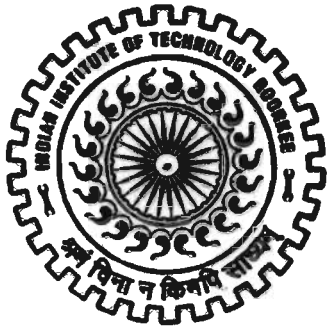
GIS-BASED LANDSLIDE SUSCEPTIBILITY STUDIES USING NEURAL AND FUZZY APPROACHES

A THESIS

Submitted in fulfilment of the
requirements for the award of the degree
of
DOCTOR OF PHILOSOPHY
in
EARTH SCIENCES

By

DEBI PRASANNA KANUNGO



DEPARTMENT OF EARTH SCIENCES
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247 667 (INDIA)

JULY, 2006

6th Annual Convocation- 2006
Degree conferred on 11.11.2006

Supdt. (PGS&R)

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **GIS-BASED LANDSLIDE SUSCEPTIBILITY STUDIES USING NEURAL AND FUZZY APPROACHES** in fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy** and submitted in the **Department of Earth Sciences** of the **Indian Institute of Technology Roorkee, Roorkee** is an authentic record of my own work carried out during a period from January 2001 to July 2006 under the supervision of **Prof. R. P. Gupta, Dr. M. K. Arora and Dr. S. Sarkar**.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other University/Institute.

(DEBI PRASANNA KANUNGO)

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(Dr. S. Sarkar)	**(Dr. M. K. Arora)**	**(Dr. R. P. Gupta)**
Scientist	Associate Professor	Professor
Geotech. Engg. Division	Deptt. of Civil Engineering	Deptt. of Earth Sciences
CBRI	IIT Roorkee	IIT Roorkee
Roorkee (INDIA)	Roorkee (INDIA)	Roorkee (INDIA)

Date:

The Ph.D. Viva-Voce Examination of **Mr. Debi Prasanna Kanungo**, Research Scholar, has been held on 27 Oct. 2006

Signature of Supervisor(s)	Signature of H.O.D	Signature of External Examiner

Abstract

Landslides are complex natural phenomena that constitute a serious natural hazard in the Himalayan region, causing damage to both property and life every year. Identification of landslide-prone areas is essential for safer strategic planning of future developmental activities. Therefore, landslide susceptibility zonation (LSZ) becomes important whereby an area may be divided into near-homogeneous domains that are ranked according to the degree of potential hazard due to landslides and mass movements. The most important task for LSZ studies is the determination of weights and ratings giving relative importance of factors and their categories respectively, for landslide occurrence. These weights and ratings can be determined by implementing different approaches, which at times are highly subjective in nature. Therefore, developing suitable approaches for determination of weights and ratings objectively and their implementation in a geographic information system (GIS) environment for LSZ mapping is extremely important. Further, most of the landslide-related studies culminate at providing LSZ maps only, therefore, procedures of landslide risk assessment (LRA) needs attention.

The main objective of this research is to explore the potential of the advanced techniques – fuzzy set theory and artificial neural network (ANN) for determination of weights and ratings of causative factors and their categories, and to devise a fully objective approach for GIS based LSZ and LRA mapping.

The study area covers a small region of about 254 km^2 in Darjeeling Himalayas (Latitude 26° 56′-27° 8′ N and Longitude 88°10′-88°25′ E). It is a part of Darjeeling district of West Bengal State of India. The area lies within the Lesser- and Sub-

Himalayan belts. Tea plantations form the most widespread land use. The area in the eastern part is dominated by thick forest.

The following datasets have been used to generate various thematic data layers:

a) Remote sensing images from IRS-1C LISS-III multispectral and IRS-1D-PAN
b) Survey of India (SOI) topographic maps at 1:50,000 and 1:25,000 scale
c) Published geological map (Geological Survey of India)
d) Extensive field data on landslides and land use land cover

The LISS-III and PAN data have been precisely co-registered with the topographic map. The LISS-III data have been corrected for atmospheric path radiance. The various thematic data layers pertaining to causative factors of landslides form the input layers for LSZ mapping and have been generated using remote sensing-GIS tools.

The digital elevation model (DEM) of the study area has been generated by digitization of contours on SOI topographic maps from which slope and aspect data layers have been derived.

The lithology layer has been prepared by digitizing polygons from the co-registered geological map. Minor modifications in lithologic boundaries at some places have also been incorporated in the vector layer after field verification. The layer was rasterized.

Lineaments have been interpreted from the PAN and LISS-III images. There is no major thrust/fault reported in the study area, but major lineaments have been identified. A lineament buffer layer was generated to deduce the influence of lineaments on the occurrence of landslide. A drainage layer has been prepared from the topographic maps and LISS-III image. The ordering of streams has been performed on the basis of Strahler's classification scheme. A drainage buffer layer

with 25m buffer zone along 1st and 2nd order drainages (only) has been generated for further analysis.

Eight dominant land use land cover classes, namely, thick forest, sparse forest, tea plantation, agriculture, barren, built up, water bodies and river sand in the area have been deciphered. The four spectral bands of LISS-III image, DEM and normalized difference vegetation index (NDVI) image have been integrated to prepare a land use land cover layer by a multi-source classification process using the most widely adopted maximum likelihood classifier.

The field data and high spatial resolution PAN and PAN-sharpened LISS-III images have been used to produce a landslide distribution map. A total of 101 landslides have been identified, the majority of which have areal extent of 500 m^2 to 2000 m^2.

For LSZ mapping, four different approaches have been implemented to determine the weights and ratings. The most commonly used conventional weighting approach involved assignment of weights and ratings to the factors and their categories based on field knowledge. A landslide susceptibility index (LSI) map has been generated by integrating the weighted layers and the range of LSI values has been categorized into five landslide susceptible zones to prepare an LSZ map.

In order to minimize subjectivity in the weight assignment process, at the first instance, ANN black box model has been attempted for LSZ mapping. A feed forward multi-layer ANN with one input layer, two hidden layers and one output layer has been designed. The input layer contained 6 neurons corresponding to 6 different causative factors. The LSZ map obtained from conventional weighting approach has been used as the reference map. The training and testing datasets consisted of 2500 mutually exclusive pixels corresponding to 500 pixels per landslide

susceptibility zone. The Levenberg-Marquardt back-propagation algorithm has been used to train the neural networks. A total of 39 neural network architectures were designed, trained and tested. The network architecture 6/13/7/1 has been found to be the most appropriate one. The connection weights obtained from this network have been captured and subsequently used to determine the network output of all the pixels in the dataset to prepare an LSZ map of the area.

However, the major limitation of the ANN black box approach is that the weights and ratings of the factors and their categories remain hidden. In this research, an attempt has been made to open the ANN black box. A novel approach to derive weights for causative factors has been proposed, which has been referred to as ANN connection-weight method. The assignment of weights in this fashion may reduce the subjectivity. Moreover, to bring objectivity in the assignment of ratings, a new fuzzy set concept has been utilized. As a result, two unique ways of LSZ map preparation have been proposed here, namely fuzzy set based and combined neural and fuzzy approaches.

In the fuzzy set based approach, ratings (r_{ij}) of each category of a given thematic layer have been determined using the cosine amplitude method. The integration of these values for various categories of thematic layers has been performed to compute LSI values in two different ways: (a) using arithmetic integration and (b) using fuzzy gamma operator. The range of LSI values thus determined has been categorized into five landslide susceptibility zones using success rate curves method to prepare the LSZ maps. The performances of the two methods have been examined and it has been found that the arithmetic integration approach has yielded better results than the fuzzy gamma operator in the present case.

The combined neural and fuzzy approach has involved three main steps: (a) determination of weights of thematic layers through ANN connection-weight analysis, (b) determination of ratings for categories using cosine amplitude method and (c) integration of ratings and weights to generate the LSZ map. A feed forward back-propagation ANN with one input, two hidden and one output layers was considered. The data for the input neurons correspond to the normalized ratings (r_{ij}) of the categories. The output corresponds to the presence or absence of landslide at the pixel. 100 neural network architectures were designed, trained and tested. The adjusted weights of input-hidden, hidden-hidden and hidden-output connections for each network were captured and analyzed to obtain the weights for thematic layers corresponding to 6 factors. The integration of these weights for causative factors and the ratings for the categories (obtained from cosine amplitude fuzzy similarity method) has been performed to obtain distribution of LSI values across the area. The range of LSI values has been categorized into five different landslide susceptibility zones using success rate curves method to produce the LSZ map.

A comparison of the LSZ maps produced from different approaches is very important. The LSZ maps have been compared and evaluated using three different approaches: a) landslide density analysis, b) error matrix analysis and c) difference image analysis.

Landslide density is defined as the ratio of the percent existing landslide area to percent area of each landslide susceptibility zone, and is calculated on the basis of the number of pixels. It has been found that the LSZ Maps produced from conventional and ANN black box approaches have a similar trend of landslide densities for various susceptibility zones. This result is on expected lines, as the conventional weighting based LSZ map has been used as the reference map to generate ANN black box based

LSZ map. The LSZ Map produced from combined neural and fuzzy approach has a much higher landslide density of VHS zone (>13) as compared to other LSZ maps. Based on the landslide density analysis, it is inferred that the LSZ map produced from combined neural and fuzzy approach is significantly better than those produced from other approaches (fuzzy, conventional and ANN black box approaches).

Three error matrices for different LSZ map combinations have been generated to understand the distribution of number of pixels in different LSZ maps. There was a high degree of matching in the pixels of LSZ maps produced from the conventional weighting and ANN black box approaches. There was also a general correspondence in the LSZ maps produced from fuzzy and combined neural and fuzzy approaches. There was a lot of mismatch in number of pixels between LSZ Maps produced from the conventional weighting and combined neural and fuzzy approaches. This mismatch is inferred to be due to the differences in weights and ratings of both the approaches.

Difference image analysis elucidates how pixels shift from one landslide susceptibility zone to another zone, based on the LSZ mapping approach adopted. A difference image of LSZ maps produced from conventional and ANN black box approaches showed a high degree of mutual correspondence and matching of landslide susceptibility zones throughout the area. A difference image of maps produced from fuzzy set based and combined neural and fuzzy approaches showed a high degree of spatial matching, with about 50% pixels having full matching, and 47% pixels exhibiting only one-zone difference. About 3.0% pixels showed two-zone difference and these mainly appeared to be related to a lithologic band in the northern part of the area. The difference image of maps produced from conventional weighting and combined neural and fuzzy approaches appeared to exhibit the widest spatial

difference, with only 37.8% pixels fully matching, 46.4% pixels exhibiting one-zone difference, 14.6% pixels exhibiting two-zone difference, and 1.2% pixels showing three-zone difference. The most important was a two-zone difference band in the northern part of the difference image marking a lithologic layer. As lithology has the highest and significantly higher weight than other factors in combined neural and fuzzy approach, the importance of lithology has been brought out in the difference image. Further, the relative importance of drainage as observed in the field vis-à-vis lineament as deduced from spatial-domain filtering is deduced and discussed.

As far as landslide risk assessment (LRA) is concerned, landslide risk is considered to be a function of landslide potential (LP) and the resource damage potential (RDP). In the present study, two different approaches namely, (1) LRA using danger pixels and (2) LRA using Fuzzy Concept have been developed and implemented to prepare LRA maps of the study area.

A concept of danger pixel has been introduced for landslide risk assessment. *Danger pixels* are considered as those pixels which lie in VHS and HS zones in all the four LSZ maps produced from different approaches, i.e., the danger pixel map is an intersection map of all the four LSZ maps with (VHS + HS) zones combined. A resource map including all the existing land use land cover patterns and also the road network of the area has been prepared. The danger pixel map and the resource map have been integrated to generate the LRA map of the study area. The LRA map shows spatial distribution of different resource categories that appeared to be under real danger due to landslides.

In the LRA using fuzzy concept approach, the LSZ map prepared using the combined neural and fuzzy approach; the best LSZ map of the area, has been used as an input to provide LP. Further, the resource map has been used as another input layer

to provide information on RDP. Linguistic rules have been framed for landslide susceptibility zones and resource categories and the fuzzy membership values representing the LP and RDP based on these linguistic rules have been assigned. Landslide risk values for different combinations of LP and RDP have been obtained by integrating LP and RDP layers and have been represented in the form of a LRA matrix. The range of landslide risk values has been segmented into five different landslide risk zones and the LRA map of the area has been prepared. It is observed that 2496 pixels (0.61% of total area) are under very high risk zone and 7204 pixels (1.77% of total area) are under high risk zone. The LRA Map has revealed that landslides pose very high risk to a few selected sites of habitation in Sonada, Darjeeling and northeastern part of Tiger hill, and high risk to a section of road from Sonada to Ghum.

It is considered that the approaches developed in this study for objective LSZ and LRA mapping can be successfully implemented in other hilly regions that are susceptible to landslides.

ACKNOWLEDGEMENTS

This space gives me the opportunity to include in humble words my gratitude and heartfelt thanks to those who have helped me see the light of this fateful day.

How do I express my indebtedness and sincere gratitude towards my mentors and supervisors **Prof. R.P. Gupta, Dr. M.K. Arora** and **Dr. S. Sarkar**, who instilled in me the interest, the courage and the will to produce this work! If only words could convey the immense respect I have towards Prof. Gupta, whose immaculate observation and inspiring guidance remain cherished, now and ever after. The persuasiveness and never-say-die attitude of Dr. Arora has left its imprints on me, and I find no word suitable enough to express my thankfulness for his able guidance and support. In this line of my mentors, I couldn't possibly thank enough my guide and senior colleague Dr. Sarkar for remaining a pillar of support and beacon of light when I had to fight the 'tyrannous and treacherous' challenges in the course of this research.

Thanks are due to **Prof. V.N. Singh**, HOD, Dept. of Earth Sciences, and the former HODs, **Prof. B. Parkash** and **Prof. A.K. Awasthi**, who so very kindly allowed me to use the resources and provided every possible help during the years of my research.

I am grateful to **Prof. K. Ganesh Babu**, Director, Central Building Research Institute, Roorkee, India for providing the necessary facilities to complete this work. The official assistance provided by the Central Building Research Institute, Roorkee, India by permitting me to carry out this research is acknowledged with full heart.

This is also an opportunity for me to convey a special word of thanks to **Dr. A.K. Sen** of the Dept. of Earth Sciences for his interest, encouragement and extraneous guidance on the subject of this research.

Many sincere thanks are due for **Mr. J.N. Vaish**, Head, Geotech. Engg. Div., CBRI, Roorkee; **Mr. A. Ghosh, Mr. Y. Pandey, Dr. R. Dharmaraju, Dr. P.K.S. Chauhan, and Mr. S.K. Srivastava**, Scientists, CBRI, Roorkee; **Mrs. Archana Sarkar**, Scientist, NIH, Roorkee and **Mr. Ashis Kr. Patra** and **Mr. Pushpendra Kumar**, Research Fellows, CBRI, Roorkee for encouragement, ready help, and moral support.

I am grateful to **Dr. Ashis Saha, Dr. Umesh Haritashya, Ashish Misra, Arundhati, Aparna** and other research scholars of the Remote Sensing Lab., Dept. of Earth Sciences, and **Dr. Saif Said**, Dept. of Civil Engg., IIT Roorkee for their enthusiastic help and ready assistance whenever and whichever way I sought it.

Thanks are due to my friends **Dr. P.K. Bhuyan**, Scientist, NIH, Roorkee; **Dr. N. P. Padhy**, Assoc. Prof., Dept. of Elect. Engg., IIT Roorkee; **Dr. P.K. Sahoo**, Asst. Prof., Dept. of Mechanical Engg., IIT Roorkee; **Dr. N. Panigrahy**, Scientist, NIH, Roorkee; **Mr. S.K. Panigrahi**, Scientist, CBRI, Roorkee; **Mr. S.K. Senapati**, Library Officer, CBRI, Roorkee and their families for being there all along the vicissitudes of research.

I cannot but express my heartfelt thanks to the people of Darjeeling and surrounding areas for their help during the field data collection.

The help extended by **Sarvesh Ji, Nair Ji** and other technical and non-technical staffers of the Dept. of Earth Sciences and Dept. of Civil Engg., IIT Roorkee is duly acknowledged.

I cannot express my feelings towards my family members - my wife **Jolly**, and my sons **Pupul** and **Lipun**. They have sacrificed a lot during this period, and it will be difficult for me to compensate for that loss ever. My children had to sacrifice their vacations for this research, and without their company and presence around me, both physically as well as mentally; this research would have never gained the shape as it has now.

Finally, I fall short of words to express how it is to have such parents as mine whose blessings and invisible support pulled me when it seemed it's the end! Imagine being away from your parents for a full two and half years completely at your own imposed will, and even then remain unquestioned and unfailingly supported and understood! I find it almost compelling to bring to the knowledge of all the future readers of this work, the very fact that brought me into this whole process – the wish of my father-in-law who always wanted to see me achieve this feat. What a feeling it is to live a wish, and that too so special! My brothers, sister, brother-in-laws and sister-in-laws remained available in every conceivable way to help me, and to cheer me when I needed it the most. Thank you so much!

Debi Prasanna Kanungo

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 General

Disasters caused by landslides are common in mountainous regions such as the Himalayas. The landslide incidences in a region have been of serious concern to the society due to loss of life, natural resources, infrastructural facilities, etc. and also posing problem for future urban development. It has been estimated that, on an average, the damage caused by landslides in the Himalayan range costs more than US$ one billion besides causing more than 200 deaths every year (Naithani, 1999). Since geologic, geomorphic and hydro-geological factors control the overall stability of slopes in an area, evaluation on a regional scale is essential to identify areas susceptible to landslides. Thus, altogether there appears to be an urgent need of landslide susceptibility zonation (LSZ) mapping and risk assessment of regions affected by landslides. The studies would help in minimizing loss of life, property and other resources of the region due to landslides.

1.2 Research Objectives and Scope of the Work

A detailed literature survey of the techniques and approaches for LSZ mapping adopted in different parts of the world has been carried out. Briefly, it is observed that several techniques for LSZ mapping have been developed which can be grouped into two broad categories, *viz.*, qualitative and quantitative. In qualitative techniques, subjective decision rules are applied to define weights and ratings based on the experience of experts. The qualitative techniques can be sub-grouped into three major categories namely distribution analysis, geomorphic analysis and map combination technique. To remove subjectivity in qualitative analysis, quantitative techniques have been employed to determine the weights and ratings of factors and their categories respectively for LSZ studies. However, it is found that this aspect needs more careful and critical evaluation and development.

Thus, the main objective of the present research is to develop an approach for LSZ mapping leading to risk assessment through the use of advanced techniques and their implementation within the domain of remote sensing and GIS. Specific objectives of the present study may therefore be enumerated as:

i) Implementation of conventional weighting approach for LSZ mapping and its evaluation with ANN black box approach

ii) Development of a fuzzy set theory based approach for computation of ratings in an objective manner

iii) Development and implementation of combined neural and fuzzy approach, a combination of ANN and fuzzy based techniques for determination of weights and ratings of factors and their categories respectively to prepare the LSZ map

iv) A comparative evaluation of LSZ maps produced from different approaches to examine their efficacy

v) Development and implementation of an approach for landslide risk assessment

1.3 Study Area

The landslide incidences in Darjeeling Himalayas (parts of West Bengal State) have been of serious concern to the preservation of natural beauties of the area, maintenance of infrastructural facilities and future urban development. Therefore, an area in the Darjeeling Himalayas which covers a part of Darjeeling district of West Bengal State has been selected for this study.

Darjeeling derives its name from the Tibetan word `Dorje' meaning a precious stone or thunderbolt of Indra which is believed to have fallen at a place where now stands Observatory Hill and `Ling' means `a place'. Hence `Darjeeling' stands for `land of thunderbolt'. Darjeeling presents an unforgettable view of the majestic amphitheatre of snow covered Himalayan panorama flanked by Singalila range on the west and Dongkya range in the east jeweled with peaks of Kanchenjunga and Mount Everest. The flaming red rhododendrons, the vast spreads of undulating green tea plantations, silvery fir forests, Yiga Cholling Monastery in Ghoom, toy train, sunrise from Tiger Hill and specks of clouds in hues of every shade justify Darjeeling as the queen of hill stations (Figure 1.1). Darjeeling is a magnificent hill resort attracting thousands of people for a leisurely respite far from the maddening crowd to all – tourists, ornithologist, geologist, botanist, artist, photographer, trekker, etc. Darjeeling tea is world famous for its superb flavour. There are many tea gardens in and around Darjeeling.

1.3.1 Location

This study focuses on Darjeeling hill which lies within the latitudes 26°56' - 27°8'N and longitudes 88°10' – 88°25'E and covers an area of about 254 km^2 (Figure 1.2). The study area is a part of Darjeeling district of West Bengal state. The study area constitutes a part of Survey of India topographic map numbers 78 A/4, A/8, B/1, B/5 (scale - 1:50,000). The main localities are Darjeeling, Sonada and Sukhiapokhri. Darjeeling is located almost at the centre of the study area.

1.3.2 Physiography and Drainage

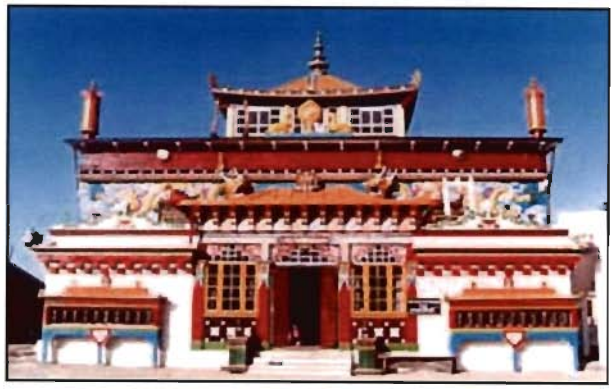
The district of Darjeeling in West Bengal lies between 26°31' and 27°13' N latitude and between 87°59' and 88°53' E longitude. It is somewhat triangular in shape. The Darjeeling Himalayas, encompassing a total area of 3149 km^2 rises abruptly from the alluvial plains of West Bengal and attains a maximum elevation of about 2500-3000 meters. The area lies between Sikkim on the North, Bhutan on the north-east, Nepal on the west, Purnea district of Bihar abutting on the south and district Jalpaiguri of West Bengal on the south-east. The southern foot hill region is characterized by East-West trending highly dissected platform of terrace deposits. The southerly flowing Tista River more or less divides the region into two parts, the eastern and western parts occupied by Kalimpong and Darjeeling hills respectively.

The principal town Darjeeling is the administrative headquarter of Darjeeling district. The town Darjeeling is situated in the lower Himalayas in 27°13' N latitude and 38°16' E longitude. Darjeeling falls into two distinct tracts, the Tarai immediately beneath the hills and the ridges and deep valleys of the lower Himalayas. The Tarai portion of the district is a low-lying belt, traversed by numerous rivers and streams rushing down from hills and by the upland ridges which mark their courses.

(a) Kanchenjunga overlooking Darjeeling town

(b) Darjeeling toy train

(c) Plucking tea leaf in Darjeeling



(d) Yiga Cholling Monastery in Ghoom

Figure 1.1: Some scenic views in Darjeeling Himalayas.

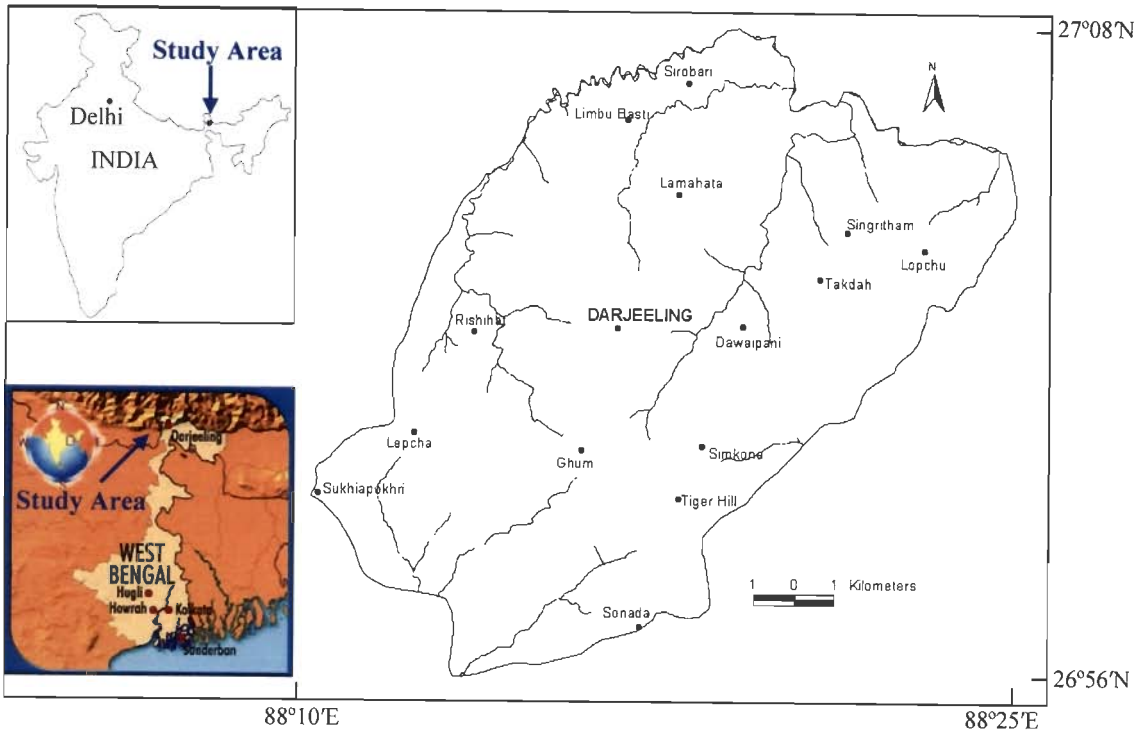


Figure 1.2: Location map of the study area

The mountains belonging to the lower Himalayan zone consist of long tortuous ranges, running generally from north to south throughout its length. The snowy range lies far beyond the limits of the district to the north. This range forms the great backbone of the Darjeeling Himalayas. To the north-west tower the giant peak of Kanchanjungha and to the north-east Dongkya are situated.

The valleys of Darjeeling hills are mainly drained by the Tista and its tributaries. The southerly flowing Tista river is on the eastern side of the study area, but not falling within it. Some of the important tributaries are Ragnu Khola, Sim Khola, Jhepi Khola, Bhanjyang Khola and Jor Khola. The hill portion is like a labyrinth of ridge and narrow valleys. There are no open valleys or plains or lakes. Most of the ridges are forest clad and in the lower slopes, tea plantation and crop cultivation are done.

1.3.3 Climate

Due to altitude differences the climate within the hill areas varies greatly. In general the hill areas enjoy pleasant summer, heavy rain in rainy season due to strong monsoonal winds and cold winter with snowfall in higher altitudinal areas. In Darjeeling hill areas, April, May, September and October form the peak tourist seasons due to pleasant climatic conditions. Darjeeling receives about 3000 mm of rainfall annually. The rainfall pattern is highly seasonal with the majority falling during the monsoon months of June to October. The mean maximum temperature is approximately 11.1 degree Celsius and minimum temperature is 1.7 degree Celsius.

1.3.4 Vegetation

Density of vegetation varies in different parts of the area. Darjeeling represents a combination of temperate and sub-tropical areas of the State of West Bengal with the

tea gardens along with some agricultural land. It embodies more than 7000 species of flowering plants along with numerous small and beautiful non-flowering plants. Darjeeling represents one of the rich floras in the world. The richness and variety are due to the physiographic and climatic conditions. The main land use practice in the study area is tea plantation. The agriculture lands are mostly present around the habited areas. The area is dominated by thick forest particularly in the eastern part.

1.3.5 Geology

The lithologic units in the Darjeeling Himalayas occur in a reverse order of stratigraphic succession represented by Siwalik Group in the south followed by the Gondwana, Daling and Crystallines Group of rocks towards north. The rocks of the region have been stratigraphically classified as follows (Roy, 1976; Acharya, 1989):

Group	Formation	Rock Type
Kanchenjangha Group	----	Gneisses and quartzites
Darjeeling Group	----	Gneisses
Chungthang Group	Paro Formation	Quartzites and gneisses
Lingtse Granite	----	Granite gneisses
Daling Group	Buxa Formation	Dolomites, quartzites and slates
	Reyang Formation	Quartzites
	Gorubathan Formation	Feldspathic graywackes
Gondwana Group	Damuda Formation	Sandstones, shales and coal beds
	Talchir Formation	Slates, conglomerates and boulder beds
Siwalik Group	----	Sandstones, shales and conglomerates

In the present study area, the rocks of Daling Group, Lingtse Granite, Chungthang Group and Darjeeling Group have been found. Daling Group in the study area represents low grade metamorphic rocks namely Reyang quartzites and feldspathic graywackes. Gneisses and granite gneisses are the high grade metamorphic rocks in the area. In this study, the geological map prepared by Roy (1976) and Acharya (1989) has been referred. The study area does not contain any major thrust or fault. However, major lineaments are present there.

1.4 Organisation of the Thesis

In this chapter, the research objectives have been listed out, and a brief description of the study area has been given. Chapter 2 provides a review on landslides and various LSZ approaches, which have assisted in identifying the research gaps in this direction and developing a suitable approach for this study. In Chapter 3, the data sources used for this study and an overview of the methodologies have been described. Chapter 4 describes in detail the preparation of various thematic data layers. In Chapter 5, the details of implementation of conventional weighting approach for LSZ mapping has been presented. The details of implementation of neural and fuzzy approaches namely ANN black box approach, fuzzy set theoretic based approach and combined neural and fuzzy approach for LSZ mapping have been provided in Chapter 6. A comparative evaluation of different LSZ maps thereby four different LSZ mapping approaches have been presented in Chapter 7 to examine the efficacy of these approaches. In Chapter 8, a review on landslide risk assessment and the details of the proposed approach and its implementation for landslide risk assessment have been presented. Finally, Chapter 9 provides summary and conclusions.

Chapter 2

Landslides and Landslide Susceptibility Zonation – A Review

2.1 Introduction

Landslides are one of the most widespread and damaging natural hazards in hilly regions. The study of landslides has drawn global attention mainly due to increasing awareness of its socio-economic impacts as well as increasing pressure of urbanization on mountain environment (Aleotti and Chowdhury, 1999). Landslides constituted 4.89% of the natural disasters that occurred worldwide during the years 1990 to 2005 (www.em-dat.net). According to Schuster (1996), this trend is expected to continue in future also due to increased unplanned urbanization and development, continued deforestation and increased regional precipitation as a result of changing climatic conditions in landslide prone areas. Landslides cause loss of life and property, and damage to natural resources, developmental projects and essential commodities, etc. It has been estimated that, on an average, the damage caused by landslides in the Himalayas costs more than US$ one billion, besides causing about

200 deaths every year, which amounts to 30% of such losses occurring world-wide (Naithani, 1999). In 1998, due to massive landslides in Ukhimath area, Garhwal Himalayas, 109 people were dead and several families were affected. Also, Malpa landslide wiped out the whole Malpa village in Uttaranchal during 1998 and at least 210 people were dead (Juyal, 2002). Other major landslides namely Phata landslide of 2001, Budhakedar landslide of 2002 and Uttarkashi landslide of 2003 are burning examples in Himalayas that have caused large-scale human tragedies, resources damage and associated environmental-social hazards. Hence, landslide susceptibility studies are essential for safer strategic planning of future developmental activities in the Himalayan region.

This chapter provides a brief overview of different causative factors responsible for landslide occurrences and also a detailed review on approaches for landslide susceptibility zonation (LSZ) that are in vogue around the world.

2.2 Landslides

Landslides are the natural processes, which occur and recur in specific geo-environmental conditions. Landslide has generally been defined as the down slope movement of soil or rock masses as a result of shear failure at the boundaries of the moving mass (Skempton and Hutchinson, 1969). Another definition by Coates (1977) states landslide as abrupt short lived geomorphic event that constitutes rapid motion end of the mass spectrum.

2.3 Factors Responsible for Landslide Occurrence

A landslide is seldom attributed to a single causative factor. It is of fundamental importance to identify the causative factors for landslide occurrences in a region,

which often is difficult. It is also usually hard to establish the relationships between various causative factors. Nevertheless, it may be possible to demarcate landslide susceptible areas by identifying and analyzing the factors that have caused landslides in the past (Aleotti and Chowdhury, 1999).

There are two types of causative factors responsible for landslide occurrences; one relates to internal or preparatory, and the other to external or triggering (Crozier, 1986; Siddle et al., 1991). Internal factors assume a state which will allow the normal fluctuation of external factors to be sufficient to trigger a landslide. Although, internal factors may change over a long period of time to reduce the resistance/shear stress ratio, there is always an external factor which triggers the movement. The internal factors represent the inherent attributes of the ground which make the slopes susceptible to landslides.

Various researchers (e.g., Brabb, 1984; Carrara and Merenda, 1976; Cotecchia, 1978; Cruden and Varnes, 1996; Hansen, 1984; Hutchinson, 1996) have considered a number of causative factors that may be responsible for landslide occurrences in a region. These include (Varnes, 1984; Dikau et al., 1996; Naithani, 1999):

(A) Internal or preparatory factors:
- i. Lithology of slope material
- ii. Structural features
- iii. Geomorphology
- iv. Vegetation
- v. Hydrogeologic conditions

(B) External or triggering factors:
- i. Seismicity
- ii. Climate
- iii. Undercutting by river

iv. Anthropogenic factors:

(a) Land use change

(b) Unplanned construction

Lithology: Lithology basically involves the composition, texture, degree of weathering, as well as other details that influence the physico-chemical and engineering behaviours such as permeability, shear strength, etc. of the rocks and soils. These characteristics in turn affect the slope stability.

Structural features: In relation to landslides, the structural features include mainly the geological discontinuities such as bedding, joints, faults, folds and shear zones in the slopes. The inter-relationship between the slope and the discontinuities plays an important role particularly in rock slopes to understand the mechanism of failure. Further, the proximity of a slope to a tectonically active zone such as major faults or thrusts or lineaments influences the landslide activity to a great extent.

Geomorphology: An important geomorphologic characteristic of slope instability is to identify the nature and type of pre-existing landslides, as this governs the behaviour of the terrain. The geomorphology also includes slope morphology of the area i.e., slope angle and aspect and their physical features involving scarps, concavity/convexity, bulging toes, etc. The slope angle has a direct bearing on instability as the gravitational forces are accentuated with increasing slope angle. Aspect, which represents the direction of slope face, may have a local effect on slope stability.

Vegetation: Vegetation is an important factor in reducing the erosional activities on the slopes. A thickly vegetated slope reduces the effect of erosion because of

natural anchorage provided by the tree roots whereas barren slopes are generally more prone to erosional activity and therefore cause slope instability.

Hydrogeologic conditions: The water infiltration into the slope increases pore water pressure and decreases the shear strength, thereby causing instability to the slopes. The excessive surface run-off through drainages aggravates the erosional activity on the slopes. Therefore, the hydrogeologic conditions indicating the drainage network and the nature of distribution of surface and sub-surface water are also important for landslide occurrences.

Seismicity: The earthquake shocks may be responsible for triggering new landslides and reactivating old landslides. The vibrations due to earthquake may induce instability, particularly in loose and unconsolidated material on steep slopes.

Climate: The climatic pattern due to change in geographic location may influence landslide activities. High rainfall in tropical and sub-tropical climatic regions may trigger landslides, as in the Himalayas.

Undercutting action of river: The undercutting action of river removes the toe support to the slope thereby causing slope instability.

Land use change: The land use change, such as deforestation, exploitation of natural resources, conversion of vegetated slopes into built up area, etc. may result into landslide occurrences.

Unplanned construction: The overloading of slopes or removal of lateral support by human interference is a prime concern for slope failures in many areas. The ill-planned construction activities related to hill development programme such as road cutting, housing, quarrying, mining, etc. aggravate the problem of slope instability in hilly regions.

The effective selection of these causative factors is important and will depend on the study area, mapping scale, reliability as well as accuracy of the data (Aleotti and Chowdhury, 1999).

2.4 Landslide Susceptibility Zonation (LSZ) Mapping

Spatial prediction of landslide is termed as landslide susceptibility, which is a function of landslide and landslide related internal factors. The aim is to identify places of landslide occurrence over a region on the basis of a set of internal causative factors. This is specifically known as landslide susceptibility zonation (LSZ), which can formally be defined as the division of land surface into near-homogeneous zones and then ranking these according to the degrees of actual or potential hazard due to landslides (Varnes, 1984). Due to conceptual and operational limitations, landslide hazard zonation is conceptually stated as landslide susceptibility zonation (Brabb, 1984). However, LSZ maps do not directly incorporate time and magnitude (Fell, 1994; Cruden and Varnes, 1996; Hsli, 1975; Sassa, 1988). LSZ maps also differ from landslide inventory or distribution maps, which represent a database or catalogue of existing landslides over a region.

2.4.1 Basic Assumptions

All the available approaches for LSZ mapping are based upon some widely accepted assumptions (Varnes, 1984; Carrara et al., 1991; Hutchinson and Chandler, 1991; Hutchinson, 1996; Turner and Schuster, 1996), which can be stated as:

i) The past and present are keys to the future (Varnes, 1984; Carrara et al., 1991; Hutchinson, 1996). This implies that landslides in future will more likely to occur under similar geological, geomorphological, hydrogeologic and climatic

conditions, which were and are responsible for the occurrence of past and present landslides. Hence, experiences on existing landslides will be more helpful for landslide susceptibility assessment.

ii) Landslides with distinct geomorphological features can be identified, classified and mapped both through field surveys and remote sensing image interpretations (Rib and Liang, 1978; Varnes, 1978; Hansen, 1984; Hutchinson, 1988; Dikau et al., 1996).

iii) Landslides are controlled by identifiable internal factors (i.e., inherent attributes of the ground) known as causative factors, which can also be mapped from field surveys and remote sensing image interpretations (Dietrich et al., 1995).

Nevertheless, a number of obstacles may be faced while producing LSZ maps (Aleotti and Chowdhury, 1999). For example,

i) The discontinuous nature of landslides in space

ii) The difficulty in identifying the causative factors, which often is subjective

iii) Lack of complete historical data related to landslide occurrences.

2.4.2 Mapping Scale

The scale of LSZ mapping depends on three basic factors (Aleotti et al., 1996a):

i) The purpose of the study

ii) The extent of the study area and

iii) Data availability

The choice of the mapping scale affects the selection of the approach (Aleotti and Chowdhury, 1999). Thus, for example, geotechnical investigation based approach may be suitable for studies concerning individual slopes or small areas, whereas LSZ

approach may be suitable for a regional scale study. Further, the mapping scale for a landslide susceptibility zonation study will control the selection of different causative factors and also the level of detailed mapping. A scale of 1:25,000 to 1:50,000 is generally used for delineation of landslide susceptibility zones in hilly regions.

2.4.3 Mapping Unit

A mapping unit is a land surface that is homogeneous in itself and show heterogeneity with adjacent units (Hansen, 1984). LSZ requires the selection of a suitable mapping unit, which depends on a number of factors. These include type and degree of details of landslides to be studied; the scale of study; the quality, resolution, scale and type of input data; and the availability of analysis tools such as GIS and remote sensing. For example, in raster-based GIS approach for LSZ mapping is applied whereby the study area is divided into regular grids of pre-defined size depending on the data availability. These grid-cells or pixels serve as the mapping units of reference (Carrara, 1983; Bernknopf et al., 1988; Pike, 1988; van Westen, 1993, 1994; Mark and Ellen, 1995). In this approach, each pixel in the study area is assigned a value of importance or weight corresponding to each causative factor and the weights are integrated in GIS environment to generate a raster output layer.

2.5 Landslide Susceptibility Zonation (LSZ) Approaches – A Review

The landslide susceptibility zonation is a complex task (Brabb, 1991). Several approaches for LSZ mapping have been proposed. These approaches can be grouped into two broad categories; qualitative and quantitative. The taxonomy of different approaches for LSZ mapping is given in Figure 2.1. These LSZ approaches, which have been reviewed in detail by Hansen (1984), Varnes (1984), van Westen (1994),

Carrara and Guzzetti (1995), Hutchinson (1996), Mantovani, et al. (1996), Aleotti and Chowdhury (1999), Guzzetti et al. (1999) and Saha et al. (2005), are summarized in this section.

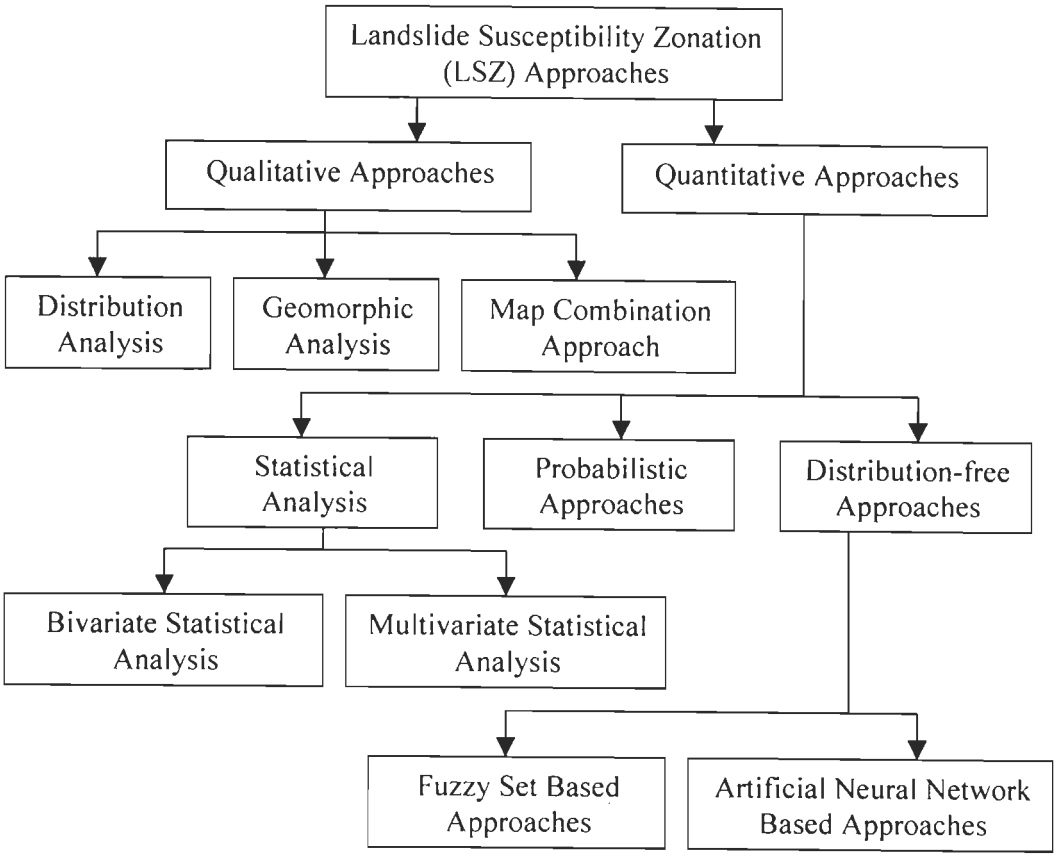


Figure 2.1: Flow chart showing taxonomy of LSZ approaches

2.5.1 Qualitative Approaches

In qualitative approaches, lot of subjectivity is introduced in preparation of various thematic data layers contributing for landslide occurrences, which are integrated to the generation of LSZ map of the area.

2.5.1.1 Distribution Analysis

Distribution analysis is a straightforward approach for landslide susceptibility zonation, which otherwise is known as landslide inventory. This approach shows the distribution of existing landslides mapped from aerial photographs, field surveys and/or historical data of landslide occurrences. These landslide inventory maps, in most of the cases, provide a basis for other landslide susceptibility zonation approaches. The main features and different contributors of this approach are listed in Table 2.1.

