Generation of event-based landslide inventory maps in a data-scarce environment; case study around Kurseong, Darjeeling district, West Bengal, India

S. Ghosh
Geological Survey of India, Kolkata, India

C.J. van Westen, E. Carranza & V. Jetten
International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, Netherlands

ABSTRACT: In order to be able to express the spatial and temporal probability of landslides in a regional landslide hazard zonation, it is important to have insight in the location of landslides and their behavior through time. One of the techniques used to generate hazard maps following a combined heuristic/statistical method is to generate so-called event-based landslide maps, displaying the landslides triggered by a single major rainfall event, and use the temporal probability of the event in the hazard assessment, combined with the spatial probability resulting from the statistical analysis. This paper demonstrates the generation of such event-based landslide inventory maps for a region with very limited historical information on landslides. Various data sources were used to collect information on 7 different periods, spanning 39 years, such as topographic maps, remote sensing data and extensive fieldwork. The landslide distribution has been analyzed through time, and with the use of topographic terrain units, derived from a digital elevation model, the variations in landslide density and magnitude were also investigated. The known landslides were related with the intensity of triggering rainfall events for the assessment of temporal probability. The resulting event-based landslide inventory maps will be used subsequently in landslide hazard zonation and semi-quantitative landslide risk assessment.

1 INTRODUCTION

The incorporation of landslide hazard and risk into regional and local planning is an important tool to reduce the impact of landslides in mountainous regions. 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. At best they have only susceptibility maps, with qualitative legends that are difficult to translate into actual expected impacts of landslides. In order to convert such susceptibility maps into hazard maps, information on landslide distribution and its evolution in time through inventory maps are needed (Hansen, 1984; Wieczorek, 1984; Guzzetti et al., 2004; van Westen et al., 2008). In India however, there is no centrally organized landslide database, although some initiatives have been taken at the local level that cover localized areas. Due to security reasons, the use of large-scale aerial photographs is restricted in all the border zones of India, including the entire Himalayan part of the country. This makes the generation of event-based landslide inventory maps for quantitative landslide hazard assessment a difficult task.

The aim of the research presented here was to generate event-based multi-temporal landslide inventories using all available sources of information between 1968 and 2007 in a highly landslide-prone area around Kurseong in Darjeeling Himalaya (see Fig. 1a), with its inherent data limitations, and to attempt to link the temporal probability of triggering events with the spatial probability of the landslides generated during these events. The source data were unfortunately incomplete, containing data gaps, differences in scale and resolution etc. Through the multi-temporal event-based landslide database, it was possible to study the changes in landslide patterns and distribution during the last four decades and despite uncertainties, can also be used for hazard assessment.

We accomplished the above task by i) analysing the distribution of past landslides and their basic attributes (type, failure mechanism, depth, areal extent etc.), for each time period for which data was available ii) evaluating the changes in the distribution (both space and time) of landslides over different terrain units, iii) identification of triggering rainfall thresholds, and relating them to known landslide events and calculating its exceedance probability.

1.1 Study area

The study area is part of a continuous tectono-stratigraphic sequence of metamorphic rocks of the Eastern Himalayan Fold-Thrust Belt (FTB) from north to the foreland molasse basin in the south (see Fig. 1 b & c). Towards the north, high-grade metamorphic rocks (migmatites) are present whereas the southern boundary is marked by a high-strain ductile shear zone, called the Main Central Thrust (MCT), coinciding with an ensemble of high to low grade metamorphic rocks (Hubbard, 1996; Searle and Szule, 2005). Tectono-stratigraphically, the study area represents the southern part of the Darjeeling klippe, where high grade metamorphic rocks of the Darjeeling and Chungthang Groups are thrusted over low grade metamorphic rocks of the Daling Group along the MCT (Mallet, 1875; Sinha-Roy, 1982). Towards the south, the foreland molasse sediments of the Siwalik Group are underlain by an intra-thrusted slice of Gondwana sediments. Toward the north, the Gondwana sediments are thrusted over by the Daling Group of metasediments along the abrupt southern-most front of Himalayan FTB known as the Main Boundary Fault (MBT).

