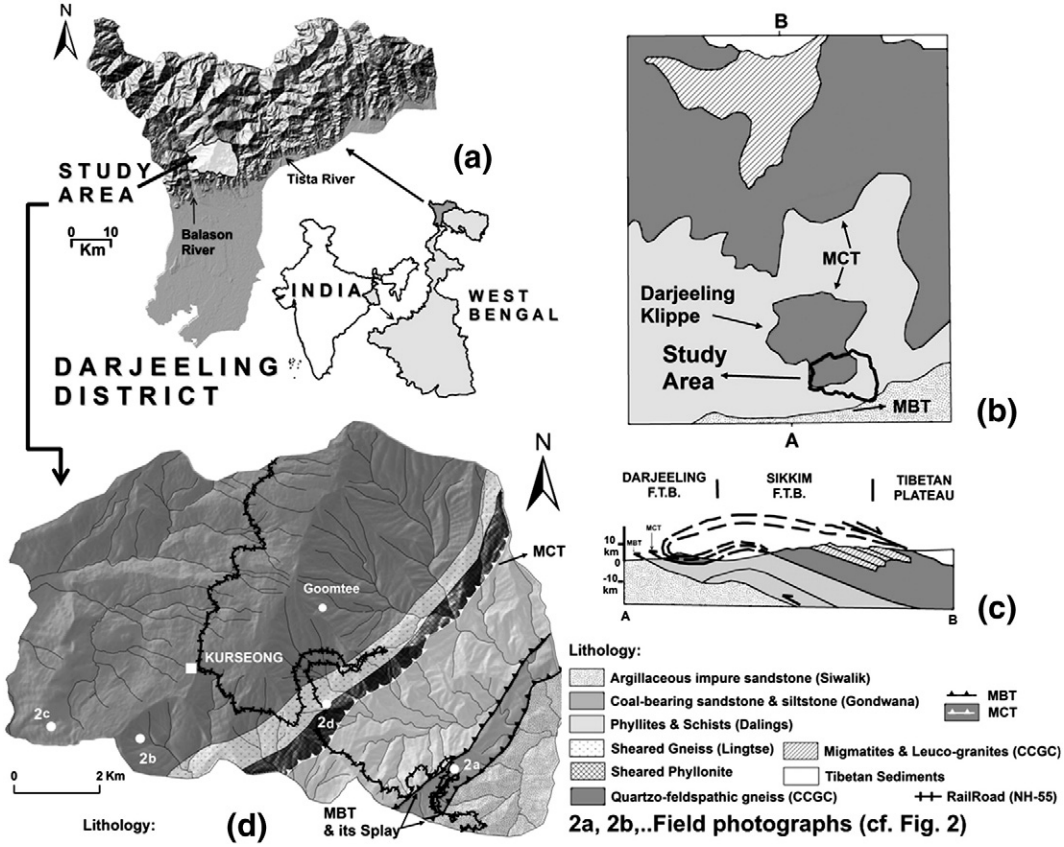

Fig. 1. Location of the study area. (a) Geographic location of the study area. (b) Schematic regional geological map and (c) section of Darjeeling-Sikkim Himalayas (after Searle and Szule, 2005). (d) Geological map of the study area showing various lithologies and trends of regional thrust planes — Main Central Thrust (MCT) and Main Boundary Thrust (MBT). The map in (d) also shows prominent places, railroad and locations of field photographs shown in Fig. 2.

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

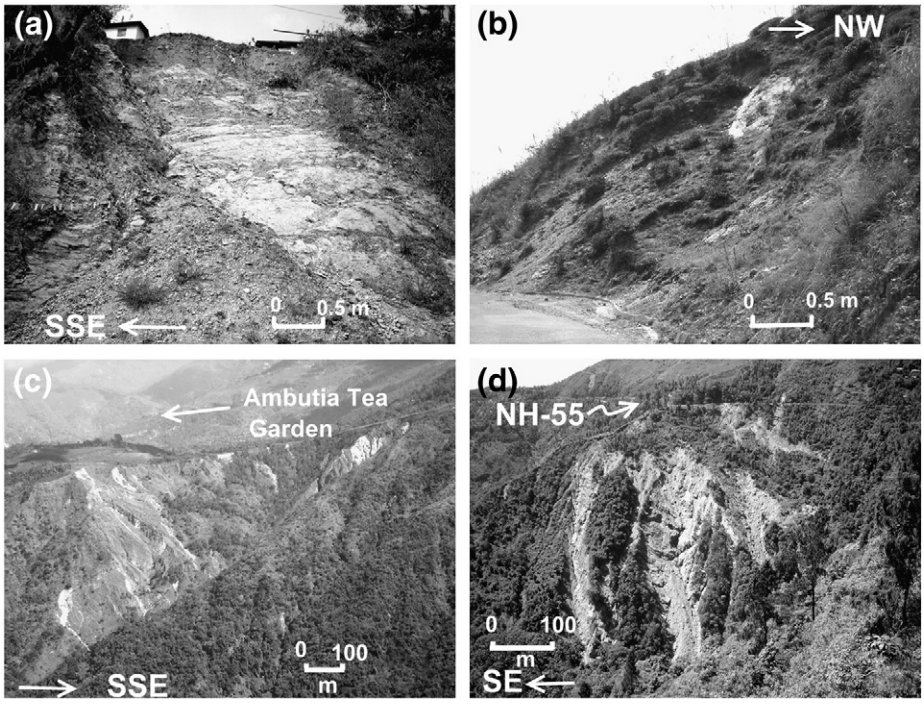

Fig. 2. Examples of landslide-types in different lithologies of the study area. (a) A shallow translational rockslide in Gondwana sediments. (b) A shallow translational debris slide in quartzo-feldspathic gneiss (CCGC). (c) A large landslide complex containing both shallow translational and deep-seated rock slides in quartzo-feldspathic gneiss (CCGC) (known as Ambutia landslide complex). (d) A deep-seated rockslide (14 Mile Slide) in the Paglajhora slide complex. For locations of the above field photographs see Fig. 1d. Scale bars shown are approximate.

3. Generation of event-based landslide inventory

3.1. Source data

The source data that were used in this study for landslide inventory mapping consisted of high-resolution satellite images, aerial photographs, topographic maps, old landslide inventory maps and reports of field investigations (Table 1). In this section these data sets are described in chronological order. The oldest data set consists of topographic maps prepared by the Survey of India (SOI) in 1969, one year after a major rainfall event that occurred between 2 and 4 October 1968, with a rainfall amount of 1100 mm (Basu and De, 2003). During 1969, the SOI updated their topographic survey and prepared new 1:25,000 topographic maps which included the locations of prominent and active landslides of the 1968-event. The next data set consists of 1:50,000 and 1:10,000 scale black-and-white stereo-air photos from 1980. The third set is a field-based landslide inventory map from 1993 prepared by the Geological Survey of India (GSI) soon after a landslide event that occurred between 1 and 3 July 1993 (Sengupta, 1995). Unfortunately, this map covers only 56 km^2 in the southeastern part of the study area. The fourth data set is another event-based landslide inventory map prepared by GSI after a landslide event that occurred between 6 and 8 July 1998. Also this inventory map covers only a part of the area (~20 km^2) in the center of the study area, and does not overlap with the previous inventory map (Bhattacharya et al., 1998). For the period 2002 to 2006 four high-resolution Indian satellite images were available (Table 1). The most recent data set is a detailed landslide inventory map prepared by the first author through field surveys soon after a recent landslide event in 2007.

3.2. Landslide inventory mapping

Since the source data vary in resolution, type and extent, different methods were adopted for extracting landslide information. For each period, a landslide inventory was made, and landslides were mapped as polygons, with a separation of their erosional and depositional area (preferably for large landslides). Each landslide was characterized by attributes such as movement type, material involved, activity, depth, failure mechanism, and date of occurrence, according to international standards (cf. Varnes, 1978; UNESCO-WP/WLI, 1990, 1993). We considered a depth-to-failure surface of 5 m as the main criterion to differentiate between deep-seated and shallow landslides. The source data of the 1968-event (topographic map sheets of 1969) contain no information on landslide types, depth and failure mechanisms. Thus, we used the 1980 airphotos to stereoscopically examine and characterize the 1968-landslides (LI68).

Stereoscopic interpretation of the 1:10,000 and 1:50,000 scale black-and-white stereo airphotos from 1980 was carried out using a mirror stereoscope, and landslide types were identified based on texture, tone of photo-elements, association of objects, morphometry, depth and freshness of scarps. We also used the airphotos from 1980 to map large scarps and depositional areas representing old/inactive landslides,

Table 1
Types, extent and period (DOA = date of acquisition) of source data and associated rainfall events for landslide inventory (LI).

Source data (scale/resolution)	Extent (km^2)	DOA/period	Rainfall event	LI period
Topographic map (1:25,000)	90	1969–1970	Oct. 1968	LI68
B×W stereo airphotos (1:50,000 and 1:10,000)	90	1980	1979	LI79
Field-based landslide inventory map (1:25,000)	56	1993	July 1993	LI93
Field-based landslide inventory map (1:25,000)	20	1998	July 1998	LI98
IRS 1-D PAN image (5.8 m)	90	2000	Unknown	LI99-02
IRS 1-D LISS-3 MX image (23.5 m)	90	2002	Unknown	LI99-02
IRS P-6 LISS-4 MX (5.8 m)	90	2004	July 2003	LI03
IRS P5 Stereo Cartosat – 1 (2.5 m)	90	2006	Unknown	LI04-06
Field-based landslide inventory map (1:25,000)	90	2007	July/Sept. 2007	LI07

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

ARTICLE IN PRESS

which occurred prior to 1968 and that cannot be linked to any known landslide event.

The landslides from the field-based inventory maps of 1993 and 1998 (LI93 and LI98) were converted to base map, and additional non-spatial attributes, such as depth and failure mechanism, were interpreted from later images.