Table 2.1: Distribution analysis approach for LSZ mapping

LSZ approach		Main features	References	Approach used
QUALITATIVE	Distribution Analysis	Direct mapping of landslides and mass movements and representing in a map, which demarcates areas of past landslide occurrences.	Wright and Nilsen (1974)	Landslide isopleths based on density map of landslide distribution
			Canuti et al. (1979)	Landslide activity maps based on multi-temporal aerial photo interpretation
			Wieczorek (1984)	Landslide inventory map showing spatial distribution of landslides as point feature or as affected areas
			Espizua and Bangochea (2002)	LSZ mapping based on landslide inventory

The landslide inventory provides a spatial distribution of existing landslides represented on a map either as the affected areas (polygons) or as point events (Wieczorek, 1984). In another alternative, the landslide distribution was represented as a density map (Wright et al., 1974). Landslide isopleths were drawn by interpolating these density values. This method did not reflect the relationship

between the landslides and their causative factors, but it was useful in presenting landslide densities quantitatively (Turner and Schuster, 1996).

Espizua and Bengochea (2002) prepared susceptibility and risk zonation maps based on an inventory of landslides generated through field work and interpretation of aerial photographs. The purpose was to provide a practical basis for rational land use planning. Landslide susceptibility and risk zones were mapped, in view of the natural hazards and the degree of loss to elements at risk along roads and routes because of a given magnitude of landslide.

The landslide inventory maps do not provide information on the temporal changes in landslide distribution. Therefore, a modification in the inventory maps was done in the form of landslide activity maps, which were based on multi-temporal aerial photo interpretation (Canuti et al., 1979). These activity maps are useful to study the effect of temporal changes in land use on landslide activity.

The distribution analysis approaches are very time consuming, cumbersome and costly, but maps based on these approaches may be useful in providing first hand information on the landslide activities of the area. These maps do not provide information on the degree of susceptibility of future landslide activity.

2.5.1.2 Geomorphic Analysis

Geomorphological mapping of landslide susceptibility is a direct, qualitative approach that relies on the ability of the investigator or expert to estimate actual and potential slope failures (Guzzetti et al., 1999). The main features and different contributors of this approach are listed in Table 2.2.

In this approach, the LSZ is carried out directly in the field by scientists/geomorphologists, based on their experience in the subject, about the area

and in other similar situations, without describing any rules which have led to this assessment. The LSZ maps are directly evolved from detailed geomorphological maps.

Table 2.2: Geomorphic analysis approach for LSZ mapping

LSZ approach		Main features	References	Approach used
QUALITATIVE	Geomorphic Analysis	Direct qualitative approach of geo-morphological mapping of landslides which relies on the ability of investigator.	Carmassi et al. (1992)	Landslide susceptibility map based on field geomorphic analysis
			Hearn (1995)	Proposed a hazard classification sheet for mapping compiled directly in the field for different geomorphological units
			Soeters and van Westen (1996) and van Westen et al. (2003)	Digital geomorphological LSZ mapping

A lot of work on LSZ mapping using this approach has been carried out since 1970s (e.g., Carrara and Merenda, 1976; Kienholz, 1978; Fenti et al., 1979; Ives and Messerli, 1981; Kienholz et al., 1983, 1984; Zimmerman et al., 1986; Rupke et al., 1988; Seeley and West, 1990; Hansen et al., 1995; Soeters and van Westen, 1996; van westen et al., 2000; van Westen et al., 2003). One of the most comprehensive projects reported in the literature was the French ZERMOS maps (Humbert, 1977) which involved analysis of active and inactive landslides with respect to the factors responsible for landslide susceptibility and then extrapolation of similar physical conditions for preparation of LSZ maps. These maps generally showed three different classes with varying degrees of susceptibility (i) null or low susceptibility, (ii) potential or uncertain susceptibility, and (iii) ascertained susceptibility. The ZERMOS map of the Moyenne Vesubie region, France, prepared by Meneroud and Calvino

(1976) showed four zones of instability defined on the basis of five factors such as lithology, structures, slope, morphology and hydrology. Another ZERMOS map prepared by Landry (1979) identified seven classes of susceptibility on the basis of the factors like geological nature of the soil and sub-soil, slope angle, drainage and local history of landslides. The LSZ map of Carmassi et al. (1992) was used to identify the most favorable sites for construction of power plants. Hearn (1995) developed an LSZ map compiled directly in the field based on geo-morphological features at 1:10,000 scale. Soeters and van Westen (1996) and van Westen et al. (2003) reported LSZ mapping based on the geomorphological criteria for slope instability.

This approach allows a rapid assessment of landslide susceptibility in a given area. The main disadvantages of such approaches are (Leroi, 1996): (i) the subjective decision rules that govern the landslide occurrences; this fact makes it difficult to compare the LSZ maps prepared by different experts; (ii) difficult in updating the susceptibility assessment as new data becomes available; and (iii) extensive field surveys are required.

2.5.1.3 Map Combination Approach

The map combination approach for LSZ mapping involves a number of steps (Soeters and van Westen, 1996):

(i) Selection and mapping of the causative factors

(ii) Thematic data layer preparation with relevant categories of the factors

(iii) Assignment of weights and ratings to factors and their categories respectively

(iv) Integration of thematic data layers and

(v) Preparation of LSZ map showing different zones

The main features and some recent developments on this approach are listed in Table 2.3.

Table 2.3: Map combination approach for LSZ mapping

LSZ approach		Main features	References	Approach used
QUALITATIVE	Map Combination Approach	Direct or semi-direct approach in which several causative factor maps are combined together to generate LSZ map using subjective decision rules based on experience of the earth scientist.	Nagarajan et al. (1998)	Qualitative map combnation using GIS
			Anbalagan (1992)	Landslide Hazard Evaluation Factor (LHEF) subjective rating scheme for LSZ mapping
			Gupta et al. (1999)	Ordinal weighting-rating system based on relative importance of causative factors and map combination in GIS to generate LSZ map
			Saha et al. (2002)	Ordinal scale relative weighting-rating technique to assign landslide susceptibility index based on relative importance of causative factors and map combination in GIS to generate LSZ map
			Sarkar and Kanungo (2004)	Qualitative map combination in GIS using an ordinal rating system based on relative importance of factors and their categories

A review of literature reveals that the pre-requisite for LSZ mapping is the preparation of thematic data layers pertaining to different causative factors. Commonly these factors include lithology, lineament, slope, aspect, land use land cover, and drainage etc.

Early efforts on LSZ mapping considering lithology and slope as the causative factors were made by Blanc and Cleveland (1968) and Radbruch-Hall and Varnes (1973) in California, Bowman (1972) in Australia, Dobrovolny and Schmoll (1974) in

Alaska, Radbruch-Hall and Crowther (1976) in United States, Rodriguez Ortiz et al. (1978) in Spain and Obermeir (1979) in Virgenia. Brabb et al. (1972) first introduced the landslide frequency analysis with respect to litho units (geology) and slope categories by a simple superimposition method and produced an LSZ map. Varnes (1984) prepared an LSZ map considering slope, soil thickness, land use practice and drainage as the causative factors. Takei (1982) prepared a debris flow susceptibility map in Japan considering rock types, fracturing, weathering characteristics, springs, vegetation cover, valley slopes and historical records of large landslides as the contributory factors. In New Zealand, Eyles (1983) identified different types of erosion and their severity based on lithology, structure, slope and topography. In the last two decades, LSZ mapping was conventionally carried out based on manual interpretation of a variety of thematic data layers and their superimposition (Seshagiri and Badrinarayana, 1982; Anbalagan, 1992; Choubey and Litoria, 1990; Pachauri and Pant, 1992; Gupta et al., 1993; Sarkar et al., 1995; Mehrotra et al., 1996; Virdi et al., 1997; Turrini and Visintainer, 1998).

This approach is time consuming, laborious and uneconomical with data collected over long time intervals. Also, error may be introduced at each stage, as the process is largely manual.

In recent times, due to the availability of a wide range of remote sensing data together with data from other sources in digital form and their analysis using GIS, it has now become possible to prepare different thematic data layers corresponding to the causative factors responsible for the occurrence of landslides in a region (Gupta and Joshi, 1990; McKean et al., 1991; van Westen, 1994; Nagarajan et al., 1998; Gupta, 2003). The integration of these thematic layers with weights assigned according to their relative importance in a GIS environment leads to the generation of

an LSZ map (Carrara et al., 1991; van Westen, 1994; Lakhera and Champatiray, 1996; Nagarajan et al. 1998; Gupta et al., 1999; Rautela and Lakhera, 2000; Saha et al., 2002; Sarkar and Kanungo, 2004; Saha et al., 2005).

However, in this approach, the weights were assigned on the basis of the experience of the experts on the subject and about the study area. The weights may vary from expert to expert and also from region to region. The subjectivity in assigning weights to each thematic data layer and to its categories is the major limitation of this approach. Also, there is a difficulty in extrapolating a model developed for a particular area to other areas.

2.5.2 Quantitative Approaches

In order to minimize subjectivity in the weight assignment process, quantitative approaches, objective ways of quantifying the relative importance of various causative factors, can be deployed to produce an LSZ map. A number of approaches have been developed, which are summarized in the following sections:

2.5.2.1 Statistical Analysis

The statistical approaches have been adopted for LSZ studies to minimise the subjectivity in weight assignment procedure associated with qualitative approaches. The statistical approach compares the spatial distribution of existing landslides in relation to different causative factors (Aleotti and Chowdhury, 1999). GIS tools are quite useful in this analysis. Statistical approaches can broadly be classified into two types: bi-variate and multivariate.

a) Bi-variate statistical analysis

In bi-variate statistical analysis, each individual thematic data layer is compared to the existing landslide distribution layer. The weight value of each category of causative factors is assigned based on landslide density. This involves the overlay of landslide distribution layer on each of the thematic data layers, and calculation of respective landslide density values. Different approaches such as frequency analysis approach (Pachauri and Pant, 1992, 1998; Sarkar et al., 1995; Mehrotra et al., 1996), information value (InfoVal) approach (Yin and Yan, 1988; van Westen, 1997; Lin and Tung, 2003; Saha et al., 2005), landslide nominal risk factor (LNRF) approach (Gupta and Joshi, 1990) and land hazard evaluation factor (LHEF) approach (Anbalagan, 1992) can be adopted. The main features and some recent contributions on this approach are listed in Table 2.4.

The frequency analysis approach involves determination of normalized frequency distribution of landslides per unit area in each category of individual factors. This is achieved by overlaying the landslide layer on each thematic data layer manually or in GIS environment. These frequency values are used as the ratings of the respective categories of causative factors. Constant or arbitrary weights are assigned to the causative factors. These ratings and weights for the factors and their categories are integrated to produce the LSZ map.

The Information Value (InfoVal) approach for LSZ mapping considers the probability of landslide occurrence within each category of thematic data layer. The rating of a particular category of a thematic data layer is determined as:

$$W_i = \ln\left(\frac{Densclas}{Densmap}\right) = \ln\frac{Npix(S_i)/Npix(N_i)}{\sum_{i=1}^{n} Npix(S_i) \Big/ \sum_{i=1}^{n} Npix(N_i)} \qquad (2.1)$$

where, W_i denotes the weight given to the i^{th} category of a particular thematic data layer; *Densclas* denotes the landslide density within the category; *Densmap* denotes the landslide density within the thematic data layer; $Npix(S_i)$ denotes the number of pixels, which contain landslides, in a category; $Npix(N_i)$ denotes the total number of pixels in a category and n is the number of categories in a thematic data layer. The thematic data layers are overlaid and the ratings in the form of InfoVal are added to prepare a landslide susceptibility index (LSI) map, which is later categorized into five different landslide susceptibility zones to prepare an LSZ map.

Table 2.4: Bi-variate statistical analysis for LSZ mapping

LSZ approach		Main features	Reference	Approach Used
QUANTITATIVE	Bi-variate Statistical Analysis	Indirect methods in which statistical approach compares the spatial distribution of existing landslides with that of causative factors and the results are applied to areas currently free of landslides but conditions exist for susceptibility mapping.	Yin and Yan (1988)	InfoVal method for LSZ mapping
			Gupta and Joshi (1990)	Landslide nominal risk factor (LNRF) rating scheme based on statistical relationships of factors
			Carrara et al. (1991)	Statistical models for landslide hazard evaluation
			Pachauri and Pant (1992; 1998)	Weighted rating system based on relationships between factors
			Jade and Sarkar (1993)	LSZ mapping using Information value method
			Mehrotra et al. (1996)	Bi-variate statistics for category rating
			Turrini and Visintainer (1998)	Bi-variate statistics for category rating
			Lin and Tung (2003)	InfoVal method for potential analysis of factors for landslide and structural equation model for data integration/prediction
			Saha et al. (2005)	LSZ mapping using InfoVal method and modified m-LNHF technique

Another approach, known as the landslide nominal risk factor (LNRF) approach, was developed by Gupta and Joshi (1990), which determines the rating of each category of thematic data layers. The LNRF is determined using the following equation:

$$LNRF_i = \frac{Npix(S_i)}{\left(\sum_{i=1}^{n} Npix(S_i)\right)/n} \tag{2.2}$$

where, $Npix(S_i)$ denotes the number of pixels containing landslides in i^{th} category and n is the number of categories present in the particular thematic data layer. A higher value of LNRF (i.e., LNRF >1) implies more susceptibility to landslides than the average; an LNRF value <1 indicates less susceptibility to landslides; whereas, an LNRF value =1 indicates a category with an average landslide susceptibility. The LNRF values were regrouped broadly into three classes for each thematic data layer, and were assigned ratings 0, 1 and 2 for LNRF<0.67 (low susceptibility), 0.67<LNRF<1.33 (medium susceptibility) and LNRF>1.33 (high susceptibility) respectively. The thematic data layers were overlaid and the values were added to prepare an LSI map. The LSI values were classified into three susceptibility zones: low, medium and high. However, it has been observed that regrouping of LNRF values into ordinal numbers (0, 1, 2) leads to coarsening of approach and reduction in the relative importance of various categories. Therefore, Saha et al. (2005) proposed a modified LNRF approach known as modified landslide nominal hazard factor (m-LNHF), where the computed ratings were directly used without any regrouping.

The bi-variate statistical approaches are based on the observed relationships between each category of factors and the existing landslide distributions in the area. Although, the bi-variate statistical approaches are considered to be a quantitative approach for LSZ mapping, a certain degree of subjectivity exists, particularly in the

weight assignment procedures for different causative factors. In all cases, constant weights or arbitrary weights have been assigned to the causative factors for LSZ mapping.

b) Multivariate statistical analysis

Multivariate approaches consider relative contribution of each thematic data layer to the total susceptibility within a defined area. The procedure involves several important steps (Aleotti and Chowdhury, 1999):

i) Identification of percentage of landslide affected areas in each pixel and their classification into stable and unstable zones,

ii) Preparation of an absence/presence matrix of a given category of a given thematic layer,

iii) Multivariate statistical analysis (e.g., discriminant and regression analyses) and

iv) Reclassification of the area based on the results and their classification into susceptibility classes.

These approaches involve analysis of large volume of data and are time consuming. External statistical packages are generally used to support the GIS packages. The statistical analyses most frequently used for LSZ mapping are discriminant analysis (Carrara, 1983; Carrara et al., 1990) and multiple regression analysis (Bernknopf et al., 1988; Yin and Yan, 1988; Jade and Sarkar, 1993; Wieczorek et al., 1996; Atkinson and Massari, 1998; Chung and Fabbri, 1999; Clerici et al., 2002). The main features and some recent contributions on this approach are listed in Table 2.5.

Table 2.5: Multivariate statistical analysis for LSZ mapping

LSZ approach		Main features	Reference	Approach used
QUANTITATIVE	Multi-variate Statistical Analysis	Indirect methods in which statistical approach compares the spatial distribution of existing landslides with that of causative factors and the results are applied to areas currently free of landslides but conditions exist for susceptibility mapping.	Kawakami and Saito (1984)	LSZ mapping using Quantification method II of Hayasi (1952) for quantification of qualitative data
			Yin and Yan (1988)	Regression analysis technique used as prediction model
			Jade and Sarkar (1993)	LSZ mapping using regressive multiple analysis technique
			Dhakal et al. (2000)	Quantification scaling type II discriminant analysis (multivariate statistics) for LSZ
			Clerici et al. (2002)	Conditional analysis method (multivariate statistics) for LSZ

Carrara (1983) applied multivariate analyses approaches (e.g., discriminant analysis and multiple regression analysis) for LSZ mapping in Southern Italy. These approaches proved to be useful in predicting actual and potential landslide susceptibility. In this study, a group of geological-geomorphological attributes, which are directly or indirectly correlated with slope instability, were used in the discriminant functions and in the regression equation. The slope units were discriminated successfully into stable and unstable areas. It was reported that in multiple regression analysis, lithology and its interaction with slope angle contributed significantly in predicting the percentage of unstable areas. However, the result of these statistical approaches underlined the need of other factors capable of improving the efficiency of the approach.

Yin and Yan (1988) analysed 21 categories of different factors based on data collected from field investigation and landslide mapping. Regression analysis was

used to establish different degrees of instability for the preparation of LSZ map of the area. Clerici et al. (2002) applied conditional analysis approach for LSZ mapping which simultaneously took into account all the factors contributing to instability. The landslide density of each pixel was computed in correspondence to different combinations of causative factors and an LSZ map was prepared based on the landslide density values. It has been observed that this approach is difficult to implement and requires complex operations. Further, to achieve satisfactory results, the procedure has to be repeated few times changing the combination of factors and their categories.

The limitations of multivariate statistical approach can be listed as follows:

i) Discriminant and regression analyses require data derived from a normally distributed population that is frequently violated.

ii) A mixture of continuous (i.e., slope, aspect, etc.) and categorical (i.e., lithology, land use land cover, etc.) factors leads to incorrect solution.

iii) Some of the factors may bear weak physical relationship with landslide occurrences. Combination of such factors with other factors may generate data which is very difficult to interpret.

2.5.2.2 Probabilistic Approach

The probabilistic approaches have also been used for LSZ studies to minimise the subjectivity in weight assignment procedure. This approach compares the spatial distribution of landslides in relation to different causative factors within a probabilistic framework. Some of methods based on this approach include conditional probability model, weight of evidence method under Bayesian probability model,

certainty factor method under favorability mapping model, etc. The main features and some recent contributions on these approaches are listed in Table 2.6.

Table 2.6: Probabilistic approaches for LSZ mapping.

LSZ approach		Main features	Reference	Approach used
QUANTITATIVE	Probabilistic Approach	Indirect methods of analysis within a probabilistic framework	Chung and Leclerc (1994)	Favorability modeling (FM) approach using certainty factors (CF) for LSZ mapping
			Chung and Fabbri (1999)	Joint conditional probability model with five different estimation procedures for LSZ mapping
			Lee et al. (2002a; 2002b)	Bayesian probability model using weight of evidence method for landslide susceptibility analysis
			Chung and Fabri (1998) Lee and Min (2001)	Probabilistic prediction model based on likelihood ratio function in combination with Bayesian combination rule for LSZ mapping
			Lan et al. (2004)	Certainty Factor Model for LSZ

Favourability modeling (FM) approach is a good compromise, offering a valid quantitative method, where subjectivity or expert knowledge can be incorporated in the analysis, particularly when data are not sufficient or reliable. With FM, thematic data can be transformed into continuous data, by considering the degree of relationship between the landslides and the categories of each thematic data layer. Each continuous or non-continuous category can be transformed into a value, called favourability value. The certainty factor (CF) approach is one of the possible proposed favorability functions (FF) to handle the problem. The CF, defined as a function of probability, originally proposed by Shortliffe and Buchanan (1975) and later modified by Heckerman (1986) can be given as:

$$CF = \begin{cases} \dfrac{pp_a - pp_s}{pp_a(1 - pp_s)} & if\ pp_a \geq pp_s \\ \dfrac{pp_a - pp_s}{pp_s(1 - pp_a)} & if\ pp_a \geq pp_s \end{cases} \tag{2.3}$$

where pp_a is the conditional probability of having a number of landslide event occurring in category a and pp_s is the prior probability of having the total number of landslide events occurring in the study area. The range of CF values varies from -1 to 1. A positive value means an increasing certainty in landslide occurrence, while a negative value corresponds to a decreasing certainty in landslide occurrence. A value close to zero means that the prior probability is very similar to the conditional one. By integrating the CF values of the categories of thematic data layers, an LSZ map can be prepared. This CF model was considered and experimentally investigated by various researchers (e.g., Chung and Fabbri, 1993, 1998; Chung and Leclerc, 1994; Binaghi et al., 1998; Luzi and Pergalani, 1999; Remondo et al., 2003; Lan et al., 2004).

Chung and Fabri (1999) proposed a conditional probability model for LSZ mapping. Five different procedures namely direct estimation, Bayesian estimation under conditional independence, regression model, modified Bayesian model and modified regression model were adopted for estimating conditional probability of landslide susceptibility. GIS-based existing landslide distribution layer and various thematic data layers were used to prepare the LSZ map. The LSZ maps were validated by comparing with the later landslides. It was observed that multivariate regression analysis generated better results than other probability methods.

Lee et al. (2002a, 2002b) applied Bayesian probability model using the weight-of-evidence method of Bonham-Carter (1994) for LSZ mapping. Using the location of landslides and topographic factors, the method was used to calculate the weights

(positive and negative) and contrast (difference of positive and negative weights) for each category of different causative factors. The contrast was used as the rating of each category. The contrast is positive for a higher influence on landslide occurrences and negative for a lower influence on landslide occurrences. The ratings of the thematic data layers were summed to calculate the landslide susceptibility index (LSI). The LSI values were categorized into different susceptibility zones to prepare an LSZ map. van Westen et al. (2003) also used the weights of evidence approach to generate statistically derived ratings for all categories of thematic data layers. On the basis of these ratings, a judicious choice of relevant thematic data layers was made for preparation of an LSZ map.

The application of probabilistic prediction model based on likelihood ratio function for LSZ mapping was discussed by Chung and Fabri (1998) and Lee and Min (2001). The existing landslide locations and different thematic data layers were used to implement the model. The probability frequency distribution functions of the landslide affected and non-affected areas should be distinctly different. The likelihood ratio function, which is the ratio of the two frequency distribution functions, can highlight this difference. For each category of thematic data layers, two empirical distribution functions for the landslide affected and non-affected areas were computed and the likelihood ratio for all the categories were determined. The LSZ map was prepared using the likelihood ratio values as the ratings of the categories.

The probabilistic approaches are based on the observed relationships between each category of factors and the existing landslide distributions in the area within a probabilistic framework. The thematic data (continuous and categorical) can be transformed into continuous data, by considering the degree of relationship between the landslides and the categories of each thematic data layer. Although, the

probabilistic approaches are considered to be a quantitative approach for LSZ mapping, a certain degree of subjectivity in the weight assignment procedures for different causative factors exists.

2.5.2.3 Distribution-free Approaches

Generally, qualitative approaches are highly based on experts experience and knowledge and can be considered as subjective (conventional). On the other hand, the quantitative approaches, such as statistical (bi-variate and multivariate) and probabilistic approaches, can be considered as more objective due to their data-dependent character. However, success of these approaches is highly affected by the number, quality and reliability of data (Ercanoglu and Gokceoglu, 2004). Therefore, to overcome these limitations, some new approaches such as fuzzy logic, artificial neural networks (ANNs), etc. may be adopted for LSZ mapping on a regional scale. Recently, fuzzy set theory, neural networks and combined neural and fuzzy approaches have been used to generate LSZ maps.

Fuzzy set theory can provide us with a natural method of quantitatively processing multiple datasets. Fuzzy relations play an important role in fuzzy modeling and in the context of LSZ mapping; fuzzy relations can be established based on the philosophy that landslides are related to some extent or unrelated to the causative factors. On the other hand, the most attractive aspect of ANN approaches is the ability to express the nonlinearities in the process to solve the problem similar to the human brain reasoning. Due to uncertainties in the causative factors used in LSZ mapping and the nonlinear character of landslides, utilization of these approaches can be considered as useful alternatives. The fuzzy and ANN approaches are also free from any distributional assumptions or bias of the data and the weights are computed

in an objective manner. The main features and some recent contributions on these approaches are listed in Table 2.7.

Table 2.7: Distribution-free approaches for LSZ mapping

LSZ approach		Main features	Reference	Approach used
QUANTITATIVE	Distribution-free Approaches	Artificial neural networks (ANNs) and neuro-fuzzy based approaches, which do not depend on the distributional assumptions of the data. The weights here are computed in an objective manner.	Chi et al. (2002b)	Fuzzy inference network using combination of fuzzy OR and fuzzy gamma operator for LSZ mapping
			Gorsevski et al. (2003)	Combination of fuzzy k-means classification and a Bayesian approach for spatial prediction of landslide hazard
			Tangestani (2003)	Numerical ratings of LHEF scheme of Anbalagan (1992) were fuzzified for factor maps and were combined to generate susceptibility map using a fuzzy gamma operator
			Ercanoglu and Gokceoglu (2004)	Fuzzy relation concept to determine the strength of relationship between landslide and factor categories and fuzzy max operator for landslide susceptibility mapping
			Elias and Bandis (2000)	Neurofuzzy system for LSZ mapping
			Arora et al. (2004)	ANN Black Box approach for landslide hazard zonation
			Lee et al. (2004)	Neurofuzzy approach for LSZ mapping where the ratings for factor classes were determined using a probability method and the weights of factors by ANN method
			Gomez and Kavzoglu (2005)	ANN Black Box approach for assessment for LSZ mapping
			Yesilnacar and Topal (2005)	ANN and logistic regression based landslide susceptibility mapping and their comparison

Chi et al. (2002b) discussed the effectiveness of fuzzy set theory for landslide susceptibility mapping. The relationships between input causative factors and past landslides in terms of likelihood ratio functions of each thematic data layer were computed and used as fuzzy membership values. These membership values were able to highlight the difference between areas affected by past landslides and areas not affected by past landslides. Fuzzy inference networks using a variety of different fuzzy operators, especially combination of fuzzy OR and fuzzy Gamma operator were used for data integration to prepare the LSZ map. It was observed that fuzzy Gamma operator with high value could effectively integrate most datasets for LSZ mapping. Tangestani (2003) also performed LSZ mapping using land hazard evaluation factor (LHEF) rating scheme of Anbalagan (1992) for determination of fuzzy membership values and fuzzy gamma operator for thematic data layer integration. The LSZ map was validated based on past landslides. It was suggested to evaluate the efficacy of fuzzy gamma operator for data integration in LSZ mapping.

Gorsevski et al. (2003) demonstrated that LSZ mapping can be achieved through an integration of GIS, fuzzy *k*-means and Bayesian modeling approaches. In the modeling approach, the optimal number of categories was derived by iterative classification for a range of categories or from expert knowledge. The continuous fuzzy k-means classification provided significant amount of information about the character and variability of data and proved to be a useful indicator for landslide susceptibility mapping. The probabilities were revised with Bayes theorem after the categories with similar characteristics were grouped together by fuzzy *k*-means approach. A broad range of causative factors were integrated through continuous fuzzy k-means classification to prepare an LSZ map. It was observed that the LSZ mapping using the integrated fuzzy/Bayesian approach produced better spatial

prediction of existing landslide locations than qualitative models. It was suggested to analyze each individual model in greater detail to improve the understanding between the processes.

Ercanoglu and Gokceoglu (2004) developed a model based on fuzzy relation concept for preparation of LSZ map. The landslide distribution layer was analyzed in relation to the categories of various thematic data layers to compute the fuzzy membership values for each category. By integrating the fuzzy membership values, the LSZ map was prepared. The LSZ map was validated with the existing landslides in the area. The fuzzy relation concept is an objective approach for determination of fuzzy ratings of different categories based on actual landslide data. Hence, this approach introduces relativity concept in rating determination. However, other quantitative approaches such as statistical and probabilistic ones consider the actual landslide data for determination of rating in a crisp manner without employing the relativity.

Arora et al. (2004) proposed an ANN black box approach for LSZ mapping. This approach determines the weights objectively in an iterative process, but the weights in this case remain hidden. The neural network training and testing datasets were prepared using the attributes of various thematic data layers representing the input neurons and the existing LSZ map (Saha et al., 2002) representing the single output neuron. After successful training and testing of different neural network architectures, the best architecture for this specific problem was selected based on the highest training and testing accuracies. The adjusted connection weights of the best network were used to generate the LSZ map of the area. The distribution of landslide susceptibility zones derived from ANN showed similar trends as that observed with the existing landslide locations in the field. A comparison of the results was made

with an earlier produced GIS-based LSZ map of the same area and indicated that ANN results were better than the earlier method.

Gomez and Kavzoglu (2005) also used ANN black box approach for LSZ mapping. In this process, a multilayer perceptron with back propagation learning algorithm was used. This approach used a wide range of causative factors and the existing landslide distribution layer derived from digital elevation model, remote sensing imagery and documentary data for neural network training and testing data preparation. Neural network architecture of 9/28/1 (9 input neurons, 28 hidden neurons and one output neuron) was used for training and testing. After the training and testing process, an LSZ map was generated for the whole area. The existing landslides were considered to validate the LSZ map. It was observed that the predictions were close to reality, indicating a satisfactory performance of the model.

Yesilnacar and Topal (2005) prepared landslide susceptibility maps using both logistic regression analysis and ANN approaches. For this purpose, 19 different thematic data layers were used. In ANN approach, a feed forward back propagation algorithm was adopted. They used single hidden layer neural network architecture. The connection weights of neural networks have been used to determine the weights for the chosen input thematic layers. The landslide susceptibility map produced using the ANN approach predicted higher percentage of landslides, especially in high and very high zones than the logistic regression analysis method.

Elias and Bandis (2000) proposed a neuro-fuzzy approach for LSZ mapping. Fuzzy linguistic rules were used to assign fuzzy membership values to different categories of thematic data layers. The fuzzy membership values were used to provide data to the input neurons for neural network model. A single output neuron with values from 0 to 1 was considered to represent the degree of landslide susceptibility

based on actual landslide data. The back error propagation neural network was used for training and an LSZ map was prepared for the area. The trained network was also used for another area to generate the LSZ map. The existing landslides in both the areas were considered to validate the LSZ maps. It was observed that the predictions were close to reality, indicating a satisfactory performance of the model.

Lee et al. (2004) attempted the development, application and assessment of probabilistic and artificial neural network approaches for LSZ mapping. Landslide locations and causative factors were used for analyzing landslide susceptibility. A probabilistic method was used for determination of rating of each category and an artificial neural network approach was used for determination of weights of causative factors. The rating of each category was determined using the likelihood ratio function (Lee and Min, 2001). The weight of each factor was determined after artificial neural network training (Hines, 1997). The existing landslide locations and no-landslide areas were used to randomly generate ten sets of training data. The back error propagation neural network was used to train the networks for all the training datasets used. Neural network architecture of 7/15/2 (7 input neurons, 15 hidden neurons and 2 output neurons) was considered for the study. The initial connection weights between the neurons were assigned random values. After successful training of the network, the weights of the factors were determined based on the weight matrices analysis for all the 10 training datasets. The normalized average value of ten different weights for a particular factor was considered as the weight of the corresponding factor. The LSZ maps were prepared by integrating the ratings of the categories only and also by integrating the ratings and the weights together. The two LSZ maps were verified using the existing landslide locations. The verification results were reasonable and acceptable.

It can be observed from the above review that the distribution-free approaches (i.e., fuzzy, ANN and neuro-fuzzy) are able to determine the weights and ratings of the causative factors and their categories in an objective manner. These approaches also have the ability to handle continuous, categorical and binary data pertaining to various causative factors for LSZ mapping. However, the fuzzy set based approach addresses the determination of ratings of the categories only. In most of the ANN approaches for LSZ mapping, single neural network architecture has been attempted. However, an optimal architecture exists for each specific problem of LSZ mapping, as pointed out by Arora et al. (2004). It can also be observed that the weights for the causative factors remain hidden in case of ANN black box approach.

2.6 Summary

The literature review presented in this chapter suggests that broadly there are two groups of approaches: qualitative and quantitative approaches for LSZ mapping. The qualitative approaches, such as distribution analysis, geomorphic analysis, map combination methods, etc., were very popular at late 1970s among engineering geologists and geomorphologists. The quantitative approaches became popular in the last decades depending on the advancements in the developments of remote sensing and GIS technologies. Advantages or disadvantages of different LSZ mapping approaches have been commonly discussed by the experts in the field of landslide studies in the literature. The qualitative approaches rely on expert knowledge or experience which dictates the selection, the weighting and the combination function of the factors and therefore, can be considered as conventional or subjective. The quantitative models involve the use of mathematics and statistics to express the relationships between the existing landslide distribution and the categories of factors.

Therefore, these can be considered as more objective than conventional approaches due to the fact that data-dependent character and much less experience is needed. However, success of these approaches is highly affected by the quantity, quality and reliability of data. Statistical and probabilistic approaches require the collection of huge amount of data to produce good results. Also, these approaches contribute in determining the ratings of the categories of factors only, but consider constant or arbitrary weights for all the factors to generate the LSZ maps. Therefore, some distribution-free approaches such as fuzzy set based and ANN based approaches have been attempted to evaluate the landslide susceptibility in recent years. The fuzzy set based approach addresses the determination of ratings of the categories only. In most of the ANN black box approaches for LSZ mapping, single neural network architecture has been attempted. However, an optimal architecture exists for each specific problem of LSZ mapping. It can also be observed that the weights for the causative factors remain hidden in case of ANN black box approach. The connection weight analysis seems to be an alternative approach for determination of weights of the causative factors. Moreover, a combination of ratings determined through fuzzy set based approach and weights obtained through ANN connection weight analysis, can be explored for LSZ mapping.

Chapter 3

Data Sources and Overview of Methodology

3.1 General

This chapter describes different data sources with their specific use in the present research study. Broad overview of the methodology and the theory behind different LSZ approaches developed and implemented in the present study, have also been discussed.

3.2 Data Sources

A variety of data from three different sources have been collected and processed to achieve the research objectives. These sources are:

i) Remote sensing data

ii) Ancillary data

iii) Field data

3.2.1 Remote Sensing Data

Remote sensing data from Indian Remote Sensing (IRS) satellite have been acquired from the National Remote Sensing Agency (NRSA), Hyderabad, India. The IRS systems are under the umbrella of National Natural Resources Management System (NNRMS) and the programme is coordinated at national level by the Planning Committee of NNRMS (PC-NNRMS).

In recent times, data from sensors on-board IRS satellites are being used for a diverse range of applications such as crop acreage estimation, drought monitoring, flood risk zone mapping, hydro-geomorphological maps preparation for locating underground water resources, land use and land cover mapping, , mineral prospecting, etc.

In the present study, the preparation of thematic data layers such as land use land cover, lineaments, drainage, landslide etc. is based on data collected by LISS and PAN sensors placed in IRS-1C and IRS-1D satellites. The satellite payload consists of three sensors, namely Panchromatic Sensor (PAN), Linear Imaging and Self-Scanning Sensor (LISS-III) and Wide Field Sensor (WiFS) (Navalgund and Kasturirangan, 1983).

LISS-III sensor provides multi-spectral data in 4 bands. The spatial resolution for visible (two bands) and near infrared (one band) bands is 23.5m with a ground swath of 141km whereas the fourth band (short wave infrared band) has a spatial resolution of 70.5m with a ground swath of 148km. The PAN sensor provides data in a single broad band with a spatial resolution of 5.8m and a ground swath of 70km at nadir view. This data is highly useful in recognizing ground features. In this study, the LISS-III and PAN data have been used for land use land cover classification,

lineaments mapping, drainage network mapping and landslide mapping. Brief description of data used and sensor characteristics are presented in Table 3.1.

Table 3.1: Brief description of IRS-1C LISS-III and IRS-1D PAN data and sensor characteristics.

Characteristics	IRS-1C LISS-III	IRS-1D PAN
Path/Row	107/52	107/52
Date of acquisition	22.03.2000	03.04.2000
Spectral Bands	4	1
Spectral ranges	B2: 0.52 - 0.59μm	0.52 – 0.75μm
	B3: 0.62 - 0.68μm	
	B4: 0.77 - 0.86μm	
	B5: 1.55 - 1.70μm	
Spatial resolution	B2-B4 : 23.5m B5 : 70.5m	5.8m
Swath	B2-B4 : 141km B5 : 148km	70km
Quantization	7 bit	6 bit

3.2.2 Ancillary Data

The ancillary data such as topographic maps and geologic map have been collected from the relevant sources (Table 3.2). The Survey of India topographic maps have been used to create the base map of the study area. A number of information such as contours, drainage network, roads, *etc.* have been extracted from these maps. The contours have been digitized to prepare digital elevation model (DEM), slope and aspect data layers. Geological map has been used to generate data on lithologic features in the study area.

Table 3.2: Ancillary data used in this study.

Data type	Specification	Source
Topographic Maps	Toposheet Nos. 78A/4, A/8, B/1 Scale: 1:50,000	Survey of India (1960's)
	Toposheet Nos. 78A/8/2, A/8/3, A/8/6, B/5/1 Scale: 1:25,000	
Geologic Map	Scale: 1:250,000	Geological Survey of India (Roy, 1976; Acharya, 1989)

3.2.3 Field Data

Field data are very important for any remote sensing and GIS based study to evaluate the results. Extensive field data have been collected during the years 2001 to 2003 to get information on existing landslide distribution, lithologic contacts and to collect training and testing areas for land use land cover classification. The months of March and April were preferred for field data collection, as this season corresponded to the date of acquisition of satellite pass.

3.3 Overview of Methodology

3.3.1 General

This study has been carried out using an integrated remote sensing-GIS approach. The overview of the broad methodology adopted in the present study is outlined in Figure 3.1. The input data sources have been described in this chapter. The detailed descriptions of other steps (thematic data layer preparation, LSZ approaches, their implementation and evaluation and landslide risk assessment) have been given at relevant places in the chapters to follow. However, in the following section, a

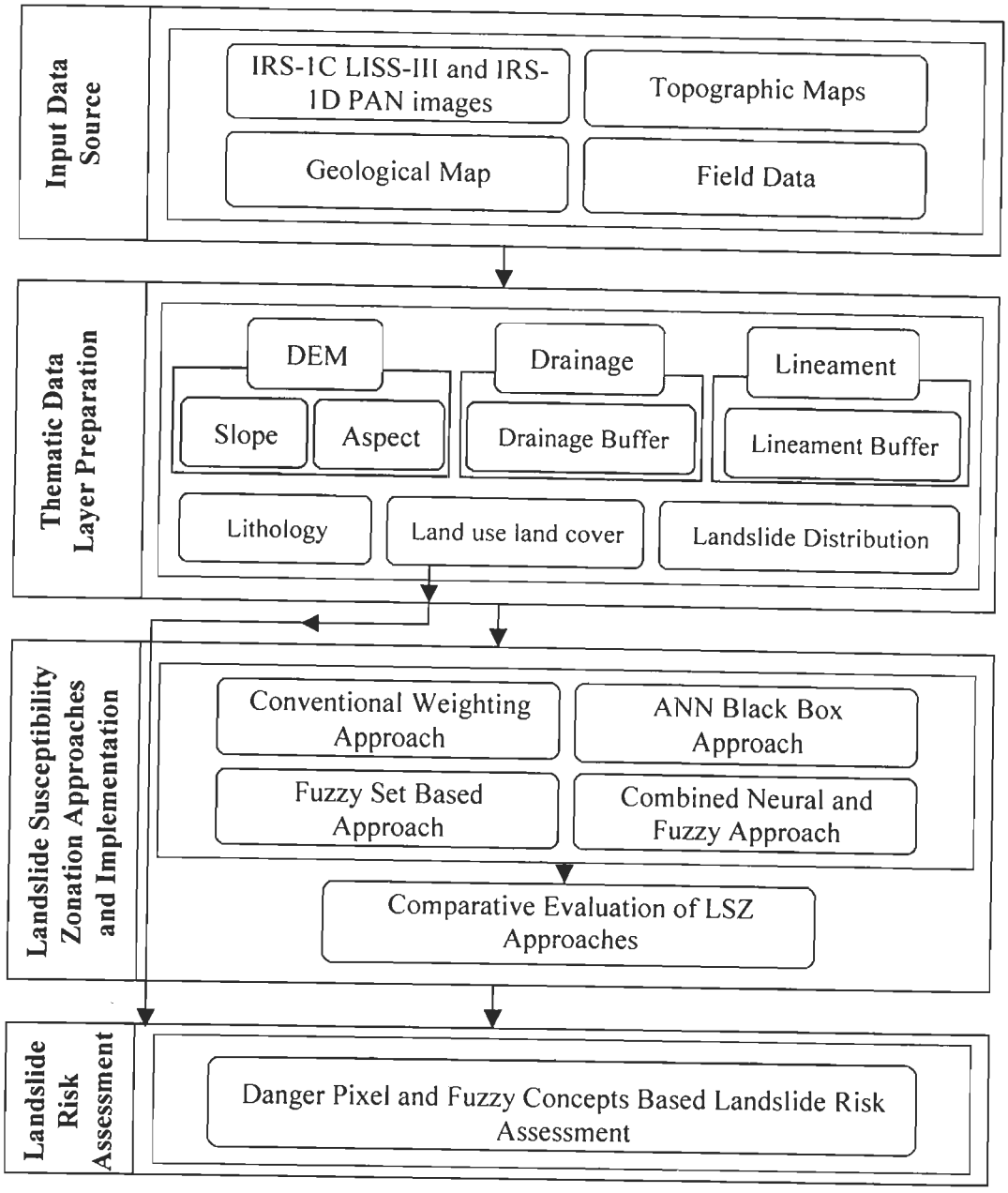


Figure 3.1: Flow diagram showing an overview of the methodology adopted in this study.

description of the LSZ approaches developed and adopted in the present study has been outlined.

3.3.2 LSZ Approaches

The LSZ mapping has attracted disaster specialists, geomorphologists, engineering geologists, scientists, academicians and others, and several approaches for LSZ have been developed. A detailed review of these approaches has already been given in Chapter 2.

Although several approaches have been developed for LSZ mapping, a number of issues that affect the performance of these approaches have been identified, including the difficulties in handling both continuous (slope and aspect) and categorical data together. However, only a few studies have been attempted and that too at preliminary stage that consider the application of fuzzy logic and ANN approaches to resolve these issues. The principal aim of the present study is to develop and implement different approaches of LSZ for predicting areas susceptible to landslides in parts of Darjeeling Himalayas using conventional, fuzzy logic, ANN and combined neural and fuzzy concepts.

The following sections highlight the salient features of different approaches adopted and developed in the present study to produce LSZ maps.

3.3.2.1 Conventional Weighting Approach

The weights of the causative factors (lithology, lineaments, drainage, land use land cover, slope and aspect in the present case) and ratings of the categories can be combined to generate an LSZ map of the area. The assignment of these weights and ratings can be achieved by adopting a conventional (subjective) weighting approach.

In this approach, the weights and ratings to the causative factors and their categories are assigned based on the experience of the experts on the subject and about the study area. These weights and ratings generally take numerical values at an ordinal scale of 0 to 9. Higher is the numerical value of weight or rating, greater is its influence on the occurrence of landslide.

The weighted thematic data layers are generated by arithmetically multiplying the weight of the layer with the ratings of the corresponding categories of each layer. These weighted data layers are then added together to generate an LSI map. Hence, the LSI is a combined effect of both weights and ratings. Higher is the LSI, greater is the landslide susceptibility. The LSI values, thus obtained for the study area, are categorized into very high susceptibility (VHS), high susceptibility (HS), moderate susceptibility (MS), low susceptibility (VHS) and very low susceptibility (VLS) zones to prepare an LSZ map.