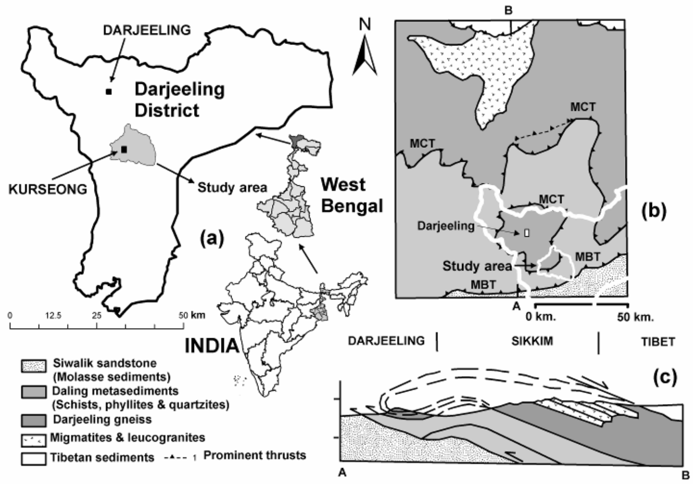

Figure 1. Upper left (a) Location map of the study area. Upper right (b): Regional Geological sketch map of Darjeeling-Sikkim Himalaya (after Searle and Szule, 2005). Below right (c): Schematic geological section of Darjeeling-Sikkim Himalaya (after Searle and Szule, 2005).

1.2 Landslide distribution types and triggers

In general, three types of landslides are mostly observed in the study area. Shallow translational rock slides and most frequent, followed by shallow translational debris slides and flows, while there are only few deep-seated rock slides, which are mostly larger in dimension than the other two groups. The study area receives a substantial amount of monsoon rainfall every year (June-October). The average annual precipitation varies from 2000 to 5000 mm (Soja and Starkel, 2007) and due to this heavy monsoon precipitation, rainfall-triggered shallow landslides are quite predominant.

2 MATERIALS AND METHODS

The first step towards the generation of a multi-temporal and event-based landslide inventory was the collection of all available data on past landslide occurrences, and all spatial data in the form of high resolution satellite images, topographic sheets, old landslide maps and reports of field investigations.

During 1969-70, just after a major landslide event in 1968, the Survey of India (SOI) updated their topographic survey and prepared new 1:25,000 topographic maps. In these topographic maps, the locations of the large active landslides of 1968 were included.

The next available data source was a field-based landslide inventory map from 1993 prepared by the Geological Survey of India (GSI) just after the landslide events that happened during the end of June and middle of July, 1993 (Sengupta, 1995). Unfortunately the field map of 1993 only covered the southeastern part of the study area (56 km^2).

The third data source represents another event-based landslide inventory map prepared by GSI through field investigations just after a prominent landslide event of 5-8 July, 1998. Also this landslide inventory map covers only a part (central portion ~ 20 km^2) of the entire study area, along a national highway (NH-55) and around Kurseong town (Bhattacharya et al., 1998).

Apart from these maps, several high resolution satellite images were available from different periods: an IRS LISS III PAN image (5.8 m resolution) from 2002, an IRS LISS IV P6 MX image (5.8 m resolution) from 2004 and a set of stereo Cartosat-1 images (2.5 m resolution) from 2006. The most recent data source was a detailed fieldwork carried out in 2007, which allowed to map landslides that happened as a result of a rainfall event in 2007.

The attributes recorded for each landslide mapped included the movement type, material involved, activity, failure mechanism, and date of occurrence, following the method by Varnes (1978) and UNESCO-WP/WLI (1990; 1993). Landslides were mapped stereoscopically from anaglyph images prepared using ortho-rectified imagery of Cartosat-1 (UTM Projection; Datum – WGS 84; Zone – 45 N) and the DEM of the area. During the stereoscopic interpretation, the shape, morphometry and association of old landslides were compared with the similar type of recent landslides mapped from the high resolution (2.5 m) stereo imagery of Cartosat 1 (2006) and/ or with the similar landslides mapped during field investigation of 2007.

Apart from the known dates of landslide events in 1968, 1993, 1998, 2003 and 2007 none of the source data products contained information on the exact dates of landslide events, thus, for some inventory maps; we could not link the landslides to a part date of a triggering event. In between 1968 and 1993, we received information on some more landslide event years (1984, 1985, 1986 and 1991) from geological reports and interviews with local people, but the spatial distribution of landslides for those events could not be reconstructed.