Landslides were mapped from the four high-resolution satellite images of 2000, 2002, 2004 and 2006 by digital stereo image interpretation with a specialized software (Stereo-Analyst in ERDAS Imagine 9.x and ILWIS 3.3) using ortho-rectified images and a digital elevation model (DEM), derived from stereo Cartosat-I images of 2006. Within the period of 2000 to 2006 only one major landslide event was known that occurred during the monsoon of 2003. The landslides for this event (LI03) were mapped by comparing the images of 2002 and 2004. By comparing the landslides mapped for the period 1968 to 1998 with those in the images of 2000 and 2002, we generated an inventory of landslides that occurred between 1999 and 2002 (LI99-02). Unlike the other inventories (e.g. LI68 and LI79), this inventory lacks the exact date of the landslide triggering event. Following a similar method, we compared further the landslides mapped from the images of 2006 and 2004 to generate another landslide inventory (LI04-06), also lacking a specific landslide event-date. After a major landslide event that occurred during the monsoon of 2007, we mapped all associated landslides (LI07) by field survey.

We also gathered information about four landslide event years (1984, 1985, 1986 and 1991) from old geological reports and interviews with local people, for which no post-event landslide maps were available. A summary of all known landslide events and the event-based inventories is given in Fig. 3. All landslide inventories were compiled in a GIS using UTM projection parameters.

4. Analysis of the landslide inventories

4.1. Landslide types and frequencies

The eight event-based landslide inventories were compared to examine variations in landslide characteristics (frequency, area, density, type and activity). We spatially compared landslides associated with different events via buffer analysis in GIS to identify likely reactivated landslides. Landslides from a later inventory were considered to be reactivated ones if they were located within 50 m of landslides from an earlier one, otherwise they were considered as new landslides. We adopted the buffer distance of 50 m based on our field experiences and scale of mapping (1:25,000 or smaller).

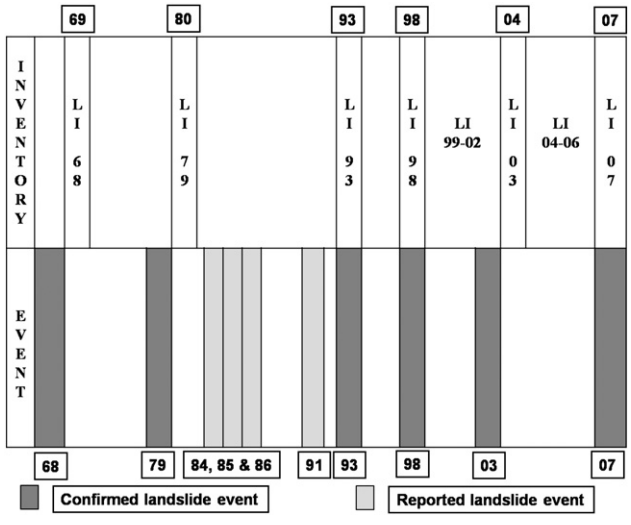

Fig. 3. Known landslide event-years and landslide inventories (LI). Gaps in inventory for the 1970–1978 and 1981–1992 periods are shown.

The descriptive statistics for the eight event-based landslide inventories are presented in Table 2. The landslide density (number per km^2) shows a variation through time, with the highest density in 1979 (6.2). The medians of areas of mapped landslides per event-based inventory vary between 519 m^2 (LI79) and 3385 m^2 (LI68), excluding the large undated landslides that were already present before 1968 (Table 2). Those large landslides older than 1968 occupy about 21% of the study area and are mostly deep-seated rockslides with areas that range up to 2.14 km^2. Since the inventory of pre-1968 landslides is not event-based, we did not compare their density and dimension with the landslides in our eight event-based inventories.

Between 1968 and 2007, shallow translational rockslides appear to be the most frequent slope movements, followed by shallow translational debris slides (Table 2). No analysis of shallow translational debris slides is made for LI98 because it only contains 31 debris slides, which were exclusively mapped along a road-cut slope, and landslides of other types occurring on natural slope were not mapped for this event. Compared to the above two landslide types, the density frequency of deep-seated rockslides is rather low, although they occupy larger areas (Table 2). Fig. 4 shows the overall landslide inventory for the study area, with indication of the most recent date of landslide activity.

4.2. Landslide activity analysis

Table 3 shows the number and area of both reactivated and new landslides in each of the landslide inventories. For the inventory LI79, the analysis is done with respect to the LI68 inventory. One has to take into account that two of the inventories (LI93 and LI98) do not cover the entire study area. Therefore it is better to evaluate the relative numbers (e.g. percentage of landslides) rather than the absolute numbers. The inventory of 1979 (LI79) shows the highest absolute number of landslides (562) and the highest percentage of reactivated landslides (90%). This is partly due to the effect of the large difference in source data from which the inventories are derived, as the inventory of 1979 was obtained from 1:10,000 scale airphotos and the one from 1968 from a generalized topographic map. The relative number of new landslides does not show a clear trend. There are years when the fraction of new landslides increases, periods when it remains constant. The event of 2003 is clearly an outlier, as all 185 landslides that were mapped in the previous event (LI99-02) were reactivated. The percentage of deep-seated rockslides that are reactivated is higher than the percentage of reactivated shallow landslides, and ranges from 54% (LI79) to 100% (LI03). On average 61% of the recent landslides that occurred in the period 1968 to 2007 were located either within or very close to the deep-seated large landslides that were already present before 1968 (see Figure 4).

We calculated the landslide density within terrain units to study changes in landslide activity. A map of terrain units was obtained via a semi-automatic technique for delineating ridge crests, streams and spur axes using a 10 m × 10 m digital elevation model (DEM) (cf. Carrara et al., 1991). Since both LI93 and LI98 inventories cover only small and separate parts of the study area (see Table 2), we merged them into a single inventory for the purpose of this analysis. In view of mapping and digitization errors, we considered terrain units with landslide density equal to or less than 2 to be stable (cf. Galli et al., 2008).

Fig. 5 illustrates the variation in area-based landslide density per terrain unit during the 40-year period. Table 4 lists the number of unstable terrain units in different periods. Initially 44% of the terrain units (i.e. 443 out of 1001 terrain units) are covered by the large pre-1968 landslides. There is a steady increase of unstable terrain units from 1968 (66) to 2007 (301). About 68% (or 205) of these 301 unstable terrain units in 2007 are located within/proximal to pre-1968 landslides. The number of new terrain units affected by landslide activity is decreasing from 1980 (94) to 2007 (8) (Table 4). The number of terrain units with increasing landslide activity is highest in 2004 (106) but the number reduces in 2007 (31). However, about 87% of

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

Table 2
Summary statistics of landslides in individual event-based landslide inventory (NA = not available).

Landslide statistics	Pre-1968 landslides	Event-based landslide inventory (LI)							
		LI68	LI79	LI93	LI98	LI99-02	LI03	LI04-06	LI07
Area of inventory (km^2)	90	90	90	56	20	90	90	90	90
Number of landslides	200	83	562	108	31	185	242	164	85
Landslide area (km^2)	9.5	0.49	0.64	0.5	0.05	0.84	1.18	0.65	0.11
Minimum slide area (m^2)	192	776	45	372	185	271	221	45	42
Maximum slide area (m^2)	2,141,500	70,253	55,815	40,906	9573	79,157	92,155	119,285	8265
Mean area (m^2)	101,455	5986	1136	4634	1713	4525	4898	3985	1357
Median area (m^2)	20,345	3385	519	2616	824	2301	1866	732	628
Number density (Nr/km^2)	2.2	0.9	6.2	1.9	1.6	2.1	2.7	1.8	0.9
Area density (%)	10	0.5	0.7	0.9	0.25	0.93	1.31	0.67	0.12
Shallow translational rock slides									
Total landslides (Nr)	0	59	374	86	NA	123	167	116	63
Landslide area (km^2)	0	0.22	0.32	0.28	NA	0.31	0.36	0.17	0.08
Shallow translational debris slides									
Total landslides (Nr)	0	13	175	13	31	48	53	34	22
Landslide area (km^2)	0	0.04	0.12	0.05	0.05	0.16	0.17	0.04	0.03
Deep-seated rock slides									
Total landslides (Nr)	200	11	13	9	NA	14	22	14	0
Landslide area (km^2)	19	0.23	0.20	0.17	NA	0.37	0.65	0.44	0

the unstable terrain units in 2007 appear to have a constant landslide activity (meaning that the absolute values of area based landslide density in such terrain units remained unchanged), as those in the inventory for the immediately preceding period.

5. Estimation of temporal probability of landslides

Because of the incompleteness of the landslide inventories we could not use them independently for a direct calculation of the temporal probability of landslides. Instead, we used the available information about the known days (or events) of landslides to establish a link with days (or events) of extreme rainfall (cf. Gabet et al., 2004; Petley et al., 2007; Dahal and Hasegawa, 2008). Unlike methods of linking rainfall and landslides using a bivariate linear relations between daily and an antecedent rainfall amounts (Chleborad et al., 2006; Jaiswal and van Westen, 2009), we preferred to use a multivariate classification technique (discriminant analysis, hereafter denoted as DA) to predict landslide-triggering events. The latter method facilitated us to objectively understand the possible non-linear relationship that exists between triggering rainfall events and various rainfall predictors, such as daily rainfall as well as a set of different antecedent rainfall amounts. The frequencies of such predicted landslide events were then used in a Poisson distribution model to calculate the temporal probability of similar landslide events in the future.