3.3.2.2 ANN Black Box Approach

ANN is a useful approach for regression and classification problems and promises to be suitable for the delineation of landslide susceptibility zones as this approach has the capability to analyse complex data patterns. Also, ANN can process data at varied measurement scales such as continuous (slope, aspect, etc.), categorical (lithology, land use land cover, etc.) and binary (presence or absence of landslide) data, a scenario which is often encountered in LSZ mapping.

An ANN comprises of three layers such as input (I), output (O) and one or two hidden (H_A and H_B) layers in between these two (Paola and Schowengerdt, 1995). Each layer in a network contains sufficient number of neurons depending on the specific problem. In the present case, six different thematic data layers (t) provide

data to the six neurons (i_1 to i_6) in the input layer and a single neuron (o) in the output layer represents the desired output (d) of either existing landslide locations or any of the five landslide susceptibility zones (i.e., VHS, HS, MS, LS and VLS). The number of hidden layers and their neurons are often determined by trial and error (Gong, 1996). In this case, two hidden layers (H_A and H_B) with varying number of neurons (j) have been considered. The neurons in a layer are connected to the neurons of the next successive layer and each connection carries a weight (c_{ij}) (Atkinson and Tatnall, 1997). Each neuron responds to the weighted inputs it receives from the connected neurons from the preceding layer (Lee et al., 2004).The ANN architecture is illustrated in Figure 3.2.

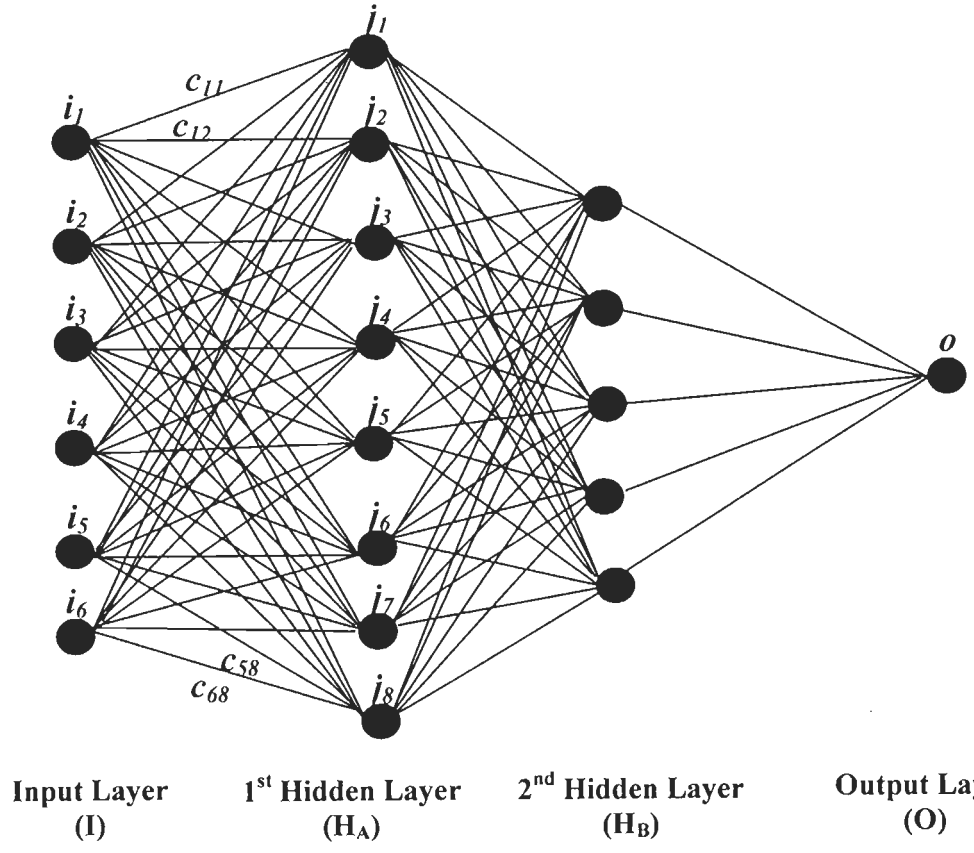


Figure 3.2: A schematic diagram of artificial neural network (6/8/5/1).

In the present study, two different cases with respect to preparation of datasets have been examined. In the first case, pixels corresponding to existing landslide locations have been considered whereas in the second case, pixels corresponding to one of the landslide susceptibility zones of the LSZ map produced from conventional weighting approach have been considered for data preparation of ANN processing. Two independent training and testing datasets have been prepared for each case. Each dataset consists of mutually exclusive pixels. The training dataset is used to train various neural network architectures while the testing dataset is used to evaluate the accuracy of neural network outputs. The testing dataset has also been used as verification dataset to control the overtraining of the networks.

Data from input neurons is processed through hidden neurons to generate an output in the output neuron. In this process, a feed-forward multilayer network is generally used. Generally, the input that a single neuron j in the 1st hidden layer (H_A) receives from the neurons (i) in its preceding input layer (I) can be expressed as:

$$net_j = \sum_{i=1}^{t} c_{ij} p_i \qquad (3.1)$$

where c_{ij} represents the connection weight between input neuron i and hidden neuron j, p_i is the output from input neuron i and t is the number of input neurons (i.e., 6 different thematic data layers in the present case). The output value produced by the hidden neuron j, p_j, is the transfer function, f, evaluated as the sum produced within neuron j, net_j. Hence, the transfer function f can be expressed as:

$$p_j = f(net_j) \qquad (3.2)$$

The function f is usually a non-linear function that is applied to the weighted sum of the inputs and then the combined effect proceeds to the next layer. Any differentiable nonlinear function can be used as a transfer function, but a sigmoid function is generally used (Schalkoff, 1997). The sigmoid function constrains the outputs of a network between 0 and 1. As in case of LSZ mapping, the desired output (d_o) is represented by the presence/absence (1 or 0) of landslides and one of the five landslide susceptibility zones from VLS to VHS (0.2 to 1.0), the sigmoid transfer function has been used for input-hidden, hidden-hidden and hidden-output layers. The sigmoid function in case of LSZ mapping can be expressed as:

$$p_j = f(net_j) = \frac{1}{1 + e^{-net_j}} \tag{3.3}$$

In this feed-forward process of neural network, the network output value for the output neuron o, p_o is obtained. The error is calculated as the difference of desired output (d_o) and network output (p_o). The error function, e, is a measure of network's performance for the processing elements in the output layer and can be expressed as follows:

$$e = 0.5 \sum_{o=1}^{s} (d_o - p_o)^2 \tag{3.4}$$

where d_o is the desired output vector, p_o is the network output vector and s is the number of training samples (Arora et al., 2004).

The error gets back-propagated through the neural network and is minimized by changing the connection weights between neurons of different layers. This can be

executed through a number of learning algorithms based on back propagation learning (Ripley, 1996; Haykin, 1999; Zhou, 1999; Arora et al., 2004; Lee et al., 2004; Gomez and Kavzoglu, 2005; Yesilnacar and Topal, 2005). The most widely used back propagation algorithms are gradient descent and gradient descent with momentum. These are often too slow for the solution of practical problems. The faster algorithms use standard numerical optimizers such as conjugate gradient, quasi-Newton and Levenberg-Marquardt approaches. Levenberg-Marquardt algorithm was designed to approach the second-order training speed like quasi-Newton methods without having to compute the Hessian Matrix (i.e., the second derivatives of the performance index of weights). This algorithm uses an approximation to the Hessian matrix in the following manner:

$$c_{ij+1} = c_{ij} - \left[J^T J + \mu I\right]^{-1} J^T e \tag{3.5}$$

where c_{ij} is a vector of current connection weights, J is the Jacobian Matrix which contains first derivatives of the network errors with respect to connection weights, e is a vector of network errors and μ is a scalar. Unlike gradient descent algorithms, the Levenberg-Marquardt algorithm does not consider learning rate and momentum factor as its parameters. It takes into account important training parameters such as mu (μ), mu_dec and mu_inc. The parameter μ is decreased by multiplying it with mu_dec after each successful step (reduction in error) and is increased only when the error is increased. The main scalar parameter μ is modified in an adaptive fashion after giving an initial random value. In this study, Levenberg-Marquardt algorithm (implemented as TRAINLM in MATLAB software) has been used for training the neural network.

The details of this algorithm can be found in Hagan and Menhaj (1994) and Hagan et al. (1996).

The process of back-propagating the error is repeated iteratively until the error is minimized to an acceptable value over the training dataset and the adjusted connection weights are captured. The adjusted weights obtained from the trained network have been subsequently used to process the testing dataset to evaluate the generalization capability and accuracy of the network. The performance of the network depends on the accuracy obtained over the training and testing datasets. A number of neural network architectures has been designed, trained and tested. The optimal neural network architecture for the dataset is identified based on the training and testing accuracies. The connection weights obtained from this network are captured and subsequently used to determine the network output for the whole dataset. The range of network output values, thus obtained, have been categorized into VHS, HS, MS, LS and VLS zones to prepare the LSZ map of the study area.

In this process, the network connection weights between different layers have not been analyzed to determine the weights and ratings of the causative factors and their categories. Thus, in this approach, the weights and ratings remained hidden. Hence, this approach is called as ANN black box approach.

3.3.2.3 Fuzzy Set Based Approach

Fuzzy set theory was firstly introduced by Zadeh (1965) but has now been commonly used in various research fields in different disciplines. Fuzzy relation concept defined by Zadeh (1973) is based on the theory of fuzzy sets. In the fuzzy set theory, membership degrees of elements have varying degrees of support or confidence in (0,1) interval. Thus, a fuzzy set can be explained as a set containing

elements that have varying degrees of membership in the set (Ross, 1995). Fuzzy relations play an important part in fuzzy modeling. Fuzzy relations are based on the philosophy that everything is related to some extent or unrelated (Dubois and Prade, 1980). There are different ways such as Cartesian product, closed-form expression, linguistic rules of knowledge, similarity methods in data manipulation, etc. to develop numerical values characterizing a fuzzy relation. All these methods attempt to determine some sort of similarity in data (Ross, 1995).

In this research, one of the well known similarity methods, cosine amplitude method, has been used to determine the relationship between the landslide occurrence and the factors responsible for such activity. The membership values of the categories of each factor are calculated by the strength of the relationship (r_{ij}) between the existing landslides and the factors.

Let n be the number of categories of the thematic layers represented as an array $X = \{x_1, x_2, \ldots, x_n\}$, each of its elements, x_i, is a vector of pixels p (i.e., the size of the image in the present context) and can be expressed as,

$$x_i = \{x_{i1}, x_{i2}, \ldots, x_{ip}\} \tag{3.6}$$

Each element of a relation, r_{ij}, results from a pair-wise comparison of a category of a thematic layer i (i.e., layer corresponding to a causative factor) with a category of thematic layer j (i.e., landslide distribution layer), say x_i and x_j containing elements x_{ik} and x_{jk} respectively. In the cosine amplitude method, r_{ij} (membership grades) between categories of a thematic data layer and that of landslide distribution layer are computed by the following equation with its values ranging from 0 to 1 ($0 \leq r_{ij} \leq 1$),

$$r_{ij} = \frac{\left[\sum_{k=1}^{p} x_{ik} \times x_{jk}\right]}{\sqrt{\left(\sum_{k=1}^{p} x_{ik}^2\right) \times \left(\sum_{k=1}^{p} x_{jk}^2\right)}} \qquad (3.7)$$

Values of r_{ij} close to 0 indicate dissimilarity whereas values close to 1 indicate the similarity between two datasets. The equation (2) leads to n-1 r_{ij} images corresponding to each category of the thematic layers under consideration. These images show r_{ij} values at the pixels belonging to the category in question whereas rests of the pixels indicate 0 values. The corresponding r_{ij} images for various categories of a thematic layer are added together to generate an image for that thematic layer, where l varies from 1 to t thematic layers belonging to each causative factor (e.g., 6 thematic layers in the present case). Thus, six different R_l images are produced. These R_l images have been arithmetically added to generate an LSI map of the area. The range of LSI values in the LSI map has been categorized into VHS, HS, MS, LS and VLS zones to prepare an LSZ map.

In this approach, only the ratings of the categories have been determined through cosine amplitude fuzzy method. However, the weights of the causative factors have been considered as constant or unity. Hence, a combined neural and fuzzy approach has been attempted, where both the weights and ratings can be determined objectively.

3.3.2.4 Combined Neural and Fuzzy Approach

The combined neural and fuzzy approach for LSZ mapping involved the processes of rating determination of the categories through cosine amplitude fuzzy similarity method (as discussed in Section 3.3.2.3) and weight determination of

causative factors through ANN connection weight analysis. In this approach, the artificial neural network connection weights are used to characterize the input data sources (e.g., the causative factors) in terms of ranks or weights. The procedures for design of ANN architecture, ANN parameters and ANN training and testing are kept same as discussed in Section 3.3.2.2. The output neurons in this approach represent the landslide and no-landslide locations in the study area. Three independent training, testing and verification datasets are formed. Each dataset consists of mutually exclusive pixels corresponding to landslide and no-landslide pixels. The training dataset is used to train various neural network architectures while the testing dataset is used to evaluate the accuracy of neural network outputs. The verification dataset has been used to control the overtraining of the networks.

In this process, the connection weight matrices for input-hidden, hidden-hidden and hidden-output layers are obtained for each two-hidden layer network. Simple matrix multiplications of these weight matrices give rise to the final weight matrix corresponding to the factors. For example, for a network of 6/14/8/1 architecture considered, connection weight matrices of 6×14, 14×8 and 8×1 are obtained. The product of 6×14 and 14×8 matrices gives a resultant matrix of 6×8. Subsequently, the product of 6×8 and 8×1 matrices gives an output matrix of 6×1 which corresponds to the weights of 6 causative factors. The absolute values of these weights are considered in the present work to rank the factors. The factor with maximum absolute weight is assigned as rank 1 and the factor with the minimum absolute weight as rank 6. This process of weight and rank determination is illustrated in Figure 3.3.

Several neural network architectures are experimented with to finalize the rank of each factor based on the rankings observed by these networks by applying the majority rule. Finally, the absolute weights of corresponding networks for a particular

|1| Input(I) – Hidden A (HA) Connection Weights

	HA_1	HA_2	HA_3	HA_4	HA_5	HA_6	HA_7	HA_8	HA_9	HA_{10}	HA_{11}	HA_{12}	HA_{13}	HA_{14}
I_1	20.45	-12.98	3.52	-0.98	23.70	9.68	-1.92	-22.60	12.37	-3.26	5.35	12.36	16.72	9.68
I_2	-49.57	-16.84	51.54	87.70	-20.19	52.71	-79.29	-24.83	43.90	-91.70	-61.46	-46.21	-43.25	15.49
I_3	-31.63	-48.57	-52.19	47.39	-11.27	52.44	-5.31	-12.58	44.33	3.84	-46.30	24.22	44.39	-66.34
I_4	12.83	-9.89	18.44	0.15	21.10	-3.93	-16.33	-24.43	-20.15	14.86	-15.81	3.08	-23.33	13.79
I_5	-2.28	9.72	6.48	3.14	-0.06	8.78	8.53	3.95	5.98	1.89	-3.46	-4.09	-4.07	-0.30
I_6	13.46	-12.42	14.70	5.39	-11.54	9.72	2.81	-3.41	10.30	6.19	17.86	25.66	1.28	20.66

|2| Hidden A (HA) – Hidden B (HB) Connection Weights

	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
HA_1	-1.41	2.15	-3.63	-1.70	2.96	1.06	-1.71	-2.60
HA_2	3.10	-1.31	0.69	-1.21	-0.90	0.60	-1.92	2.65
HA_3	1.42	-0.88	-0.07	-1.85	-0.90	-0.20	0.19	2.53
HA_4	-1.17	-0.25	-0.66	1.76	2.04	1.55	-0.72	0.60
HA_5	2.56	-2.74	0.08	-5.36	-0.76	-1.32	1.86	-2.10
HA_6	1.98	1.50	-0.45	3.35	-1.65	1.91	-3.33	1.71
HA_7	-1.54	2.76	2.04	-3.07	0.60	-1.84	-2.96	2.43
HA_8	1.84	-1.46	1.27	0.90	-2.89	-1.83	0.76	-0.08
HA_9	-2.55	-0.45	-0.96	-1.43	-3.94	-0.21	-3.50	-0.20
HA_{10}	1.42	-1.84	-2.44	0.59	0.42	-4.59	-3.45	-1.41
HA_{11}	0.01	-2.68	0.48	1.37	1.55	-0.52	-0.13	-0.71
HA_{12}	-0.30	-1.71	-1.84	2.01	1.34	0.27	-1.09	-0.04
HA_{13}	-2.62	-1.19	1.75	2.32	-1.97	-2.60	1.65	-2.55
HA_{14}	4.22	0.96	-1.32	-2.72	-2.99	-2.59	0.61	-2.55

|3| Connection Weight Products of |1| & |2|

	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
I_1	-64.53	-10.19	-125.47	-111.32	10.13	-11.87	-21.14	-184.95
I_2	66.70	259.49	29.64	234.11	-377.06	884.93	225.86	446.47
I_3	-686.94	81.93	48.73	741.85	-52.38	229.35	-242.08	24.89
I_4	193.99	11.19	-204.81	-211.31	151.55	13.25	70.24	-83.25
I_5	48.23	21.67	19.55	-36.32	-60.27	3.15	-96.88	95.38
I_6	-4.84	0.50	-154.42	69.02	13.00	-42.62	-138.08	-59.90

|4| Hidden B – Output Connection Weights

	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
Output	3.43	-1.61	-3.87	4.83	2.71	-2.14	-3.67	-3.25

|5| Connection Weight Products of |3| & |4|

Y / X	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
I_1	-221.34	16.41	485.57	-537.68	27.45	25.40	77.58	601.09
I_2	228.78	-417.78	-114.71	1130.75	-1021.83	-1893.75	-828.91	-1451.03
I_3	-2356.2	-131.91	-188.59	3583.14	-141.95	-490.81	888.43	-80.89
I_4	665.39	-18.02	792.61	-1020.63	410.70	-28.36	-257.78	270.56
I_5	165.43	-34.89	-75.66	-175.43	-163.33	-6.74	355.55	-309.99
I_6	-16.60	-0.81	597.61	333.37	35.23	91.21	506.75	194.68

$$I_X = \left| \sum_{Y=1}^{8} HB_{XY} \right|$$

	Weight	Rank	Normalized Weight
I_1	474.48	5	0.544
I_2	4368.48	1	5.007
I_3	1081.22	3	1.239
I_4	814.47	4	0.933
I_5	245.06	6	0.281
I_6	1741.44	2	1.996

Figure 3.3: Steps for computations of connection weight matrices of ANN to characterize input data layers in terms of ranks and weights (in present study, I_1-Land use land cover, I_2-Lithology, I_3-Slope, I_4-Aspect, I_5-Drainage buffer and I_6–Lineament buffer). Note: In step |5|, *X*: 1-6; *Y*: 1-8.

factor are averaged. The average weights of each factor are then normalized at a scale of 0 to 10. The weight thus obtained was treated as the final weight value assigned to that causative factor.

The ratings of the categories obtained through cosine amplitude fuzzy similarity method and the weights of the causative factors determined through ANN connection weight analysis are integrated to generate an LSI map of the area. The LSI values thus obtained has been categorised into VHS, HS, MS, LS and VLS zones to prepare an LSZ map of the study area.

3.4 Computing Resources

Implementation of these approaches for LSZ mapping mainly includes thematic data layers preparation, weight and rating determination through fuzzy and ANN concepts and their raster-based integration. The manual (conventional) way of implementing different phases of activities would have been just impossible keeping in view the quantum of tasks involved in this study. To overcome this problem, mostly an integrated remote sensing and GIS approach has been adopted.

The computing resources used for implementation of every phase of activities involved in the present case are outlined in Table 3.3. Various thematic data layers were prepared from the input data sources using an image processing (Erdas Imagine) and a GIS (ArcView) software. The calculations for implementing cosine amplitude fuzzy similarity concept are performed in MS Excel spreadsheet. The artificial neural network processing has been implemented in Neural Network Tool Box of MATLAB Software. Each of the LSZ approach has been implemented in ArcView GIS software to prepare the LSZ maps.

Table 3.3: Computing resources used for implementation of different approaches for LSZ mapping.

Activities		Computing resources
Thematic Data Layer Preparation		Erdas Imagine ArcView GIS
LSZ Mapping	Conventional Weighting Approach	ArcView GIS
	ANN Black Box Approach	MATLAB ArcView GIS
	Fuzzy Set Based Approach	MS Excel spreadsheet ArcView GIS
	Combined Neural and Fuzzy Approach	MATLAB MS Excel spreadsheet ArcView GIS

Chapter 4

Thematic Data Layer Preparation

4.1 Introduction

Preparation of thematic data layers or maps is the most crucial step for GIS-based landslide susceptibility zonation mapping in this research. A thematic data layer represents the spatial distribution of an attribute or the spatial relationships among several selected attributes of a particular theme, such as slope, land use land cover, lineaments, etc. (Demers, 2000). Successful generation of an accurate LSZ map depends upon the quality of input thematic data layers. Therefore, appropriate methods and precautions have been used here in generating the thematic data layers, which act as spatial database for this study. The spatial database mainly consists of two parts: i) existing landslide distribution layer and ii) layers corresponding to causative factors affecting landslides. These factors include lithology, DEM and its attributes (slope and aspect), lineament, land use land cover and drainage. The selected factors represent the inherent ground conditions or parameters of the terrain

which make the slope susceptible to landslides and thus, have been considered responsible for the occurrence of landslides in the region. Soil types and soil depth may also be important causative factors for landslide studies; however, in this study area, the depth of weathering is very shallow and there is relatively thin soil cover. As the bedrock is almost invariably exposed in road cuttings and as natural outcrops, soil types and soil depth have not been considered in this study.

Moreover, rainfall and earthquakes are external (triggering) factors with heavy dependence on temporal phenomena. Further, data on past records of landslide occurrence in relation to rainfall and earthquakes are not available. Therefore, these have also not been considered, as the aim here is to assess the probability of occurrence of landslides at spatial level.

A thematic data layer corresponding to the existing landslide distribution has been carefully prepared using remote sensing and field data. This is used to establish spatial correlations between existing landslides and the factors for LSZ mapping of the area.

This chapter deals in various aspects of data capturing, data pre-processing and processing techniques to generate different thematic data layers. Both conventional and advanced data processing techniques have been adopted. The tasks involved in the preparation of different thematic data layers in this study can be enumerated as:

i) Pre-processing of remote sensing data

ii) Generation of DEM and extraction of attributes, such as slope, aspect

iii) Preparation of lithology data layer from existing geological map

iv) Generation of lineament and lineament buffer maps

v) Generation of drainage order and drainage buffer maps

vi) Land use land cover map preparation using multi-source classification of spatial data

vii) Preparation of landslide distribution map using remote sensing images and field data

4.2 Pre-processing of Remote Sensing Data

The remote sensing images need to be pre-processed adequately before proceeding for extraction of useful information from them. The images often contain distortions with respect to its geometry and radiometry (e.g., atmospheric effects). Therefore, it is necessary to rectify these distortions. The rectified image is then subjected to a number of digital image processing operations, such as contrast enhancement, image ratioing, image classification, etc. An overall framework of various digital image processing operations adopted here for various tasks is shown in Figure 4.1.

4.2.1 Geo-referencing

The digital images acquired from remote sensing satellites are fraught with geometric distortions, which render them unusable, as these may not be directly correlated to ground locations (Gupta, 2003). Geo-referencing involves the process of assigning map coordinate information to the image data so that the geometric integrity of the map in the image is achieved.

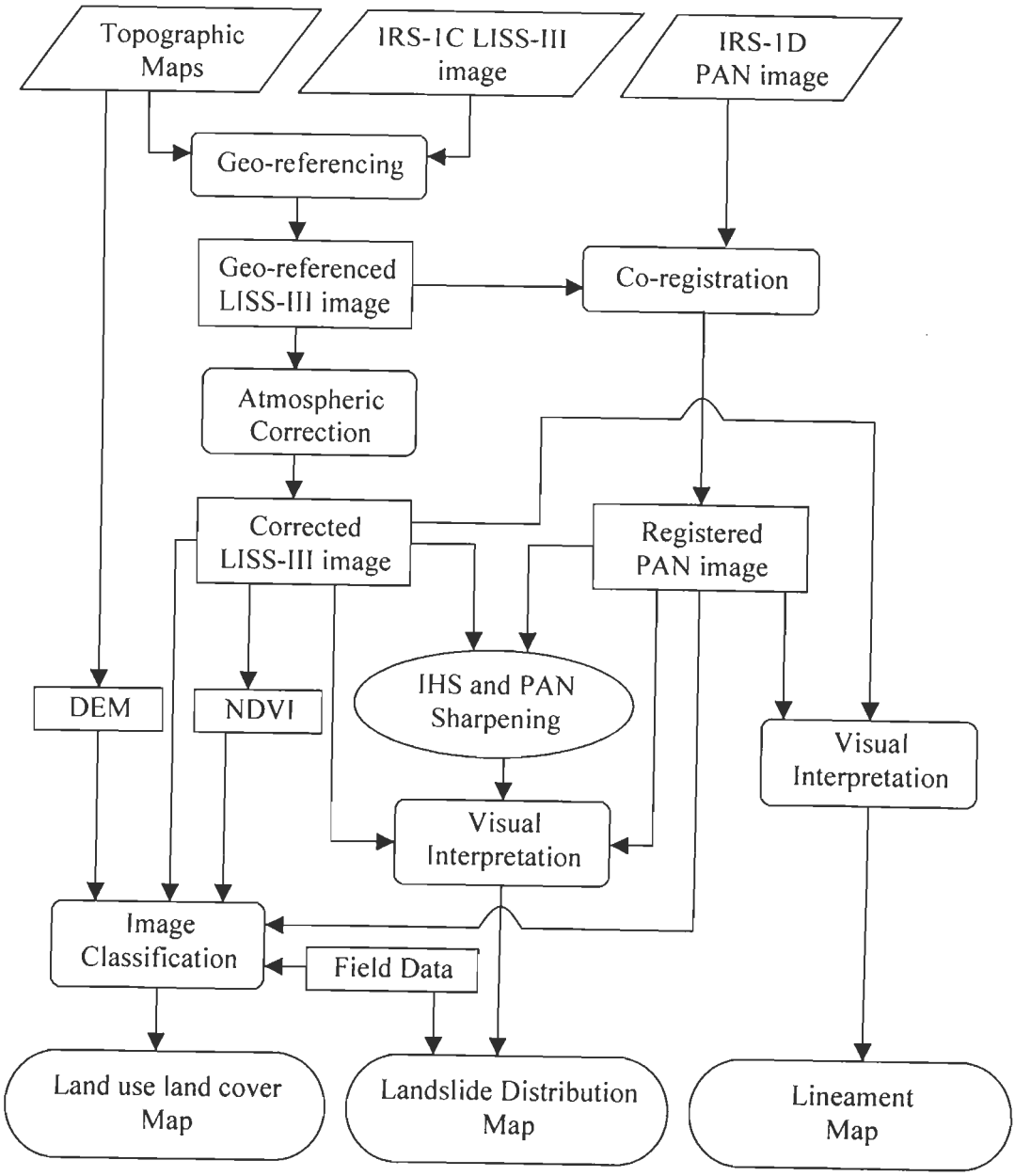


Figure 4.1: An overall framework of various image interpretation operations adopted in this study.

Geo-referencing is commonly performed using the method of rubber-sheet stretching that has been explained in many standard texts (e.g., Mather, 1999; Gupta, 2003). A number of ground control points (GCPs) uniformly distributed over the entire study area are collected such that these can easily be located on both the image to be geo-referenced and the reference image or map, such as a topographic map.

In this study, the geo-referencing of remote sensing data (IRS-1C LISS-III and IRS-1D PAN images) has been performed. In the first step, the scanned topographic maps have been geo-referenced to geographic coordinate system. In the process, input ground control points (GCPs) have been taken on the grid points of the topographic map and reference coordinates have been put through key board. To have a nominal measurement in meters, the georeferenced topographic maps have been reprojected to polyconic projection system with Modified Everest spheroid and datum. These maps have been later used as reference maps for geo-referencing of satellite images and geological map.

4.2.1.1 Geo-referencing of LISS-III image

The IRS-1C LISS-III image has been geo-referenced with the topographic maps by taking input GCPs from the LISS-III image and reference GCPs from topographic maps. The intersections of higher order drainages/rivers are mostly considered as the GCPs in view of the absence of other distinct features in the area. As the time difference between the survey time for preparing topographic maps (1960s) and the remote sensing data (2000) is about four decades, in such a hilly terrain the shifting of the high order drainage/river can be assumed to be negligible within this span of time. Therefore, it is believed that the GCPs identified for geo-referencing will be accurate to a large extent. A total of 35 well distributed GCPs have been considered for

registration and an RMS control point error of 0.83 pixel is obtained. Also, the registration is checked with another set of independent 11 GCPs, which yielded an RMS error of 0.96. It is found that the RMS errors obtained using 1st order polynomial model for geometric correction is within the acceptable limit of one pixel. The nearest neighbor resampling method has been adopted to produce the geo-referenced LISS-III image (Figure 4.2), as this preserves the original brightness values in the output image.

4.2.1.2 Registration of PAN image with LISS-III image

Co-registration of IRS-1D PAN with IRS-1C LISS-III image is essential in view of the fact that the PAN image is used for selecting training and testing data samples for multispectral (LISS-III) image classification. Therefore, the PAN image has been registered with LISS-III image by taking input GCPs from the PAN image and reference GCPs from LISS-III image. In this co-registration process, a set of 50 well distributed GCPs produced an RMS control point error of 0.71 pixel and 15 check GCPs yielded an RMS error of 0.68 pixel – both of which are within the acceptable limit of less than a pixel. The 1st order polynomial model with nearest-neighbor resampling method is used for this purpose. The registered PAN image thus obtained is shown in Figure 4.3.

4.2.2 Atmospheric Corrections

The optical remote sensing images invariably contain the effect of selective atmospheric scattering and absorption of the solar radiation. In the visible – near infrared region of the electromagnetic spectrum, scattering is the most dominant process leading to path radiance. This has an additive role and affects the brightness

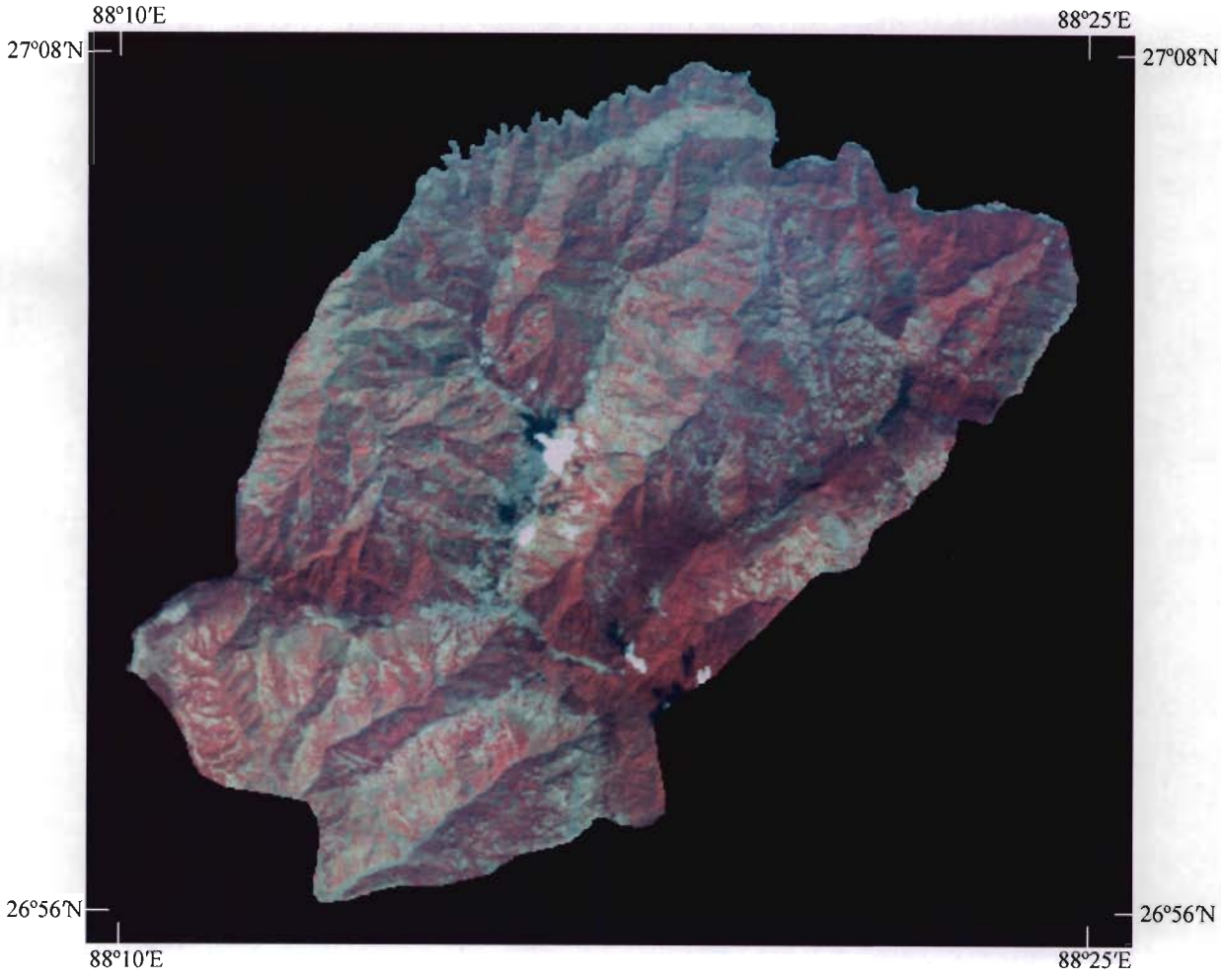


Figure 4.2: IRS-1C LISS-III false colour composite (NIR=R, Red=G, Green=B).

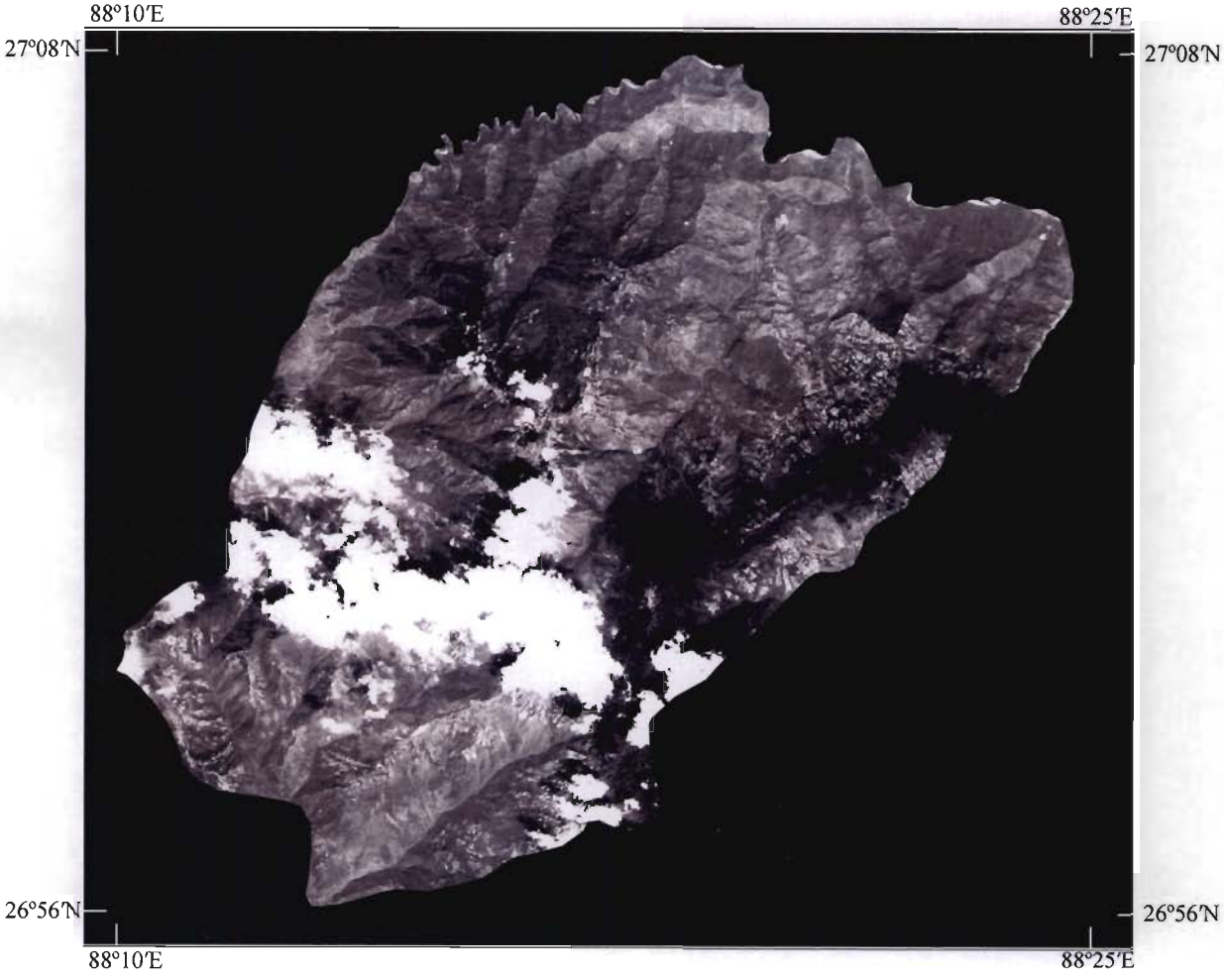


Figure 4.3: IRS-1D PAN image of the area. (Note: A portion of image is covered with clouds).

values (Jensen, 1996). The remote sensing data, therefore, need to be corrected. Although, there are many techniques to perform this correction, the most widely used, the 'dark object subtraction' technique (Chavez, 1988) has been adopted to correct the atmospheric scattering. An empirical approach has been used to rectify the remote sensing images for path radiance. The minimum DN values for green, red, near infrared (NIR) and shortwave infrared (SWIR) bands were extracted and were expected to be due to path radiance. These values were subtracted from DN values of pixels in the respective bands to generate a path radiance corrected image.

The corrected LISS-III and PAN images formed the data sources for preparation of different thematic data layers required for landslide susceptibility studies.

4.3 Landslide Distribution Layer

Mapping of existing landslides is essential to understand the relationships between the landslide distribution and the causative factors. As, it is just not possible to map each and every landslide via field surveys in such a rugged terrain, a comprehensive mapping of landslide was undertaken through remote sensing image interpretation, aided by field verifications at sampled locations.

The identification of landslides on remote sensing image is based on the spectral characteristics, shape, contrast and the morphological expression. In general, there is a distinct spectral contrast between landslides and the background area. High spatial resolution IRS-1C-PAN and PAN-sharpened LISS-III images have been used for landslide identification, recognition and mapping.

On the PAN image, landslides appear as features of very light tones due to rock debris without any vegetation on the slope. After enhancing the contrast of the PAN

image, landslides occurring in barren areas could also be identified. A few old landslides were identified on the basis of their shape, landform and drainage.

Feature extraction and interpretation is highly effective by using PAN-sharpened multi-spectral image products (Welch and Ehler, 1987; EOSAT, 1994; Sabins, 1996; Sharma et al.. 1996; Saraf, 2000; Prakash, et al., 2001; Sanjeevi et al., 2001; Shanmugam and Sanjeevi, 2001; Gupta, 2003). This product has the advantage of having spatial characteristics from the PAN image and spectral information from the multispectral image. For generating such product, the co-registered PAN and LISS-III images were used. The colour infrared composite of multispectral LISS-III image was split or transformed into intensity-hue-saturation (IHS) components. The Intensity component was replaced with the high resolution PAN image, which was followed by a reverse transformation. The resultant false colour composite (FCC) is the PAN-sharpened image (Figure 4.4). On the PAN-sharpened LISS-III image, the landslides appear as bright-white features (due to high reflectance) that are easily distinguished from other features. Further, landslides are also characterized by fan shape, sharp lines of break in topography and sometimes due to local drainage anomaly. Often, the toe part of the slide gives rise to a debris flow channel.

Many of the landslides identified on both PAN and PAN-sharpened LISS-III images have also been verified in the field. A total of 101 landslides of varying dimensions (180 m^2 to 27400 m^2) were identified from remote sensing images (Figure 4.5) and field surveys. Field photos of some landslides are shown in Figure 4.6. Majority of landslides have an areal extent of 500 m^2 - 2000 m^2. Most of the observed landslides are rock slides. However, in some cases, complex types of failure are also present.