2.1 *Methods for analysis of inventory data*

We compared in a GIS, the locations of landslides of different time periods to know the frequency and pattern of new and reactivated landslides. Through a buffer analysis in GIS, the landslides of a younger period, which are within 50 m buffer distance of the landslide polygons of an older time period were identified as reactivated landslides. All other landslides which happened further than 50 m away from older landslides were considered new landslides. The buffer distance of 50 m was based on our field experiences as the zone in which landslides could be reactivated. Also an analysis of landslide densities within topographic terrain units was carried out, using inventories of different periods.

2.2 *Methods for estimation of temporal probability*

Since the landslide inventories are affected by incompleteness and data gaps, it was quite difficult to use them directly for the calculation of temporal probability. For assessing the temporal probability of events, we analysed the past daily rainfall data (1968-2007) for the known landslide event days. To establish a possible relationship between landslide events and rainfall amounts, we applied a stepwise discriminant function analysis following the method proposed by Dai & Lee (2001) in SPSS 15.0 using various rainfall parameters (in mm) such as daily rainfall (DR) and different antecedent rainfall amounts (1-day, 2-day, 3-day, 5-day, 7-day and 10-day antecedent) as different predictor or explanatory variables. As grouping or response variable, we used the known landslide event/non-event-days. The objective of this multivariate analysis was to develop a quantitative method using the known landslide events and the triggering rainfall intensities to calculate the exceedance probability (Crovelli, 2000; Coe et al., 2004) of predicted events. The exceedance probability of events can be used as a measure of temporal probability in such data-constrained environment.

3 RESULTS AND DISCUSSION

The landslide inventories (see Fig. 2) that were generated from the available sources were analyzed in a GIS and for each data source, a number of descriptive statistics were calculated (See Table 1 & 2). Among the available event-based landslide inventories, the maximum landslide density (2.55 slides/km^2) was observed in LI_03, followed by LI_93 (1.9), LI_98 (1.6), LI_68 (1.1) and LI_07 (0.85) respectively (see Table 1). The average landslide area varies between 628 m^2 (LI_07) and 3393 m^2 in LI_1968.

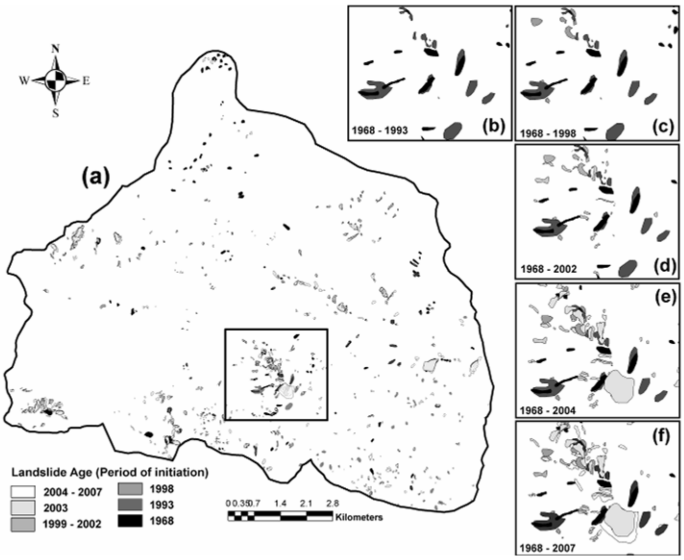

Figure 2 (a). Period of initiation of landslides in the study area based on the comparison of multi-temporal inventories; (b to f) enlarged views of the selected area for different periods.

Table 1: Summary of landslide information from the different data sources.

Inventory (LI_)	68	93	98	99-02	03	04-06	07
Area (km^2)	100	56	20	100	100	100	100
Slide number (Nr)	108	108	32	190	255	165	85
Cum.area (km^2)	0.6	0.5	0.0	0.8	1.2	0.6	0.1
Min. area (m^2)	538	372	185	271	221	45	42
Max.area (km^2)	0.1	0.0	0.0	0.1	0.1	0.1	0.01
Mean area (m^2)	5324	4634	1687	4466	4759	3962	1357
Median area (m^2)	3393	2616	840	2312	1876	722	628
Std. dev. area (m^2)	7714	5548	2038	8983	10688	12437	1649
Landslide density (Nr/km^2)	1.1	1.9	1.6	1.9	2.5	1.6	0.8
Landslide area %	0.6	0.9	0.2	0.8	1.2	0.6	0.1

The largest landslide (0.1 km^2) was mapped from landslide inventory LI_04_06. Given the large variation in sizes of landslides for most of the periods, we concluded that LI_68 did not contain all the smaller landslides. The topographic map sheet of 1969, which was the source for this, contained only the large landslides. Amongst the mapped landslides, shallow translational rock slides are predominant, varying between 59% (LI_68) and 82% (LI_99_02) (see Table 2). The proportion of deep-seated landslides is less (maximum 10%) and they often result from the retrogression of smaller landslides.