5.1. DA modeling of relationship between landslide events and triggering rainfall events

The aim of the discriminant analysis was to define a relationship between days with and without landslide events as response variable and daily variation in rainfall amounts as predictor variables. We could then predict landslide event-days by using thresholds derived from the DA model. For the DA modeling, we selected a rainfall station that is centrally located in the study area (Goomtee Tea Garden's rain gauge station, see Figure 1) and used its daily rainfall data (in mm) for the same 40-year period as our landslide inventories (1968–2007). For DA, we considered various rainfall indicators such as daily rainfall (*DR*) and different antecedent rainfall amounts (1, 2, 3, 5, 7 and 10-day antecedent rainfall, hereafter referred to as AR_1, AR_2, AR_3, AR_5, AR_7 and AR_{10}, respectively) as predictors or explanatory variables (cf. Dai and Lee, 2001). As response or grouping variable, we used the known days (or events) of landslides (24 days) in the 40-year period. To each of those 24 days of landslides, we assigned a Landslide Occurrence Score (LOS) of "1", and to all other days we assigned a LOS of "0". Since landslides in the study area were triggered only during monsoon rainfall, we used only the rainfall data in the monsoon period (June to October) for each year, resulting in a data set of 6120 days. To calibrate the DA model, we randomly selected 4862 (or about 80%) of these 6120 days, of which 20 days have LOS of 1. The other 1258 (or about 20%) days, of which 4 days have a LOS = 1, were kept aside for cross-validation of the DA model.

The DA model with seven rainfall predictors (*DR*, AR_1, AR_2, AR_3, AR_5, AR_7 and AR_{10}) resulted in a statistically significant discriminant function (Wilk's Lambda = 0.938 and significance level = 0.000), which was able to explain all the variances of the data in the model and successfully classified 80% of known landslide event-days (LOS = 1). The group centroids of the DA model have wide separation between the two response groups (LOS 0 = −0.017 and LOS 1 = 4.005) suggesting that the separation potential of the discriminant function between the two groups is strong. The highest coefficient value of the discriminant function pertains to daily rainfall (*DR* 0.849), followed by AR_1 (0.383), AR_2 (0.271) and AR_5 (−0.338) respectively, signifying their ability to classify the dichotomous response variable (i.e., LOS = 1 and LOS = 0). The predictors AR_3, AR_7 and AR_{10} are non-correlated with the discriminant function; and, therefore, contribute nothing to the model classification. The structure matrix of the discriminant function further reveals that both AR_7 (0.247) and AR_{10} (0.237) are highly correlated with AR5 (0.267) and the model statistics used in our DA model ("*F* to enter as 3.84" and "*F* to remove as 2.71") successfully resulted in the removal of AR_7 and AR_{10} from the analysis. The variables remaining at the final step of the DA are reasonably uncorrelated as depicted by the values of correlation in their structure matrix (*DR* = 0.886, AR_1 = 0.542, AR_2 = 0.480 and AR_5 = 0.267), meaning that they ultimately satisfy the basic assumption of independence among predictor variables in DA. Based on the results of the above DA model, the following equation was derived to calculate DS for each case:

$$DS = -0.637 + (0.021DR) + (0.01AR_1) + (0.004AR_2) - (0.003AR_5). \quad (1)$$

The DA model successfully classified 80% (or 16) of the 20 known landslide event-days and 95.2% (4610) of the known non-landslide event-days (4842) in the calibration set, which indicates that the

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

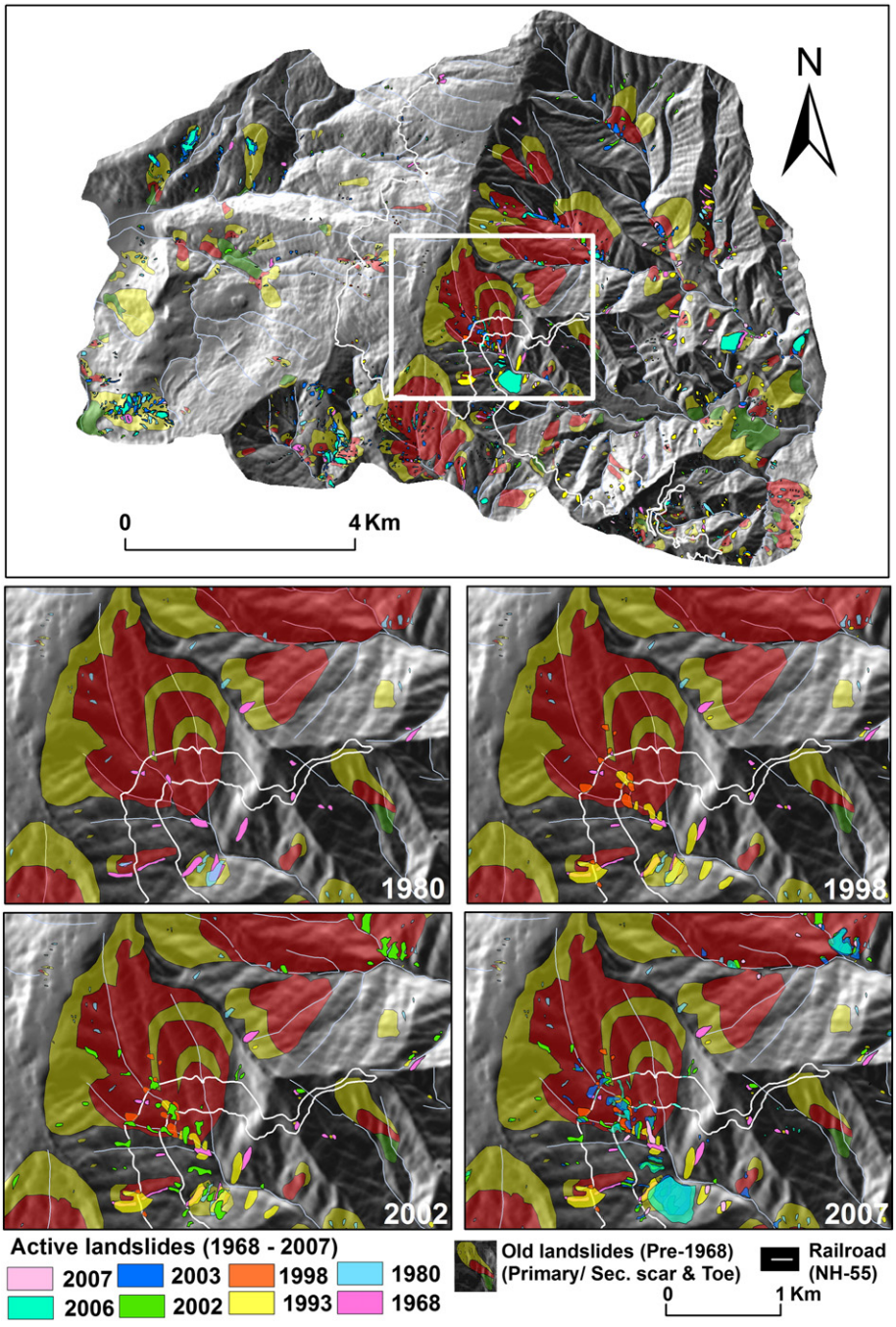

Fig. 4. Multi-temporal landslide inventory map.

overall success rate of the model is 95.1% (100*(4610 + 16)/4842) (see Table 5). The validation of the DA model is obtained from the classification of cross-validation cases (1258) that were not used to calibrate the model. The DA model successfully classified 100% (four) known landslide event-days and 95.4% (1196 out of 1254) of the non-landslide event-days in the validation data set, which indicates that the overall prediction rate of the DA is 90.8% (see Table 5).

5.2. Prediction of landslide events using thresholds of DA function scores

We used discriminant function scores (*DS*) derived from our DA model (Eq. 1) as a threshold to identify the predicted landslide event-years with varying degrees of severity. We used three different thresholds: four, ten and twenty. A discriminant score of four was selected to separate years with no landslide events from those with landslide events. Years with a cumulative discriminant score (DS) between four and ten were classified as "minor landslide event year", those with a DS between 10.0 and 20.0 as "moderate landslide event year" and those with a DS above 20.0 were designated as "major landslide event year". These thresholds were subjectively determined by examining and comparing the cumulative *DS* of the confirmed landslide events (e.g. LI68 and LI93) with the severity and damage potential of the corresponding landslide events, as expressed by the landslide density in the terrain units. In Fig. 6, we show that for the period between 1968 and 2007 there are 30 predicted landslide event-years with a cumulative DS above the minimum threshold of four. The LI68 event in Darjeeling Himalaya with DS of 21.89 (Figure 6) corresponds to the most severe landslide event year (1968) within the analyzed period. During this event a very large number of landslides occurred throughout the entire Darjeeling-Sikkim Himalayas, that

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

Table 3
Frequency (Nr) and area of reactivated and new landslides in individual event-based landslide inventories (results of GIS-based buffer analysis). (NA = not mapped).

Landslide inventory (LI)	LI79	LI93	LI98	LI99-02	LI03	LI04-06	LI07
All landslides:							
Number of landslides	562	108	31	185	242	164	85
Landslide area (km^2)	0.64	0.5	0.05	0.84	1.18	0.65	0.12
Reactivated (Nr)	55	60	13	76	185	98	53
New (Nr)	507	48	18	109	57	66	32
Reactivated area (km^2)	0.2	0.32	0.03	0.52	1.10	0.59	0.09
New area (km^2)	0.44	0.18	0.02	0.32	0.08	0.06	0.03
Shallow translational rockslides:							
Number of landslides	374	86	NA	123	167	116	63
Landslide area (km^2)	0.32	0.28	NA	0.31	0.36	0.17	0.08
Reactivated (Nr)	42	44	NA	55	124	71	38
New (Nr)	332	42	NA	68	43	45	25
Reactivated area (km^2)	0.04	0.15	NA	0.17	0.30	0.14	0.05
New area (km^2)	0.28	0.13	NA	0.14	0.06	0.03	0.03
Shallow translational debris slides:							
Number of landslides	175	13	31	48	53	34	22
Landslide area (km^2)	0.12	0.05	0.05	0.16	0.17	0.04	0.03
Reactivated (Nr)	6	9	13	12	30	14	15
New (Nr)	169	4	18	36	18	20	7
Reactivated area (km^2)	0.003	0.04	0.03	0.07	0.14	0.02	0.02
New area (km^2)	0.12	0.01	0.02	0.09	0.03	0.02	0.01
Deep-seated rockslides:							
Number of landslides	13	9	NA	14	22	14	0
Landslide area (km^2)	0.20	0.17	NA	0.37	0.65	0.44	0
Reactivated (Nr)	7	7	NA	9	22	13	0
New (Nr)	6	2	NA	5	0	1	0
Reactivated area (km^2)	0.16	0.13	NA	0.29	0.65	0.43	0
New area (km^2)	0.04	0.04	NA	0.08	0	0.01	0

destabilized the infrastructure and communication network in this region for more than a year (Starkel and Basu, 2000). Another confirmed from LI93 (DS = 4.36) was comparatively the least-severe one that occurred, causing a limited number of landslides in a few localized areas. Accordingly, in Fig. 6, we have classified 16 event-years as "minor", 10 event years as "moderate" and 4 event years as "major" within the 40 year period (1968–2007). Among the confirmed events, LI93 is predicted as a "minor" event, LI79, LI03 and LI07 as "moderate" events and both LI68 and LI98 as "major" events.