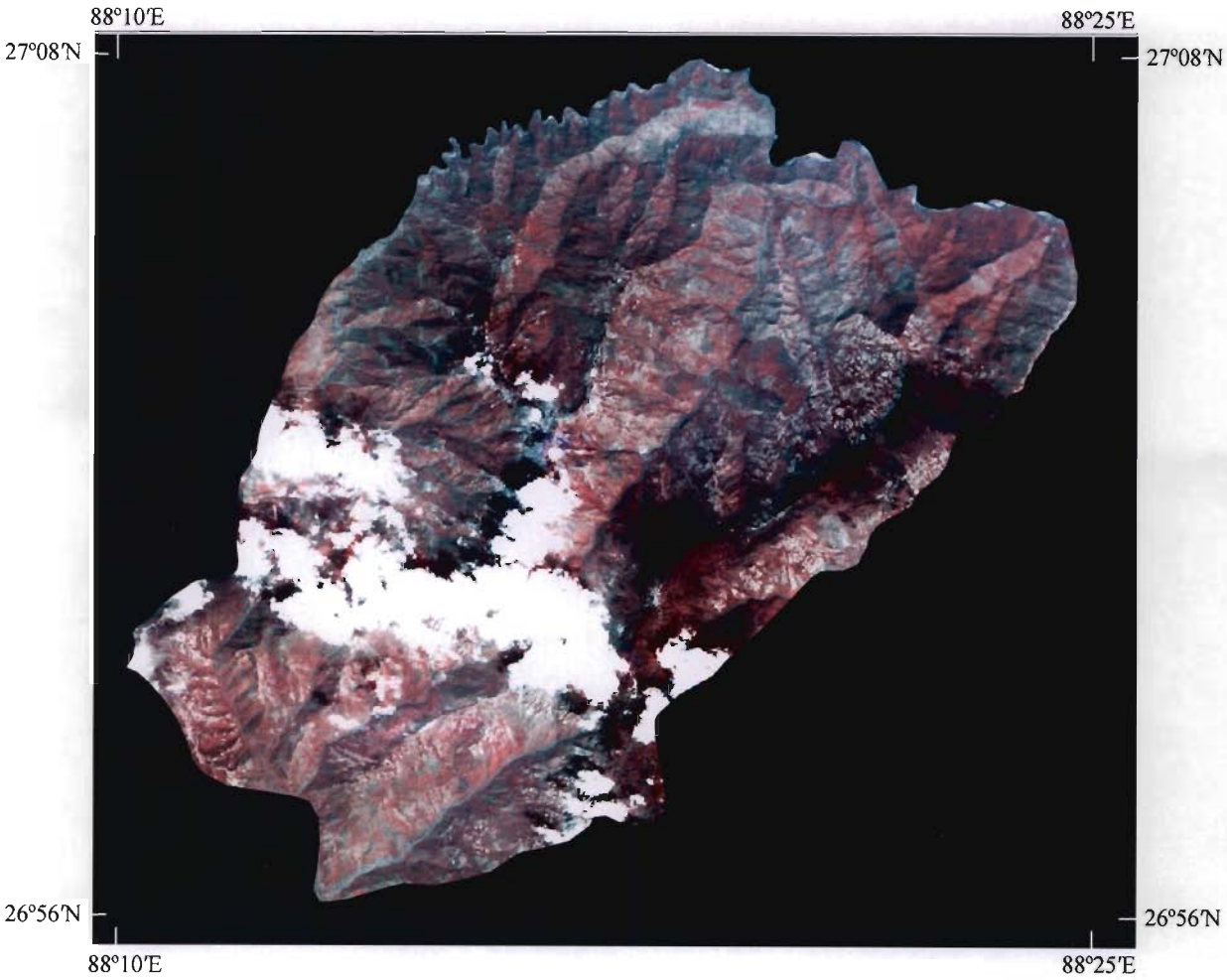
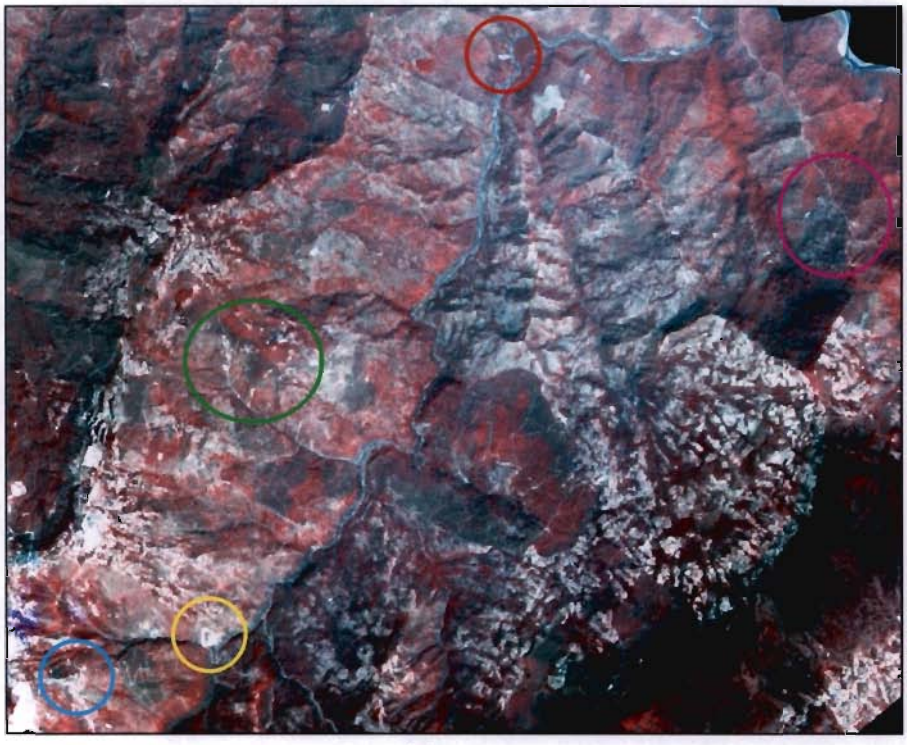
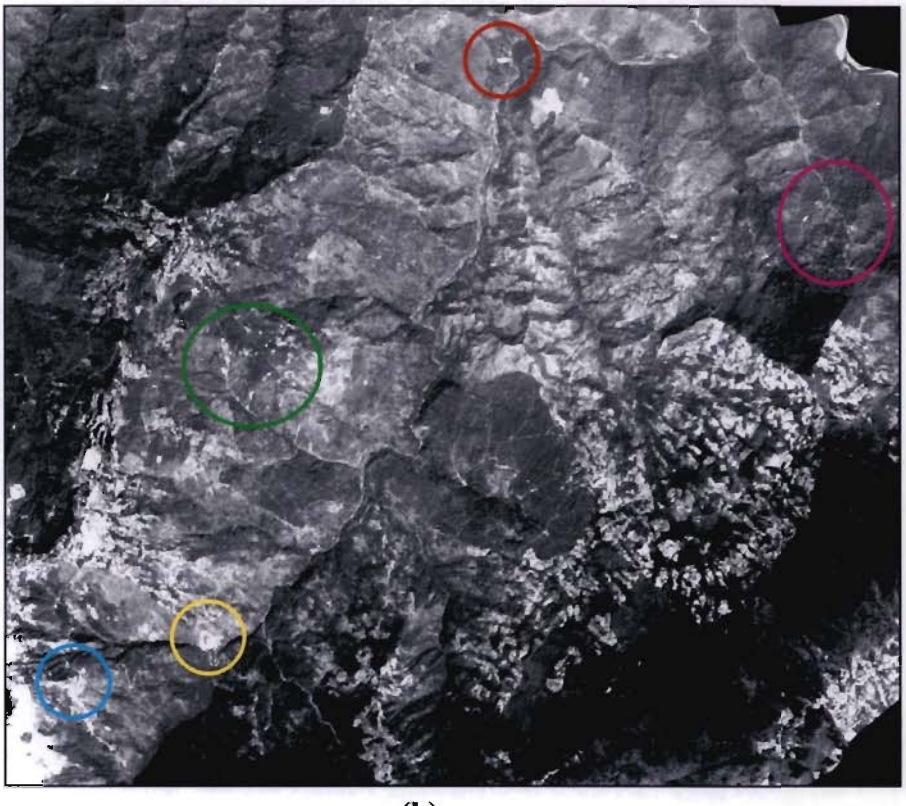


Figure 4.4: PAN-sharpened LISS-III image of the area. (Note: A portion of image is covered with clouds. This portion was masked out in the analysis).



(a)



(b)

Figure 4.5: Comparison of (a) PAN-sharpened LISS-III image and (b) IRS-PAN image for interpreting landslides. The same landslides are encircled with same colour in both the images.

(a) Landslide at Dandagaon (Lopchu)

(b) Landslide in Phubsering Tea Garden

(c) Landslide on Darjeeling-Rock Garden road

(d) Landslide in Badamtam tea garden

(e) Landslide in Ging tea garden

(f) Landslide and road sinking on Ghum-Sukhiapokhri road

Figure 4.6: Field photographs of landslides in the study area.

The landslides thus identified on the remote sensing images are digitized as polygons in one separate vector layer. Further, the landslides mapped from field data are also digitized in one vector layer. The two layers are then merged into single landslide distribution layer (Figure 4.7). This layer is then converted to a rasterized landslide distribution layer with a pixel size of 25m × 25m. This pixel size is considered to match with the nominal spatial resolution of LISS-III image. In total, there were 339 landslide pixels in the study area.

4.4 Digital Elevation Model (DEM)

DEM represents the spatial variation of elevation over an area. It is an important basic component in LSZ studies. In this study, DEM has been used:

i) To derive terrain related attributes like slope and aspect used as inputs for preparation of LSZ map.

ii) As an input layer to multi-source classification of remote sensing data for land use land cover map preparation to minimize the error in classification due to topographic variations.

The DEM has been prepared using the conventional and most prevalent method of digitising contours from Survey of India topographic maps.

4.4.1 DEM from Topographic Map

The contours at 10m and 20m intervals on 1:25,000 scale topographic maps and 40m interval on 1:50,000 scale maps respectively have been digitized onscreen. The triangulated irregular network (TIN) method has been used to generate the DEM from these digitized contours. A DEM at spatial resolution corresponding to pixel size of

25m × 25m has been generated to match the nominal spatial resolution of LISS-III image.

4.4.2 DEM based Derivatives

The DEM at 25m spatial resolution is used to derive slope and aspect data layers, which are later used for the generation of LSZ map of the area.

4.4.2.1 Slope

Slope is the angle formed between any part of the earth's surface and a horizontal datum. Slope angle is one of the key factors in inducing slope instability of the terrain. It has been widely shown that landslides tend to occur more frequently on steeper slopes (Cooke and Doornkamp, 1990; McDermid and Franklin, 1995), but this is not always the case. Slope failures tend to increase with slope angle but when the slope becomes almost vertical, landslide occurrence is generally found to be very scarce or absent (Selby, 1993).

Slope can be designated as the first derivative of DEM. A slope map is a raster map in which the attribute of each pixel denotes the maximum slope at a particular location. The slope map has been prepared from the DEM with 25m spatial resolution. The slope map thus derived from the DEM shows the range of variation of slopes from 0° to 90°, which have been further classified into 5 categories (Table 4.1) as per slope classification given in other studies (Anbalagan, 1992; Gupta et al., 1999; Dhakal et al., 2000). Figure 4.8 represents the slope data layer of the study area.

The areal distribution of different slope categories has been derived using GIS and is listed in Table 4.1. It is observed that in the present study area, the most frequent is the slope category of 15°-25°, followed by 25°-35°, 0°-15°, 35°-45°

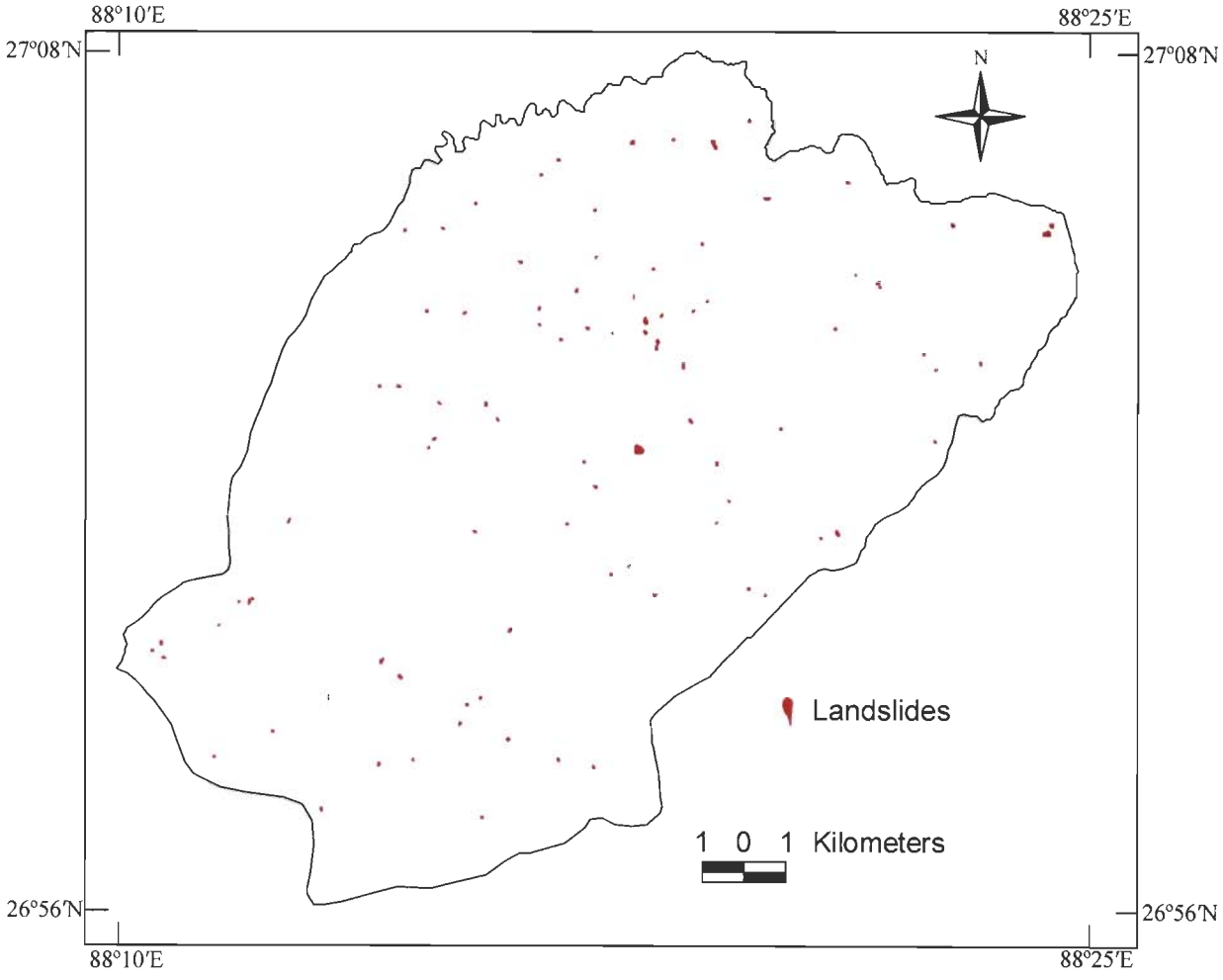


Figure 4.7: Existing landslide distribution layer of the area.

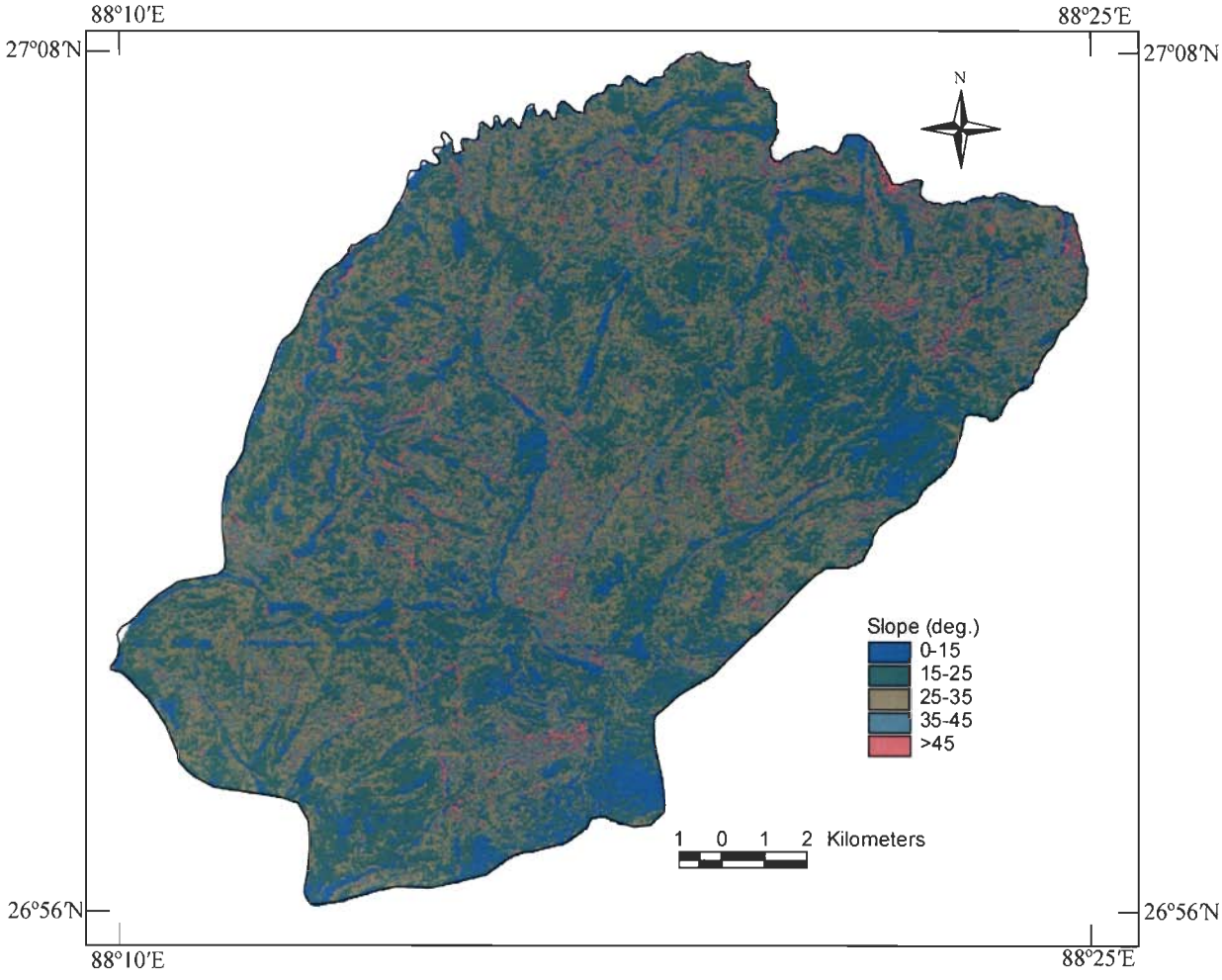


Figure 4.8: Slope layer of the area.

categories, the least frequent being the >45° slope category. It is also observed that the slope category of 0°-15° exists mostly on ridge tops, whereas the other slope categories are found to be well distributed all over the study area.

These slope categories are spatially correlated with the landslide pixels in the area (Table 4.1). It is observed that the slope category of 25°-35° has the highest incidence of landslides, followed by slope categories 15°-25°, 35°-45°, 0°-15° with least landslides occurring in the slope category >45°.

Table 4.1: Distribution of existing landslides in different slope categories.

Slope Categories	Area of slope categories (km^2)	Percent area (%)	Landslide area per category (km^2)	Percent Landslide area per category (%)
0°-15°	32.1	12.6	0.014	6.6
15°-25°	91.8	36.1	0.073	34.6
25°-35°	90.3	35.5	0.082	38.9
35°-45°	31.4	12.3	0.036	17.1
>45°	8.9	3.5	0.006	2.8
Total	254.5	100.0	0.211	100.0

4.4.2.2 Aspect

Aspect is defined as the direction of maximum slope of any location on surface of the earth. It is related to the distribution and density of vegetation on mountainous slopes. As vegetation provides anchorage to the ground, the stability of slope is also related to the aspect. Hence, aspect has an indirect influence on slope instability. In general, south facing slopes have lesser vegetation density as compared to north facing slopes and hence, the erosional activity is relatively more in former case (Sinha et al, 1975). Marston et al. (1998) reported that soil exposed on south-facing slopes

were subjected to several wetting and drying cycles, thus increasing landslide activity in the Himalayas. Based upon the landslide distribution, south and east facing slopes were considered susceptible to landslides by Dhakal et al., 2000. According to Lin and Tung (2003), the southeast aspect displayed higher potential landslide risk than other directions.

Aspect can be designated as the second derivative of DEM. The aspect map generated from the DEM has shows slope directions from 0° to 360° with respect to the north. These have been classified here into nine aspect categories (Table 4.2) as per the classification given in other studies (Sarkar and Kanungo, 2004; Saha et al., 2005). Figure 4.9 shows the aspect data layer of the study area. It is observed from the aspect map that the northern part of the study area is dominated by N, NE and E aspects, whereas the southern part is dominated S, SW, SE and W aspects. The central portion is occupied by E, SW and NW aspect categories.

The areal distribution of different aspect categories has been derived and is listed in Table 4.2. It is observed in the present area that the most frequent is the NW aspect, followed by N, W, E, SE, NE, S, SW aspects, the least frequent being the category of flat which ought to be the case in hilly terrain.

The spatial distribution of landslides in different aspect categories has been obtained (Table 4.2). It is observed that the SE, E and S aspect categories have higher percentage (22.6%, 21.7% and 14.6% respectively) of landslide occurrences in comparison to other aspects. This also indicates that the southern, southeastern and eastern facing slopes are more susceptible to landslide occurrences.

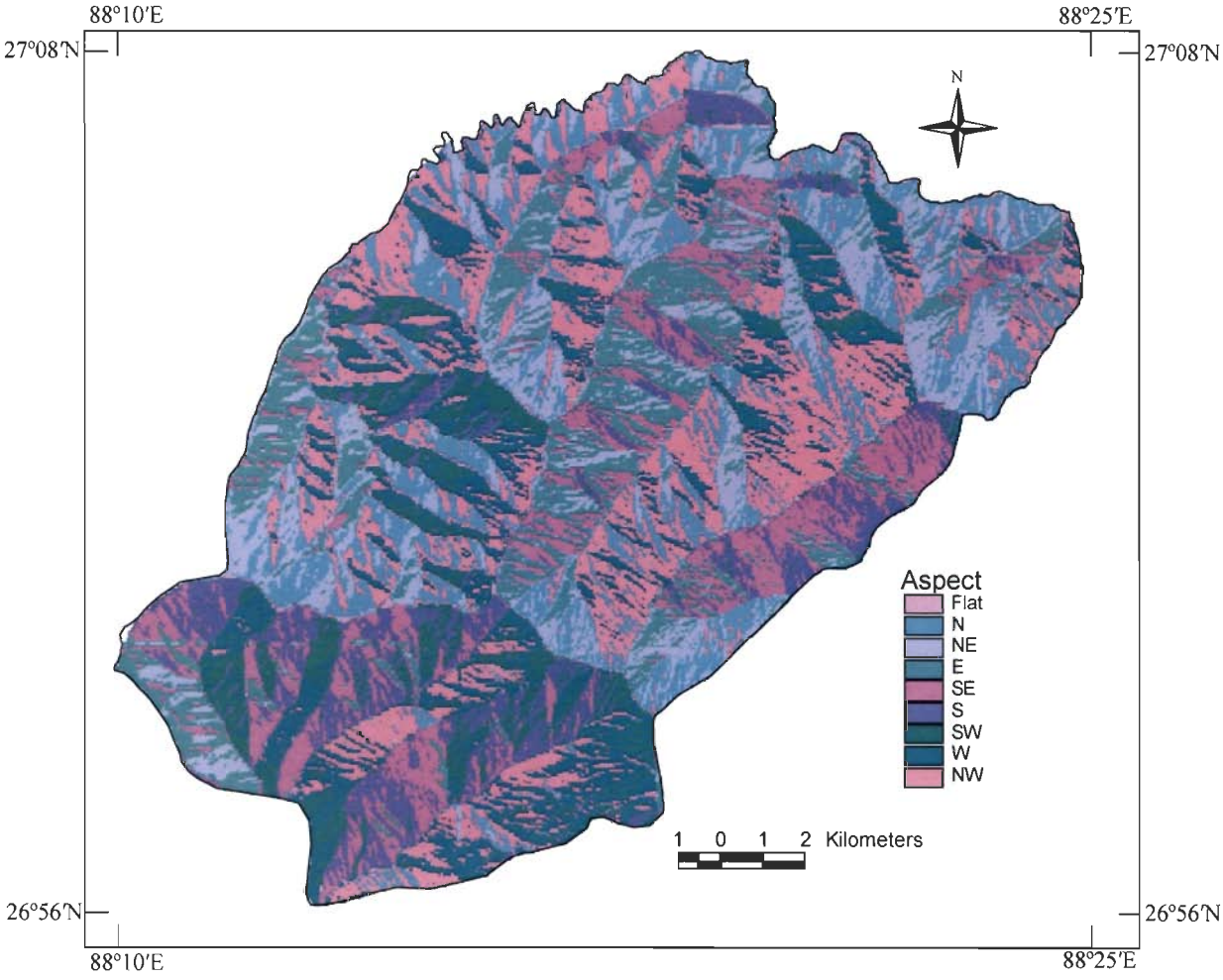


Figure 4.9: Aspect layer of the area.

Table 4.2: Distribution of existing landslides in different aspect categories.

Aspect Categories	Slope directions with respect to north	Area of slope categories (km^2)	Percent area (%)	Landslide area per category (km^2)	Percent Landslide area per category (%)
Flat	Undefined	1.3	0.5	0.000	0.0
N	0°-22.5°, 337.5°-360°	37.4	14.7	0.014	6.6
NE	22.5°-67.5°	28.2	11.1	0.020	9.4
E	67.5°-112.5°	33.0	13.0	0.046	21.7
SE	112.5°-157.5°	28.5	11.2	0.048	22.6
S	157.5°-202.5°	23.5	9.2	0.031	14.6
SW	202.5°-247.5°	18.7	7.3	0.012	5.7
W	247.5°-292.5°	34.4	13.5	0.016	7.6
NW	292.5°-337.5°	49.5	19.5	0.025	11.8
Total		254.5	100.0	0.212	100.0

4.5 Lithology

The distribution of lithology in a region reflects different lithologic groups, which indirectly represent the structure, strength and stress distribution of slopes as well as the topography (Lan et al., 2004). Different rock types (or lithology) have varied inherent characteristics such as composition, structure and compactness, which contribute to the strength of the material. The stronger rocks give more resistance to the driving forces as compared to weaker rocks, and hence are less susceptible to landslides and vice versa. Therefore, lithology is an important factor for LSZ mapping.

The geological map (scale 1:250000) of Sikkim-Darjeeling area published by Roy (1976) and Acharya (1989) is used as the base map for preparing the lithology

data layer. The rock types present in the area are Darjeeling gneiss, Paro gneiss, Lingtse granite gneiss, feldspathic greywacke, and quartzites of the Paro sub-group and the Reyang group. The Lingtse granite gneiss rocks are strongly foliated and sheared showing very high weathering at most locations. The Paro gneisses are stronger and coarse grained than the Darjeeling gneiss of higher metamorphic grade. The quartzites are highly jointed and fractured. However, all the rocks are folded and sheared to varying degrees and they have been subjected to high levels of weathering along the drainages. The competent rocks such as quartzite, greywacke are less susceptible to landslides than the gneisses as more numbers of landslides are observed in gneissic rock in the field.

The geological map is geo-referenced with the topographic maps. The boundaries between different rock types have been digitized on a vector layer and transformed to polygons. Minor modifications in lithological boundaries at some places have also been incorporated using field observations. Finally, the vector layer has been rasterized (25m × 25m pixel size) to produce a lithology data layer (Figure 4.10). It is observed from the lithology data layer that the northern part of the study area is occupied by feldspathic greywacke, Lingtse granite gneiss, Reyang quartzite and few patches of Paro quartzite. Darjeeling gneiss is present on the southeastern part of the study area. One patch of Paro quartzite is present on the southern part and rest of the area is covered by Paro gneiss.

The areal distribution of different rock types has been derived and is listed in Table 4.3. It is observed from Table 4.3 that the most frequent category of lithology is the Paro gneiss, followed by Darjeeling gneiss, feldspathic greywacke, Lingtse granite gneiss, Paro quartzite, the least being the Reyang quartzite.

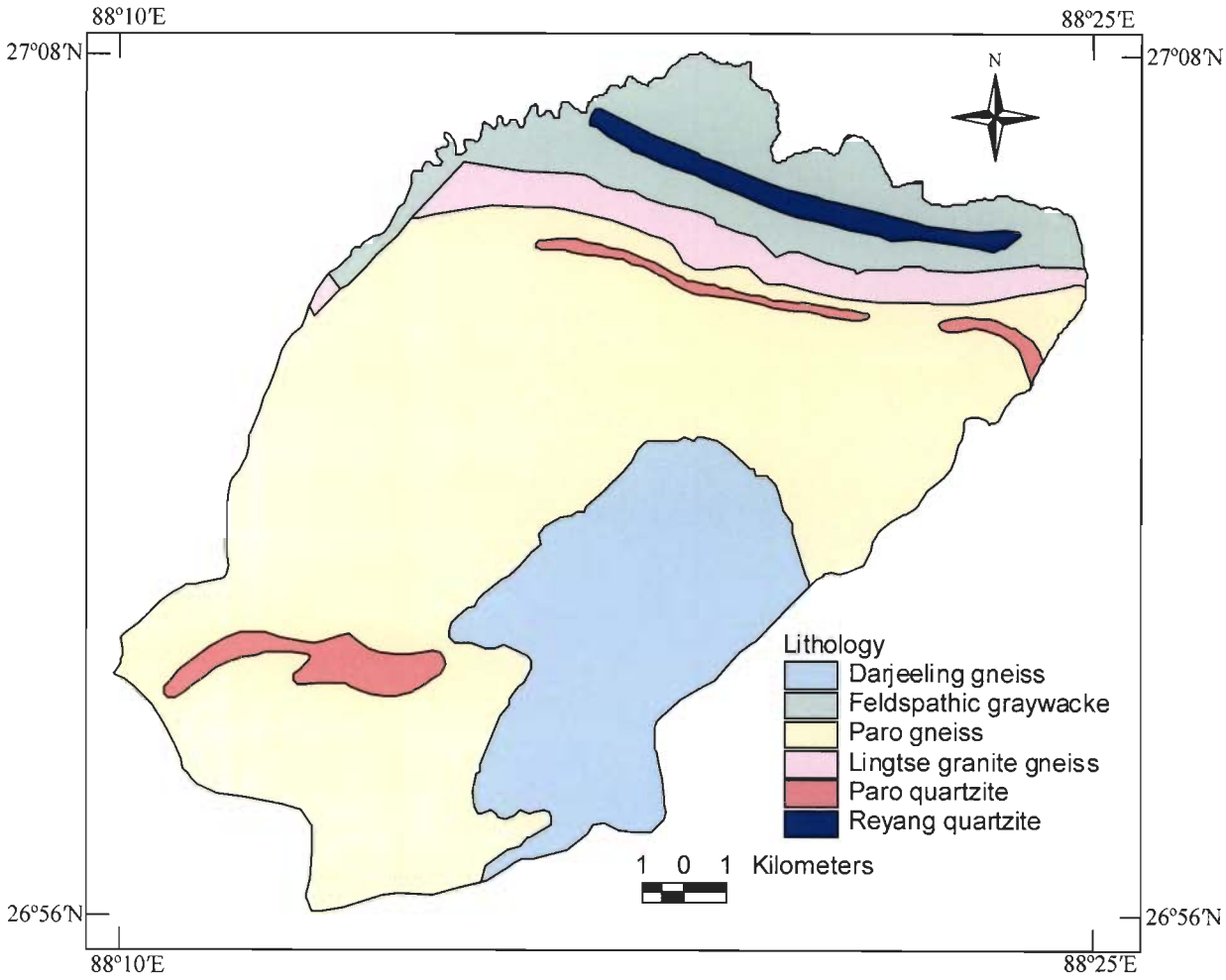


Figure 4.10: Lithology layer of the area.

The spatial distribution of landslides in different rock types has been obtained (Table 4.3). It is observed that Paro gneiss has maximum incidence of landslides in comparison to other categories and quartzites and Lingtse granite gneiss have the least occurrences of landslides.

Table 4.3: Distribution of existing landslides in different rock types (lithology categories).

Rock types (Lithology categories)	Area of lithology categories (km^2)	Percent area (%)	Landslide area per category (km^2)	Percent Landslide area per category (%)
Darjeeling Gneiss	45.8	18.0	0.048	22.6
Feldspathic Graywacke	28.7	11.3	0.038	17.9
Paro Gneiss	154.5	60.7	0.099	46.7
Lingtse Granite Gneiss	13.0	5.1	0.009	4.3
Paro Quartzite	7.5	2.9	0.009	4.3
Reyang Quartzite	5.0	2.0	0.009	4.3
Total	254.5	100.0	0.212	100.0

4.6 Lineaments

Lineaments are the structural features which describe the zone/plane of weakness, fractures and faults along which landslide susceptibility is high. The term lineament is used as "a mappable simple or composite linear feature of a surface, whose parts are aligned in a rectilinear or slightly curvilinear relationship and which differs distinctly from the pattern of the adjacent features, and presumably reflects subsurface phenomena" (O'Leary et al., 1976). Hence, the lineaments include all structural, topographical, vegetational, soil and lithological trends, which are likely to be the surface expressions of buried fractures and structures. The lineaments are

related to the fracture systems, discontinuity planes, faults and shear zones in the rocks. It has been generally observed that the probability of landslide occurrence increases at sites which are close to lineaments (Greenbaum et al., 1995). The lineaments, not only affect surface material structures but also make contribution to the terrain permeability, favoring slope instability. On a photograph or an image, lineaments (photo-lineaments) can be easily mapped by visual interpretation using various image interpretation elements, such as tonal variation, texture, pattern, association, etc. Remote sensing techniques have proven to be very successful for the detection of lineaments (Suzen and Toprak, 1998; Gupta, 2003).

In this study, the lineaments are interpreted from the PAN and LISS III images. The individual bands of LISS III image are processed using linear contrast stretching and then 3 × 3 edge enhancement filters to enhance the high frequency features. Then, all the four bands are stacked to produce the edge-enhanced image which has been used to visually interpret the lineaments on the basis of tonal contrast, structural alignments and rectilinear trends of morphological features and linear stream courses that are conspicuous by their abrupt changes in the course. There is no major thrust/fault reported in the study area, but mega lineaments are identified. The interpreted lineaments are digitized onscreen and are subsequently rasterized on a spatial resolution of 25m to produce the lineament data layer (Figure 4.11).

4.6.1 Preparation of Lineament Buffer Layer

Generally, the incidence of landslide decreases with the increase in distance from the lineaments. The 'distance' function of ArcView GIS has been used to find out the shortest distance of each pixel to any of these lineaments and a buffer zone is

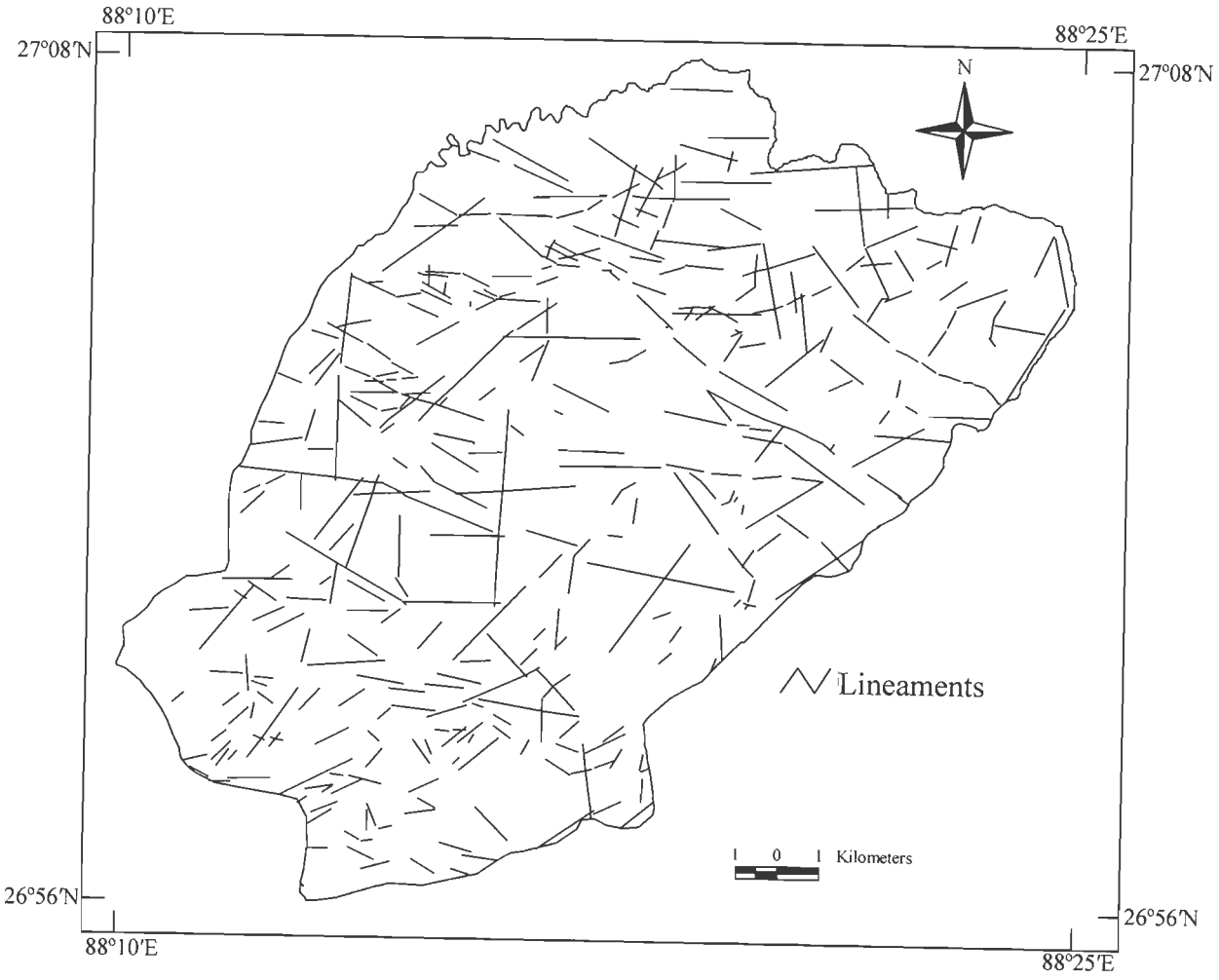


Figure 4.11: Lineament layer of the area.

created around each lineament. The lineaments make the materials more susceptible to landslide because of material weakening and variations in stress accumulation or tectonic activity at different distances from lineaments. Therefore, there is not a consensus amongst the researchers about the distances to the lineaments (width of buffer zone) to be considered for LSZ studies. As a result, different researchers have used different distances (i.e., 500m to 2km) with respect to the closeness to the lineaments (Gupta and Joshi, 1990; van Westen and Bonilla, 1990; Mehrotra et al., 1991; Anbalagan, 1992; Pachauri and Pant, 1992; Maharaj, 1993; Mejia-Navarro et al., 1994; Gupta et al., 1999; Luzi and Pergalani, 1999; Donati and Turrini, 2002; Saha et al., 2005).

Keeping these in view, buffer zones at 250 m intervals are initially created. The spatial distributions of existing landslides in these buffer zones are examined and it is found that 98% of landslide pixels occur in 1st two buffer zones (up to 500 m). Hence, it is decided to consider four buffer zones at 125 m intervals up to 500 m and another buffer zone beyond 500 m to establish the influence of lineaments on landslide occurrence. Thus, a lineament buffer layer consisting of five buffer categories namely 0-125 m, 125-250 m, 250-375 m, 375-500 m and >500 m has been prepared. The lineament buffer layer of the study area is shown in Figure 4.12. The spatial distributions of existing landslides in these five categories have been determined and are given in Table 4.4. It is observed from this table that the lineament buffer category of 0-125m has the highest incidence of landslides, followed by categories 125-250m, 250-375m, 375-500m with least landslides occurring in the lineament buffer category of >500m. Hence, it can be stated that as the distance from lineaments increases, the landslide susceptibility decreases.

Table 4.4: Distribution of existing landslides in different lineament buffer zones.

Lineament Buffer Categories	Area of Lineament Buffer Categories (km^2)	Percent area (%)	Landslide area per category (km^2)	Percent Landslide area per category (%)
0-125m	91.7	36.0	0.152	71.7
125-250m	68.0	26.7	0.022	10.4
250-375m	45.2	17.8	0.022	10.4
375-500m	25.7	10.1	0.011	5.2
>500m	23.9	9.4	0.005	2.3
Total	254.5	100.0	0.212	100.0

4.7 Drainage

Most of the landslides in hilly areas occur due to the erosional activity associated with drainage. Therefore, a drainage layer has been prepared by digitizing the drainage lines from the topographic maps in a vector layer. Later, this layer is overlaid on LISS-III image for modifying the changes in drainage lines. This is felt necessary as most of the 1st order drainages, which are generally not present in the topographic maps, are easily interpreted from the LISS-III image, which also shows change in the course of the river and other major drainages at some places. This layer is used for drainage ordering and preparation of drainage buffer layer. Drainage ordering of all the drainage lines is done to consider the lower order drainages for LSZ mapping as these play important role in initiating erosional activities in the form of landslides in hilly terrains.

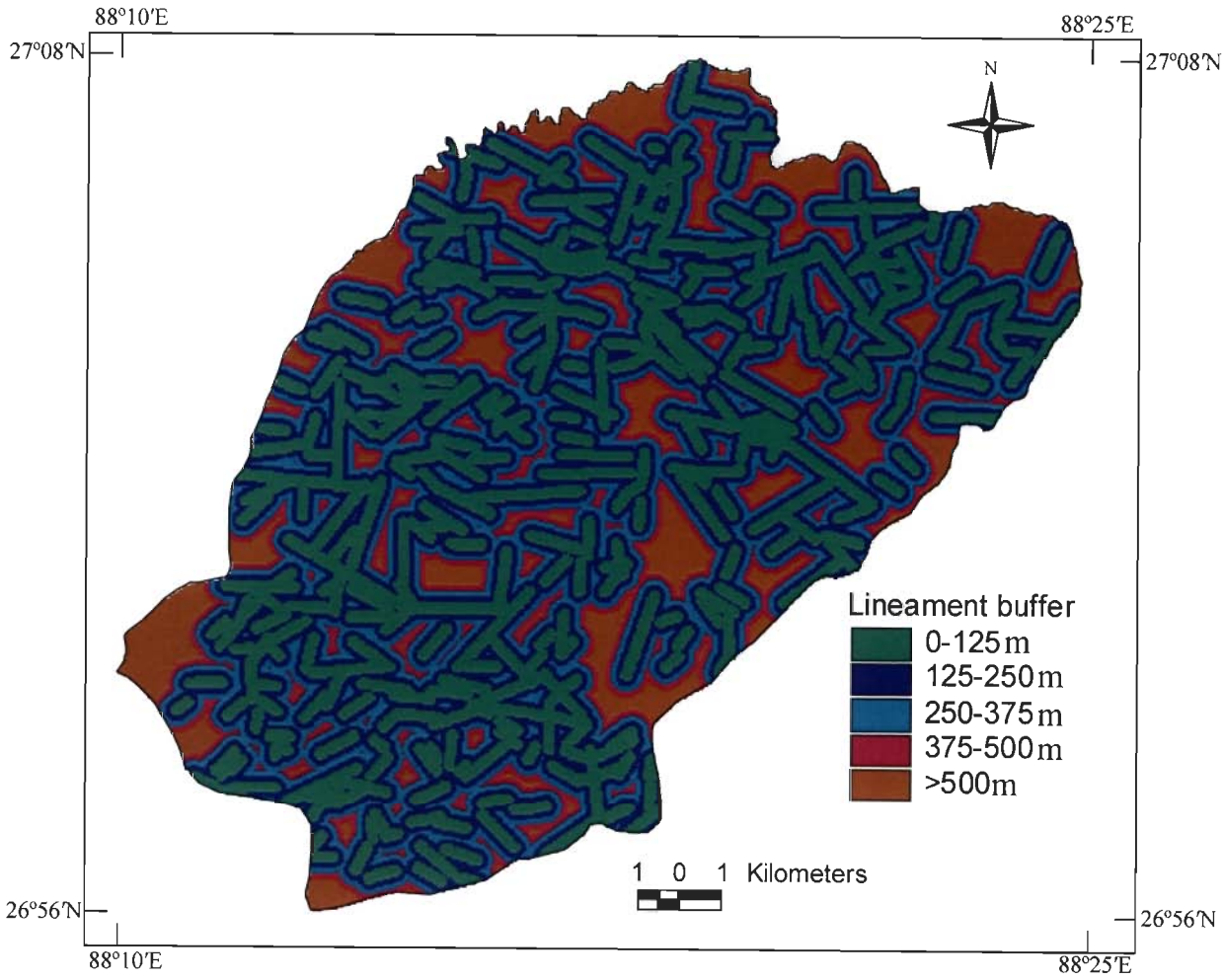


Figure 4.12: Lineament buffer layer of the area.

4.7.1 Drainage Ordering

The ordering of the drainage lines has been performed on the basis of Strahler classification (Strahler, 1964). It is a simple method of classifying stream segments based on the number and type of tributaries in the upstream. A stream with no tributaries (headwater stream) is considered as a first order stream. A segment, downstream of the confluence of two first order streams is a second order stream. Thus, an n^{th} order stream always starts downstream of the confluence of two $(n-1)^{th}$ order streams. This has been illustrated in the Figure 4.13. As per Strahler's classification, the drainage lines in the study area are classified into different drainage orders. The drainage order layer thus prepared (Figure 4.14) shows that the drainages up to 6^{th} order are present in the study area. Subsequently, this layer has been rasterised at 25m spatial resolution for its use in preparing the drainage buffer layer which will serve as an input thematic data layer for LSZ mapping.

4.7.2 Preparation of Drainage Buffer Layer

Buffer zones of 25m width on either side of the drainage lines for all the drainage orders are created to consider the influence of drainages on landslide occurrences. It is observed from the spatial distribution of existing landslides in these buffer zones (Table 4.5) that majority of landslides occur in the 1^{st} and 2^{nd} order drainage buffers only. This is also the fact that in hilly terrains, the lower order drainages play an important role in initiating erosional activities in the form of landslides. Therefore, 25m buffer zones along 1^{st} and 2^{nd} order drainages only are considered to create a drainage buffer layer (Figure 4.15) for further analysis.