Table 2. Summary of landslide type information from the different data sources.

Inventory (LI_)	68	93	98	99-02	03	04-06	07
Landslides (Nr)	108	108	32	190	255	165	85
Landslide area (km^2)	0.5	0.5	0.0	0.8	1.2	0.6	0.1
Shallow translational debris slides and flows							
% Nr	39	29	100	15	27	18	31
% area	24	24	100	7	11	2	17
Shallow translational rock slides (planar and wedge)							
% Nr	59	69	0	82	70	72	69
% area	60	64	0	81	76	48	83
Deep-seated rock slides							
% Nr	2	2	0	3	3	10	0
% area	16	12	0	12	13	50	0

3.1 Spatio-temporal landslide evolution.

In our inventory the maximum frequency of *reactivated* landslides (75%) were mapped in *LI_04_06,* followed by *LI_03* (62%) and *LI_07* (54%) (Table 3). Signs of reactivation were less in older inventories such as *LI_93* (27%) *and LI_99_02* (24%). In contrast the maximum number of landslides (76%) occurring at *new* locations were found in the inventory *LI_99_02* (see Table 3).

This indicates that there could be a substantial lack of landslide information prior to 1999 as signs of past landslides are quickly obliterated, due to rapid land use changes in the Himalayas, and thus, most of landslides mapped from the data source of 1999-2002 apparently occurred at new locations with respect to the previous inventory, from 1993.

To analyze further the temporal changes in landslide abundance, we compared the landslide area percentage values of each topographic terrain units of the two subsequent time periods. On the basis of a 10 by 10 m digital elevation model, the study area was divided into 1126 terrain units following the method proposed by Carrara *et al* (1991)**.** In this analysis, we considered the terrain units with landslide densities less than 2.0 (landslide per km^2), as *stable* terrain units, and those with higher values were considered *unstable* (Galli et al., 2008). The above threshold value of 2.0 was considered keeping in view the probable uncertainty of mapping and digitization errors. Based on this criteria, a maximum of 229 terrain units out of 1126 (22.4% of total area) have been affected by any form of slope failures in the last 39 years. The prominent landslide event of 2003 caused a sharp increase in the number of affected terrain units from 180 (16%) to 228 (20%) for the pre_2004 period.

In contrary, a negligible change in landslide density (0.1% increase) was noticed between 2004 and 2007, which confirmed a reduction of landslide activity in comparison to pre_2004 period (Fig. 3).

Table 3. Size and frequency distribution of reactivated and new landslides of different temporal frames (Results of buffer analysis in a GIS).

Landslide inventory	LI_93	LI_99_02	LI_03	LI_04_06	LI_07
Number of landslides (Nr)	108	190	255	165	85
Total landslide area (km^2)	0.5	0.85	1.2	0.6	0.1
Shallow translational debris slides and flows					
Re-act (% of Nr)	10	3	15	6	16
Re-act (% of area)	14	2	7	1	8
New (% of Nr)	18	12	12	12	14
New (% of area)	10	5	4	1	8.4
Shallow translational rock slides (planar and wedge)					
Re-act (% of Nr)	15	18	44	59	38
Re-act (% of area)	22	25	67	46	49
New (% of Nr)	55	64	25	13	32
New (% of area)	42	56	9	2	35
Deep-seated rock slides					
Re-act (% of Nr)	2	3	3	10	0
Re-act (% of area)	12	12	12	50	0
New (% of Nr)	0	0	0.4	1	0
New (% of area)	0	0	1	0.3	0
All landslide types					
Re-act (% of Nr)	27	24	62	75	54
Re-act (% of area)	63	76	38	25	46
New (% of Nr)	48	39	86	97	57
New (% of area)	52	61	14	3	43