5.3. Calculation of exceedance probability

We estimated the mean landslide recurrence interval (μ) by dividing the period of analysis (40 years) by the number of predicted landslide event-years (major, moderate and minor). We later used the estimated μ to calculate the exceedance probability of landslide events by using the Poisson distribution model (cf. Crovelli, 2000; Coe et al., 2004). In the Poisson distribution model, exceedance probability or probability of experiencing one or more landslide events during a period 't' can be estimated by the following equation:

$$P[N_L(t)\geq 1] = 1 - P[N_L(t) = 0] = 1 - e^{-\lambda t} = 1 - e^{-t/\mu} \quad (2)$$

where $P[N_L(t)\geq 1]$ is the exceedance probability or probability of occurrence of one or more landslide events in period t and λ is the reciprocal of the mean recurrence interval (μ), that is,

$$\lambda = \frac{1}{\mu}. \quad (3)$$

A graph of exceedance probability against period (in years) as illustrated in Fig. 7 can be used as an indication of the temporal probability of landslide events. For example, within a 10-year period, the exceedance probability of "moderate" and "minor" predicted events are very high (0.91 and 0.98 respectively) compared to a "major" event (0.63).

6. Estimation of landslide magnitude probability

6.1. Landslide magnitude–frequency analysis

The quantitative estimation of the probability of occurrence of landslides of a given magnitude is a key issue for any regional landslide hazard assessment method (Fell et al., 2008). In the magnitude–frequency analysis performed, we considered the area of landslide (m^2) as a proxy for landslide magnitude (cf. Aleotti, 2004; Dapporto et al., 2005; Guzzetti et al., 2005). We applied frequency–size analysis of landslide area to all the eight event-based landslide inventories covering the period 1968 to 2007. We performed this analysis by calculating the probability density function (hereafter, denoted as pdf) of landslide area using the maximum likelihood estimation method (Fisher, 1922a, 1922b) assuming two distribution functions: (i) the Inverse-Gamma distribution function (Malamud et al., 2004), and (ii) the Double-Pareto distribution function (Stark and Hovius, 2001). For each inventory we further calculated raw probability density using the histogram estimation method and considering logarithmic bin width. Starting from the estimated probability density functions, we further calculated the cumulative density function (hereafter, denoted as cdf). The above estimation of landslide area–frequency distribution was performed using a script developed in R (R Development Core Team, 2010), a free software environment for statistical computing and graphics (http://www.R-project.org). The R package "bbmle" was used in the script to implement the maximum likelihood estimation method.

The results of the magnitude–frequency analysis of landslide areas are shown in Figs. 8, 9 and 10 and in Table 6. Column A in Figs. 8 and 9 shows the box plots of landslide area; column B shows the estimates of the Inverse-Gamma pdf and column D, the estimates of the Double Pareto pdf. Columns C and E show the cumulative probability through cdfs calculated considering the corresponding Inverse-Gamma and the Double Pareto fitted distributions (columns B and D) respectively. Table 6 gives the related values for the slope of the pdfs (α) and the rollover points, where the distribution changes direction for smaller landslides (in m^2). The assumption in the analysis was that larger landslide events would show a higher value for the rollover point, whereas the slope of the right part of the curve (α) would also be slightly steeper. The value for α as reported by Malamud et al. (2004) in their Inverse-Gamma distribution (which they indicate as ρ) is 1.40 for cumulative distribution, and for non-cumulative distribution as 2.40 ($\alpha+1$). They reported the rollover at 1280 m^2. Stark and Hovius (2001) presented equivalent α values using a Double Pareto distribution ranging between 1.11 (for medium and large landslides in Taiwan) and 1.48 (for large landslides in New Zealand). From the data in Table 6 and Fig. 10, we can conclude that the inventories show different values for the rollover point and for α. From the magnitude–frequency distributions illustrated in columns B and D of Figs. 8 and 9, we infer that except for LI93, LI98 and LI07, the rest of the inventories can satisfactorily be described by both standard distribution functions available in the literature. This could be an indication that those inventories are more or less complete for landslide areas that are greater than the rollover sizes, and their respective cumulative probability estimates (cdfs in columns C and E) can be used for the estimation of the magnitude of different landslide events. However, of the eight inventories, only 5 were able to actually map individual triggering events (see Table 7). Inventory LI79 is made from a very good source data; in fact it is the only inventory for which large scale (1:10,000) stereo aerial photographs were available. However, it is the first data source available since 1968, and therefore the landslides that were mapped might have been produced by several triggering events between 1968 and 1979 (see also Figure 6). The inventories

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

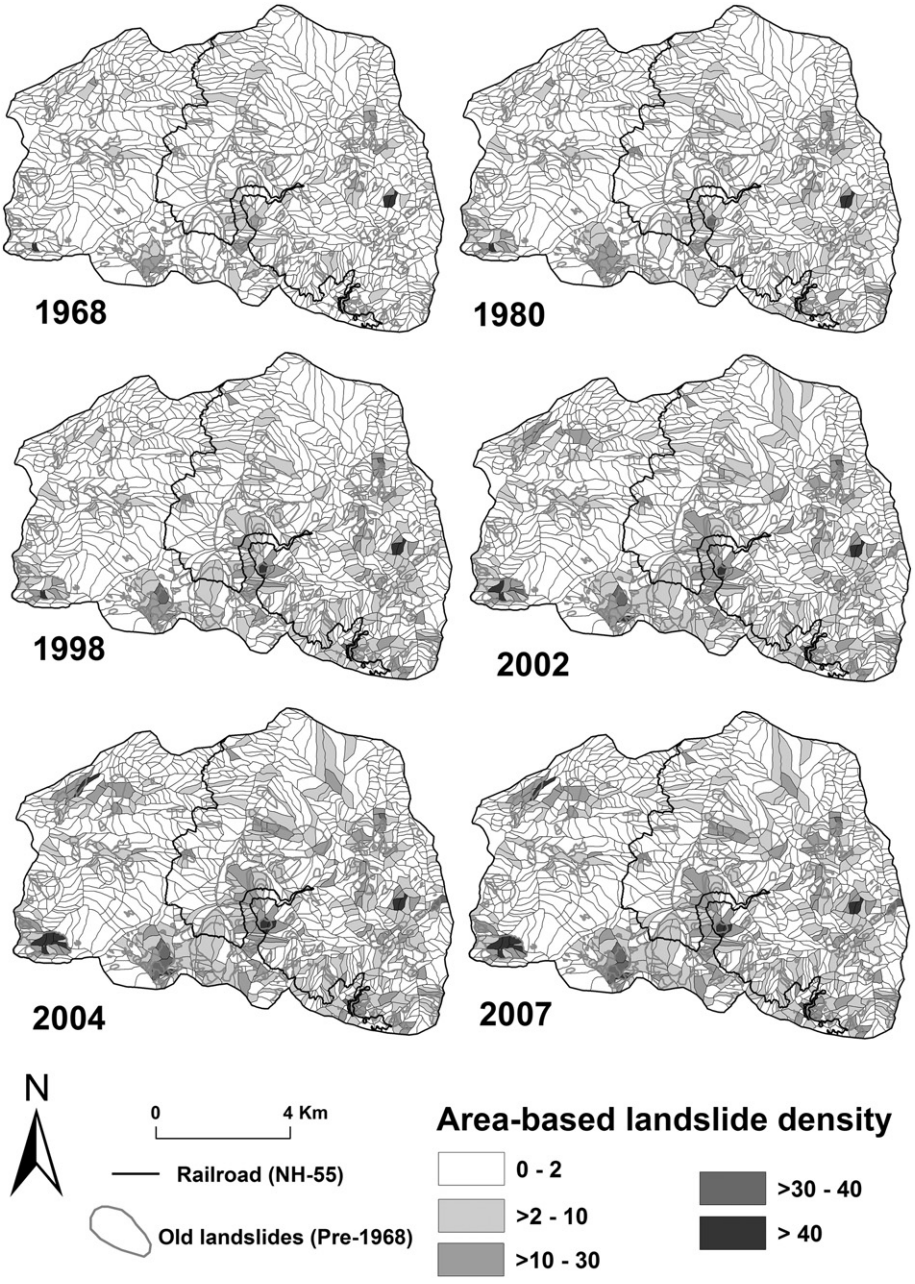

Fig. 5. Distribution of area-based landslide density per terrain unit (landslide area × 100/area of terrain unit) for periods from 1968 to 2007.

LI99-02 and LI04-06 also might include several triggering events, as no data was available for the intermediate years.

The cumulative distribution curves shown in columns C and E allow determining the probability to have an area smaller or larger than a particular size, which can be used as the quantitative estimate for landslide size (or, magnitude) probability (Guzzetti et al., 2005). For an example, based on the LI68 event and the Inverse-Gamma fit of the landslide area–frequency data, the probabilities that a future landslide event of a similar magnitude could result in landslides that have an area larger than 1000 m^2 or larger than 10,000 m^2 are 0.98 and 0.15 respectively (Figure 8 and Table 6). The same estimate, based on the Double Pareto fit of LI68 data gives somewhat similar probability values (0.95 and 0.12 respectively) (Figure 8 and Table 6). Table 6 presents the probability of future landslide sizes based on the

Table 4
Temporal trend of landslide activity in affected terrain units. There are 1001 terrain units in the study area.