Table 4.5: Distribution of existing landslides in different drainage buffer zones.

Drainage Buffers (25m)	Number of pixels	Area (km^2)	Landslide pixels per category	Landslide area per category (km^2)
1st Order	116168	72.6	102	0.064
2nd Order	27690	17.3	44	0.027
3rd Order	10294	6.4	4	0.002
4th Order	6025	3.8	11	0.007
5th Order	4708	2.9	18	0.011
6th Order	2766	1.7	0	0.000
Total	167651	104.7	179	0.111

4.8 Land Use Land Cover Classification

The land use land cover information of an area is quite essential for proper planning, management and monitoring of natural resources. It is an important input for many geological, hydrological, ecological and agricultural models. Land use land cover map generally shows distribution of forest cover, water bodies and types of land use practices. Many studies (e.g., Coppin and Richards, 1990; Selby, 1993; Mehrotra et al., 1996) have revealed a clear relationship between vegetation cover and slope instability. Parameters, such as cohesion and friction angle of soil and pore-water pressure, tend to get modified substantially by the presence of vegetation. The incidence of landslide is inversely related to the vegetation density. Hence, barren slopes are more prone to landslide occurrence as compared to the vegetated surfaces. Therefore, the land use land cover information has been is considered as one of the factors responsible for landslide occurrences, as considered in other studies (e.g., Gupta et al., 1999; Saha et al., 2002; Sarkar and Kanungo, 2004; Saha et al., 2005).

Remote sensing images help in gathering quality land use land cover information at local, regional and global scales because of its synoptic view, map like

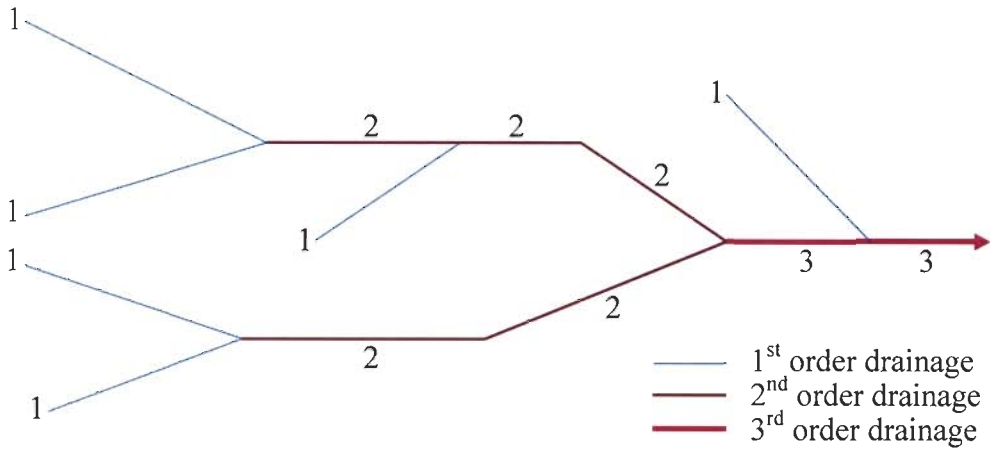


Figure 4.13: The drainage order scheme (Strahler, 1964).

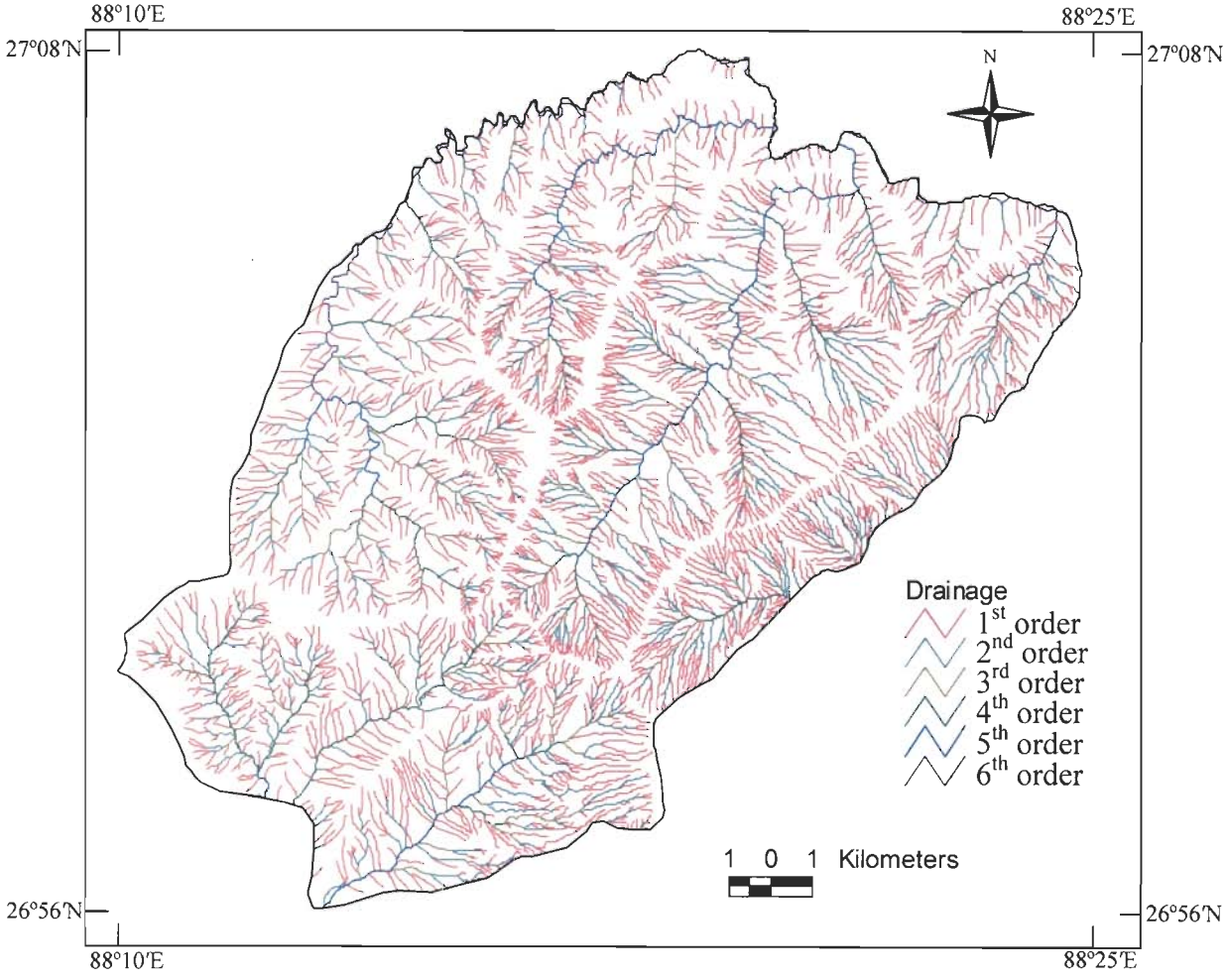


Figure 4.14: Drainage order layer of the area.

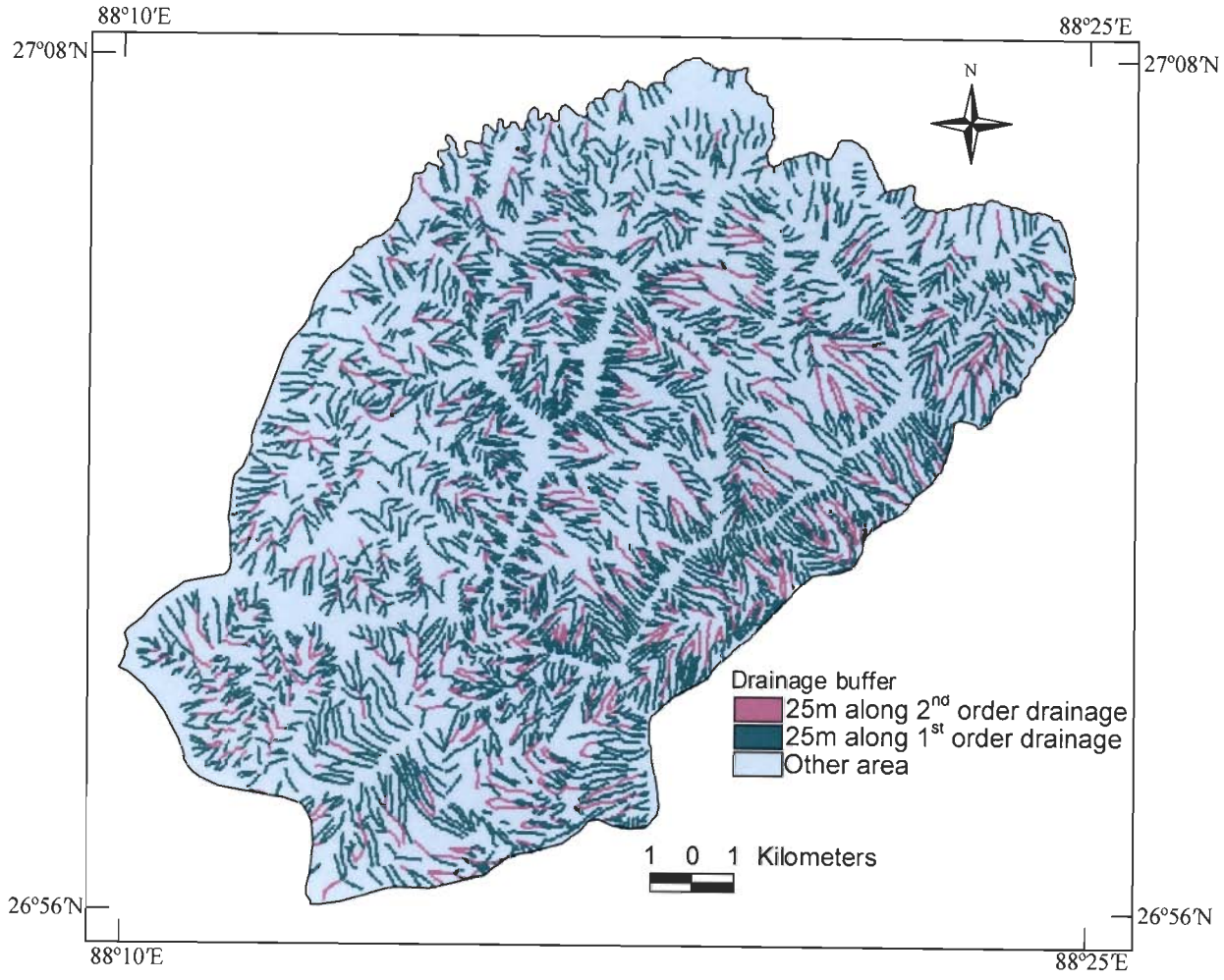


Figure 4.15: Drainage buffer layer of the area.

format and repetitive coverage (Csaplovics, 1998; Foody, 2002). Further, in mountainous regions like the Himalayas, particularly in the inaccessible areas due to high altitudes and ruggedness in the terrain, remote sensing images are quite useful for mapping. Over the years, land use land cover mapping in these areas using remote sensing data have been reported with varying degrees of success. This may be attributed to the factors that influence the remote sensing data processing and interpretation, particularly in mountainous terrains. These factors include the presence of cloud cover, shadows due to high altitudes, steep slopes, low sun angle and differential vegetation cover. Hence, due to changes in topographical and environmental conditions, spectral characteristics also change from region to region (Arora and Mathur, 2001). Therefore, the approach for land use land cover classification that incorporates ancillary data from other sources may be more effective than that is based solely upon multi-spectral data from one sensor. The data from other sources include the topographic maps (Bruzzone et al., 1997), geological (Gong, 1996) and other maps. The topographic maps are useful in generating the DEM, which alongwith its attributes, such as slope and aspect, provide the basis for multi-source classification (Strahler et al., 1978; Jones et al., 1988; Frank, 1988; Janssen et al., 1990; Saha et al., 2005). Multi-sensor classification approach for vegetation mapping can also be adopted by combining data from different remote sensing sensors (Michelson et al., 2000). Furthermore, the derivatives of multispectral images like principal components analysis (PCA) and normalised difference vegetation index (NDVI) may also be useful to improve the land use land cover classification from remote sensing data in mountainous regions (Eiumnoh and Shrestha, 2000; Saha et al., 2005).

In mountainous terrain, such as Himalayas, shadow is the major problem in achieving the accurate land use land cover classification from remote sensing data. The effect of shadow in mountainous region can be minimized. Majority of the shadow suppressing methods are based on shaded relief models that are produced from DEM. But, many studies (e.g., Kawata et al., 1988; Civco, 1989; Colby, 1991; Curran and Foody, 1994), have shown that the correction for shaded slopes may get over-estimated due to the errors in creating DEM and also the way the DEM is applied in the classification process. The use of NDVI image as an additional layer for classification has been recommended to overcome this problem, since the band ratio derivatives may help in nullifying the topographic effect to some extent (Holben and Justice, 1981; Apan, 1997). However, NDVI alone may not be able to eliminate the shadow effect completely. Later, Eiumnoh and Shrestha (2000) and Saha et al. (2005) incorporated both NDVI and DEM images as additional layers in the classification process and found a significant improvement in the classification accuracy.

In this study, the IRS-1C LISS-III image has been used as the primary data source along with NDVI and DEM images as additional data layers to implement multi-source land use land cover classification process. Separability analysis using transformed divergence is performed to examine the significance of various spectral bands in the classification process. Most widely used maximum likelihood classifier (MLC) is used to perform the classification. A very small portion covered by the cloud and its shadow in the original LISS-III image has been masked and then used as the primary data to perform land use land cover classification. The PAN image is used as the reference data for generating training and testing datasets. This is in accordance with other studies on land use land cover classification of remote sensing data, where finer resolution data have also been used as reference data (Fisher and Pathirana,

1990; Foody and Arora, 1996; Shalan et al., 2003; Saha et al., 2005). The preparation of reference data is ably supported with field data as well as information from topographic maps.

4.8.1 Methodology

A multi-source image classification involves a number of steps which include generation of ancillary data layers (NDVI and DEM), image classification and accuracy assessment (Figure 4.16).

4.8.1.1 Normalised Difference Vegetation Index (NDVI)

During field surveys, different types of vegetation were observed in the study area. Hence, NDVI has been used as an ancillary data layer in the classification process to enhance the separability among various vegetation classes, and also to reduce the effect of shadow due to topography. The NDVI can be stated as,

$$NDVI = (NIR\ band - Red\ band) / (NIR\ band + Red\ band) \qquad (4.1)$$

The DN values of pixels of the NDVI image thus produced range from 0.00 to 0.83 with higher values indicating increasing biomass. The positive values represent various types of vegetation classes. Near zero values indicate non-vegetation classes, such as water, river sand and barren land.

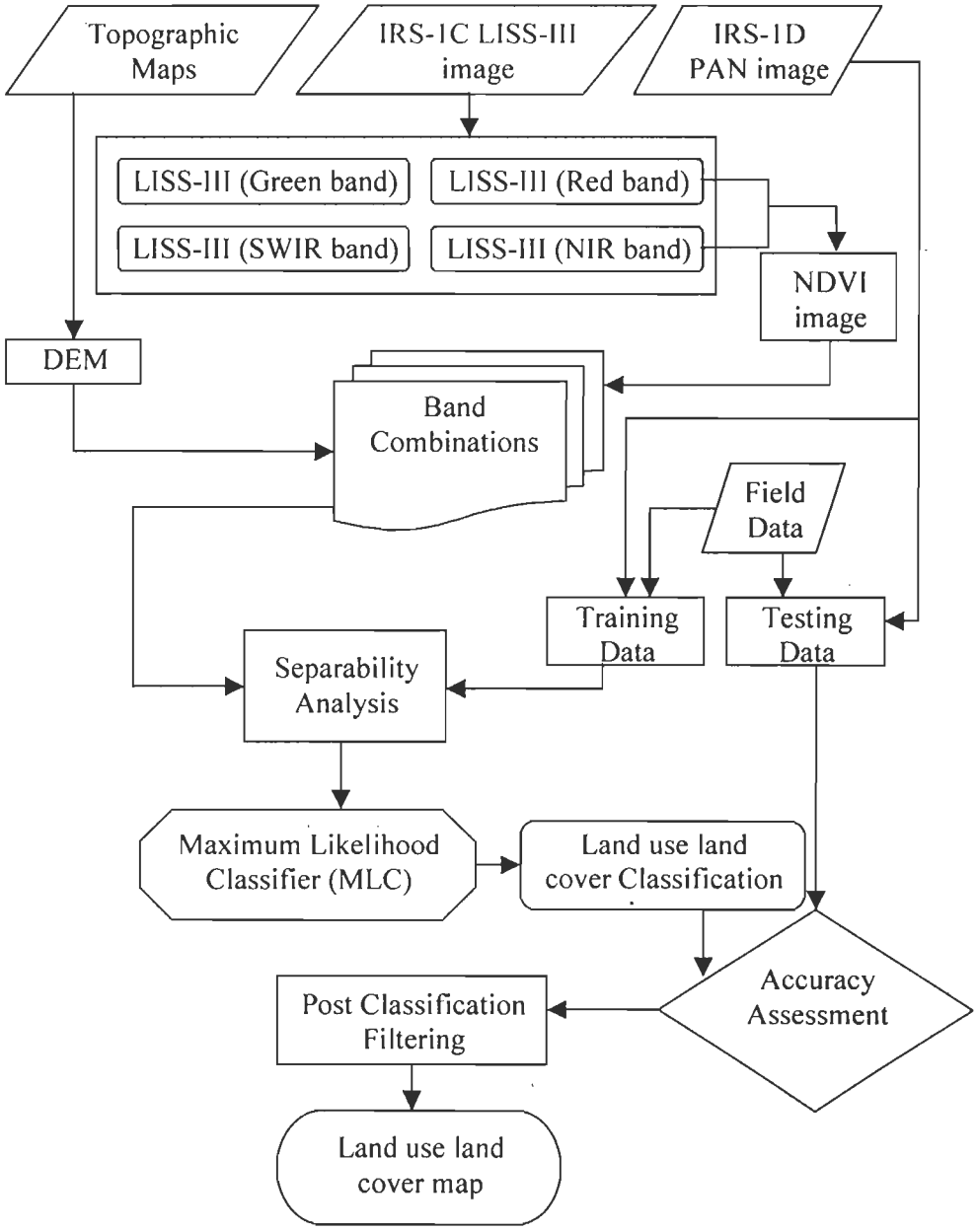


Figure 4.16: Steps for multi-source land use land cover classification.

4.8.1.2 Image Classification

Image classification process is based on several steps, as listed below:

a) Selection of land use land cover classification scheme
b) Formation of training dataset
c) Separability analysis and
d) Maximum likelihood classification (MLC)

a) Selection of land use land cover classification scheme

A classification scheme defines the land use land cover classes to be considered to prepare land use land cover map from remote sensing image data. Sometimes, a standard classification scheme, such as Anderson's land use land cover classification system (Anderson et al., 1976) may be used, while at other times the number of land use land cover classes may be chosen according to the requirements of the specific project for a particular application (Arora and Mathur, 2001; Saha et al., 2005). In this study, during field visits, eight land use land cover classes were identified in the study area. These classes are thick forest, sparse forest, tea plantation, agriculture, barren, built up, water bodies and river sand. This scheme of classification was adopted keeping in view the land use land cover scenario in the region and also the schemes adopted in other studies (e.g., Sarkar and Kanungo, 2004; Saha et al., 2005) for the purpose of LSZ mapping. Some of these classes as they appear in the field are shown in Figure 4.17. Detailed description of all these classes along with their interpretative characteristics on the FCC of LISS-III image and PAN image (Figure 4.18) is given in Table 4.6. These information are used to identify the training and testing areas on the image for carrying out supervised classification and accuracy assessment.

Table 4.6: Characteristics of land use land cover classes.

Land use land cover class	Description	Interpretation on LISS-III false colour composite	Interpretation on PAN image
Thick forest	Tall dense trees	Dark red with rough texture	Dark tone with rough texture
Sparse forest	Scanty tall trees and low vegetation density with exposed ground surface	Dull red to pinkish	Light tone with dark patches
Tea plantation	Tea plants with moderate vegetation density	Pink and smooth appearance	Light tone with smooth patches
Agriculture	Crops on hill terraces as step cultivation	Dull red and smooth appearance	Step like arrangement of fields and bright tone with smooth texture
Barren land	Exposed rocks/soils without vegetation	Yellowish	Very bright tone
Built up area	Towns and villages; block like appearance	Bluish	Typical blocky appearance with light tone
Water bodies	Rivers and lakes	Cyanish blue to blue according to the depth of water and sediment content	Dark tone
River sand	River sediments on the bank	Cyanish	Bright tone

b) Formation of training dataset

Training dataset preparation is a key step in supervised image classification process. The success of image classification highly depends on the quality of training dataset which in turn depends on the capability of image interpretation and knowledge on the land use land cover patterns of the study area (i.e., field verification). Hence, the training areas for different land use land cover classes must be selected from the regions that are representative of the land use land cover classes under investigation and also from relatively homogeneous areas containing these classes. The collection of training data is generally a tedious, time consuming and costly affair and, therefore,

Figure 4.17: Field photographs showing various land use land cover classes. (Tf – Thick forest; Sf – Sparse forest; Tp – Tea plantation; Ag – Agriculture; Hb – Habitation; Ls – Landslide; Bl – Barren land).

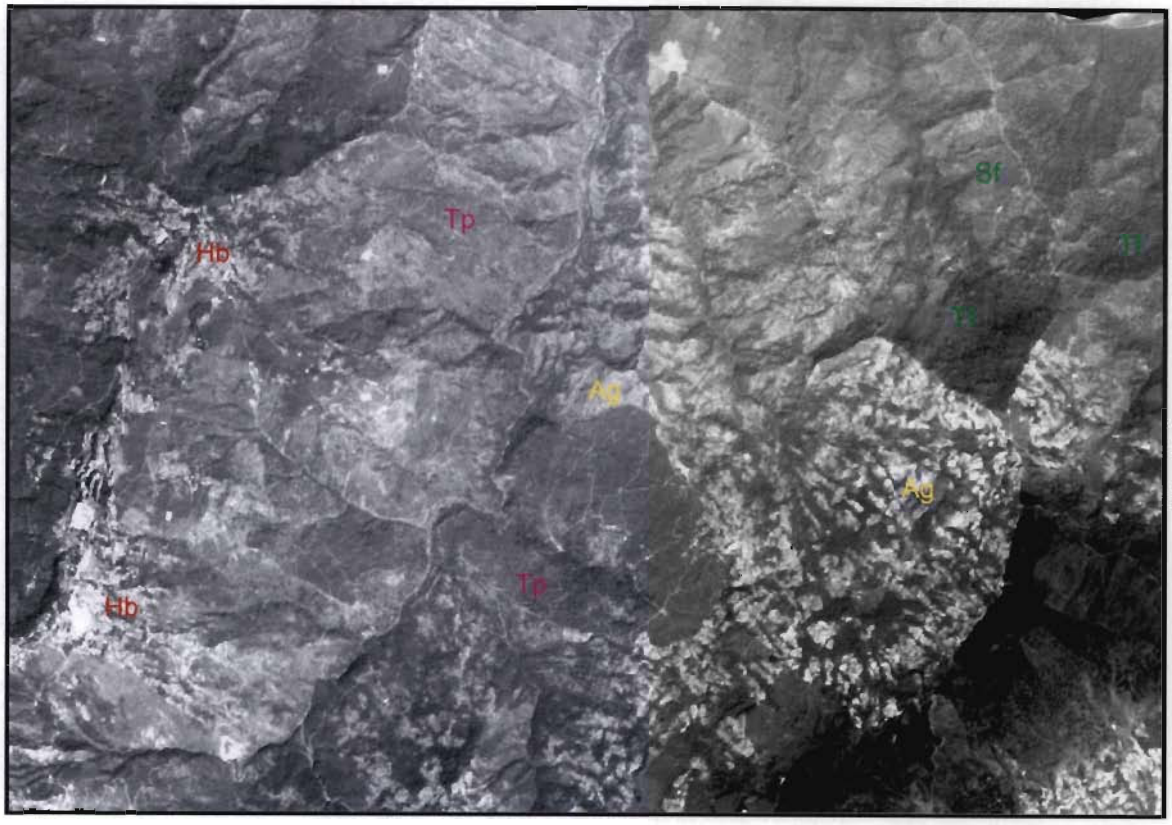


Figure 4.18: A sub-scene of IRS-1D PAN image showing various land use land cover classes. (Tf – Thick forest; Sf – Sparse forest; Tp – Tea plantation; Ag – Agriculture; Hb – Habitation).

the size of the training dataset must be kept small. On the other hand, the number of pixels in each class constituting the training dataset must be large enough to accurately characterize and classify the land use land cover classes (Saha et al., 2005). As a thumb rule, the number of training pixels for each class may be kept at 30 times the number of bands under consideration (Mather, 1999). In this study, the number of training pixels for each class (Table 4.7) was defined in accordance with the proportion of the area covered by the respective classes on the ground. Similar to other studies, the fine spatial resolution PAN image and topographic maps were used as reference data for ground truths to demarcate the training pixels on the LISS-III image. All the eight land use land cover classes were visually interpreted on the PAN image based on the characteristics defined in Table 4.6. The PAN image derived land use land cover information and ground truth data from field survey were used to demarcate training areas on LISS-III image for all the classes. Majority of training areas were normally distributed, which is a basic requirement of the maximum likelihood classifier used in this study.

Table 4.7: Number of training pixels for each land use land cover class used in image classification.

Land use land cover class	Number of training pixels
Thick forest	1129
Sparse forest	622
Tea plantation	1127
Agriculture	639
Barren land	204
Built up area	288
Water bodies	193
River sand	273

c) **Separability analysis**

A separability analysis was performed with multi-source data layers using the training dataset of all the eight land use land cover classes to observe the spectral discrimination between these classes. In this study, a combination of six data layers comprising of Green, Red, NIR and SWIR bands of multispectral LISS-III image, NDVI and DEM data layers were used as the input dataset for multi-source classification. Separability is a statistical measure devised on the basis of spectral distances computed from a combination of bands. Out of a number of separability measures, the transformed divergence (TD) measure (Janssen et al., 1990) has been adopted in this study. The separability analysis through transformed divergence is a widely used measure and computes the spectral distance between the mean vectors of each pair of signatures. This separability analysis is related to the maximum likelihood decision rule and helps the user to predict the results of a maximum likelihood classification. The TD values range from 0 to 2000. A value close to 2000 indicates the best separability between the classes. The values between 1800 and 2000 are generally considered adequate to proceed for classification. The separability analysis for the present study has been performed using a combination of six data layers (four bands of LISS-III image, NDVI and DEM) and the TD values obtained are listed in the form of a matrix (Table 4.8). The average TD value is 1977, which has been considered appropriate for performing classification. The lowest TD value of 1723 is obtained for the signatures of barren land and agriculture. This is on expected lines as the agriculture lands without cultivation appear to be barren lands. Hence, a low separability between these two classes is observed.

d) Maximum likelihood classification (MLC)

A number of classifiers have been developed and tested for remote sensing image classification. Each of these classifiers has its own merits and demerits in terms of efficiency and accuracy. The maximum likelihood classifier was found to be the most accurate and most widely used for image classification, when the data distribution assumptions are met. The MLC is based on the decision rule that pixels of unknown class membership are allocated to those classes with which they have the highest likelihood of membership (Foody et al., 1992). The details on this classifier may be found in Richards and Jia (1999).

In this study, the MLC has been used to produce land use land cover map, as it takes the variability of classes into account via covariance matrix.

Table 4.8: Transformed divergence (TD) matrix for land use land cover classes using combination of Green, Red, NIR and SWIR bands of LISS-III image, NDVI and DEM data layers.

Land use land cover classes	Tf	Sf	Tp	Ag	Bl	Bu	Wb	Rn
Thick forest (Tf)	0	1870.3	1999.6	2000	2000	2000	2000	2000
Sparse forest (Sf)	1870.3	0	1987.7	2000	2000	2000	2000	2000
Tea plantation (Tp)	1999.6	1987.7	0	1982.7	1996.9	1999.6	2000	2000
Agriculture (Ag)	2000	2000	1982.7	0	1723	1881.9	2000	2000
Barren land (Bl)	2000	2000	1996.9	1723	0	1916.9	2000	2000
Built-up area (Bu)	2000	2000	1999.6	1881.9	1916.9	0	2000	2000
Water body (Wb)	2000	2000	2000	2000	2000	2000	0	2000
River sand (Rn)	2000	2000	2000	2000	2000	2000	2000	0

4.8.1.3 Accuracy Assessment of Land Use Land Cover Classification

A testing dataset has been prepared with the help of reference data (PAN image and field data). The class allocation of each pixel in the classified image is compared with the corresponding class allocation on reference data to determine the classification accuracy. The pixels of agreement and disagreement are compiled in the form of an error matrix where the rows and columns represent the number of land use land cover classes and the elements of the matrix represent the number of pixels in the testing (reference) dataset. A number of accuracy measures, such as overall accuracy, user's accuracy and producer's accuracy can be estimated from the error matrix (Congalton, 1991). The overall accuracy indicates the accuracy of classification as a whole, where as user's and producer's accuracy measures indicate the accuracy of individual land use land cover classes. The overall accuracy can be defined as the ratio of the number of correctly classified pixels to the total number of pixels in the error matrix. User's accuracy is defined as the probability that a pixel classified on the map actually represents that class on the ground or reference data. The producer's accuracy is defined as the probability of a pixel on reference data being correctly classified (Congalton, 1991).

a) <u>Formation of testing dataset</u>

The choice of suitable sampling scheme and the determination of appropriate sample size for testing dataset play a significant role in the accuracy assessment of image classification (Arora and Agarwal, 2002). In the present study, field data on land use land cover classes and finer resolution PAN image have been used as reference data to prepare testing dataset for accuracy assessment. The testing pixels for each land use land cover class have been randomly selected. These pixels are

distributed all over the study area and are larger than 75 to 100 pixels per class as recommended by Congalton (1991) for accuracy assessment purposes. The number of testing pixels considered for each class in this study is listed in Table 4.9.

Table 4.9: Number of testing pixels for each land use land cover class used in accuracy assessment of image classification.

Land use land cover class	Number of testing pixels
Thick forest	368
Sparse forest	254
Tea plantation	366
Agriculture	387
Barren land	113
Built up area	114
Water bodies	111
River sand	151

b) Assessment of classification accuracy

The classified image has been compared with the class identity of testing data (reference data) and the agreement/disagreement of the pixels have been compiled (Table 4.10).

The user's and producer's accuracies(Table 4.11) have been computed from the error matrix as the percentage of the number of pixels correctly classified to the classified totals (C) and the reference totals (R) of each individual class respectively. The overall classification accuracy of 94.7% has been obtained as the percentage of the total number of correctly classified pixels (1766 pixels) to the reference or classified totals (1864 pixels).

Table 4.10: Error matrix of the classified image with respect to the reference data (Diagonal elements of the matrix (as underlined) represent the number of pixels in the classified image correctly matching with those in the reference data).

Land use land cover classes in the classified image	Classes on reference data								Total number of pixels in classified image (C)
	Tf	Sf	Tp	Ag	Bl	Bu	Wb	Rn	
Thick forest (Tf)	363	6	0	0	3	0	0	0	372
Sparse forest (Sf)	4	247	3	1	3	5	3	2	268
Tea plantation (Tp)	0	1	363	9	5	8	0	5	391
Agriculture (Ag)	0	0	0	371	7	7	0	0	385
Barren land (Bl)	0	0	0	3	93	5	0	0	101
Built-up area (Bu)	1	0	0	3	2	89	0	0	95
Water body (Wb)	0	0	0	0	0	0	99	3	102
River sand (Rn)	0	0	0	0	0	0	9	141	150
Total number of pixels in the reference data (R)	368	254	366	387	113	114	111	151	1864

Table 4.11: Producer's accuracy and user's accuracy of individual land use land cover classes derived from accuracy assessment of classification.

Land use land cover classes	Producer's accuracy	User's accuracy
Thick forest	98.6%	97.6%
Sparse forest	97.2%	92.2%
Tea plantation	99.2%	92.8%
Agriculture	95.9%	96.4%
Barren land	82.3%	92.1%
Built-up area	78.1%	93.7%
Water body	89.2%	97.1%
River sand	93.4%	94.0%

A glance at the values of the producer's accuracy values (Table 4.11) and the error matrix (Table 4.10) shows that most of the land use land cover classes except barren lands, built-up areas and water bodies have shown more than 90% producer's accuracy. The class barren land has been misclassified to some extent with the classes agriculture and tea plantation whereas the class built-up area has been misclassified with the classes tea plantation, agriculture, barren land and sparse forest. The class water body has been considerably misclassified with the class river sand.

4.8.1.4 Post-classification Filtering

The classified image, thus, produced may contain some stray pixels. To generate a smooth image by removing these stray pixels, a 3 × 3 pixels majority filter has been applied which assigns the most dominant class to the central pixel.

Subsequently, the land use land cover information of the masked portion in the original LISS III image has been replaced by the information collected from the topographic maps and some field data. The land use land cover layer thus prepared is shown in Figure 4.19.

It can be observed from Figure 4.19 that the northeastern, southeastern and southwestern parts of the area are dominated by thick forests. Tea plantation and sparse forests are the major land use land cover categories, which are distributed all over the area.

The areal distribution of different land use land cover categories has been derived and is listed in Table 4.12. It is observed from Table 4.12 that the most frequent categories of land use land cover are tea plantation and sparse forest, followed by thick forest, agriculture land, barren land, habitation, the least being water bodies and river sand.

Table 4.12: Distribution of existing landslides in different land use land cover categories.

Land use land cover categories	Area of land use land cover categories (km^2)	Percent area (%)	Landslide area per category (km^2)	Percent Landslide area per category (%)
Agriculture	22.3	8.8	0.053	25.0
Tea Plantation	89.1	35.0	0.052	24.6
Thick Forest	45.4	17.8	0.024	11.3
Sparse Forest	81.1	31.9	0.041	19.3
Barren Land	8.9	3.5	0.036	17.0
Habitation	6.5	2.6	0.006	2.8
Water	0.6	0.2	0.000	0.0
River Sand	0.6	0.2	0.000	0.0
Total	254.5	100.0	0.212	100.0

The spatial distribution of landslides in different land use land cover categories has been obtained (Table 4.12). It is observed that agriculture and tea plantation categories have maximum incidence of landslides in comparison to other categories and water bodies and river sand categories are devoid of landslides.

4.9 Summary

In the foregoing sections, a detailed description of thematic data layer preparation was presented. It can be observed from Figure 4.7 that a total of 101 existing landslides of varying dimensions are identified in the area. Most of the observed landslides are rock slides.

It is also observed that the most frequent is the 15°-25° slope category and the least frequent being the >45° slope category in the present area. Further, the slope category of 25°-35° has the highest incidence of landslides with least landslides

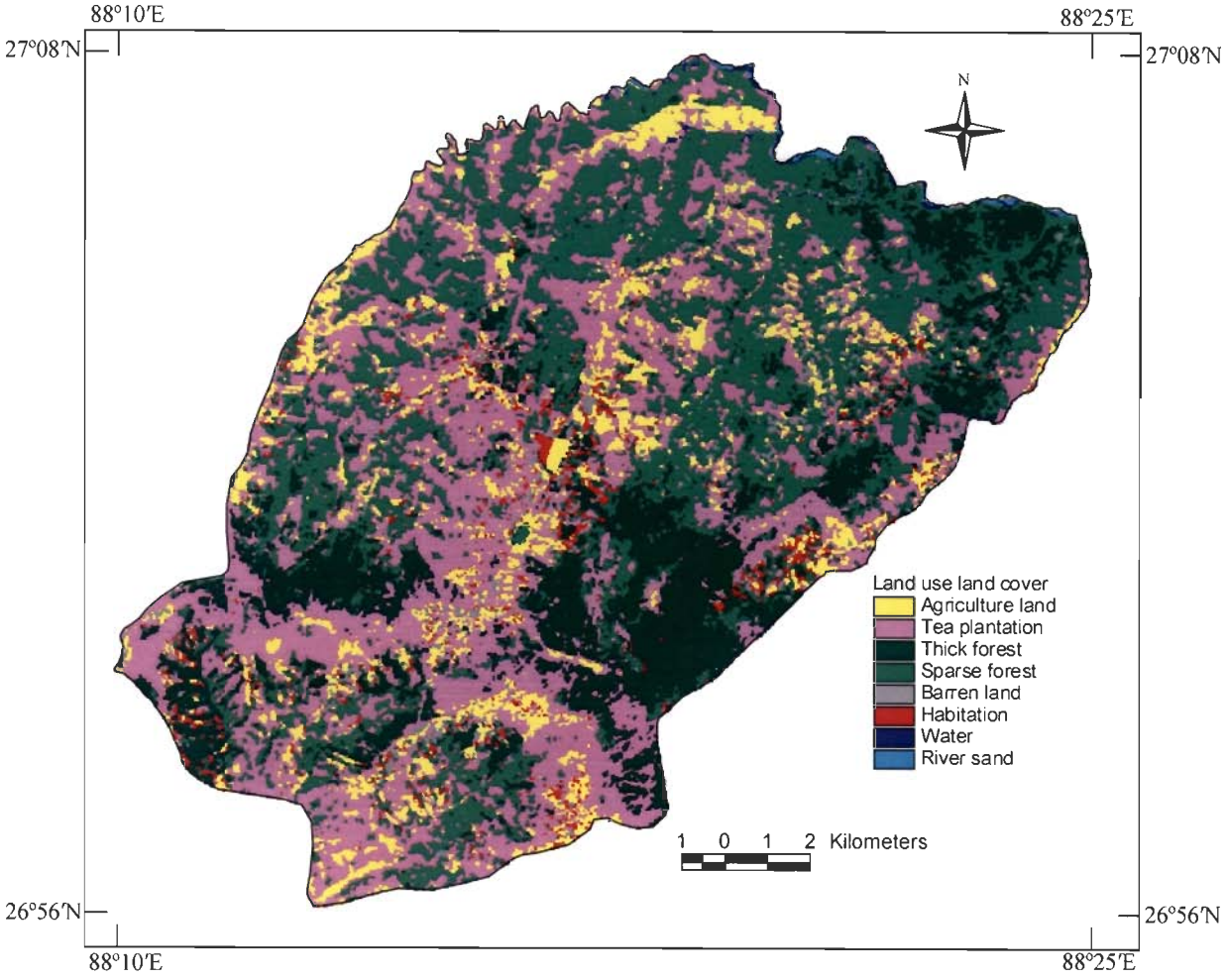


Figure 4.19: Land use land cover layer of the area.

occurring in the slope category >45° (Table 4.1). It can be observed from the spatial distribution of different aspect categories in the present area that the most frequent is the NW aspect and the least frequent being the category 'flat' which ought to be the case in hilly terrain. It is observed that the S, SE and E aspects have the maximum incidence of landslides in comparison to other aspects (Table 4.2).

The most frequent category of lithology observed in the area is the Paro gneiss, followed by Darjeeling gneiss, feldspathic greywacke, Lingtse granite gneiss, Paro quartzite, the least being the Reyang quartzite. It is observed that Paro gneiss has maximum incidence of landslides in comparison to other categories and quartzites and Lingtse granite gneiss have the least occurrences of landslides (Table 4.3).

A lineament buffer layer has been derived from the lineament data layer with a buffer width of 125m. It is observed that the lineament buffer category of 0-125m has the highest incidence of landslides with least landslides occurring in the lineament buffer category of >500m (Table 4.4). Hence, it can be inferred that the landslide susceptibility decreases as the distance from lineaments increases.

A drainage buffer layer has been prepared with 25m buffer width along all the drainage lines. It is observed from the spatial distribution of existing landslides in these buffer zones (Table 4.5) that majority of landslides occur in the 1st and 2nd order drainage buffer categories only. This indicates the fact that in hilly terrains, the lower order drainages play an important role in landslide occurrences. Therefore, 25m buffer zones along 1st and 2nd order drainages only have been considered for landslide susceptibility studies.

It can be observed from the land use land cover map (Fig. 4.19) that tea plantation and sparse forests are the major land use land cover categories, which are distributed all over the area. It is observed from the spatial distribution of landslides

that agriculture and tea plantation categories have maximum incidence of landslides in comparison to other categories and water bodies and river sand categories are devoid of landslides (Table 4.12).

These thematic data layers pertain to the causative factors that influence the landslide susceptibility of the study area. In the next chapters, different approaches developed for landslide susceptibility zonation (LSZ) and risk assessment mapping have been discussed and implemented using these thematic data layers.

Chapter 5

Landslide Susceptibility Zonation Using Conventional Weighting Approach

5.1 Introduction

As mentioned earlier, the LSZ mapping has attracted disaster specialists, geomorphologists, engineering geologists, scientists, academicians and others, and several approaches have been developed. A detailed review of these approaches has already been given in Chapter 2. Although several approaches have been developed for LSZ mapping, a number of issues that affect the performance of these approaches have been identified, including the difficulties in handling both continuous and categorical data together.

The present chapter deals with the implementation of the simplest and the most widely adopted conventional (subjective or qualitative) weighting approach for LSZ mapping in Darjeeling Himalayas. This has been attempted to evaluate the performance of any advanced objective approach such as neural and fuzzy approaches with respect to the conventional one.

5.2 Implementation of Conventional Weighting Approach

The LSZ mapping is performed to classify the area into different zones of varying degrees of landslide susceptibility, based on an estimated influence of causative factors and their categories in landslide occurrences. Therefore, the basic pre-requisite for LSZ studies is to assign/determine the relative importance of causative factors (thematic layers) and their categories in terms of weights and ratings respectively for landslide occurrences. This chapter highlights the implementation of conventional (subjective) weighting approach to produce the LSZ map. This approach involves assigning of weights and ratings to the causative factors and their categories respectively based on the knowledge of the study area and the experience on the subject. The integration of these subjective weights and ratings produces the LSZ map of the area. Salient steps used in this approach for LSZ mapping are given in the form of a flow diagram (Figure 5.1).