3.2 Temporal probability assessment

The step-wise discriminant function analysis utilised seven different rainfall intensity variables as predictors to derive a suitable discriminant function, using the 22 known event days. A predictor variable was allowed to enter into the model if the significance of its *F* value was more than 3.84 and it was removed if the significance of its *F* value is less than 2.71. Finally after 14 steps, the discriminant model identified the following four statistically significant predictor variables – DR, AR1, AR2 and AR5, which were used for discriminating the *grouping variables*. The standarised cannonical discriminant function coefficients for the above identified variables were 0.814, 0.328, 0.394 and (–) 0.364 respectively. From these, it is apparent that the *daily rainfall* (DR) played the most significant role, followed by AR2 and AR1 respectively. The above quantitative measure also signified that to classify the known landslide events of the study area, rainfall antecedents of more than five days were not found statistically significant. The above discriminant function analysis could successfully discriminate the *known landslide event days* and *non-landslide days* with overall 95.7% original grouped cases and 95.6% of cross-validated correctly grouped cases. The discriminant function was able to classify 19 known landslide event days out of the 22 correctly (86.4%).

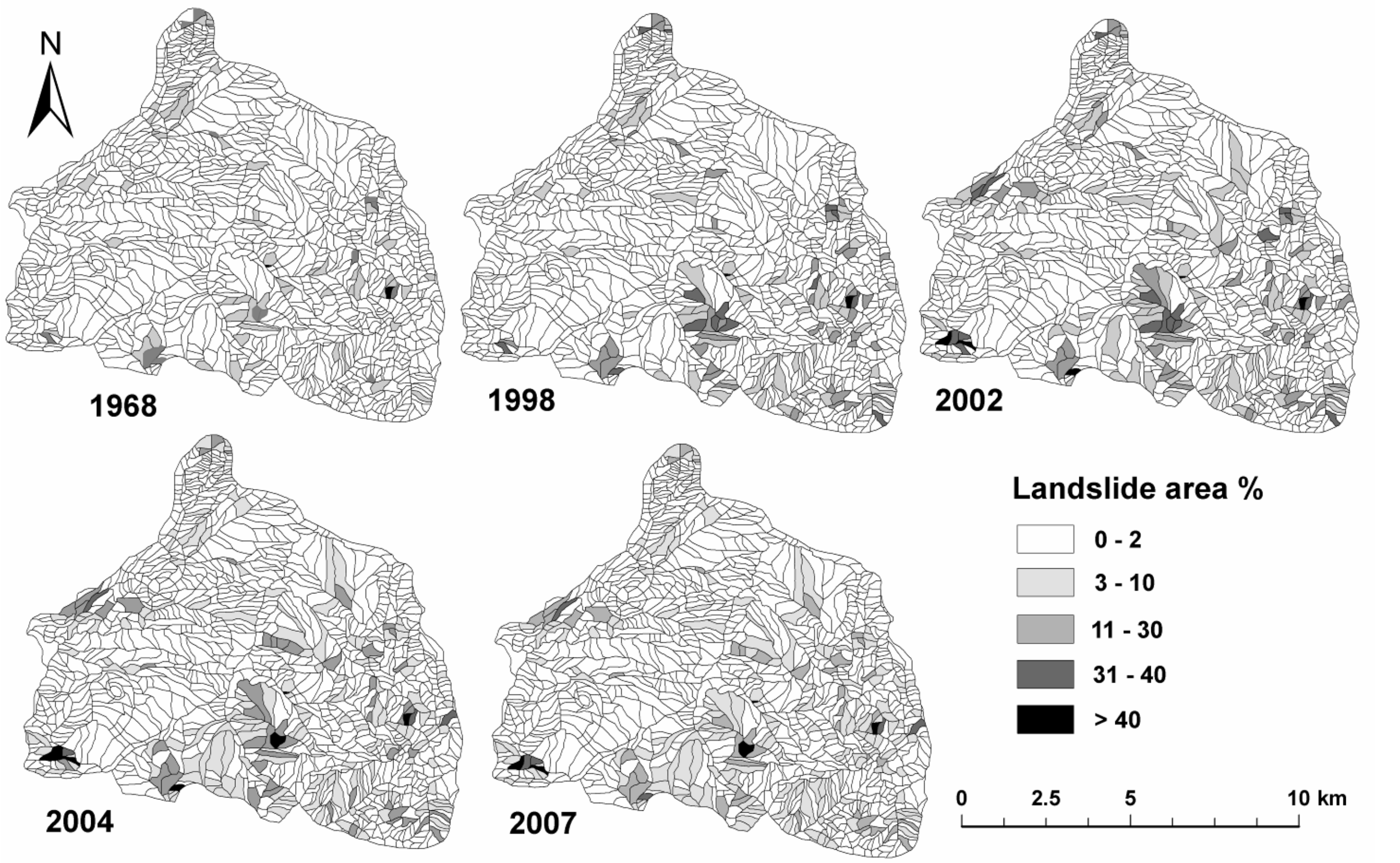

Figure 3. Spatio-temporal variation of landslide area percentage from 1968 till 2007.

Observing the above model results, and comparing them with the already known severe events, we deduced an upper and lower threshold of the discriminant scores for predicting future landslide events. The lower threshold discriminant score was fixed at 3 whereas the upper threshold has been fixed at 4 (Fig. 4). Accordingly, two threshold discriminant equations for predicting unknown landslide events are proposed using the un-standarised discriminant function coefficients, which are as follows:

$-0.645 + 0.021*DR + 0.008*AR1 + 0.006*AR2 - 0.003*AR5 \leq 3$ (Lower threshold) (1)

$-0.645 + 0.021*DR + 0.008*AR1 + 0.006*AR2 - 0.003*AR5 \leq 4$ (Upper threshold) (2)

The use of equation 2 (upper threshold) predicted 59 major landslide events which are spread over 29 landslide event years between 1968 and 2007. Excluding the already known major events, the above analysis was able to identify 20 more unknown landslide events in the study area in between 1968 and 2007.

Similarly, equation 1 can also be used to predict minor landslide events in the area.

We subsequently utilised the deduced landslide event years for the calculation of temporal probability of events by assuming that similar landslide activity will prevail in future and at least one landslide event will occur per event year during the specific return period.

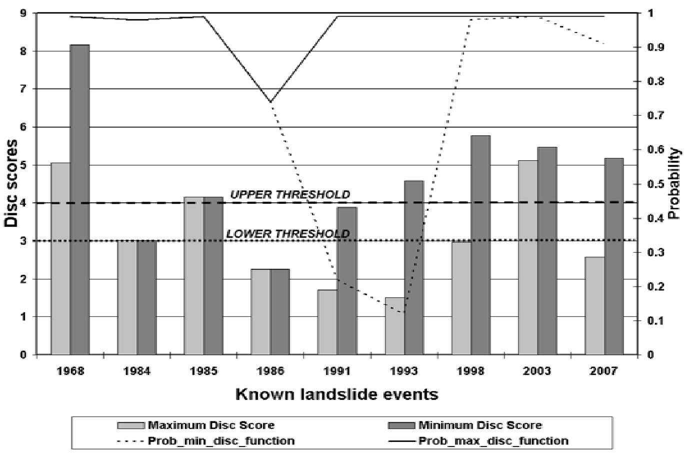

Figure 4. Threshold discriminant scores and probability of occurrence of known events.

To calculate the exceedance probability or the measure of temporal probability of major landslide events, we deduced the mean recurrence interval of the landslide event years of the predicted events, which is 1.34 (29 event years predicted within 39 years). Simultaneously, we also deduced the mean recurrence interval in years of the known major events, which was 4.33 (9 event years within 39 years) for comparison. Then we applied both the Poisson (Eq. 3) and Binomial distribution models (Eq. 4) to calculate the exceedance probability for both the above cases; the results are illustrated in Figure 5.

$$P\,[\,N_L\,(t)\ \geq\,=1\,] = 1 - P\,[\,N_L\,(t) = 0\,] = 1 - e^{-\lambda t} = 1 - e^{-t/\mu} \qquad (3)$$

$$P\,[NL\,(t)\ \geq\,=1] = 1 - P\,[NL\,(t) = 0] = 1 - (1-P)^{t} = 1 - \left(1 - \frac{1}{\mu}\right)^{t} \qquad (4)$$