Inventory year	Number of unstable terrain units			
	Total	With new activity	With increased activity	With constant activity
1968	66	–	–	–
1980	160	94	27	39
1998	206	46	56	104
2002	259	53	61	145
2004	293	34	106	153
2007	301	8	31	262

Table 5
Classification results of DA model.

			Predicted		Total
			No landslide	Landslide	
Cases for model calibration	Original count (days)	No landslide	4610	232	4842
		Landslide	4	16	20
Cases for model validation	Original count (days)	No landslide	1196	58	1254
		Landslide	0	4	4

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

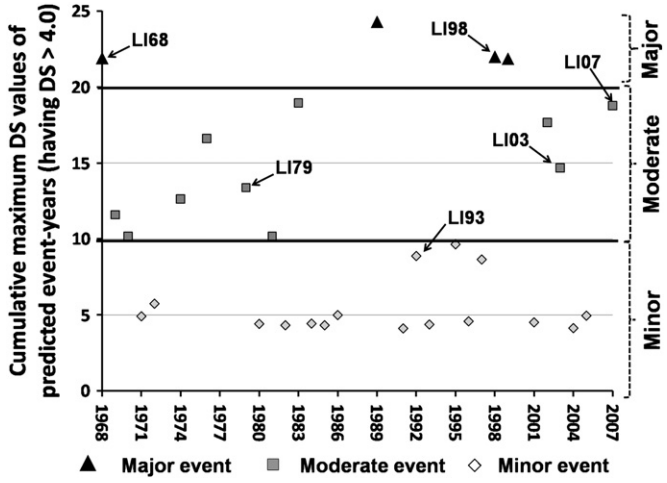

Fig. 6. Threshold cumulative discriminant scores (DS) for predicted "minor", "moderate" and "major" landslide events. The confirmed six events are also shown.

magnitude–frequency analysis of each of the eight landslide inventories, for small (smaller than 1000 m^2) and large landslides (larger than 10,000 m^2). Our aim was to make an assessment of landslide size probabilities for the three magnitude events, as obtained from the rainfall analysis (major, moderate and minor events as shown in Figure 6). However the results from the frequency–size analysis do not support this very well, as can be seen from Table 6 and Fig. 10. For major landslide events, such as the one from 1968, the magnitude–frequency analysis of LI68 indicates that the probability of landslides larger than 10,000 m^2 is as high as 0.15. The rollover point of this curve is also the highest of all inventories (1682 m^2). However, values with a similar order of magnitude are obtained for LI93 inventory, which according to the classification in Fig. 6 were rated as a minor triggering event.

Given the problems with the differences in the source data, with respect to their aerial and temporal coverage and resolution of their source data which are summarized in Table 7, we also made a frequency–size analysis using the combination of the 8 inventories as a single inventory (figures in columns A–E along the last row of Figure 9). The results shown in Figs. 9–10 and Table 6 for the combined inventory (α ranging from 1.09 to 1.17 and rollover point from 218 to 300 m^2) are on the lower range of those reported by Malamud et al. (2004) and Stark and Hovius (2001). Therefore, given the problems

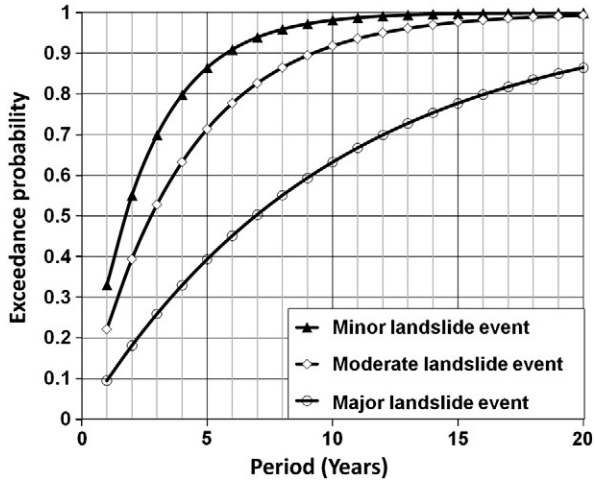

Fig. 7. Exceedance probability of predicted major, moderate and minor landslide events, resulting from the application of the Poisson distribution model.

involved in the generation of the individual inventories, it is better to use the size probability values derived from the overall inventory. Based on the Inverse-gamma fit of the overall inventory data, the probabilities of landslides with an area larger than 1000 m^2 and larger than 10,000 m^2 are 0.50 and 0.08 respectively, which, based on the Double Pareto fit also give similar probability values (0.52 and 0.08 respectively).

7. Discussion and conclusions

For quantitative landslide hazard analysis, an estimation should be made of the spatial, temporal and magnitude probabilities for future landslide events. The reliability of such estimations depends on the completeness of these event-based landslide inventories, which are only as good as the source data from which historical landslide information is extracted. The results that have been presented in this paper on the temporal and magnitude probabilities indicate that several of the landslide inventories that were used, are not complete. The generation of event-based landslide inventory maps is not a trivial task in areas where source data vary not only in resolution but also are constrained by irregular gaps between two successive events. When there is a large time gap between two successive data sources (e.g. LI68 and LI79), the landslides that have been triggered by several events are all mapped together. Therefore the resulting landslide inventory might contain too many landslides in relation to the last triggering event. This was the case for the inventory from 1979, which according to Fig. 6 actually contains the landslides caused by six moderate landslide events that occurred between 1968 and 1979. Given the unavailability of source data in the intermediate period makes it difficult to prove this phenomenon of oversampling of landslides. The quality of the event-based landslide inventories also depends very much on the quality of the input data. For instance, the landslides caused by the major event of 1968 were mapped from an updated topographic map. However, this inventory was far from complete, and many small landslides were not included. Therefore the 1968 landslide inventory contains much less landslides than could be expected given the importance of the triggering event (Figure 6). The landslide inventory of 1979, however, was made using large scale aerial photographs from which many more landslides could be mapped. This is an inherent uncertainty in landslide mapping, generally evidenced by inconsistent landslide density values of successive events.

Since our derived landslide inventory maps suffer from problems related to the quality and resolution of source data from which they are mapped, often we get such inconsistent landslide density estimates. The increased values of landslide densities (6.2 and 2.7) of LI79 and LI03 respectively are thus attributed to the ease in identification of smaller landslides (e.g., LI79, LI03, and LI04-06) due to the availability of high-resolution source data, so that smaller landslides with areas as small as 45 m^2 (LI79 and LI04-06) could be easily mapped. On the other hand the inventory LI68, which was derived from 1:25,000 scale topographic maps, only contains landslides with areas larger than 776 m^2. The smaller landslides were not on the source data and therefore this inventory represents a landslide density (0.9) that is too low given the importance of the event.

This inconsistency in landslide density caused problems in the direct utilization of landslide density values for the estimation of temporal probability as presented by Guzzetti et al. (2005). Therefore, we adapted an indirect method for the estimation of temporal probability (cf. Dai and Lee, 2001), assuming rainfall as the only landslide trigger. The importance of rainfall as a trigger is confirmed by Basu and De (2003) in the study area and also in the adjacent Nepal Himalayas by Dahal and Hasegawa (2008) and Gabet et al. (2004). Thus, we considered different rainfall amounts (daily rainfall and six antecedent rainfall amounts) as possible predictors to model landslide event days using discriminant analysis (DA). If we would have had landslide inventories with consistent landslide densities, which we lack, it would have been better to apply a multinomial logistic