5.2.1 Weight and Rating Assignment

The identification of potential landslide areas (landslide susceptibility zones) depends on the cumulative importance of causative factors and their categories for landslide occurrences. This can be achieved by developing a weighting scheme in which factors and their categories are assigned numerical values. A weighting scheme has been developed based on the importance of causative factors and their categories for landslide occurrences as observed from field survey and also based on the knowledge from previous works. In this scheme, the factors have been assigned numerical weights on a 1-9 scale in order of importance. Ratings have also been assigned to the categories of the factors on a 0-9 scale. The higher weight or rating indicates more importance towards landslide occurrences. The weights and ratings

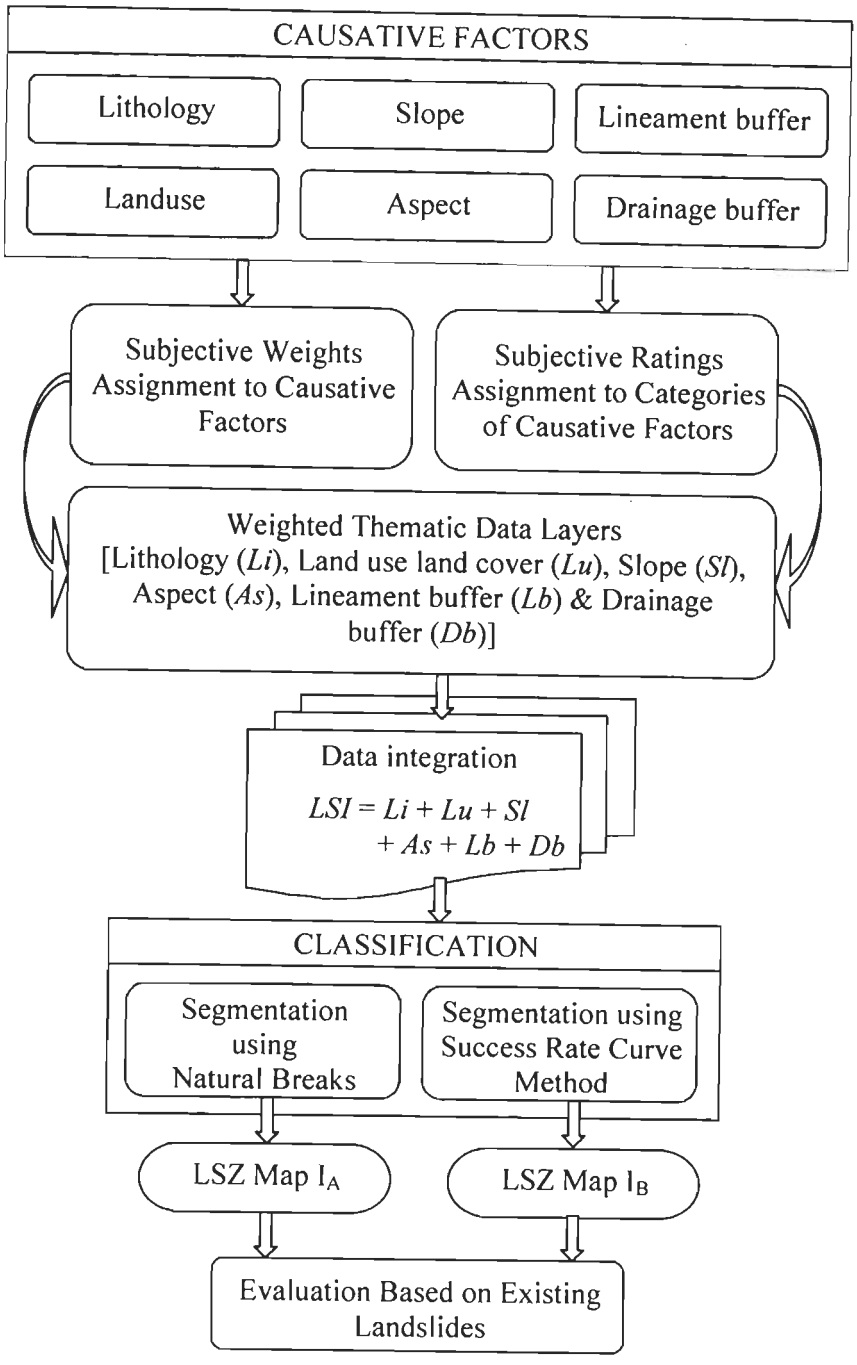


Figure 5.1: Flow diagram showing different steps of conventional weighting approach.

assigned to the causative factors and their categories in this study are given in Table 5.1.

Table 5.1: Weights and ratings for causative factors and their categories (conventional weighting approach).

Causative Factors	Weights	Categories	Ratings
Drainage Buffer	9	25m along 1st order drainage	9
		25m along 2nd order drainage	5
Lineament Buffer	8	0-125m	9
		125-250m	7
		250-375m	5
		375-500m	3
		>500m	1
Slope	7	>45°	9
		35°-45°	7
		25°-35°	5
		15°-25°	3
		0°-15°	1
Lithology	6	Lingse Granite Gneiss	9
		Darjeeling Gneiss	7
		Paro Gneiss	5
		Feldspathic Greywacke	3
		Paro Quartzite	1
		Reyang Quartzite	1
Land use land cover	4	Barren land	9
		Sparse forest	7
		Agriculture land	5
		Tea plantation	3
		Habitation	2
		Thick forest	1
		River sand	0
		Water body	0
Aspect	1	South (S)	9
		Southeast (SE)	8
		East (E)	7
		Southwest (SW)	6
		Northeast (NE)	4
		West (W)	3
		Northwest (NW)	2
		North (N)	1
		Flat	0

In the study area, it has been observed that most of the landslides are associated with drainages and hence the maximum weight has been assigned to the drainage layer. Also, maximum rating 9 has been assigned to 1st order drainage buffer category as most of the landslides initiate from the lower order drainages. The next important factor considered is lineament. Here, the maximum rating of 9 has been given to the 0-125m lineament buffer category because of the well established fact that the nearness to the lineaments controls the occurrence of landslide. The ratings for the categories of lineament buffer have been assigned in a decreasing trend as the distance increases. Since the steeper slopes are more prone to landslide, the slope classes have been given ratings in descending order of slope angles. The competent rocks such as quartzite, greywacke are less susceptible to landslides than the gneisses and hence, the ratings to lithology categories have been assigned accordingly. Occurrence of landslides also depends on the type of land use land cover. Barren slopes are more susceptible to erosion as compared to areas with thick forest and hence, maximum rating has been assigned to the barren slopes and minimum to the thick forest. The slope aspect has an indirect influence on slope instability. Based upon the existing landslide distribution, south and east facing slopes have been considered more potential for landslides (e.g., Dhakal et al., 2000). Accordingly, ratings for slope aspect categories have been assigned.

The weighted thematic data layer for each factor has been generated by multiplying the weight of the factor with the ratings of the corresponding categories (as mentioned in Table 5.1). In the present case, six different weighted thematic data layers corresponding to the causative factors have been produced.

5.2.2 Data Integration

The weighted thematic data layers have been integrated arithmetically according to the following equation to generate a landslide susceptibility index (LSI) map in GIS,

$$LSI = Li + Lu + Sl + As + Lb + Db \quad (5.1)$$

where *Li*, *Lu*, *Sl*, *As*, *Lb* and *Db* are abbreviations for the weighted thematic layers for lithology, land use land cover, slope, aspect, lineament buffer and drainage buffer respectively.

It is found from the LSI map that the LSI values range from 21 to 310 in the study area.

5.2.3 Segmentation of LSI Values into Landslide Susceptibility Zones

The LSI values need to be segmented to generate very high, high, moderate, low and very low landslide susceptibility zones. Two methods for this have been used:

(a) Segmentation using natural breaks in distribution

(b) Segmentation using success rate curve method

5.2.3.1 Segmentation Using Natural Breaks in Distribution

A judicious way for segmenting LSI values is to search for abrupt changes (Davis, 1986) in the range of LSI values. The classification procedure reported by Sarkar and Kanungo (2004) has been followed. For this purpose, a frequency distribution curve of LSI values has been prepared which shows many oscillations.

Hence, moving averages with moving intervals of 5, 9 and 13 have been considered to smoothen the curve. A moving interval of 5 means that the frequency value in the corresponding curve at any point is an average of the five consecutive values centered at that point.

The class boundaries have been drawn at significant breaks in the curve and found to occur at LSI values of 68, 137, 176 and 236 in this case (Figure 5.2). Thus, the LSZ map has been prepared with class intervals 21-68, 69-137, 138-176, 177-236 and 237-310, which are designated as very low susceptibility (VLS), low susceptibility (LS), moderate susceptibility (MS), high susceptibility (HS) and very high susceptibility (VHS) zones. The LSZ map thus prepared is referred here as LSZ Map I_A (Figure 5.3).

Visual inspection of LSZ Map I_A suggests that all five zones are distributed all over the study area. The map thus does not show any well defined pattern for the distribution of susceptibility zones. It is again observed from this map that the VHS and HS zones have represented mostly the 1^{st} and 2^{nd} order drainage buffer areas. Thus, it can be inferred that there is a major control of drainage order on landslide incidence and LSZ mapping.

The area covered by different landslide susceptibility zones and also the landslide distribution in different susceptibility zones have been calculated (Table 5.2).

It is inferred from Table 5.2 that the VHS zone occupies 6.5% of the study area, whereas the HS, MS, LS and VLS zones occupy 26.8%, 30.2%, 34.9% and 1.6% of the study area respectively. This shows that the area wise coverage of different susceptibility zones is skewed towards lower susceptibility zones, which should not be the case. The distribution of landslides in different susceptibility zones has been

compared. It has been found that 10.6% landslide incidence occurs in VHS zone, while 48.1%, 26.5%, 14.5% and 0.3% landslide area fall in HS, MS, LS and VLS zones respectively. Hence, it can be stated that 33.3% of HS and VHS areas together contain 58.7% of existing landslide areas. Also, this shows that the distribution of landslides over VHS to VLS zones is skewed towards the higher susceptibility zones, which should in fact be the case.

Table 5.2: Landslide distribution in landslide susceptibility zones of LSZ Map I_A (segmentation using natural breaks in distribution of LSI in conventional weighting approach).

Landslide Susceptibility Zones	Area (km^2)	Percent Area (%) (a)	Landslide Area per Class (km^2)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	16.6	6.5	0.023	10.6	1.63
HS	68.4	26.8	0.102	48.1	1.79
MS	77.1	30.2	0.056	26.5	0.88
LS	88.8	34.9	0.031	14.5	0.41
VLS	4.0	1.6	0.001	0.3	0.19

Further, the landslide density for each susceptibility zone has been determined as the ratio of percent existing landslide area to percent area of that susceptibility zone (Table 5.2). Ideally, the landslide density should be highest for VHS zone and should attain a successively decreasing trend from VHS to VLS zones. In Map I_A, it is observed that the landslide densities for VHS and HS zones are almost equal. This map has a systematic and reasonable trend of variation in landslide density values only from HS to VLS zones. For VHS and HS zones combined, the landslide density is 1.76. Hence, it is inferred from Table 5.2 that these higher susceptibility zones do not have much higher landslide density values as compared to other susceptibility zones.

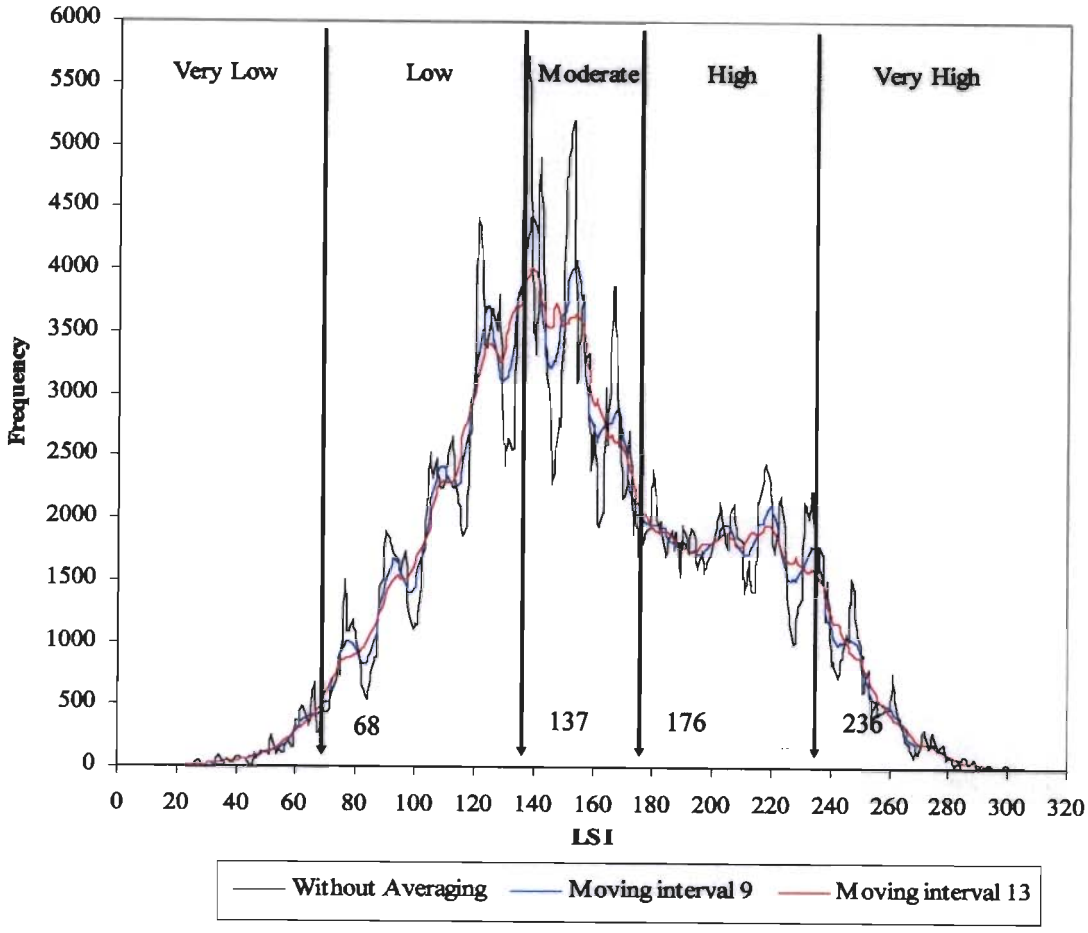


Figure 5.2: Segmentation of LSI values using natural breaks for LSZ classification.

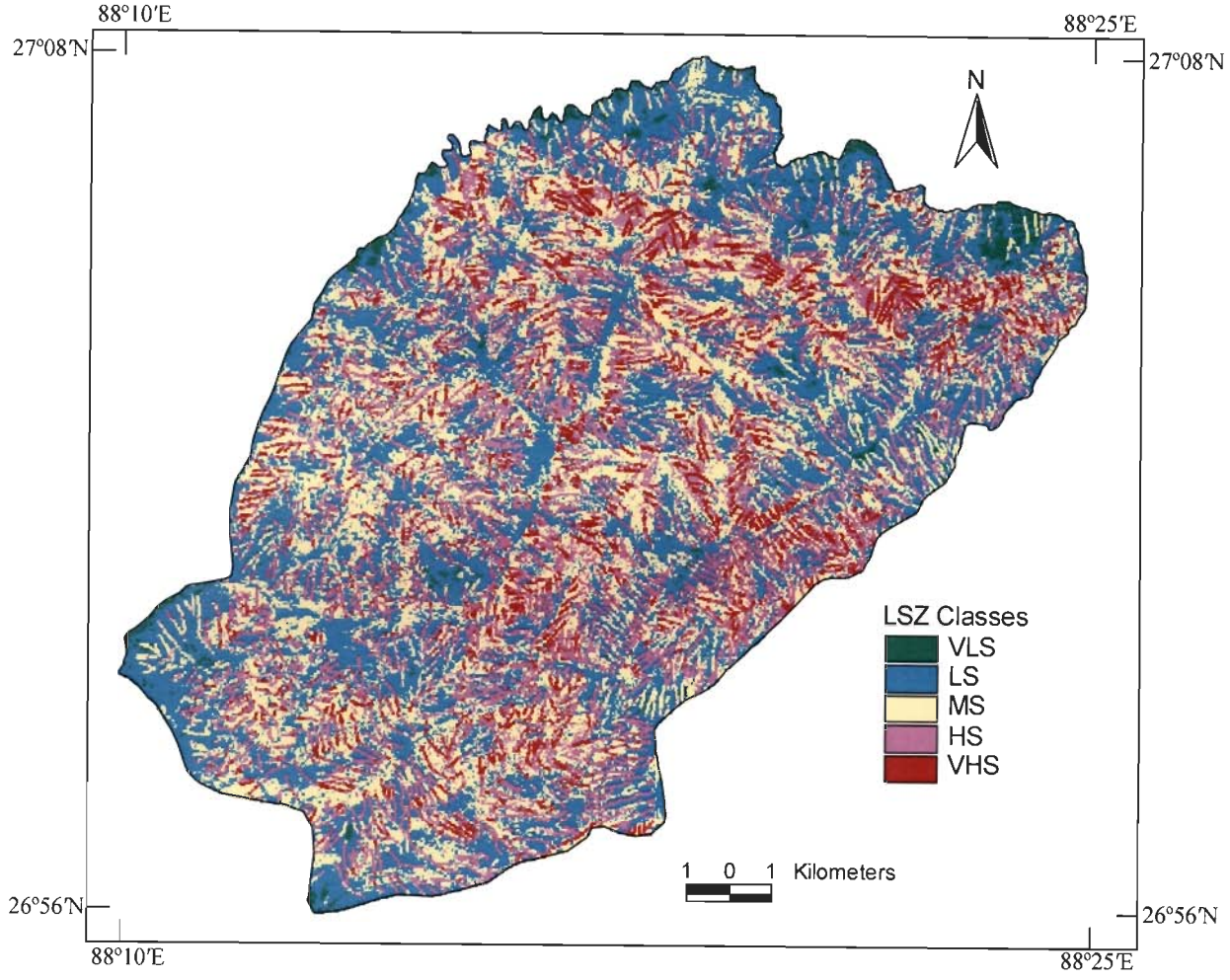


Figure 5.3: LSZ Map I_A using segmentation of LSI values by natural breaks in conventional weighting approach.

5.2.3.2 Segmentation Using Success Rate Curve Method

Another way for segmenting the range of LSI values statistically is by using success rate curve method. This method has been adopted to avoid the subjectivity, if any, in segmentation based on natural breaks. Success rate is defined as percentage of landslide occurrence in any susceptibility zone.

To implement this method, the observed mean (μ_o) and standard deviation (σ_o) from the probability distribution curve of the range of LSI values (21 to 310) have been determined and are found to be 158.7 and 47.9 respectively. The LSI values have been divided into five distinct classes (susceptibility zones) with boundaries at ($\mu_o - 1.5m\sigma_o$), ($\mu_o - 0.5m\sigma_o$), ($\mu_o + 0.5m\sigma_o$) and ($\mu_o + 1.5m\sigma_o$) where m is a positive, non-zero value. Several LSZ maps of the area have been prepared for different values of m. The suitability of any LSZ map can be judged by the fact that more percentage of landslides must occur in VHS zone as compared to other zones. Therefore, the cumulative percentage of landslide occurrences in various susceptibility zones ordered from very high susceptibility (VHS) to very low susceptibility (VLS) have been plotted against the cumulative percentage of area of the susceptibility zones for LSZ maps with different values of m. These curves have been defined as the success rate curves (Chung and Fabbri, 1999; Lu and An, 1999; Lee et al., 2002b) and have been used to select the appropriate value of m to decide the suitability of an LSZ map.

Five representative success rate curves corresponding to m = 1.1, 1.2, 1.3, 1.4 and 1.5 have been produced (Figure 5.4). It has been observed that for 10% of the area in VHS zone, the curves corresponding to m=1.1, 1.2, 1.3, 1.4 and 1.5 show the landslide occurrences of 16.3%, 17.5%, 16.5%, 15.5% and 13.0% respectively. Hence, for the first 10% area in VHS zone, the curve corresponding to m=1.2 has the highest success rate. Based on this analysis, the LSZ map corresponding to m = 1.2

appears to be the most appropriate one for the study area. Accordingly, with m = 1.2, the landslide susceptibility zone boundaries have been fixed at LSI values of 72, 129, 187 and 225. The LSZ map has been prepared with class intervals 21-72, 73-129, 130-187, 188-225 and 226-310, which are designated as VLS, LS, MS, HS and VHS zones. The LSZ map thus produced is referred here as LSZ Map I_B (Figure 5.5).

Visual inspection of LSZ map I_B suggests that all five zones are distributed all over the study area. The map thus does not show any well defined pattern. It is again observed that the VHS and HS zones have represented mostly the 1st and 2nd order drainage buffer areas. It can be inferred that there is a major control of drainage order on landslide incidence and LSZ mapping.

The area covered by five different landslide susceptibility zones and the landslide distribution in different susceptibility zones have also been determined (Table 5.3).

Table 5.3: Landslide distribution in landslide susceptibility zones of LSZ Map I_B (success rate curve method of segmentation in conventional weighting approach).

Landslide Susceptibility Zones	Area (km^2)	Percent Area (%) (a)	Landslide Area per Class (km^2)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	10.6	4.1	0.016	7.4	1.80
HS	61.1	24.0	0.08	37.7	1.57
MS	108.1	42.4	0.088	41.6	0.98
LS	69.9	27.4	0.027	13.0	0.47
VLS	5.2	2.1	0.001	0.3	0.14

It is inferred from Table 5.3 that the VHS zone occupies 4.1% of the study area, whereas the HS, MS, LS and VLS zones occupy 24.0%, 42.4%, 27.4% and 2.1% of the study area respectively. This shows that the area wise coverage of different susceptibility zones is normally distributed, which should be the case. The distribution

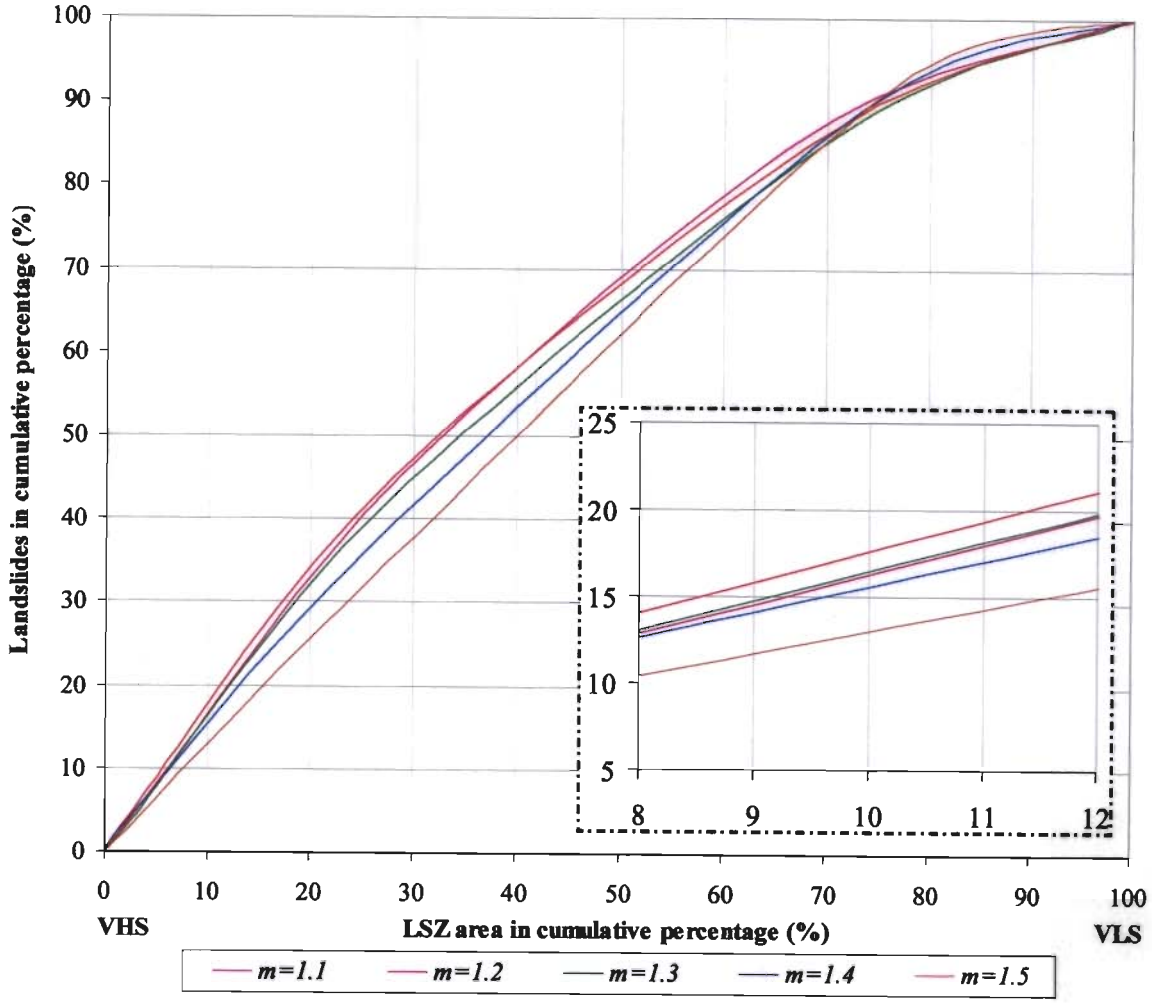


Figure 5.4: Success rate curves for choosing the best segmentation in LSI values for LSZ classification in conventional weighting approach.

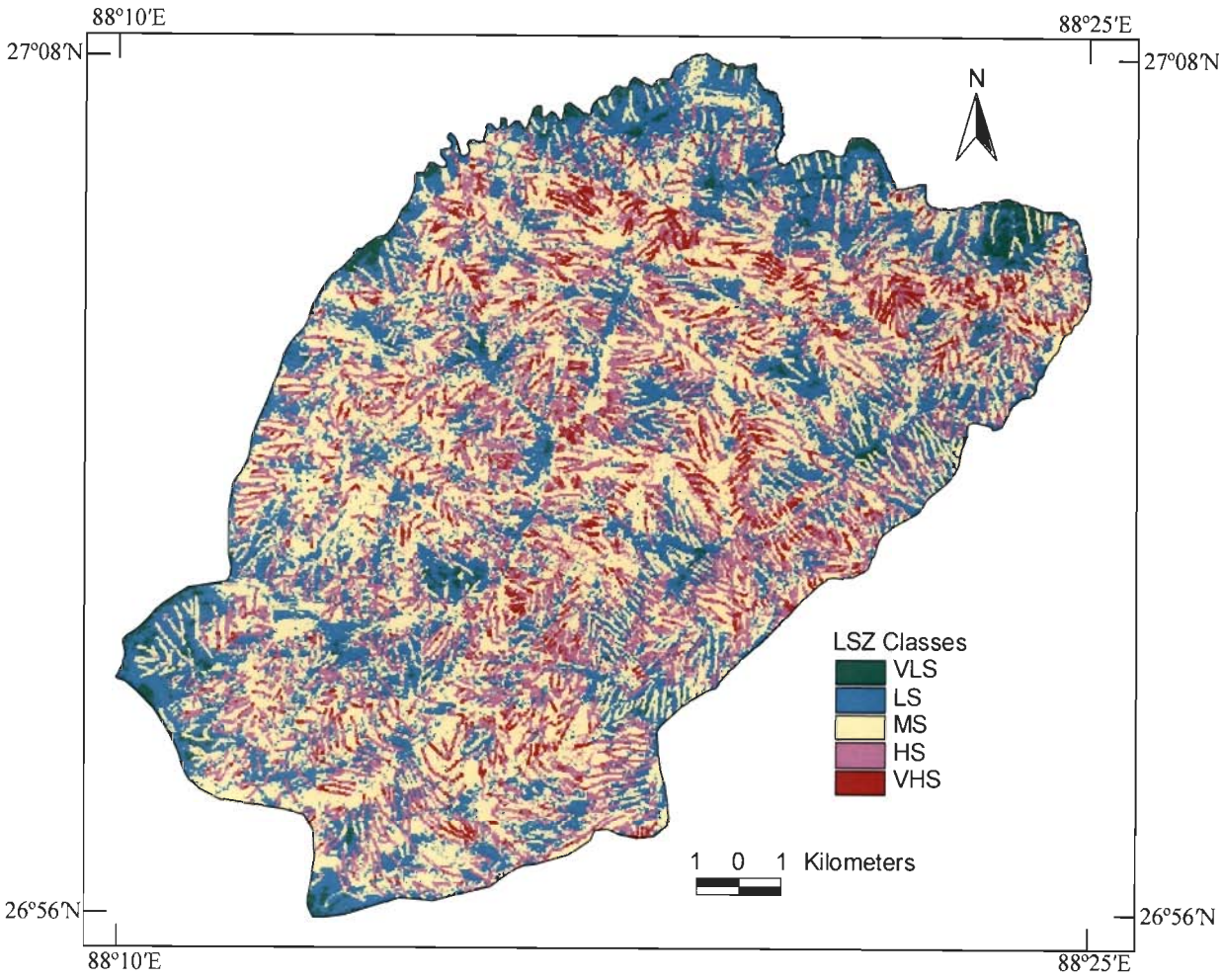


Figure 5.5: LSZ Map I_B using segmentation of LSI values by success rate curve method in conventional weighting approach.

of landslides in different susceptibility zones has been compared. It has been found that 7.4% landslide incidence occurs in VHS zone, while 37.7%, 41.6%, 13.0% and 0.3% landslide area fall in HS, MS, LS and VLS zones respectively. Hence, it can be stated that 28.1% of HS and VHS areas together contain 45.1% of existing landslide areas. Also, it is inferred that the existing landslides are normally distributed over VHS to VLS zones, which should not be the case.

Further, the landslide densities of all the five susceptibility zones have been determined (Table 5.3). In case of Map I_B, it is observed that the landslide densities for VHS and HS zones are almost equal. This map has a systematic and reasonable trend of variation in landslide density values only from VHS to VLS zones. For VHS and HS zones combined, the landslide density is 1.60. Hence, it is inferred from Table 5.3 that these higher susceptibility zones do not have much higher landslide density values as compared to other susceptibility zones.

5.3 Summary

Visual inspection of LSZ Maps I_A and I_B suggests that all five susceptibility zones are distributed all over the study area. The maps thus do not show any well defined pattern for the distribution of susceptibility zones. It is again observed from these maps that the VHS and HS zones have represented mostly the 1st and 2nd order drainage buffer areas. It can be inferred that there is a major control of drainage lines on landslide incidence and LSZ mapping.

The area wise coverage of different susceptibility zones in case of LSZ Map I_B is normally distributed, whereas that in case of LSZ Map I_A is skewed towards lower susceptibility zones. The distribution of landslides over VHS to VLS zones is skewed

towards the higher susceptibility zones in case of LSZ Map I_A, whereas in case of LSZ Map I_B it is normally distributed.

In general, it is observed that there is not much difference in landslide densities of different susceptibility zones for both the LSZ maps. It is also observed that for VHS and HS zones combined, the landslide density is better for LSZ Map I_A (1.76) than LSZ Map I_B (1.60). Also, the LSZ Map I_A has a skewed landslide distribution towards higher susceptibility zones.

In conventional weighting approach, the weights and ratings are assigned based on the field observation, where field of view is limited. Considering the distribution of landslides in field along drainage lines, we presumably considered drainage line to be the most important causative factor for LSZ and generated the LSZ maps. However, the causative factors in terms of their importance may have a different order, which may not be decided with a limited field of view. Further, the weights and ratings in this approach have been assigned on 1-9 and 0-9 scales respectively. However, it may happen to be a different level of scale differences between the causative factors (weights) as well as their categories (ratings). Hence, there is a lot of subjectivity in the order and value of importance in terms of weights and ratings of the causative factors and their categories.

Therefore, to overcome this subjectivity and bias in assigning weights in conventional weighting approach, advanced techniques like neural and fuzzy approaches have been attempted for LSZ mapping. The LSZ Map I_A has been used as reference map for evaluating the output of neural and fuzzy approaches (Chapter 6). But, success rate curve method has been used for segmenting the range of LSI values statistically to generate the LSZ maps in the later cases.

Chapter 6

Neural and Fuzzy Set Theoretic Approaches for Landslide Susceptibility Zonation

6.1 Introduction

Conventional weighting approaches for LSZ mapping have been widely adopted worldwide. In Chapter 5, a detailed description of this approach as implemented in the present study area was provided. As has been understood earlier, this approach is highly subjective in assigning weights and ratings to the causative factors and their categories respectively. In order to minimize the subjectivity in the weight assigning process, an objective approach needs to be developed for LSZ mapping.

Recently, the use of artificial neural networks (ANN) have been found to be advantageous for LSZ mapping (Arora et al., 2004; Lee et al., 2004; Gomez and Kavzoglu, 2005; Yesilnacar and Topal, 2005) as these have the capability to analyze complex data patterns and can also handle continuous and categorical data. In addition, fuzzy set theory has also been found useful for such studies (Chi et al., 2002b; Gorsevski et al., 2003; Tangestani, 2003; Ercanoglu and Gokceoglu, 2004).

Here, a fuzzy set can be utilised to assign varying degree of memberships to the categories of causative factors according to their importance in the process. These membership values can be determined using fuzzy relation concept. A fully objective approach can be evolved using the neural and fuzzy set theoretic concepts. In this study, three different approaches, namely ANN black box approach, fuzzy set based approach and combined neural and fuzzy approach have been developed. The aim is to demonstrate the utility of these approaches in mapping landslides in an objective manner. The detail theoretical background of these was provided in Chapter 3. In this Chapter, the implementation of these approaches for LSZ has been described.

6.2 LSZ Using ANN Black Box Approach

The basic concept of ANN black box approach has been discussed in Section 3.3.2.2 of Chapter 3. A flow diagram showing different steps followed in implementation of ANN black box approach is given in Figure 6.1.

6.2.1 Implementation

A feed forward back-propagation multi-layer artificial neural network with one input layer, two hidden layers and one output layer has been considered. The input layer contains 6 neurons each representing a thematic data layer (i.e., causative factor) that contributes to the occurrence of the landslide. The output layer contains a single neuron that represents either the landslide pixels (Case I) or the landslide susceptibility zones (Case II). Thus, two different datasets have been considered for implementing the respective cases.

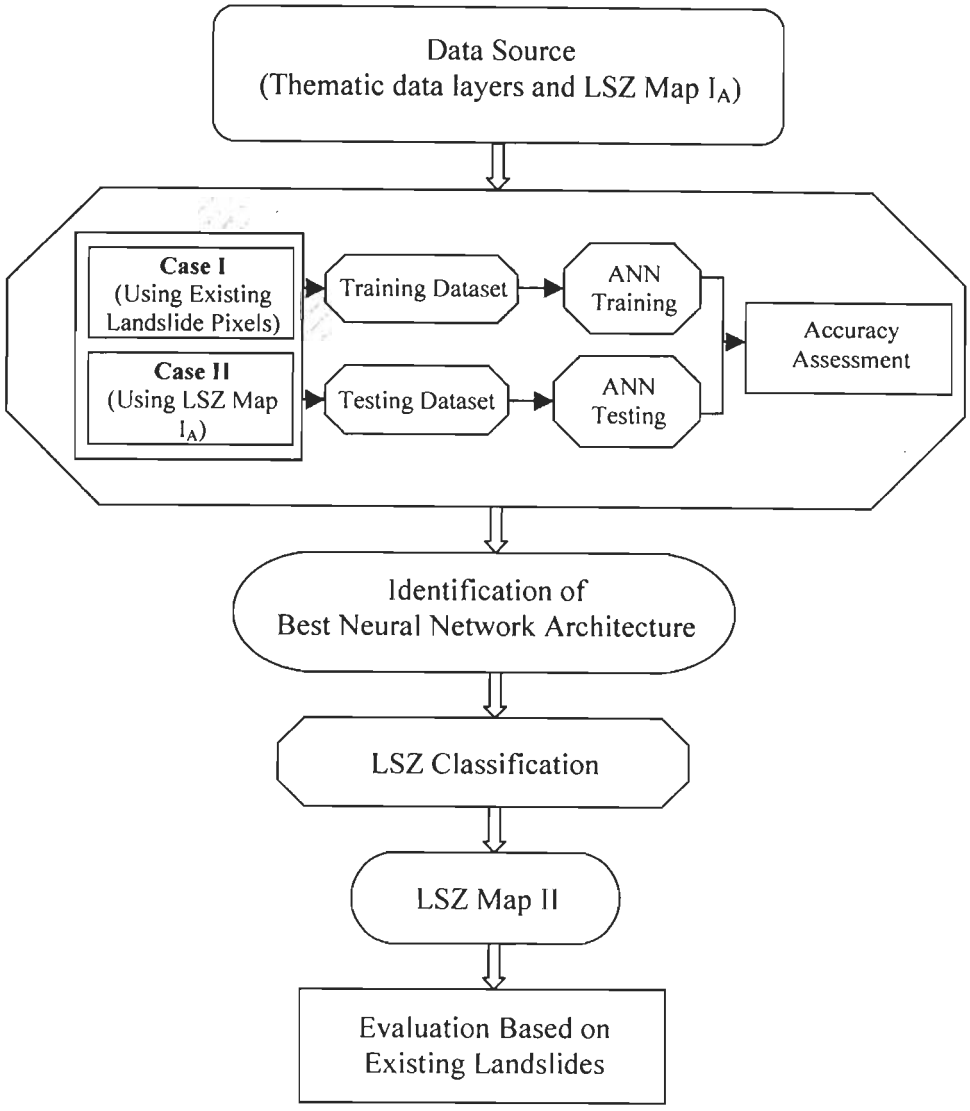


Figure 6.1: Flow diagram showing different steps of ANN black box approach.

6.2.1.1 Case I – Using Existing Landslide Locations

ANN Architecture

A schematic representation of ANN architecture, adopted in this case, is given in Figure 6.2. The ANN architectures with one input layer, two hidden layers and one output layer have been considered. The input layer contains 6 neurons corresponding to 6 different causative factors and the output layer contains a single neuron corresponding to existing landslide locations. The data supplied to input neurons correspond to the normalized attributes, as given in Table 6.1. The attribute of each category has been normalized with respect to the highest attribute within the corresponding thematic data layer (causative factor). The desired output data supplied to output neuron corresponds to 1.0 for existing landslide location. Initially, a few ANN architectures were designed by varying the number of neurons in hidden layers. These networks were trained and tested to evaluate their performance. Subsequently, the number of networks was increased and a total of 24 ANN architectures were designed, trained and tested to evaluate their generalization capabilities and accuracies in the present case.

Data Preparation

The dataset consisted of 339 existing landslide pixels (see Section 4.3). The complete dataset was divided into two parts, one each for training and testing data consisting of 170 and 169 independent pixels respectively. The testing dataset was also used as the verification dataset to control the overtraining of networks.

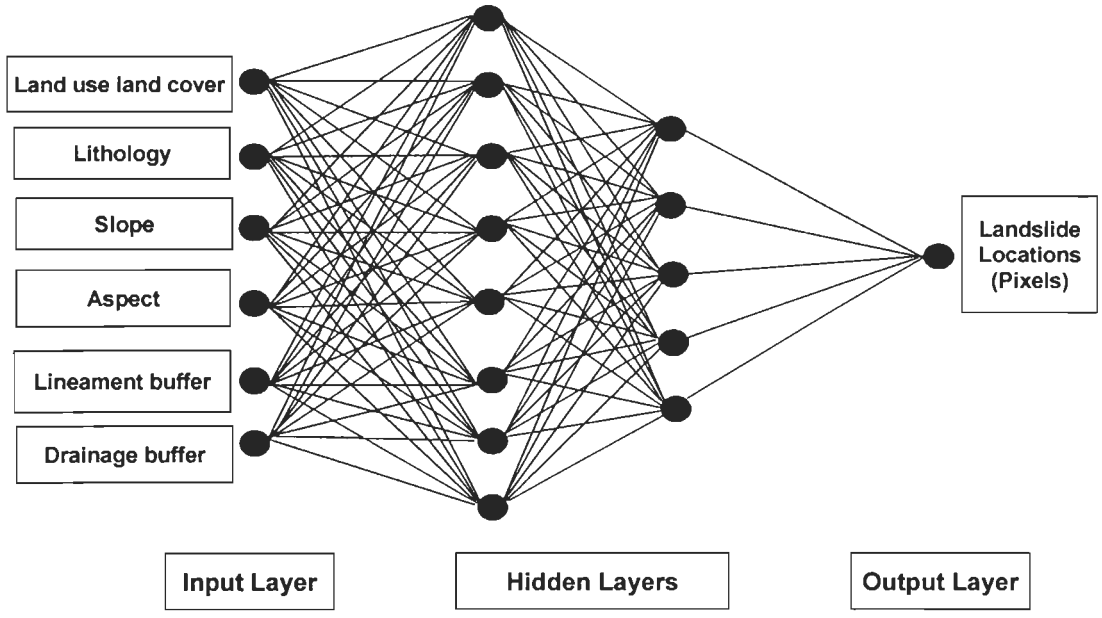


Figure 6.2: A schematic diagram of ANN architecture in ANN black box approach (Case I - using existing landslide locations).

Training and Testing of ANN

The Levenberg-Marquardt back-propagation learning algorithm (implemented as TRAINLM in MATLAB software) has been used to train the neural networks. Unlike conventional gradient descent algorithm, it does not consider parameters like learning rate and momentum factor. In contrast, it takes into account mu (μ), mu_dec and mu_inc as training parameters. The parameter μ is decreased by multiplying it with mu_dec after each successful step (reduction in performance function) and is increased only when the performance function is increased. The detail description on these parameters has been provided in Hagan and Menhaj (1994) and Hagan et al. (1996). The important training parameters considered in this study are listed in Table 6.2. The default values for these parameters as listed in this table have been used and assumed fixed. No attempt has been made to change the values of these training parameters for effective comparison across various networks.

Table 6.1: Normalized attributes of categories of thematic layers used as input data to ANN black box approach.

Thematic Layers	Categories	Attributes	Normalized Attributes of Categories Used as Input to ANN
Drainage Buffer	25m along 1st order drainage	1	0.500
	25m along 2nd order drainage	2	1.000
Lineament Buffer	0-125m	1	0.200
	125-250m	2	0.400
	250-375m	3	0.600
	375-500m	4	0.800
	>500m	5	1.000
Slope	0°-15°	1	0.200
	15°-25°	2	0.400
	25°-35°	3	0.600
	35°-45°	4	0.800
	>45°	5	1.000
Lithology	Reyang Quartzite	1	0.167
	Paro Quartzite	2	0.333
	Feldspathic Greywacke	3	0.500
	Paro Gneiss	4	0.667
	Darjeeling Gneiss	5	0.833
	Lingse Granite Gneiss	6	1.000
Land use land cover	Water body	1	0.125
	River sand	2	0.250
	Thick forest	3	0.375
	Habitation	4	0.500
	Tea plantation	5	0.625
	Agriculture land	6	0.750
	Sparse forest	7	0.875
	Barren land	8	1.000
Aspect	Flat	1	0.111
	North	2	0.222
	Northeast	3	0.333
	East	4	0.444
	Southeast	5	0.555
	South	6	0.667
	Southwest	7	0.778
	West	8	0.889
	Northwest	9	1.000

Table 6.2: Values of different training parameters as used in ANN training.

Training Parameters	Value as used
Epochs	10000
min_grad	1×10^{-10}
max_fail	5
mu (μ)	0.001
mu_dec	0.1
mu_inc	10
mu_max	10^{10}

The training process is initiated by assigning arbitrary initial connection weights, which are constantly updated until an acceptable training accuracy is reached. The adjusted weights obtained from the trained network have been subsequently used to process the testing data to evaluate the generalization capability and accuracy of the network.

The performance of the networks has been evaluated by determining both training and testing data accuracies in terms of correlation coefficient, root mean squared error (RMSE) (Freund, 1992) and percent correct or overall classification accuracy (Congalton, 1991). The training and testing accuracies for all the 24 networks are given in Table 6.3. The overall accuracy plot is shown in Figure 6.3.

From Table 6.3, it is observed that the training accuracies are quite high, as can be gauged from high correlation coefficients (0.533 to 0.971), low RMSE values (0.118 to 0.033) and high overall accuracy (73.5% to 97.1%). However, the testing data accuracies are very low, as can be seen from low correlation coefficients (0.018 to 0.365), high RMSE values (0.358 to 0.131) and low overall accuracy (45.0% to 72.8%). The trend in increase in training data accuracy and decrease in testing across various ANN architectures can clearly be seen from Figure 6.3. As the number of neurons in both the hidden layers is changed, the overall training accuracy has

increased up to the order of 97%. However, the overall testing accuracies have decreased to the order of 50%. Thus, there exists a significant difference in training and testing accuracies, which should not happen. This shows that the networks are able to get trained to a high accuracy, but could not attain the generalization capability. This may be due insufficient number of training and testing samples because of paucity of field data. Therefore, similar to earlier studies (Arora et *al.*, 2004), the LSZ Map I_A obtained from conventional weighting approach has been considered as reference map, to derive representative training and testing samples with known LSZ class. This corresponds to Case II, as explained in the next section.