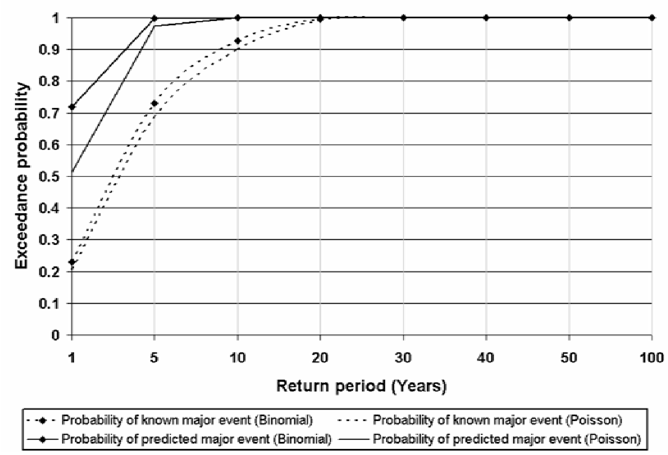

Figure 5: Exceedance probability of major landslide events.

From Figure 5, it can be concluded that the study area is quite active and considering the mean recurrence interval on the basis of our predicted landslide events, the occurrence of a major landslide event with 100 percent certainty can be expected once in every five years. In contrast, if we consider the mean recurrence interval of events based only on the known and available landslide event records, we can expect the occurrence of a major landslide event with 100 percent certainty once in every twenty years. This might be a serious underestimation due to large scale inadequacy in our source data.

4 CONCLUSIONS

Different event-based and temporal landslide inventories can be prepared and spatially combined in a GIS to derive the knowledge about type, failure mechanism, frequency and magnitude of past landslide events. The above analysis can be coupled with a continuous time-frame to trace the spatio-temporal evolution of landslides. The above task is quite difficult in case of large scale uncertainties, data gaps and incompleteness of the source data. The research presented here has successfully demonstrated how event-based multi-temporal landslide inventory can be generated despite such uncertainties, and how best the same can be used for the quantitative estimation of hazard.

In this research, we presented a method, where despite data scarcity; the event-based multi-temporal inventory was optimally used for the calculation of temporal probability of events. Temporal probability estimation in those cases largely depends on the identification of unknown landslide events through statistical analysis of triggering rainfall intensities and the related known landslide events. In this research, we demonstrated that in the study area, occurrence of a major landslide event once in 5 years has a very high level of certainty. Our multivariate statistical model also demonstrated that incompleteness of inventory information or lack in complete knowledge on known events can seriously underestimate such temporal probability calculation. The above spatio-temporal landslide inventory database will further be used to calculate the spatial probability of landslides in our future course of research.

ACKNOWLEDGEMENTS

This work was accomplished as a part of the joint Ph.D. research project of Geological Survey of India (GSI), National Remote Sensing Agency (NRSA), India and ITC, the Netherlands. We acknowledge GSI and ITC for allowing us to publish this research. We also thank the local administrative authority of Kurseong Sub-division, Darjeeling district, Government of West Bengal, and Goomtee tea garden authorities for providing us the necessary infrastructure and rainfall data at field. We are indebted to all the anonymous reviewers for their comments and suggestions to improve the manuscript. This research is carried out in the framework of the United Nations University – ITC School for Disaster Geo-Information Management (www.itc.nl/unu/dgim/)

REFERENCES

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

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V. and Reichenbach, P., 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth surface processes and landforms, 16(5): 427-445.

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

Crovelli, R.A., 2000. Probability models for estimation of number and costs of landslides, USGS, Denver, Colorado.

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

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

Guzzetti, F., Cardinali, M., Reichenbach, P., Cipolla, F., Sebastiani, C., Galli, M. and 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.

Hansen, A., 1984. Landslide hazard analysis. In: D. Brunsden and E. Prior (Editors), Slope Instability. John Wiley & Sons., New York, pp. 523-602.
Hubbard, M., 1996. Ductile shear as a cause of inverted metamorphism: example from Nepal Himalaya. Journal of Geology, 194: 493-499.
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 pp.
Searle, M.P. and 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.
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. (Editor). 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. and Starkel, L., 2007. Extreme rainfall in Eastern Himalaya and southern slope of Meghalaya plateau and their geomorphologic impacts. Gemorphology, 84: 170-180.
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., Castellanos, E. and Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology, In Press.
Varnes, D.J., 1978. Slope movements types and processes. In: R.L. Schuster and R.L. Krizek (Editors), 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.