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

ARTICLE IN PRESS

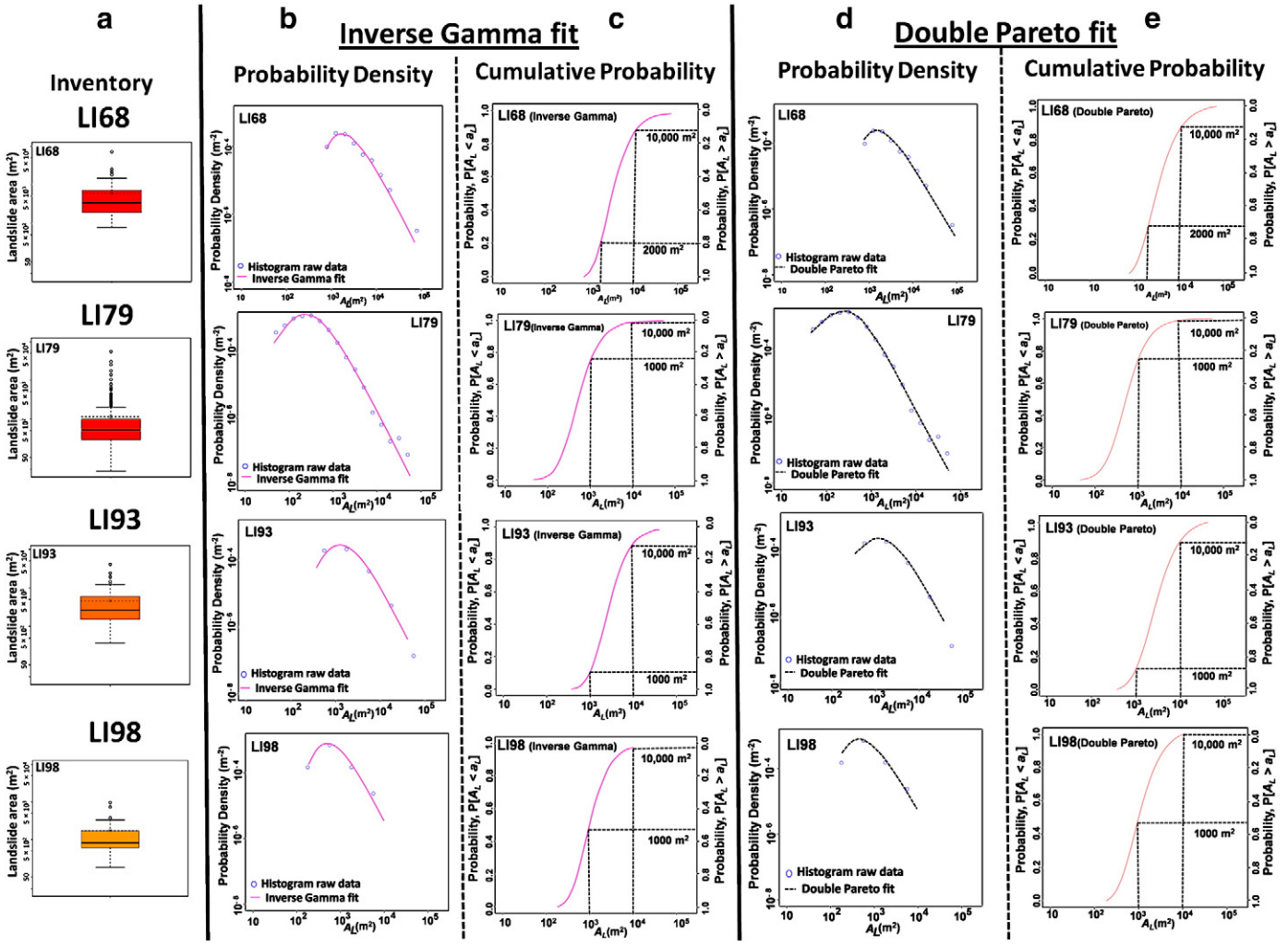

Fig. 8. Probability density functions (pdfs) of landslide area (m^2) for historic landslides of 1968–1998. (a) Box plot showing distribution of landslide area (m^2) and (b) pdfs fitted with Inverse-Gamma distribution function of Malamud et al., 2004. (c) Probability estimates as per the fitted Inverse-Gamma distribution functions and (d) pdfs fitted with the Double Pareto distribution function of Stark and Hovius, 2001. (e) Probability estimates as per the fitted Double Pareto distribution functions.

regression using ranges of landslide density estimates as multiple response variables to estimate temporal probability of events linked to different levels of severity. To overcome this problem, we used DA function scores of our known landslide events as a basis for defining thresholds, assuming a linear relation between DA function scores and the severity of events. It allowed us to differentiate three different classes of landslide events (major, moderate and minor) based on decreasing levels of cumulative DA function scores. Their recurrence intervals in 40 years (1968–2007) were used for the calculation of temporal probability of future similar events using the well-established Poisson's distribution model (Crovelli, 2000; Coe et al., 2004).

Based on this method, we established that LI68 belongs to a "major" landslide event having a cumulative DA function score of 21.89 (Figure 6), which is supported by additional information about this event from the literature (cf. Starkel and Basu, 2000). However, the landslide density of LI68 is not the highest one because of the problems mentioned earlier with the source data. In contrary, the inventory LI79 was classified as a "moderate" landslide event based on the rainfall analysis, but shows the highest landslide density (6.2) (Figure 6). Compared to all the derived landslide inventories, the predicted "moderate" event, LI03 with a landslide density of 2.7, represents a more or less realistic situation since it was made from high-resolution source data (5.8 m multi-spectral satellite image, LISS-4), and is preceded and followed by substantially complete landslide inventories, which were mapped from similar source data. Therefore, it becomes apparent from this research that it is important to reduce the inconsistency and uncertainty of landslide inventories, by analyzing the results of our magnitude–frequency analysis (Figures 8–10).

Both the fitted Inverse-Gamma (Malamud et al., 2004) and Double Pareto (Stark and Hovius, 2001) probability density functions of all the inventories show power–law area–frequency distributions for medium to large landslides in all the event-based landslide inventories and a distinct exponential rollover for smaller landslides (excepting the Double Pareto fitted pdf of LI07). This indicates that the nature of distribution of landslide area (m^2) is quite similar across all the landslide events of the study area. Due to the widely variable minimum landslide areas caused by the different quality of the source data (42 m^2 for LI07 to 776 m^2 in LI68), though, rollover estimates vary between 45 m^2 for Double Pareto fit of LI98 and 1682 m^2 for Inverse-Gamma fit for LI68 (Table 6). The widely variable rollover estimates observed in different event-based landslide inventories (Figures 8–10) indicate that these could be used together with the slope of the curve for medium to large landslide size (α) as an indication to separate distributions resulting from major, moderate and minor triggering events. It can further be inferred that the rollover phenomena are not always an artifact of censoring, where small landslides are simply not mapped (Guthrie and Evans, 2004) since we observed rollover almost in all our inventories and at different

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

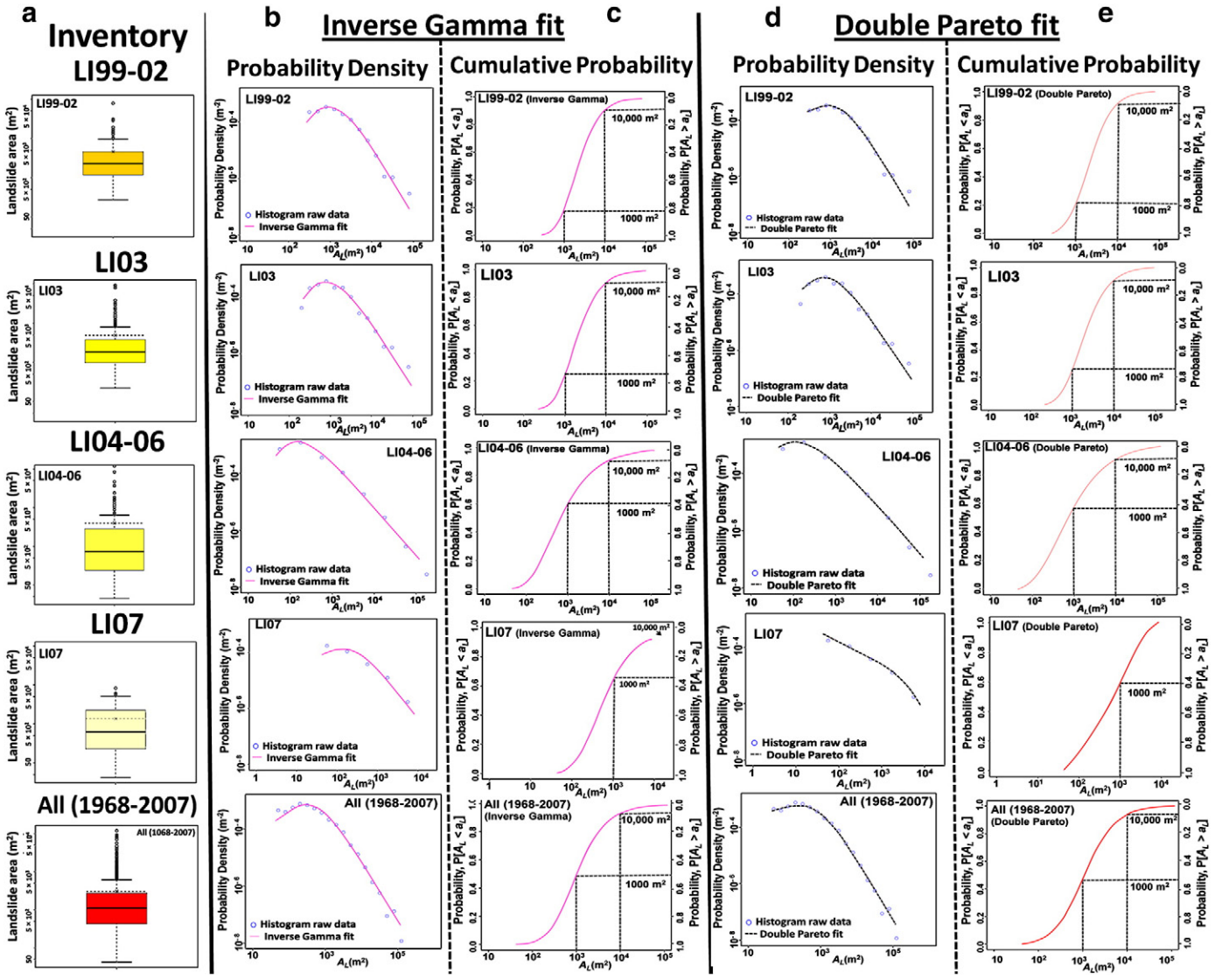

Fig. 9. Probability density function (pdfs) of landslide area (m^2) for historic landslides of 1999–2007 and for all inventories together. (a) Box plot showing distribution of landslide area (m^2) and (b) pdfs fitted with Inverse-Gamma distribution function of Malamud et al., 2004. (c) Probability estimates as per the fitted Inverse-Gamma distribution functions and (d) pdfs fitted with the Double Pareto distribution function of Stark and Hovius, 2001. (e) Probability estimates as per the fitted Double Pareto distribution functions.

rollover estimates. But nevertheless, it can also be argued that the variable rollover changes observed in the present analysis could be caused due to the use of landslide source data of variable resolutions which support the connotations of Stark and Hovius (2001) that rollover changes can be attributed to survey resolution (Brardinoni and Church, 2004).

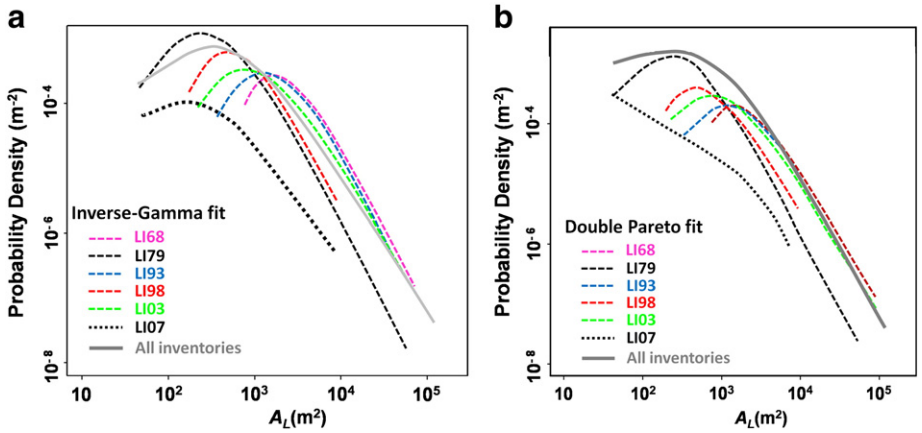

Fig. 10. Comparison of the probability density functions (pdfs) of all historic landslides that occurred in the period between 1968 and 2007. (a) Using Inverse-Gamma fit, and (b) using Double Pareto fit.