Table 6.3: Training and testing accuracies for Case I in ANN black box approach.

ANN Architecture	Correlation Coefficient		RMSE		Overall Accuracy (%)		
	Training	Testing	Training	Testing	Training	Testing	Difference in training and testing accuracies
6/3/1/1	0.533	0.365	0.118	0.131	73.5	68.0	5.5
6/3/2/1	0.749	0.306	0.092	0.144	84.7	72.8	11.9
6/3/3/1	0.817	0.199	0.080	0.213	85.3	68.6	16.7
6/3/6/1	0.861	0.177	0.071	0.277	90.0	65.1	24.9
6/3/9/1	0.869	0.119	0.069	0.236	90.0	59.2	30.8
6/4/2/1	0.743	0.309	0.093	0.156	81.8	59.8	22.0
6/4/3/1	0.828	0.357	0.078	0.162	87.6	63.9	23.7
6/5/2/1	0.844	0.266	0.074	0.198	90.0	61.5	28.5
6/5/3/1	0.709	0.272	0.098	0.171	82.9	69.8	13.1
6/5/4/1	0.871	0.292	0.068	0.185	88.8	64.5	24.3
6/6/2/1	0.810	0.099	0.082	0.241	88.8	60.9	27.9
6/6/3/1	0.881	0.316	0.066	0.177	92.3	68.0	24.3
6/6/4/1	0.932	0.106	0.050	0.298	95.3	51.5	43.8
6/6/9/1	0.971	0.233	0.033	0.244	97.1	66.3	30.8
6/7/3/1	0.924	0.224	0.053	0.292	94.1	50.9	43.2
6/7/6/1	0.966	0.274	0.036	0.298	97.1	52.1	45.0
6/7/12/1	0.971	0.195	0.033	0.358	97.1	50.9	46.2
6/8/2/1	0.922	0.248	0.054	0.260	94.7	71.6	23.1
6/8/5/1	0.970	0.234	0.034	0.289	97.1	60.4	36.7
6/9/4/1	0.971	0.324	0.033	0.248	97.1	61.5	35.6
6/12/5/1	0.971	0.274	0.033	0.254	97.1	58.6	38.5
6/12/10/1	0.971	0.018	0.033	0.327	97.1	45.0	52.1
6/13/7/1	0.971	0.236	0.033	0.311	97.1	53.8	43.3
6/14/5/1	0.971	0.284	0.033	0.306	97.1	49.7	47.4

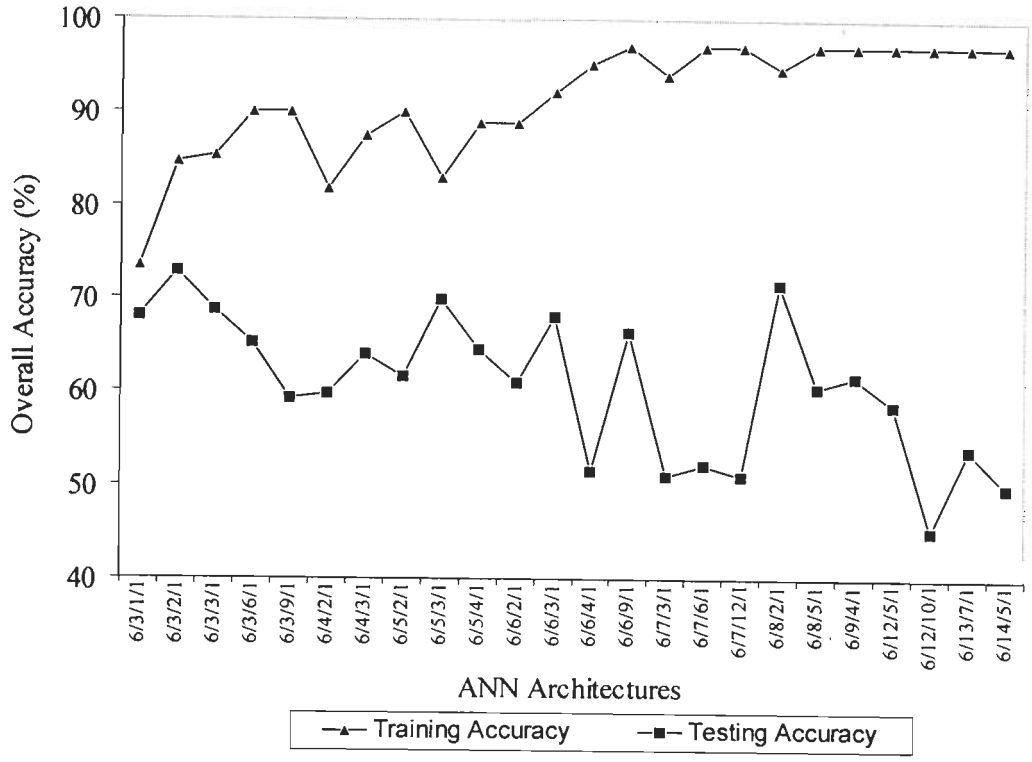


Figure 6.3: Graph showing training and testing accuracies for Case I in ANN black box approach.

6.2.1.2 Case II – Using LSZ Map I_A

ANN Architecture

A schematic representation of ANN architecture, adopted in this case, is given in Figure 6.4. The ANN architectures with one input layer, two hidden layers and one output layer have been considered. The input layer contains 6 neurons corresponding to 6 causative factors and the output layer contains a single neuron representing one landslide susceptibility zone. The data supplied to input neurons correspond to the normalized attributes, as given in Table 6.1. The desired output data supplied to output neuron corresponds to the normalized attribute values 0.2, 0.4, 0.6, 0.8 and 1.0

of different landslide susceptibility zones (VLS, LS, MS, HS and VHS) in order of increasing susceptibility.

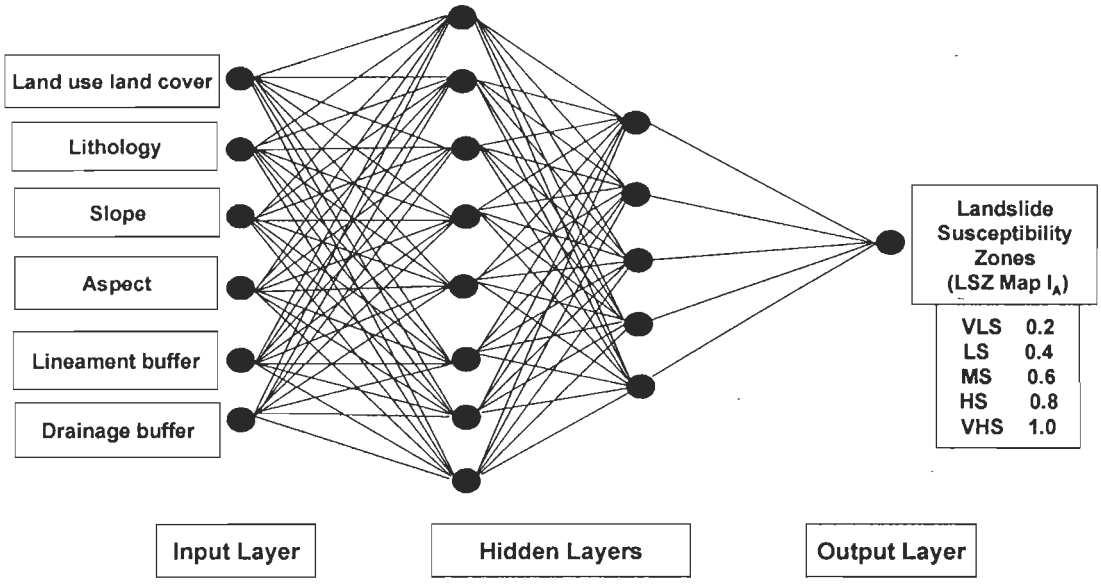


Figure 6.4: A schematic diagram of ANN architecture in ANN black box approach (Case II - using LSZ Map I_A).

Initially, a few neural network architectures were designed by varying the number of neurons in hidden layers. These networks were trained and tested to evaluate their performance. Subsequently, the number of networks was increased and a total of 39 network architectures were designed, trained and tested to evaluate their generalization capabilities and accuracies in the present case.

Data Preparation

In this case, the LSZ Map I_A obtained from conventional weighting approach has been considered as reference map. Two independent training and testing datasets are formed. Each dataset consists of 500 mutually exclusive pixels corresponding to one landslide susceptibility zone as identified from LSZ Map I_A. Since, there are five zones, in total 2500 pixels as training and another 2500 pixels for testing datasets

have been considered. The pixels in both the datasets were mutually exclusive. The testing dataset has also been used as verification dataset to control the overtraining of the networks.

Training and Testing of ANN

The Levenberg-Marquardt back-propagation algorithm is used to train the neural networks. The training parameters used in this case is already given in Table 6.2. The training process is initiated by assigning arbitrary initial connection weights, which are constantly updated until an acceptable training accuracy is reached. The adjusted weights obtained from the trained network have been subsequently used to process the testing data to evaluate the generalization capability and accuracy of the network. The performances of the networks are evaluated by different measures as described earlier (see Section 6.2.1.1). The training and testing accuracies for 18 networks (out of 39 networks) showing better accuracies comparative to other networks are given in Table 6.4. The overall accuracy plot is shown in Figure 6.5.

In the present case, it can be seen from this table that in general both training and testing data accuracies are high for all the networks. The noticeable feature, as can be gauged from Figure 6.5, is that both the training and testing data accuracies are more or less similar for all the networks. The overall training accuracies vary from 65% to 74% whereas the overall testing accuracies varied from 63% to 73% for different networks. This clearly shows that the network has achieved generalization capability besides getting trained at high accuracy. Further, from Figure 6.5, a variation in both training and testing data accuracies can be noticed as the neural network architecture changes. This suggests that there exists an optimal neural network architecture for a given dataset. In the present case, the highest overall

training and testing accuracies have been achieved for 6/13/7/1 network. Moreover, the difference between the training and testing accuracies for this network is the least, which in fact should be the case. Therefore, the network 6/13/7/1 has been found to be the most appropriate one for this case. Hence, the connection weights obtained from this network are captured and subsequently used to determine the network output of all the pixels in the dataset to prepare the LSZ map of the study area.

Table 6.4: Training and testing accuracies for Case II in ANN black box approach (underline indicates the best acceptable architecture).

ANN Architecture	Correlation Coefficient		RMSE		Overall Accuracy (%)		
	Training	Testing	Training	Testing	Training	Testing	Difference in training and testing accuracies
6/4/2/1	0.877	0.870	0.136	0.139	64.9	63.2	1.7
6/5/2/1	0.886	0.875	0.131	0.137	67.4	65.0	2.4
6/6/4/1	0.892	0.881	0.128	0.134	68.0	65.3	2.7
6/7/5/1	0.906	0.893	0.120	0.127	71.0	69.3	1.7
6/8/5/1	0.908	0.891	0.119	0.128	70.8	68.7	2.1
6/9/5/1	0.911	0.893	0.117	0.127	72.1	68.6	3.5
6/10/4/1	0.912	0.893	0.116	0.128	71.4	69.8	1.6
6/11/3/1	0.915	0.897	0.114	0.125	73.9	70.0	3.9
6/12/4/1	0.912	0.892	0.116	0.128	72.9	70.9	2.0
6/13/5/1	0.915	0.899	0.114	0.124	73.8	70.3	3.5
6/13/7/1	0.918	0.896	0.112	0.126	74.4	72.6	1.8
6/13/9/1	0.918	0.893	0.112	0.128	73.2	68.0	5.2
6/14/3/1	0.912	0.892	0.116	0.128	71.9	70.1	1.8
6/15/4/1	0.919	0.891	0.112	0.128	74.2	69.2	5.0
6/15/6/1	0.915	0.894	0.114	0.127	73.8	69.5	4.3
6/15/8/1	0.914	0.894	0.115	0.127	72.9	69.6	3.3
6/16/3/1	0.917	0.890	0.113	0.129	72.7	69.0	3.7
6/16/7/1	0.920	0.889	0.111	0.130	73.8	67.2	6.6

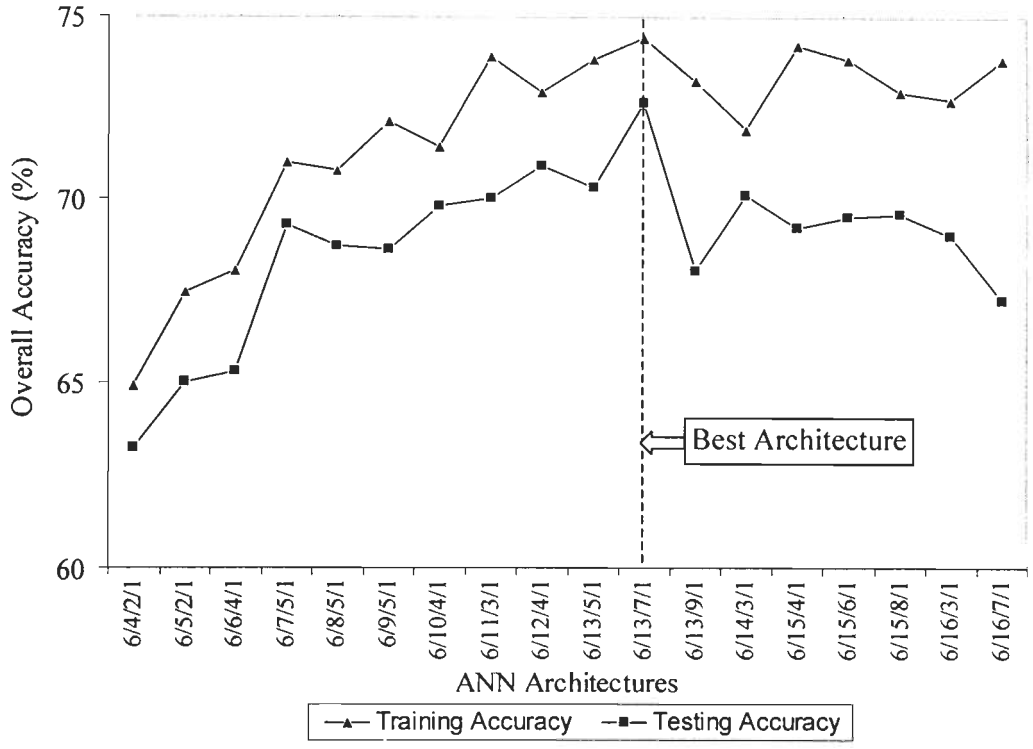


Figure 6.5: Graph showing training and testing accuracies for Case II in ANN black box approach.

LSZ Map Preparation

The neural network outputs of the study area range from 0.062 to 0.993. This range has been categorized into one of the five landslide susceptibility zones to produce the LSZ map, as per the classification criteria given in Table 6.5. The boundaries of various landslide susceptibility zones have been fixed arbitrarily.

Table 6.5: Classification criteria for categorizing neural network output values to various landslide susceptibility zones (ANN black box approach).

Range of values	Landslide Susceptibility Zones
≤0.3	VLS
>0.3 and ≤0.5	LS
>0.5 and ≤0.7	MS
>0.7 and ≤0.9	HS
>0.9	VHS

Using this classification criteria, an LSZ map (referred here as Map II) has been produced and is shown in Figure 6.6. It can be seen from this map that the various susceptibility zones have been distributed all over the area. Thus, the Map II does not show any well defined pattern for the distribution of landslide susceptibility zones. It is also observed from Map II that the VHS and HS zones have represented mostly the 1st and 2nd order drainage buffer areas. It can therefore be inferred that there is a major control of drainage lines on landslide incidence and LSZ mapping, as has been observed in case of LSZ Map I_A. This may be attributed to the fact that the LSZ Map I_A has been used as the reference map for training and testing of neural networks to generate the LSZ Map II. It may also be stated that these results are highly data dependent and may vary from one dataset to another.

Further, the area covered by different landslide susceptibility zones, the area of landslides occupied per class and the landslide densities of different zones have also been determined (Table 6.6).

Table 6.6: Landslide distribution in landslide susceptibility zones of LSZ Map II (ANN black box approach).

Landslide Susceptibility Zones	Area (km^2)	Percent Area (%) (a)	Landslide Area per Class (km^2)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	19.4	7.7	0.022	10.3	1.34
HS	68.4	26.9	0.086	40.4	1.50
MS	90.0	35.4	0.077	36.1	1.02
LS	65.6	25.8	0.027	12.7	0.49
VLS	10.7	4.2	0.001	0.5	0.12

It can be observed from this table that 7.7% of the total area is occupied by VHS zone whereas 26.9%, 35.4%, 25.8% and 4.2% area are occupied by HS, MS, LS and VLS zones respectively. This shows that the area wise coverage of different

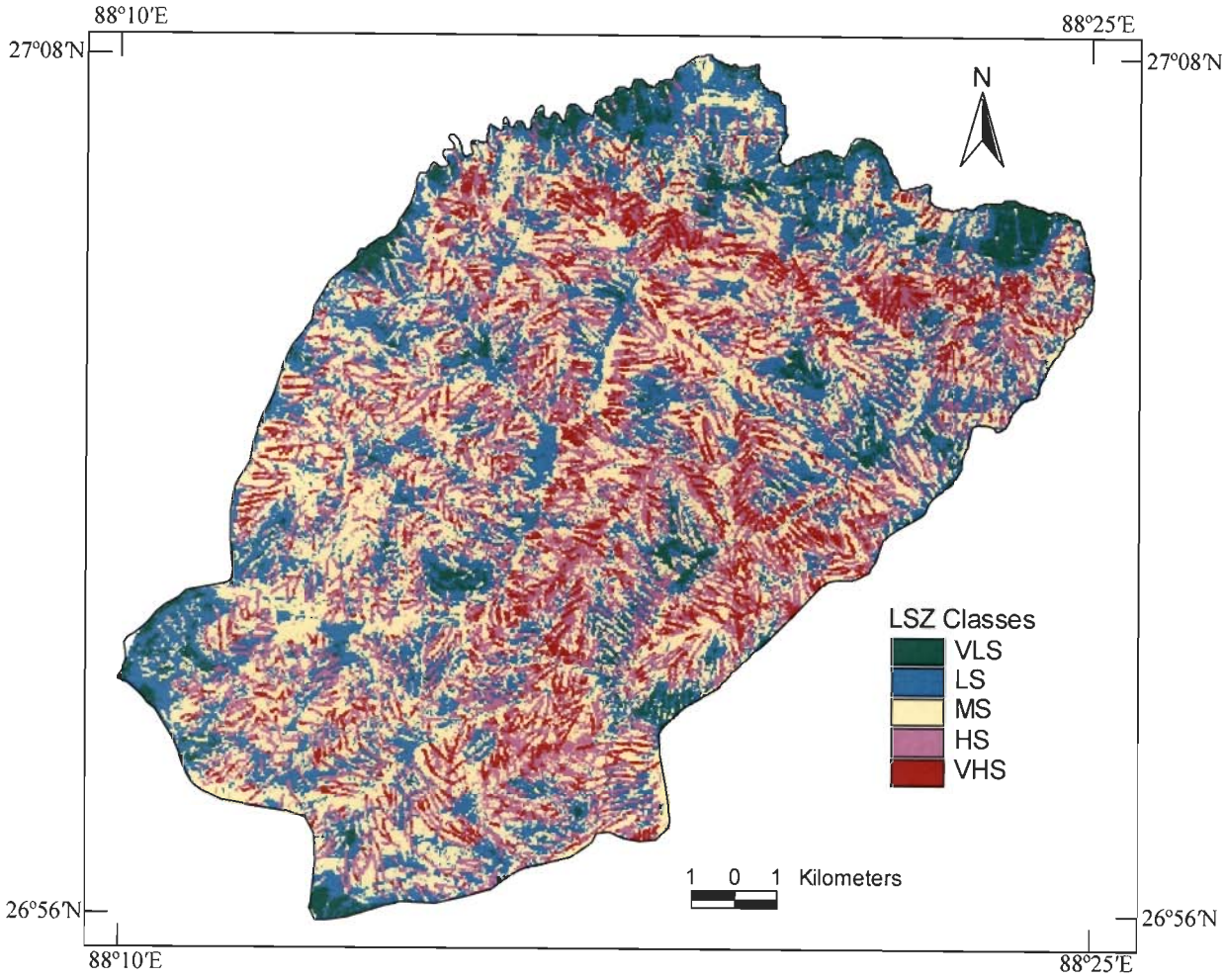


Figure 6.6: LSZ Map II produced from ANN black box approach.

susceptibility zones is normally distributed, which in fact, should be the case. The distribution of landslides in different susceptibility zones has also been compared. It has been found that 10.3% of landslide area have been predicted over VHS zone while 40.4%, 36.2%, 12.7% and 0.5% of landslide area have been predicted over HS, MS, LS and VLS zones respectively. Hence, it can be stated that 34.6% of HS and VHS areas together contain 50.7% of existing landslide areas. This shows that the distribution of landslides over VHS to VLS zones is skewed towards the higher susceptibility zones, which should in fact be the case.

Further, it is observed from Table 6.6 that the landslide densities for VHS and HS zones are almost equal. This demonstrates that the Map II has a systematic and reasonable trend of variation in landslide density values only from HS to VLS zones. For VHS and HS zones combined, the landslide density is 1.46. Hence, it is inferred from Table 6.6 that these higher susceptibility zones do not have much higher landslide density values as compared to other susceptibility zones.

From the analysis of results from ANN black box approach for the two cases, it can be stated that the nature and size of reference data for the output neuron in a neural network influence the training and testing data accuracies. Moreover, as expected there is lot of similarity between the LSZ Map I_A and LSZ Map II.

6.3 LSZ Using Fuzzy Set Based Approach

The limitation of ANN black box approach is that the importance in terms of weights and ratings of the thematic data layers and their categories respectively for LSZ mapping is hidden and thus is not known. This happens to be a key limitation of the ANN black box approach, where the weights and ratings can not be quantified and

therefore the contribution of a particular factor is not known. Alternative approaches are therefore sought. Recently, an ANN connection weight analysis has been proposed, which shall be implemented here (see Section 6.4.1.3). This method however provides information about the weights of the factors only. The contribution of each category of the factor may not be known. The fuzzy set based approach looks attractive to determine the ratings of the categories of thematic data layers (causative factors). Unlike defining crisp ratings to each category, as is done in conventional weighting approach, the fuzzy set theory determines ratings on a 0 to 1 continuous scale thereby providing more realistic values. Based only on the ratings of categories as determined through a fuzzy set based approach and keeping the weights of causative factors constant, an LSZ map of the area has been generated. It is expected that a fuzzy set based LSZ map shall be able to depict more natural variation of landslide susceptibility zones in the area in comparison to conventional weighting approach based map. The flow diagram showing different steps followed in implementation of fuzzy set based approach for LSZ mapping is given in Figure 6.7.

6.3.1 Rating Determination by Cosine Amplitude Method

The cosine amplitude method of fuzzy relation concept (see Section 3.3.2.3 in Chapter 3) has been adopted to determine the ratings of the categories of thematic data layers. The landslide distribution and different categories of thematic layers taken one at a time have been considered as two datasets for the computation of rating or strength of relationship (r_{ij}). The pixels in the landslide areas are assigned a value of 1.0, whereas rests of the pixels are assigned a value of 0.0 in the landslide distribution layer. Similarly, a value of 1 is assigned to a particular category of a thematic layer

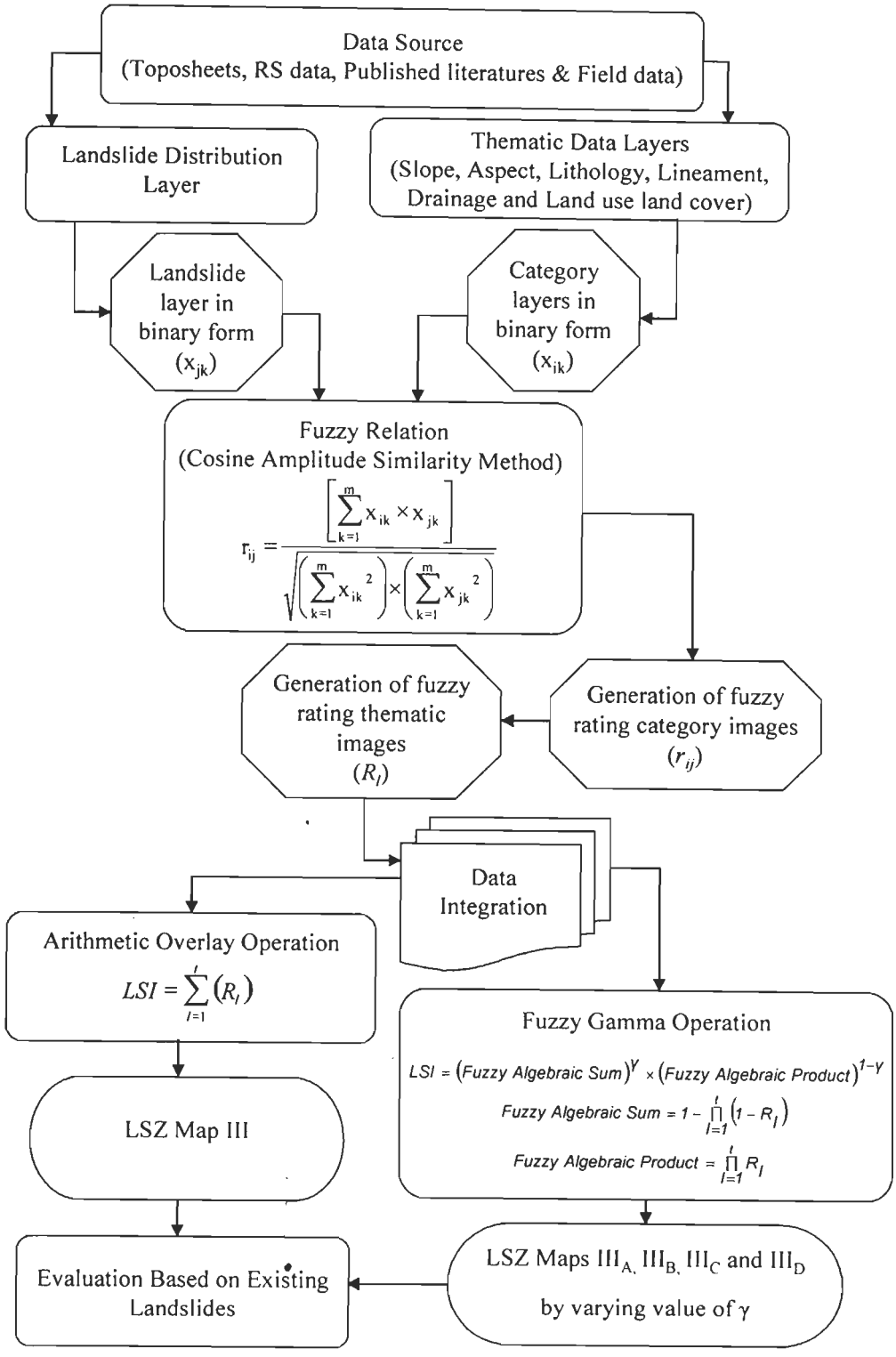


Figure 6.7: Flow diagram showing different steps of fuzzy set based approach.

and a value of 0 to rest of the pixels. Hence, in total 36 data layers in binary form are generated, which correspond to 35 layers of categories of thematic data layers and one layer of landslide distribution. These layers have been used to determine r_{ij} (see Eq. 3.7). The number of pixels and number of landslide pixels in each category of thematic data layers (Table 6.7) have been determined using these 36 binary layers. With the help of these data, ratings for all the 35 categories (r_{ij}) (Table 6.7) have been calculated in MS Excel spreadsheet using the Eq. (3.7).

It can be observed from Table 6.7 that the highest fuzzy rating (0.0638) has been attributed to the barren land category of land use land cover and the lowest (0.0144) to the lineament buffer category given as >500m. The categories namely water bodies, river sand and flat areas have zero fuzzy rating zero. This is true also since usually landslides will not occur in these areas. This also proves the performance of fuzzy set based approach. The 25m buffers along 1st and 2nd order drainage categories have fuzzy ratings of 0.0296 and 0.0399 respectively. Within the lineament buffer categories, 0-125m buffer has the highest rating of 0.0407 and >500m buffer has the lowest rating of 0.0144. This indicates that the probability of landslide occurrences is high at locations that are closer to lineaments. Further, within the slope categories, 35°-45° slopes have the highest rating of 0.0340 and 0°-15° slopes have the lowest rating of 0.0212. This indicates that steeper slopes are more prone to landslide occurrences in the area. Amongst the lithology categories, Reyang quartzites have the highest rating of 0.0416 and Paro gneiss has the lowest rating of 0.0253. In this case, quartzitic rocks have higher ratings than other rocks as they are highly jointed and fractures and more prone to landslide occurrences. Within the land use land cover categories, barren land has the highest rating of 0.0638 and thick and sparse forests have the lowest ratings of 0.0229 and 0.0223. This reflects that the incidence of

Table 6.7: Fuzzy ratings for different categories of causative factors.

Thematic layers (Causative Factors)	Categories	Number of pixels	Number of Landslide pixels	Fuzzy Rating (r_{ij})
Drainage Buffer	25m along 1st order drainage	116168	102	0.0296
	25m along 2nd order drainage	27690	44	0.0399
Lineament Buffer	0-125m along a lineament	146761	243	0.0407
	125-250m along a lineament	108929	35	0.0179
	250-375m along a lineament	72380	36	0.0223
	375-500m along a lineament	41360	17	0.0203
	>500m along a lineament	38317	8	0.0144
Slope	0°-15°	51380	23	0.0212
	15°-25°	146974	117	0.0282
	25°-35°	144495	131	0.0301
	35°-45°	50246	58	0.0340
	>45°	14329	10	0.0264
Lithology	Darjeeling Gneiss	73371	77	0.0324
	Feldspathic Graywacke	45938	61	0.0364
	Paro Gneiss	247242	158	0.0253
	Lingtse Granite Gneiss	20926	15	0.0268
	Paro Quartzite	12154	14	0.0339
	Reyang quartzite	8089	14	0.0416
Land use land cover	Agriculture Land	35692	85	0.0488
	Tea Plantation	142541	84	0.0243
	Thick Forest	72685	38	0.0229
	Sparse Forest	131088	65	0.0223
	Barren Land	14237	58	0.0638
	Habitation	10341	9	0.0295
	Water	970	0	0
	River Sand	1005	0	0
Aspect	Flat	2072	0	0
	North (N)	59880	22	0.0192
	Northeast (NE)	45077	32	0.0266
	East (E)	52868	73	0.0372
	Southeast (SE)	45689	77	0.0411
	South (S)	37630	49	0.0361
	Southwest (SW)	29860	20	0.0259
	West (W)	55132	26	0.0217
	Northwest (NW)	79148	40	0.0225

landslide is inversely related to the vegetation density. Hence, barren slopes are more prone to landslide activity as compared to the forest areas. Amongst the aspect categories, southeast (SE) facing slopes have the highest rating of 0.0411 and north (N) facing slopes have the lowest rating of 0.0192. It is observed from the ratings of aspect categories that the south facing slopes have higher ratings than north facing slopes. This also supports the fact that the south facing slopes have lesser vegetation density as compared to the north facing slopes and hence, the landslide activity is relatively more in former case (Sinha et al., 1975).

These results demonstrate that the fuzzy set based approach has been able to appropriately assign relative importance (ratings) to different categories of thematic data layers in an unbiased and objective manner. This has the advantage over conventional weighting approach, where the ratings are assigned based on the field observation, where field of view is limited and also the ratings are crisp at 0-9 categorical scale.

6.3.2 Data Integration and LSZ Map Preparation

By assigning the ratings of the 35 categories in the corresponding binary layers of categories, 35 images of r_{ij} have been generated. The corresponding r_{ij} images for various categories of a thematic layer have been composited together to generate an R_l image for that thematic layer, where l varies from 1 to t thematic layers corresponding to each causative factor (e.g., 6 thematic layers in the present case). The integration of these 6 thematic layers (R_l images) representing the ratings of the categories (r_{ij}) have been performed using two methods. First method corresponds to the conventional simple arithmetic overlay operation, which is generally applied. Another method is

based on a novel concept of fuzzy combination of rules using fuzzy gamma operator. Thus, via this method, fuzziness in the integration process can also be introduced.

6.3.2.1 Data Integration Using Arithmetic Overlay Operation

The integration of 6 different R_l images has been performed using the following equation to obtain the LSI for each pixel of the dataset.

$$LSI = \sum_{l=1}^{t} (R_l) \tag{6.1}$$

where t is the number of R_l images (i.e., 6 images in the present case).

The LSI values for the whole area have been found to lie between 0.014 and 0.252. This range of LSI values has been divided into five different susceptibility zones with boundaries located at ($\mu_o - 1.5m\sigma_o$), ($\mu_o - 0.5m\sigma_o$), ($\mu_o + 0.5m\sigma_o$) and ($\mu_o + 1.5m\sigma_o$) values where observed mean (μ_o) is 0.150, standard deviation (σ_o) is 0.024 and m is a positive, non-zero value (Saha et al., 2005). This classification is adopted to fix the boundaries of classes statistically and also, to avoid the subjectivity in arbitrarily selecting the boundaries of classes. Several LSZ maps have been prepared and different success rate curves (as described in conventional weighting approach in Chapter 4) have been plotted for various values of m. Five representative success rate curves corresponding to m = 0.8, 1.0, 1.1, 1.2 and 1.4 are shown in Figure 6.8. The suitability of any LSZ map can be judged from the fact that more percentage of landslides should occur in very high susceptibility zone as compared to other zones. It can be observed from Figure 6.8 that for 10% of the area in very high susceptibility zone, the curves corresponding to m = 0.8, 1.0, 1.1, 1.2 and 1.4 show the landslide

occurrences of 42.7%, 47.7%, 48.6%, 47.3% and 41.8% respectively. Hence, for the first 10% area, the curve corresponding to m=1.1 shows the highest success rate and the corresponding LSZ map appears to be the most appropriate one. Accordingly, putting the values of μ_o as 0.150, σ_o as 0.024 and m as 1.1, the landslide susceptibility zone boundaries have been fixed at LSI values of 0.111 ($\mu_o - 1.5m\sigma_o$), 0.137 ($\mu_o - 0.5m\sigma_o$), 0.163 ($\mu_o + 0.5m\sigma_o$) and 0.189 ($\mu_o + 1.5m\sigma_o$). The five different susceptibility zones thus categorized along with the range of LSI values are given in Table 6.8.

Table 6.8: Classification of LSI values into landslide susceptibility zones (using arithmetic overlay operation in fuzzy set based approach).

Landslide Susceptibility Zones	Range of LSI Values
VLS	$0.014 \leq LSI \leq 0.111$
LS	$0.111 < LSI \leq 0.137$
MS	$0.137 < LSI \leq 0.163$
HS	$0.163 < LSI \leq 0.189$
VHS	$0.189 < LSI \leq 0.252$

The LSZ map (referred here as Map III) thus produced is shown in Figure 6.9. The visual inspection of the LSZ Map III depicts an overall NNE-SSW landslide susceptibility zonation trend in the area. It can be observed that the southeast and east facing slopes are more susceptible to landslides than other slopes. Hence, it can be stated that there is a topographic control over this LSZ map. It is also observed from this map that the VHS and HS zones have represented mostly the 1st and 2nd order drainage buffer areas. It can therefore be inferred that there is also a control of drainage lines on landslide incidence and LSZ mapping.

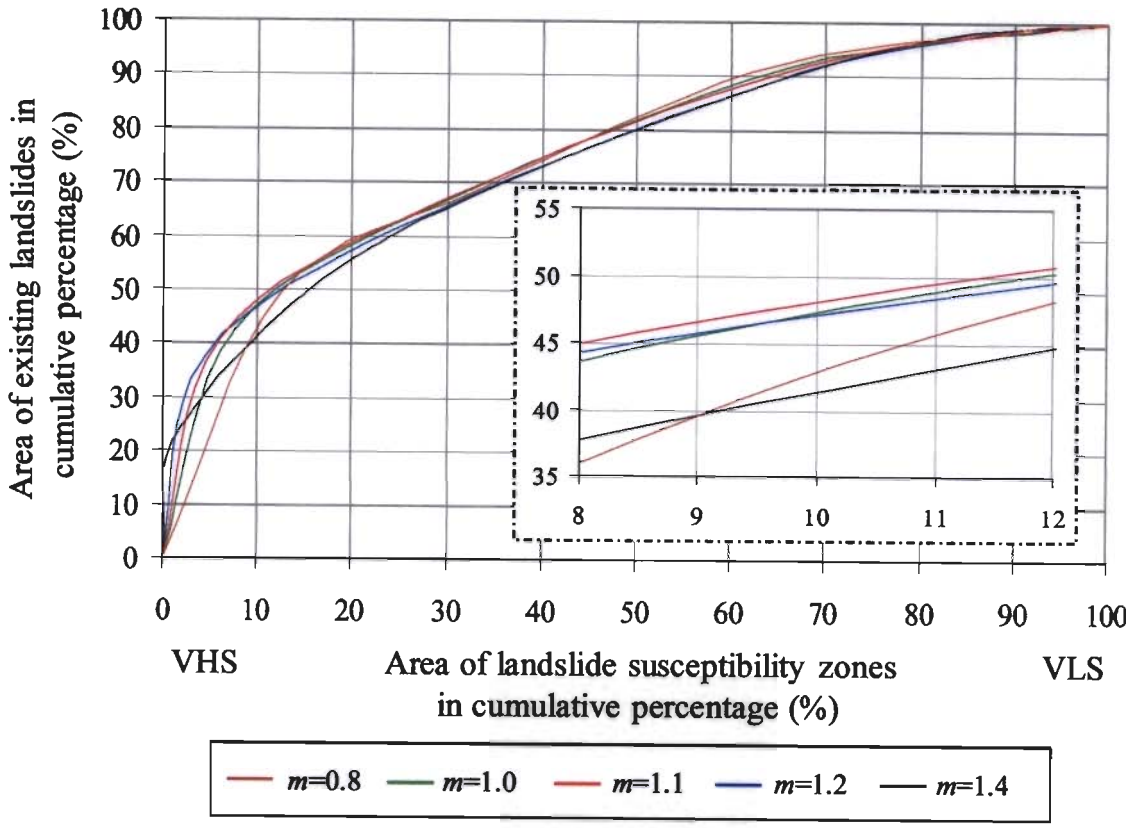


Figure 6.8: Success rate curves for choosing the best segmentation in LSI values for LSZ classification in fuzzy set based approach (arithmetic overlay operation for thematic layers integration).

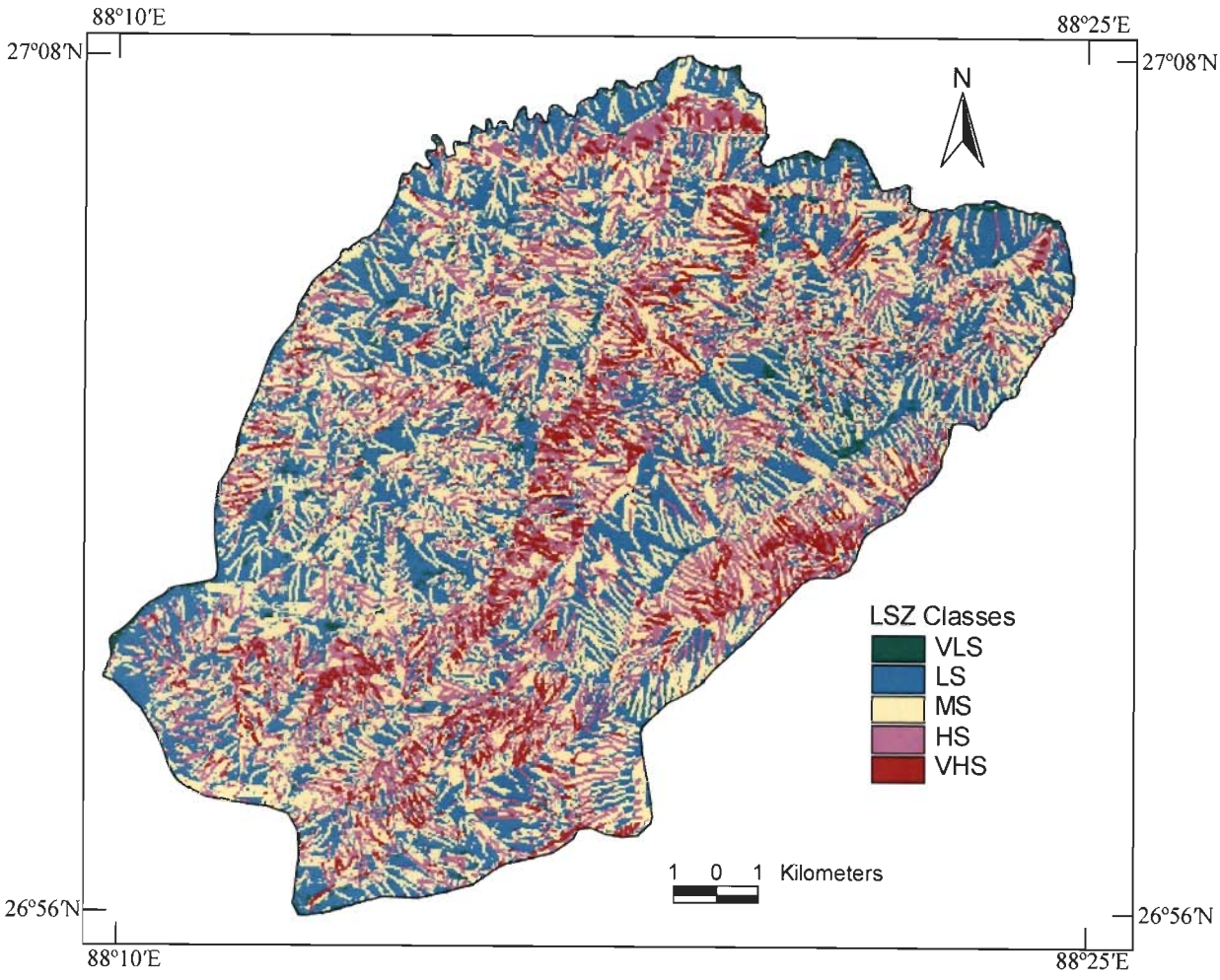


Figure 6.9: LSZ Map III produced from fuzzy set based approach (arithmetic overlay operation for thematic layers integration).

The area covered by different landslide susceptibility zones, the area of landslides occupied per class and the landslide densities of different zones have been determined and are provided in Table 6.9.

Table 6.9: Landslide distribution in landslide susceptibility zones of LSZ Map III using arithmetic overlay operation (fuzzy set based approach).

Landslide Susceptibility Zones	Area (km^2)	Percent Area (%) (a)	Landslide Area per Class (km^2)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	15.4	6.1	0.087	41.0	6.72
HS	57.9	22.7	0.053	25.1	1.10
MS	100.3	39.4	0.055	25.9	0.66
LS	77.6	30.4	0.017	8.0	0.26
VLS	3.6	1.4	0.000	0.0	0.00

It can be observed from Table 6.9 that 6.1% of the total area is occupied by VHS zone while 22.7%, 39.4%, 30.4% and 1.4% areas are occupied by HS, MS, LS and VLS zones respectively. This shows that the area wise coverage of different susceptibility zones is normally distributed, which should be the case. The distribution of landslides in different susceptibility zones has been compared. It is found that 41.0% of landslide area is predicted over VHS zone while 25.1%, 25.9%, 8.0% and 0.0% of landslide area are predicted over HS, MS, LS and VLS zones respectively. These results show that 28.8% area of VHS and HS zones could predict 66.1% landslide area, which indicate that the distribution of landslides over VHS to VLS zones is skewed towards the higher susceptibility zones. This should indeed be the case.