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

Table 6
Results of size–frequency analysis of landslide area for each event-based inventory. The table also demonstrates examples of two magnitude probability estimates, $P[A_L > 1000\ m^2]$: probability that a future landslide will have area larger than 1000 m^2 and $P[A_L > 10{,}000\ m^2]$: probability that a future landslide will have area larger than 10,000 m^2 (see Figures 8–9); IG = Inverse-Gamma and DP = Double Pareto; α = Power–law exponent or slope of the Power–law tail (both for IG and DP cumulative distributions); β = slope of the rollover segment of pdf; Rollover (m^2) = the area at which pdf changes direction for smaller landslides.

Inventory		LI68	LI79	LI93	LI98	LI99-02	LI03	LI04-06	LI07	All
Classification of event		Major	Mod.	Minor	Major	N.A.	Mod.	N.A.	Mod.	–
Min. area (m^2) — c (DP)		776	45	372	185	271	221	45	42	42
Max. area (m^2) — m (DP)		70,253	55,815	40,906	9573	79,157	92,155	119,285	8265	119,285
Number		83	562	108	31	185	242	164	85	1460
IG	α	1.66	1.64	1.72	1.75	1.53	1.35	0.68	0.51	1.09
	Rollover (m^2)	1682	232	1269	496	918	770	143	–	300
	$P[A_L > 1000\ m^2]$	0.98	0.25	0.88	0.52	0.82	0.72	0.39	0.32	0.50
	$P[A_L > 10{,}000\ m^2]$	0.15	0.01	0.15	0.02	0.09	0.10	0.09	–	0.08
DP	α	1.27	1.49	1.34	1.28	1.47	1.22	0.69	1.97	1.17
	β	13.72	2.58	3.51	5.91	1.84	3.31	9.61	0.23	1.46
	t	390.17	327.46	1125.83	253.50	1730.39	702	10.53	3960	828
	Rollover (m^2)	1500	241	1188	459	832	726	108	–	218
	$P[A_L > 1000\ m^2]$	0.95	0.25	0.85	0.55	0.79	0.73	0.42	0.40	0.52
	$P[A_L > 10{,}000\ m^2]$	0.12	0.01	0.15	0.0	0.1	0.10	0.1	–	0.08

Our area–frequency analysis (Table 6) of different landslide inventory maps also indicates wide variability of data as evidenced from variation in cumulative and non-cumulative power–law exponents or slopes (α). For example, according to the Inverse-Gamma fit, the cumulative area–frequency power–law exponent (α) varies from 0.51 (LI07) to 1.75 (LI98) and non-cumulative power–law exponent of the same ($\alpha + 1$) varies from 1.51 to 2.75. The lower range of the above exponent matches with the power–law exponent of cumulative power–law exponent of 0.75 ± 0.30, as observed in the frequency–volume statistic of the 1937 number of rock fall and rockslides in British Columbia (Hungr et al., 1999). Similarly the upper limit of cumulative power–law exponent ($\alpha = 1.75$) is quite comparable with the error estimate of α (1.88 ± 0.30) as observed by Stark and Hovius (2001). According to Guzzetti et al. (2002), such a variation over cumulative/non-cumulative power–law exponent (α) can be attributed to occurrence of landslides in different physiographic settings (lithology, geomorphology and geology). But causes of such a variation in a small study area (ca. 90 m^2) used for our present analysis could be different such as subtle changes in the intensities of climatic triggers, which justifies our assumptions to link the results of magnitude–frequency analysis with the different types of predicted landslide events that are linked to variation in rainfall parameters acting as triggers.

Table 7
Summary of problems associated with the various landslide inventories (N.A. = Not applicable).

Inventory	Type of event	Triggering event	Explanation
LI68	Major	Single event	Data source is a topographic map, which is updated just after the 1968 landslide event. Many small landslides are not mapped.
LI79	Moderate	Combination of several triggering event	Amalgamation of several landslide events in the period between 1968 and 1979. Only source data consisting of detailed aerial photographs (1:m), therefore many more smaller landslides could be mapped, as compared to the other inventories
LI93	Minor	Single event	The source data is a field map made after a triggering event, but it doesn't cover the entire study area.
LI98	Major	Single event	The source data is a field map made after a triggering event, but it doesn't cover the entire study area.
LI99-02	N.A.	Combination of several events	Amalgamation of several landslide events in the period between 1999 and 2002. Interpreted from satellite images.
LI03	Moderate	Single event	Interpreted from satellite images; could represent a near-complete landslide inventory since its source data is bounded by other high-resolution satellite imagery at close intervals.
LI04-06	N.A.	Combination of several events	Amalgamation of several minor landslide events in the period between 2004 and 2006. Interpreted from satellite images.
LI07	Moderate	Single event	Field-based maps; no high-resolution remote sensing data available after the event.

But nevertheless, compared to all the inventories, the magnitude–frequency distributions of LI07 can be termed a distinct outlier ($\alpha = 0.51$) (see Table 6). The reason is difficult to explain since LI07 is the most-recent inventory prepared through field investigation immediately after the event (within two months) and contains most of landslides triggered by that particular event. Some inconsistencies and incompleteness in the fitting of magnitude–frequency distributions are also noticed for LI93 and LI98, both of which are also old field-based landslide occurrence maps (Figures 8–10). This could be due to comparison of pdfs of landslide inventories prepared using two different techniques and two distinct groups of source data — one through field-based mapping and another through employing different remote sensing techniques using high resolution satellite imagery and air photos. Depending on the spatial and spectral resolutions, the maps from the latter source data depict near-exact spatial dimension of landslides, facilitating correct estimation of landslide area, whereas the field-based landslide maps, though contain a large number of smaller landslides having very high locational accuracy, but their spatial dimensions in most cases are exaggeratedly represented. Moreover, field-based inventory maps might lack landslide information from highly inaccessible areas. The above invokes inaccuracies in the correct estimation of landslide areas, which can be avoided, if landslide polygon boundaries, identified through field investigations are surveyed on large scale and later represented by down-scaling to the base maps, and the same is generally a cost-prohibitive exercise for any medium to small scale (1;25,000 and smaller) landslide hazard. Examples of the above inconsistencies in magnitude–frequency data in literature are rather rare, therefore, more studies comparing such types of inventories are required to unravel this anomaly, if any obtained also by other researchers.

The landslides that occurred before 1968 for which we do not know their date of occurrence were not incorporated in the analysis of the temporal or magnitude frequency. As can be seen from Table 2 and Fig. 4, these old landslides are generally large and cover a substantial part of the study area (10%). As we do not know the type and dates of triggering events or the landslides associated with them, we cannot

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

estimate the temporal and magnitude probabilities for extreme events. If we would have had such data over larger periods we expect that earthquakes would be very important for triggering the very large landslide complexes in the area.

This study emphasizes the importance of collecting event-based landslide maps, directly after the occurrence of landslide triggering events. The use of ancillary data sources, with different qualities and aerial coverage, creates serious problems in generating reliable event-based landslide inventory maps. The estimation of temporal probability of triggering events in such a data-scarce environment can be done by identification of unknown events through statistical analysis using triggering rainfall variables as predictors and known landslide events as response variables. Using the recurrence interval of such unknown predicted events, temporal probability of similar such events can be quantified. To estimate the probability of landslide magnitude, size–frequency analysis considering historic landslides is a valuable tool. Both temporal and magnitude probability measures, when combined together with estimates of spatial probability, can give a reasonable quantitative estimate of landslide hazard.

Acknowledgments

This work presented is part of the joint project for PhD research by Geological Survey of India (GSI), National Remote Sensing Centre (NRSC) and ITC, University of Twente, The Netherlands. We acknowledge GSI and ITC for allowing us to publish this research. We thank the local administrative authority of Darjeeling district, Government of West Bengal, India for providing us with the necessary logistics in the field. We thank the authorities at the Goomtee Tea Garden rain gauge station of Kurseong sub-division for sharing their rainfall data with us. We thank all the anonymous reviewers and editors of this paper for their valuable comments and suggestions to improve the manuscript. This research was carried out in the framework of the United Nations University — ITC School for Disaster Geo-Information Management (www.itc.nl/unu/dgim/).

References

Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. Engineering Geology 73 (3–4), 247–265.

Basu, S.R., De, S.K., 2003. Causes and consequences of landslides in Darjeeling-Sikkim Himalayas. Geographia Polonica 76 (2), 37–52.

Bhattacharya, A., Mishra, P., Ghoshal, T.B., Bahuguna, H., Ghatak, T., 1998. A geotechnical appraisal of landslides on 7th July, 1998 along National Highway No. 55. In: G.S.I. (Ed.), Progress Report.

Binnie, S.A., Phillips, W.M., Summerfield, M.A., Fifield, L.K., 2007. Tectonic uplift, threshold hillslopes and denudation rates in developing mountain range. Geology 35, 743–746.

Brardinoni, F., Church, M., 2004. Representing the landslide magnitude–frequency relation: Capilano River basin, British Columbia. Earth Surface Processes and Landforms 29 (1), 115–124.

Burbank, D.W., Leland, J., Fielding, E., Anderson, R.S., Brozovic, N., Reid, M.R., Duncan, C., 1996. Bedrock incision, rock uplift and threshold hillslopes in the northwestern Himalayas. Nature 379, 505–510.

Caine, N., Mool, P.K., 1982. Landslides in the Kolpu Khola drainage, Middle Mountains, Nepal. Mountain Research and Development 2 (157–173).

Cardinali, M., Reichenbach, P., Guzzetti, F., Ardizzone, F., Antonini, G., Galli, M., Cacciano, M., Castellani, M., 2002. A geomorphological approach to the estimation of landslide hazards and risks in Umbria, Central Italy. Natural Hazards and Earth System Sciences 2, 57–72.

Carrara, A., 1993. Potentials and pitfalls of GIS technology in assessing natural hazards. In: Guzzetti, F., Reichenbach, P., Carrara, A. (Eds.), Geographical Information Systems in Assessing Natural Hazards — Abstracts. CNR, Perugia, pp. 128–137.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., Reichenbach, P., 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth Surface Processes and Landforms 16 (5), 427–445.

Chleborad, A.F., Baum, R.L., Godt, J.W., 2006. Rainfall thresholds for forecasting landslides in the Seattle, Washington, area-exceedance and probability. USGS Open-File Report 2006–1064 http://www.usgs.gov/pubprod.

Coe, J.A., Michael, J.A., Crovelli, R.A., Savage, W.Z., Laprade, W.T., Nashem, W.D., 2004. Probabilistic assessment of precipitation-triggered landslide using historical records of landslide occurrence, Seattle, Washington. Environmental and Engineering Geoscience X (2), 103–122.

Crovelli, R.A., 2000. Probability Models for Estimation of Number and Costs of Landslides. USGS, Denver, Colorado.

Dahal, R.K., Hasegawa, S., 2008. Representative rainfall thresholds for landslides in the Nepal Himalaya. Geomorphology 100 (3–4), 429–443.

Dai, F.C., Lee, C.F., 2001. Frequency–volume relation and prediction of rainfall-induced landslides. Engineering Geology 59 (3–4), 253–266.

Dapporto, S., Aleotti, P., Casagli, N., Polloni, G., 2005. Analysis of shallow failures triggered by the 14–16 November 2002 event in the Albaredo valley, Valtellina (Northern Italy). Advances in Geosciences 2, 305–308.

Landslips in Darjeeling and neighbouring hill-slopes in June, 1950. In: Dutta, K.K. (Ed.), Bulletin of GSI. Geological Survey of India. 7–30 pp.

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., Savage, W.Z., 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology 102 (3–4), 85–98.

Fisher, F.A., 1922a. On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London, Series A 222, 309–368.

Fisher, F.A., 1922b. The goodness of regression formulae, and the distribution of regression coefficients. Journal of the Royal Statistical Society 85 (4), 597–612.

Gabet, E.J., Burbank, D.W., Putkonen, J.K., Pratt-Sitaula, B.A., Ojha, T., 2004. Rainfall thresholds for landsliding in the Himalayas of Nepal. Geomorphology 63 (3–4), 131–143.

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., Reichenbach, P., 2008. Comparing landslide inventory maps. Geomorphology 94 (3–4), 268–289.

Ghoshal, T.B., Sarkar, N.K., Ghosh, S., Surendranath, M., 2008. GIS based landslide susceptibility mapping — a study from Darjeeling-Kalimpong area, Eastern Himalaya, India. Journal of the Geological Society of India 72, 763–773.

Glade, T., Anderson, M., Crozier, M.J., 2005. Landslide Hazard and Risk. John Wiley & Sons, Ltd., Chichester, England. 802 pp.

Gupta, R.P., Kanungo, D.P., Arora, M.K., Sarkar, S., 2008. Approaches for comparative evaluation of raster GIS-based landslide susceptibility zonation maps. International Journal of Applied Earth Observation and Geoinformation 10 (3), 330–341.

Guthrie, R.H., Evans, S.G., 2004. Analysis of landslide frequencies and characteristics in a natural system, coastal British Columbia. Earth Surface Processes and Landforms 29 (11), 1321–1339.

Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31 (1–4), 181–216.

Guzzetti, F., Cardinali, M., Reichenbach, P., Carrara, A., 2000. Comparing landslide maps: a case study in the upper Tiber River Basin, central Italy. Environmental Management 25 (3), 247–263.

Guzzetti, F., Malamud, B.D., Turcotte, D.L., Reichenbach, P., 2002. Power-law correlations of landslide areas in central Italy. Earth and Planetary Science Letters 195 (3–4), 169–183.

Guzzetti, F., Cardinali, M., Reichenbach, P., Cipolla, F., Sebastiani, C., Galli, M., Salvati, P., 2004. Landslides triggered by the 23 November 2000 rainfall event in the Imperia Province, Western Liguria, Italy. Engineering Geology 73 (3–4), 229–245.

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72 (1–4), 272–299.

Guzzetti, F., Ardizzone, F., Cardinali, M., Galli, M., Reichenbach, P., Rossi, M., 2008. Distribution of landslides in the Upper Tiber River basin, central Italy. Geomorphology 96 (1–2), 105–122.

Hansen, A., 1984. Landslide hazard analysis. In: Brunsden, D., Prior, E. (Eds.), Slope Instability. John Wiley & Sons, New York, pp. 523–602.

Harp, E.L., Jibson, R.L., 1996. Landslides triggered by the 1994 Northridge, California earthquake. Seismological Society of America Bulletin 86, S319–S332.

Hubbard, M., 1996. Ductile shear as a cause of inverted metamorphism: example from Nepal Himalaya. Journal of Geology 194, 493–499.

Hungr, O., Evans, S.G., Hazzard, J., 1999. Magnitude and frequency of rock falls and rock slides along the main transportation corridors of southwestern British Columbia. Canadian Geotechnical Journal 36, 224–238.

Jaiswal, P., van Westen, C.J., 2009. Estimating temporal probability for landslide initiation along transportation routes based on rainfall thresholds. Geomorphology 112 (1–2), 96–105.

Kanungo, D.P., Arora, M.K., Sarkar, S., Gupta, R.P., 2006. A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. Engineering Geology 85 (3–4), 347–366.

Malamud, B.D., Turcotte, D.L., Guzzetti, F., Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29 (6), 687–711.

Mallet, F., 1875. On the geology and mineral resources of the Darjeeling district and the western Duars. Memoirs of Geological Survey of India 11, 1–50.

Martha, T.R., Kerle, N., Jetten, V., van Westen, C.J., Vinod Kumar, K., 2010. Characterising spectral, spatial and morphometric properties of landslides for (semi-) automatic detection using object-oriented methods. Geomorphology 116 (1–2), 24–36.

Mathew, J., Jha, V.K., Rawat, G.S., 2009. Landslide susceptibility zonation mapping and its validation in part of Garhwal Lesser Himalaya, India, using binary logistic regression analysis and receiver operating characteristic curve method. Landslides 6 (1), 17–26.

Petley, D., Hearn, G., Hart, A., Rosser, N., Dunning, S., Oven, K., Mitchell, W., 2007. Trends in landslide occurrence in Nepal. Natural Hazards 43 (1), 23–44.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org.

Rib, H.T., Liang, T., 1978. Recognition and identification: landslides-analysis and control special report, 176. National Academy of Sciences. Transport Research Board, Washington D.C. 34–80 pp.

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016

ARTICLE IN PRESS

Searle, M.P., Szule, A.G., 2005. Channel flow and ductile extrusion of the high Himalayan slab — the Kanchenjunga–Darjeeling profile, Sikkim Himalaya. Journal of Asian Earth Sciences 25, 173–185.

Selby, M.J., 1988. Landforms and denudation of the High Himalayas of Nepal: results of continental collision. Journal of Geomorphology, New Folge 69, 133–152 Supplementebande.

Sengupta, C.K., 1995. Detailed study of geofactors in selected hazard prone stretches along the surface communication routes in parts of Darjeeling and Sikkim Himalaya, Phase-I, Part-I (Rongtong-Kurseong road section). In: G.S.I. (Ed.), Annual Progress Report (F.S. 1993–94).

Sinha-Roy, S., 1982. Himalayan main central thrust and its implication for Himalayan inverted metamorphism. Tectonophysics 84, 197–224.

Soja, R., Starkel, L., 2007. Extreme rainfalls in Eastern Himalaya and southern slope of Meghalaya Plateau and their geomorphologic impacts. Geomorphology 84 (3–4), 170–180.

Sorriso-Valvo, M., Antronico, L., Gaudio, R., Gullà, G., Iovine, G., Merenda, L., Minervino, I., Nicoletti, P.G., Petrucci, O. and Terranova, O., 2004. Map of mass movement, erosion and flooding caused in south Calabria by the 8–10 September 2000 storms, CNR Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche Publication n. 2859, Rubettino Industrie Grafiche ed Editoriali, map at 1:50,000 scale (in Italian and English).

Stark, C.P., Hovius, N., 2001. The characterisation of landslide size distributions. Geophysical Research Letters 28 (6), 1091–1094.

Starkel, L., Basu, S. (Eds.), 2000. Rains, Landslides and Floods in the Darjeeling Himalaya. Indian National Science Academy, New Delhi. 168 pp.

UNESCO-WP/WLI, 1990. A suggested method for reporting a landslide. Bulletin of the International Association of Engineering Geology 41, 5–12.

UNESCO-WP/WLI, 1993. Multilingual Landslide Glossary. Bitech Publishers Ltd, Richmond. 34 pp.

van Westen, C.J., van Asch, T.W.J., Soeters, R., 2006. Landslide hazard and risk zonation—why is it still so difficult? Bulletin of Engineering Geology and Environment 65 (5), 167–184.

van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Engineering Geology 102 (3–4), 112–131.

Varnes, D.J., 1978. Slope movements types and processes. In: Schuster, R.L., Krizek, R.L. (Eds.), Landslides: Analysis and Control. Special Report 176. Transportation Research Board, National Academy of Sciences, Washington, D.C, pp. 11–33.

Wieczorek, G.F., 1984. Preparing a detailed landslide-inventory map for hazard evaluation and reduction. Bulletin of the Association of Engineering Geologists XXI (3), 337–342.

Please cite this article as: Ghosh, S., et al., Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities, Eng. Geol. (2011), doi:10.1016/j.enggeo.2011.03.016