Further, it is observed from Table 6.9 that the landslide density for VHS zone (i.e., 6.72) is much higher than those for other LS zones. The Map III has a systematic trend in variation of landslide density values from VHS to VLS zones. Hence, this

corroborates the belief that LSZ Map III produced from fuzzy set based approach represents a reasonably accurate LSZ map of the study area.

6.3.2.2 Data Integration Using Fuzzy Gamma Operators

Generally, the integration is performed based on simple arithmetic operations. These have limitations in the sense that the integration takes place in a crisp form, as the categories of the thematic layers are nominal. A more suitable approach may be to introduce fuzziness in the integration process also so as to depict the real situations on the ground. Fuzzy set theory suggests a variety of fuzzy operators that may be employed to combine different thematic layers once the fuzzy membership values of the categories have been defined on some basis.

In the present study, the integration of different thematic layers for the generation of LSZ map has been performed using fuzzy gamma operator. The goal is to investigate the contribution and interaction of membership values of thematic maps in the integration process.

In this process, the integration of 6 different R_I images has been performed. The membership value for each pixel in an R_I image represents the rating of the category (r_{ij}) of the corresponding thematic data layer. The integrated image represents the combined membership values, termed as Landslide Susceptibility Index (LSI) for each pixel of the dataset. The process is based on determination of fuzzy algebraic product, fuzzy algebraic sum and fuzzy gamma operators using the following relations,

$$LSI = (Fuzzy\ Algebraic\ Sum)^{\gamma} \times (Fuzzy\ Algebraic\ Product)^{1-\gamma} \qquad (6.2)$$

$$\text{Fuzzy Algebraic Product} = \prod_{I=1}^{t} R_I \tag{6.3}$$

$$\text{Fuzzy Algebraic Sum} = 1 - \prod_{I=1}^{t} (1 - R_I) \tag{6.4}$$

where gamma (γ) is a parameter chosen in the range (0,1). When γ is 1, the LSI is equal to the fuzzy algebraic sum and when γ is 0, the LSI equals the fuzzy algebraic product.

Initially, random pixels have been selected with each pixel representing 6 different membership values (R_I) for 6 different thematic layers. With different gamma values varying from 0 to 1, the combined membership values as per Eq.(6.2) for all random pixels have been determined and plotted against the gamma values to select a suitable gamma value for integration. The output for one such pixel is explained in Figure 6.10. The minimum and maximum average gamma values obtained are 0.79 and 0.94 respectively. Thus, the gamma value >0.94 has an increasive effect of the fuzzy algebraic sum and the gamma value <0.79 has a decreasive effect of the fuzzy algebraic product. A judicious gamma value should lie between 0.79 and 0.94 to ensure a compromise between the increasive and decreasive tendencies.

Thus, using gamma values of 0.79 and 0.94, the LSI for all the pixels of the dataset have been determined using Eq. (6.2). The range of LSI values has been categorized into five susceptible zones such as VHS, HS, MS, LS and VLS with boundaries fixed by using success rate curves method (see Section 6.3.2.1). The most appropriate LSZ maps for gamma values of 0.79 (LSZ Map III_A) and 0.94 (LSZ Map III_B) have been selected using the success rate curves method for further

interpretation. The area covered by different landslide susceptibility zones, the area of landslides occupied per class and the landslide densities of different zones have been determined (Table 6.10).

Table 6.10: Landslide distribution in landslide susceptibility zones of LSZ Maps III_A and III_B using fuzzy Gamma operation (fuzzy set based approach).

Landslide Susceptibility zones	LSZ Map III_A $\gamma = 0.79$			LSZ Map III_B $\gamma = 0.94$		
	Percent Area (%) (a)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)	Percent Area (%) (a)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	10.1	24.2	2.40	4.9	19.2	3.92
HS	24.4	18.9	0.77	25.3	40.4	1.60
MS	8.7	32.1	3.69	32.6	31.3	0.96
LS	56.8	24.8	0.44	36.6	9.1	0.25
VLS	0	0	0.00	0.6	0	0.00

It is observed from Table 6.10 that the area wise coverage of different susceptibility zones is not normally distributed in case of Map III_A, whereas it is almost normally distributed in case of Map III_B. Further, it is observed that 34.5% area of VHS and HS zones could predict 43.1% landslide area in case of LSZ Map III_A, whereas 30.2% area of VHS and HS zones could predict 59.6% landslide area in case of LSZ Map III_B. Also, in case of Map III_A, MS zone shows maximum landslide density (i.e., 3.69). However, the landslide density values do not have a successively decreasing trend from VHS to VLS zones. But, in case of Map III_B, VHS zone shows maximum landslide density (i.e., 3.92) and the landslide density values show a successively decreasing trend from VHS to VLS zones which should be the case.

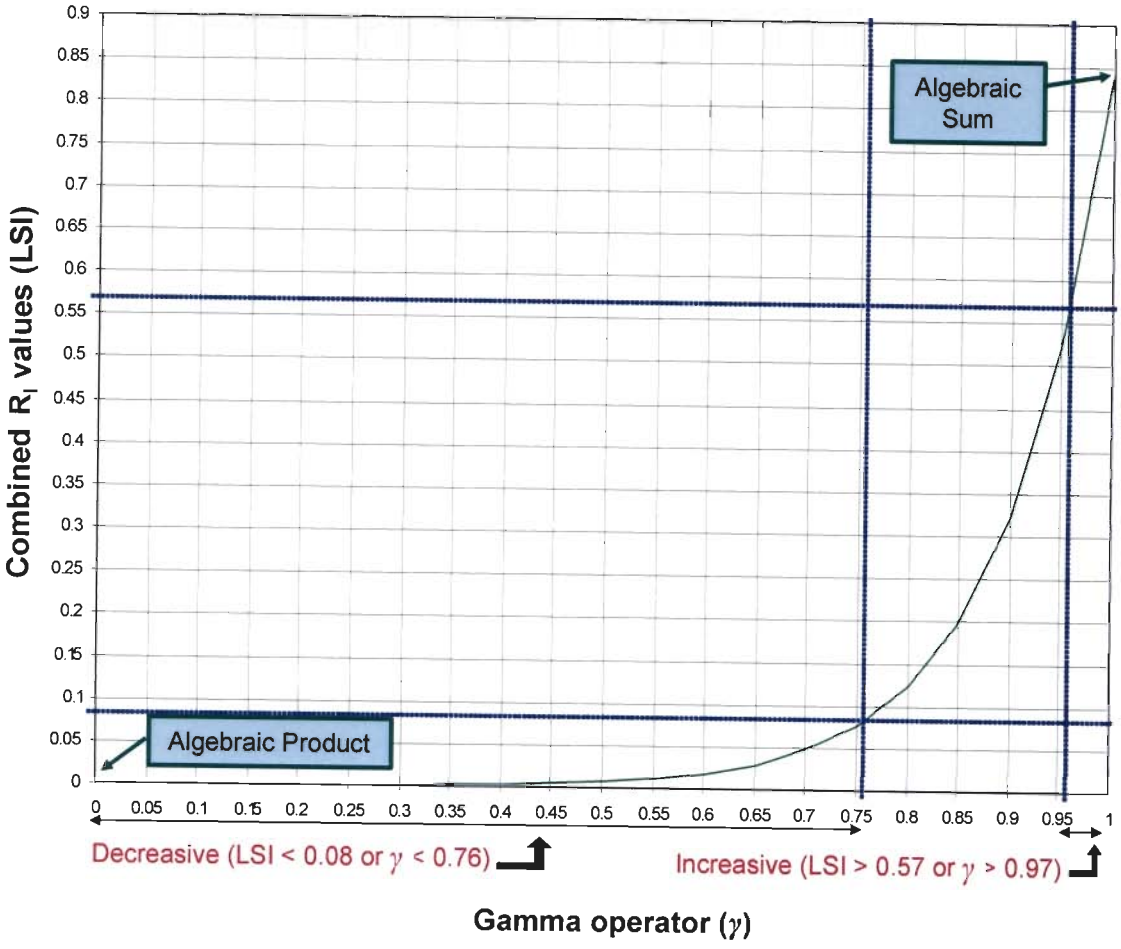


Figure 6.10: A graph showing combined membership values against different Gamma (γ) values for a pixel using fuzzy Gamma operation in fuzzy set based approach.

From these results, it is inferred that the gamma value of 0.94 has given comparatively better results in terms of landslide occurrences in different susceptibility zones than that of 0.79. These results further give an insight that the use of higher gamma values may also produce more accurate map.

Therefore, higher gamma values of 0.95 and 0.97 have also been considered to generate the LSZ maps. The LSZ maps corresponding to gamma values of 0.95 and 0.97 have been referred here as LSZ Maps III_C and III_D. The area covered by different landslide susceptibility zones, the area of landslides occupied per class and the landslide densities of different zones have also been determined (Table 6.11).

Table 6.11: Landslide distribution in landslide susceptibility zones of LSZ Maps III_C and III_D using fuzzy Gamma operation (fuzzy set based approach).

Landslide Susceptibility zones	LSZ Map III_C $\gamma = 0.95$			LSZ Map III_D $\gamma = 0.97$		
	Percent Area (%) (a)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)	Percent Area (%) (a)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	2.3	15.0	6.52	2.3	15.1	6.56
HS	24.7	41.9	1.70	23.7	45.4	1.91
MS	43.0	36.0	0.84	45.2	33.3	0.74
LS	29.6	7.1	0.24	28.4	6.2	0.22
VLS	0.4	0.0	0.00	0.4	0.0	0.00

It is observed from Table 6.11 that the area wise coverage of different susceptibility zones is normally distributed in both Maps III_C and III_D. Further, it is observed that 27% area of VHS and HS zones is able to predict 56.9% landslide area in case of LSZ Map III_C, whereas 26% area of VHS and HS zones could predict 60.5% landslide area in case of LSZ Map III_D. In both the Maps III_C and III_D, VHS zone shows maximum landslide density (i.e., 6.52 for Map III_C and 6.56 for Map III_D)

and the landslide density values show a successively decreasing trend from VHS to VLS zones. However, LSZ Map III_D shows a slightly higher landslide density than that of Map III_C and also the differences/separations in landslide density values of different LS zones are slightly higher in case of Map III_D than Map III_C. From these results, it is inferred that the LSZ Map III_D with gamma value of 0.97 has yielded better result than the LSZ maps with lower gamma values.

The LSZ map III_D is shown in Figure 6.11. The visual inspection of the LSZ Map III_D depicts an overall NNE-SSW zonation trend in the area. The southeast and east facing slopes are more susceptible to landslides than other slopes. Hence, there is a topographic control in this LSZ map also. It is also observed that the VHS and HS zones have represented mostly the 1^{st} and 2^{nd} order drainage buffer areas. It can be inferred that there is also a control of drainage lines on landslide incidence and LSZ mapping.

The integration of thematic maps (layers) for LSZ mapping using fuzzy gamma operator indicates that as the gamma value increases, the accuracy of final integrated product (i.e., LSZ map) also increases. In the present study, the LSZ map prepared after integration of various thematic maps with fuzzy gamma operator having the highest value has reflected a more real situation in terms of landslides in the study area. Hence, the fuzzy gamma operator can be a suitable approach to introduce fuzziness in the integration process so as to depict the real physical situations on the ground.

The comparative analysis also shows that there is a lot of similarity between LSZ Map III (produced by fuzzy set based approach using arithmetic overlay integration process) and Map III_D (produced by fuzzy set based approach using fuzzy gamma integration process). However, the comparison of landslide density values of

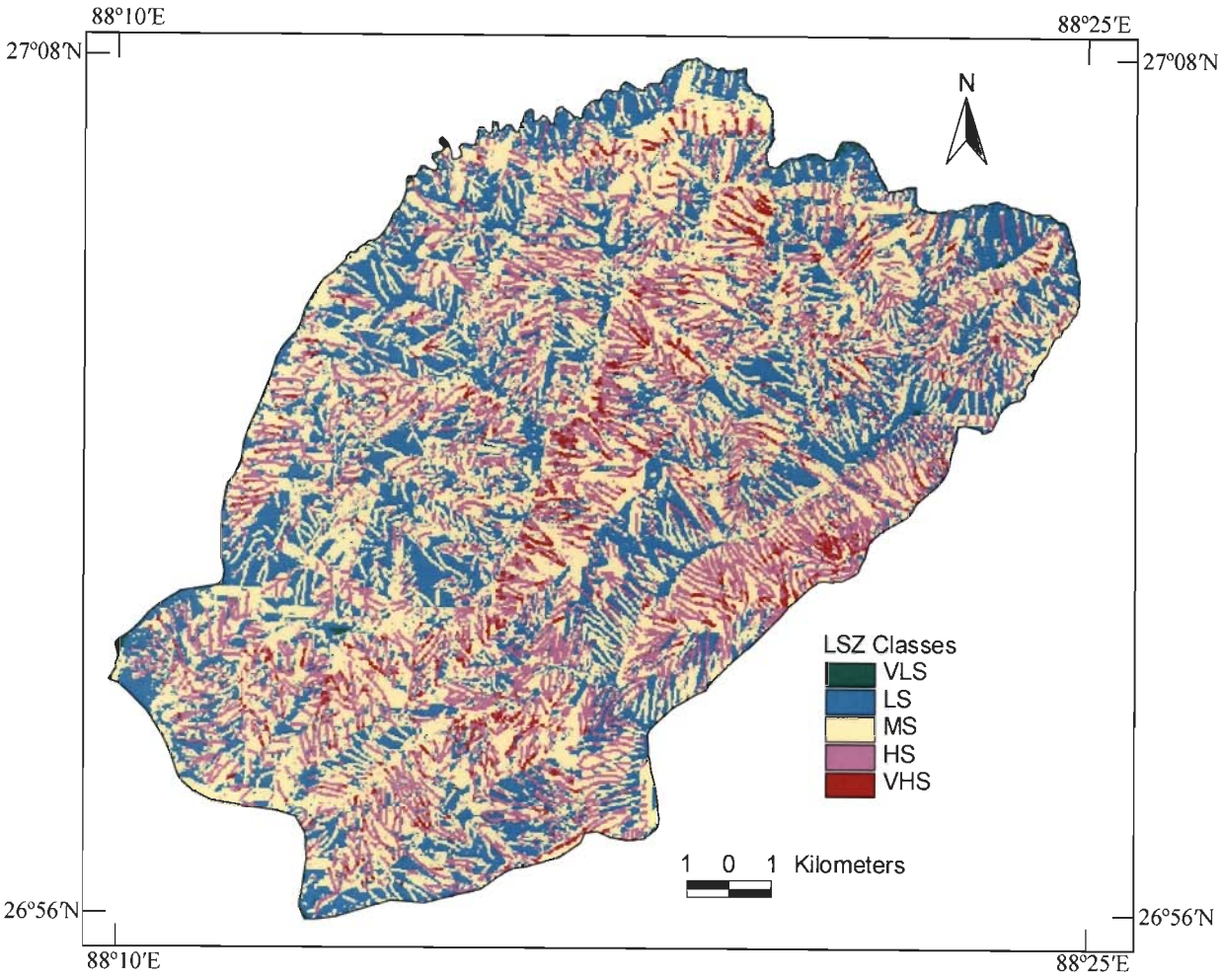


Figure 6.11: LSZ Map III_D produced from fuzzy set based approach (fuzzy Gamma operation for thematic layers integration with $\gamma = 0.97$).

these LSZ maps shows that the landslide density of VHS zone in case of Map III (i.e., 6.72) is marginally better than that in case of Map III_D (i.e., 6.56). Therefore, the LSZ Map III has later been selected used for its comparative evaluation with other LSZ maps prepared from different approaches.

6.4 LSZ Using Combined Neural and Fuzzy Approach

In conventional weighting approach, the weights and ratings have been assigned based on experts' opinions and the field situations, where field of view is limited. However, in real field situations, the relative weights and ratings of factors and their categories may differ. Thus, this approach induces lot of subjectivity in weight and rating assignment process, though it is very widely used.

The ANN black box approach is one of the advanced objective approaches for LSZ mapping. But, in this case, the weights are hidden as the connection weights between different layers of the neural network have not been opened up and analyzed, as observed in this study (see Section 6.2) also.

In the fuzzy set based approach, only the ratings of the categories of factors have been determined using cosine amplitude method. In this approach, the weights of the factors were considered unity thereby giving equal importance to all the factors. However, in real situations on the ground, different causative factors may also have varying degree of importance (i.e., weights) in initiating a landslide activity in a region.

Hence, to overcome these limitations, a novel approach named as ANN connection weight analysis has been implemented here to determine the weights of the causative factors. These weights have been combined with ratings obtained

through fuzzy set based approach. Thus, a combined neural and fuzzy approach has been developed for LSZ mapping in this study.

The combined neural and fuzzy approach involves three main steps:

i) Determination of weights of thematic layers through ANN connection-weight approach

ii) Determination of ratings for categories of thematic layers using fuzzy set based approach

iii) Integration of ratings and weights using GIS to produce the LSZ map.

The methodology for LSZ using this approach is shown in Figure 6.12.

6.4.1 ANN Implementation

6.4.1.1 ANN Architecture

A schematic representation of ANN architecture, adopted in this case, is given in Figure 6.13. The ANN architectures with one input layer, two hidden layers and one output layer have been considered. The input layer contains 6 neurons corresponding to 6 different causative factors and the output layer contains a single neuron corresponding to the presence or absence of existing landslide. The data supplied to input neurons correspond to the normalized ratings (r_{ij}) of the categories of different thematic layers, as given in Table 6.12. The desired output data supplied to output neuron corresponds to 1 or 0 for the presence or absence of existing landslides. Initially, a few ANN architectures were designed by varying the number of neurons in hidden layers. These networks were trained and tested to evaluate their performance. Subsequently, the number of networks was increased and a total of 100

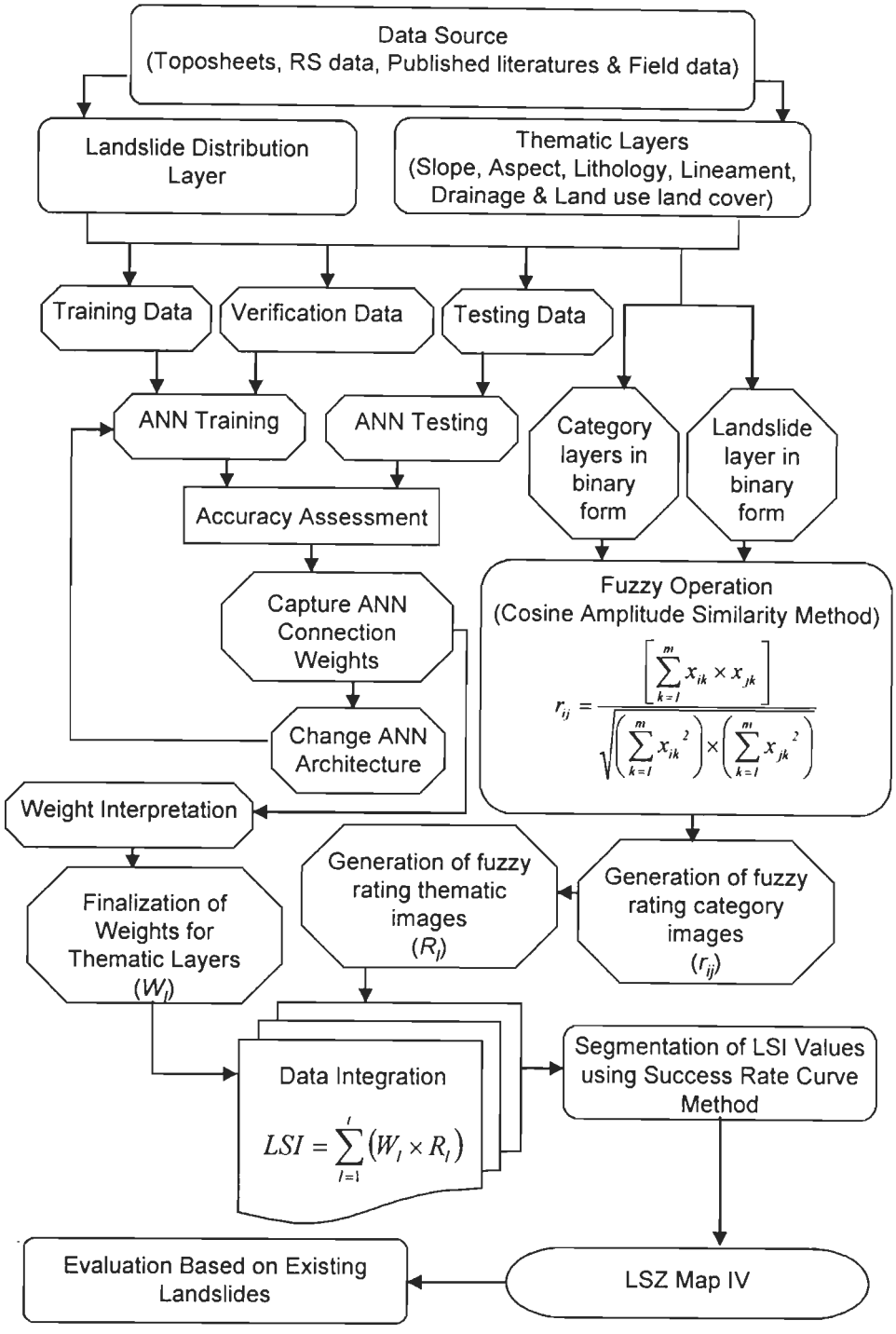


Figure 6.12: Flow diagram showing different steps of combined neural and fuzzy approach.

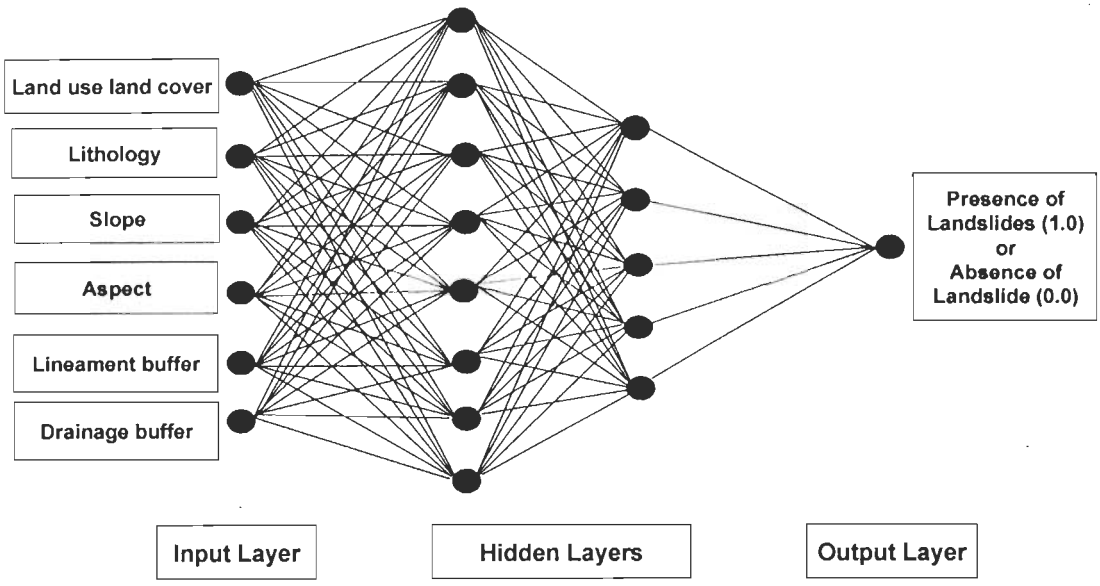


Figure 6.13: A schematic diagram of ANN architecture in combined neural and fuzzy approach.

network architectures were designed, trained and tested to evaluate their generalization capabilities and accuracies in the present case.

6.4.1.2 Data Preparation

Three independent training, testing and verification datasets are formed. Each dataset consists of 226 mutually exclusive pixels, which correspond to 113 existing landslide pixels and remaining 113 no-landslide pixels. The pixels in all the three datasets were mutually exclusive. The verification dataset has been used to control the overtraining of the networks.

6.4.1.3 Training and Testing of ANN

The Levenberg-Marquardt back-propagation algorithm is used to train the neural networks. The training parameters used in this case is already given in Table 6.2. The

Table 6.12: Normalized fuzzy ratings for different categories of causative factors used as input attributes for ANN implementation in combined neural and fuzzy Approach.

Thematic layers (Causative factors)	Categories	Fuzzy Ratings (r_{ij})	Normalized Fuzzy Ratings
Drainage Buffer	25m along 1st order drainage	0.0296	0.4259
	25m along 2nd order drainage	0.0399	0.5741
Lineament Buffer	0-125m	0.0407	0.3521
	125-250m	0.0179	0.1548
	250-375m	0.0223	0.1929
	375-500m	0.0203	0.1756
	>500m	0.0144	0.1246
Slope	0°-15°	0.0212	0.1515
	15°-25°	0.0282	0.2016
	25°-35°	0.0301	0.2151
	35°-45°	0.0340	0.2430
	>45°	0.0264	0.1887
Lithology	Darjeeling Gneiss	0.0324	0.1650
	Feldspathic Graywacke	0.0364	0.1853
	Paro Gneiss	0.0253	0.1288
	Lingtse Granite Gneiss	0.0268	0.1364
	Paro Quartzite	0.0339	0.1726
	Reyang quartzite	0.0416	0.2118
Land use land cover	Agriculture Land	0.0488	0.2306
	Tea Plantation	0.0243	0.1148
	Thick Forest	0.0229	0.1082
	Sparse Forest	0.0223	0.1054
	Barren Land	0.0638	0.3015
	Habitation	0.0295	0.1394
	Water	0	0
	River Sand	0	0
Aspect	Flat	0	0
	N	0.0192	0.0834
	NE	0.0266	0.1155
	E	0.0372	0.1615
	SE	0.0411	0.1785
	S	0.0361	0.1567
	SW	0.0259	0.1125
	W	0.0217	0.0942
	NW	0.0225	0.0977

training process is initiated by assigning arbitrary initial connection weights, which are constantly updated until an acceptable training accuracy is reached. The adjusted weights obtained from the trained network have been subsequently used to process the testing data to evaluate the generalization capability and accuracy of the network. The overall training accuracies observed for all the 100 networks are of the order of 75% to 90%, whereas the testing accuracies are of the order of 60% to 70%. The connection weights thus captured for each of the 100 networks are further analyzed to determine the weights for the causative factors.

6.4.1.4 Weight Determination of Factors by ANN Connection Weight Approach

The connection weights of the neurons from input-hidden, hidden-hidden and hidden-output layers for each of the 100 networks are further analyzed to determine the weights of the causative factors for each neural network. Thus, three different weight matrices are obtained for the connection weights from input-hidden, hidden-hidden and hidden-output layers of each network. Simple matrix multiplication has been performed on these weight matrices to obtain a final 6×1 weight matrix for each network which represents the weights of six causative factors in this study. These causative factors are ranked according to the corresponding absolute weights for each network (Appendix-I), which means higher is the value of absolute weight of a factor, more crucial is that factor for the occurrence of landslide. Considering all the 100 networks, the rank of a factor is decided based on the rank observed by maximum number of networks (majority rule). Out of 100 networks, 41 networks categorize lithology as rank 1 (most important), 31 networks for lineament as rank 2, 30 networks for slope as rank 3, 27 networks for aspect as rank 4, 33 networks for land use land cover as rank 5 and 49 networks for drainage as rank 6 (least important).

These results are summarized in Table 6.13. Subsequently, the normalized average of the weights of these networks at a scale of 0 to 10 for a particular factor is calculated and assigned as the weight of that factor (W_i) for preparation of LSZ map. The weights thus obtained through ANN connection weight approach for all the factors are listed in Table 6.14.

Table 6.13: Ranks of factors based on majority rule in combined neural and fuzzy approach (the entries in the matrix represent the number of networks categorizing a factor to a particular rank and the rank corresponding to the maximum number of networks for a factor (underlined) represents the final rank of that factor).

Factors	Number of networks						Final rank (Majority rule)
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	
Land use land cover	1	8	10	22	<u>33</u>	26	5
Lithology	<u>41</u>	21	12	10	10	6	1
Slope	23	24	<u>30</u>	9	9	5	3
Aspect	13	15	22	<u>27</u>	12	11	4
Drainage buffer	0	1	2	17	31	<u>49</u>	6
Lineament buffer	22	<u>31</u>	24	15	5	3	2

Table 6.14: Weights of factors derived through ANN (combined neural and fuzzy approach).

Factors	ANN derived weights
Lithology	4.8
Lineament Buffer	2.1
Slope	1.3
Aspect	1.1
Land use land cover	0.5
Drainage buffer	0.2

It can be observed from Table 6.14 that lithology is the most important (weight = 4.8) and drainage buffer is the least important (weight = 0.2) causative

factors for landslide occurrences in the area. Lineament buffer shows the 2nd rank with a weight of 2.1, whereas slope shows the 3rd rank with a weight of 1.3, aspect shows the 4th rank with a weight of 1.1 and land use land cover has the 5th rank with a weight of 0.5. The weights determined through ANN (Table 6.14) and the weights assigned in conventional weighting approach (Table 5.1) have also been compared. In conventional weighting approach, as believed to be important through field observations, drainage buffer was considered as the most important (weight = 9) and aspect as the least important (weight = 1) causative factors. Lithology was considered as 4th from top in order of importance (weight = 6) in this case. In the field due to limited field of view, only drainage lines could be seen. Considering the distribution of existing landslides along drainage lines, we initially considered drainage to be the most important factor for LSZ. However, the lineaments mark many drainage lines and the ANN derived weights reveal that the real feature of importance is the lineament buffer, which could not be ascertained from conventional approach. Further, the importance of lithology which was not so obvious from field data has been found to be the most important factor for LSZ mapping.

Thus, it can be inferred that in conventional weighting approach, the weights are assigned based on the field observation, where field of view is limited. The fully objective approach (combined neural and fuzzy) could bring out the relative importance (weights and ratings) of different thematic data layers in an unbiased manner.

6.4.2 Data Integration and LSZ Map Preparation

The weights for factors determined through ANN and the ratings for the categories determined through fuzzy set based approach have been considered in the

integration process to prepare the LSZ map. The integration of 6 thematic data layers representing the ratings for the categories (R_l) of the factors and weights for the factors (W_l) has been performed by using simple arithmetic overlay operation in GIS. Hence, this approach has been named here as combined neural and fuzzy approach. The LSI for each pixel of the study area is obtained by using the following equation:

$$LSI = \sum_{l=1}^{t} (W_l \times R_l) \tag{6.5}$$

where t is the number of thematic data layers (i.e., 6 causative factors in this case).

The LSI values have been found to lie in the range from 0.030 to 0.408. The range of LSI values has been categorized into five susceptible zones such as VHS, HS, MS, LS and VLS with boundaries fixed by using success rate curves method (see Section 6.3.2.1). The observed mean (μ_o) and standard deviation (σ_o) from the probability distribution curve of these LSI values are 0.276 and 0.032 respectively. Using these values, several LSZ maps of the study area have been prepared for different values of m. Five representative success rate curves corresponding to m = 1.2, 1.3, 1.4, 1.5 and 1.6 are shown in Figure 6.14. It can be seen that for 10% of the area in VHS zone the curves corresponding to m = 1.2, 1.3, 1.4, 1.5 and 1.6 show the landslide occurrences of 43.9%, 45.6%, 46.7%, 43.3% and 43.9% respectively. Hence, for the first 10% area, the curve corresponding to m=1.4 has the highest success rate. Based on this analysis, the LSZ map corresponding to m = 1.4 appears to be the most appropriate one for the study area. Accordingly, putting the values of μ_o as 0.276, σ_o as 0.032 and m as 1.4, the landslide susceptibility zone boundaries have been fixed at LSI values of 0.208 ($\mu_o - 1.5m\sigma_o$), 0.253 ($\mu_o - 0.5m\sigma_o$), 0.299 ($\mu_o +$

$0.5m\sigma_o$) and 0.344 ($\mu_o + 1.5m\sigma_o$). Hence, the five different susceptibility zones have been categorized with the range of LSI values as given in Table 6.15.

Table 6.15: Classification of LSI values into landslide susceptibility zones (using combined neural and fuzzy approach).

Landslide Susceptibility Zones	Range of LSI Values
VLS	$0.030 \leq LSI \leq 0.208$
LS	$0.208 < LSI \leq 0.253$
MS	$0.253 < LSI \leq 0.299$
HS	$0.299 < LSI \leq 0.344$
VHS	$0.344 < LSI \leq 0.408$

The LSZ map (referred here as Map IV) thus produced is shown in Figure 6.15. The map reflects a preferential distribution of higher landslide susceptibility zones along structural discontinuities (lineaments), which should indeed be the case. Owing to landslide susceptibility of the terrain, the buffer zones of lineaments ought to leave some traces, called as 'ghost-effect' (Saha et al., 2005), on the LSZ map. These 'ghost effects' are clearly indicated by the VHS and HS zones in the N and SE parts of the area. Also, the Darjeeling gneiss rock type in south-eastern part, feldspathic greywacke and Reyang quartzite in the northern part of the study area have clearly indicated moderate to very high susceptibility zones. Most of the lineaments up to 125m buffer zone in these rock types have indicated high and very high susceptibility zones. Hence, it depicts the importance of lithology (i.e., rock types) as well as lineaments on the LSZ map.

The area covered by different landslide susceptibility zones, the area of landslides occupied per class and the landslide densities of different zones have been determined (Table 6.16).

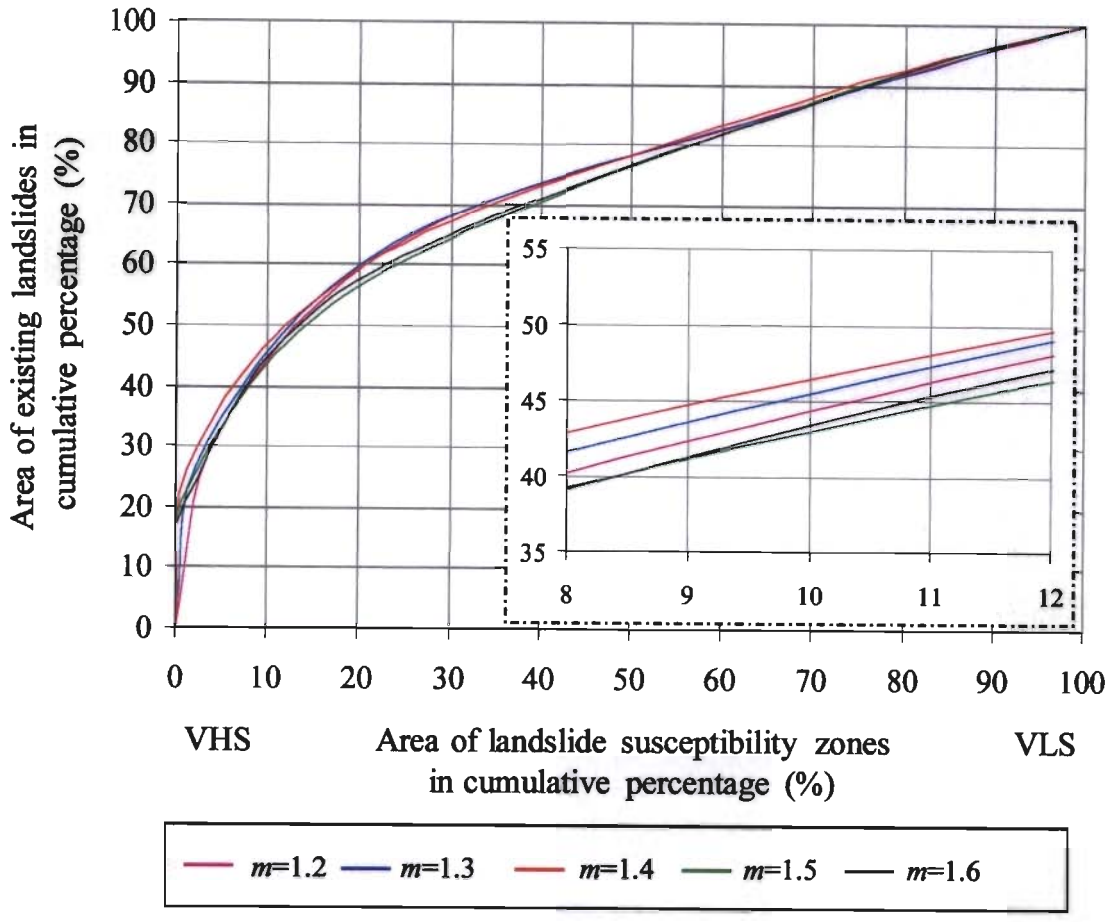


Figure 6.14: Success rate curves for choosing the best segmentation in LSI values for LSZ classification in combined neural and fuzzy approach.

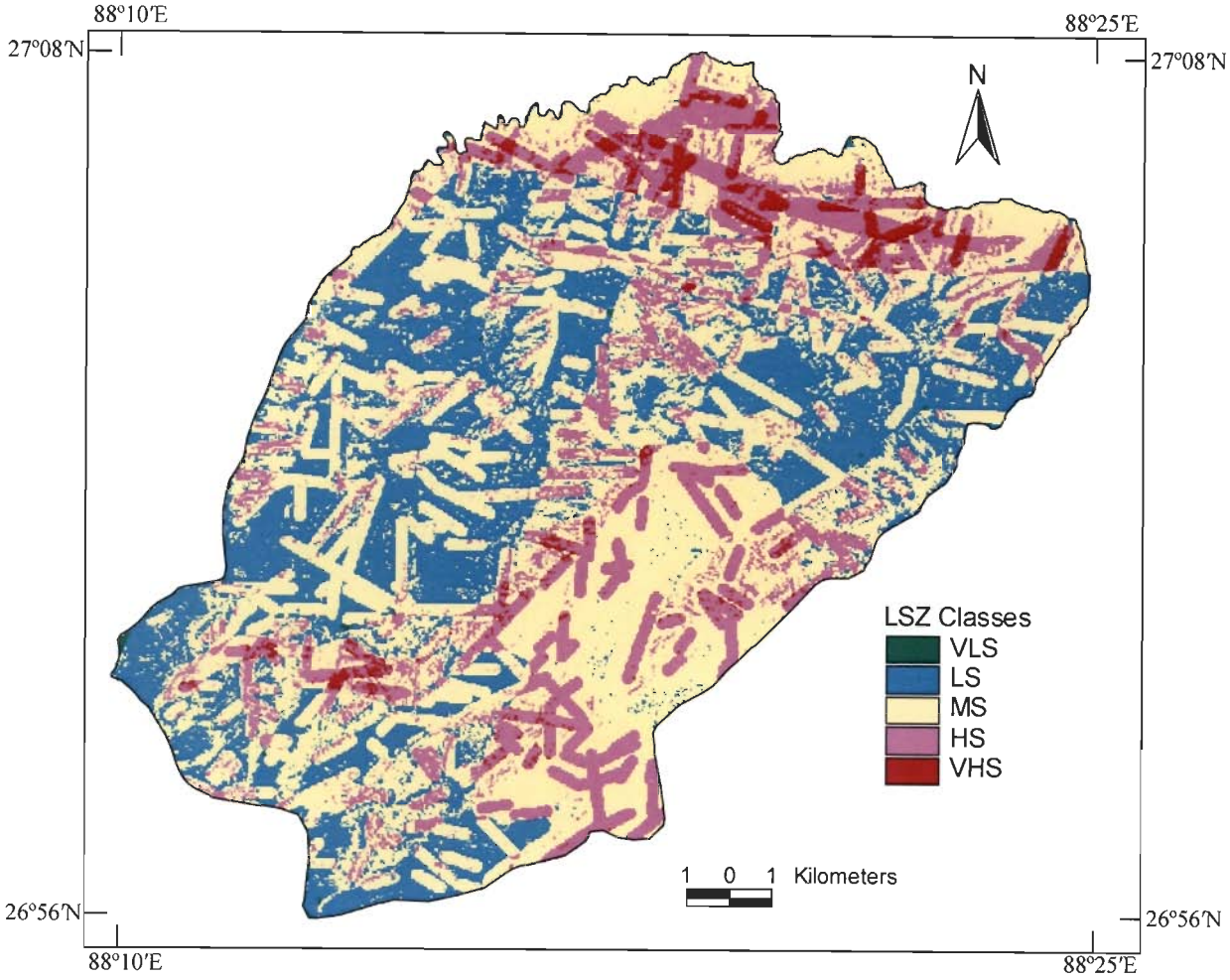


Figure 6.15: LSZ Map IV produced from combined neural and fuzzy approach.

Table 6.16: Landslide distribution in landslide susceptibility zones of LSZ Map IV (combined neural and fuzzy approach).

Landslide Susceptibility Zones	Area (km^2)	Percent Area (%) (a)	Landslide Area per Class (km^2)	Percent Landslide Area per Class (%) (b)	Landslide Density (b/a)
VHS	5.9	2.3	0.064	30.1	13.09
HS	51.5	20.2	0.068	31.9	1.58
MS	123.3	48.4	0.056	26.5	0.55
LS	73.4	28.8	0.024	11.5	0.40
VLS	0.8	0.3	0.000	0.00	0.00

It can be observed from Table 6.16 that 2.3% of the total area have been occupied by VHS zone while 20.2%, 48.4%, 28.8% and 0.3% area have been occupied by HS, MS, LS and VLS zones respectively. This shows that the area wise coverage of different susceptibility zones is normally distributed, which should be the case. The distribution of landslides in different susceptibility zones has been compared. It has been found that 30.1% of landslide area is predicted over VHS zone while 31.9%, 26.5%, 11.5% and 0.0% of landslide area are predicted over HS, MS, LS and VLS zones respectively. These results show that 22.5% area of VHS and HS zones could predict 62.0% landslide area. Further, the distribution of landslides over VHS to VLS zones is skewed towards the higher susceptibility zones, which should in fact be the case.

It can also be observed from Table 6.16 that the landslide density for VHS zone (i.e., 13.09) is much higher than those for other landslide susceptibility zones. The Map IV has a systematic and a decreasing trend of variation in landslide density values from VHS to VLS zones. Hence, it is inferred from these results that LSZ Map IV produced from combined neural and fuzzy approach is reasonably excellent representation of landslide susceptibility zones in the study area.

6.5 Summary

In this thesis, neural and fuzzy set theoretic based approaches have been developed and implemented for LSZ mapping to overcome the problem of subjectivity in weight and rating assignments to the factors and their categories. Three different objective approaches namely ANN black box approach, fuzzy set based approach and combined neural and fuzzy approach were implemented and the results discussed.

In the ANN black box approach, the weights and ratings remain hidden and are not known. The LSZ Map II produced through this approach showed a lot of similarity with the LSZ Map I_A of conventional weighting approach as was expected also. This was due to the fact that the LSZ Map I_A was used as the reference map to generate the LSZ Map II.

Therefore, a fuzzy set based approach using cosine amplitude similarity method was adopted to determine the ratings of the categories of factors for LSZ mapping. This could bring out the relative importance (ratings) of different categories of thematic data layers in terms of landslide occurrences in an unbiased manner. In this approach, the integration of different weighted thematic layers was performed using arithmetic overlay operation as well as a new method based on fuzzy gamma operators. It was observed that fuzzy gamma operator appeared to be a useful way for data integration so as to introduce fuzziness in the integration process thereby depicting the real physical situations on the ground. The LSZ Map prepared after integration of various thematic layers using fuzzy gamma operator having higher gamma value reflected a more real situation in terms of landslides in the area as compared to lower gamma values. However, in the present case, the LSZ map produced using arithmetic overlay operation for data integration (Map III) reflected a

comparatively more real situation in terms of landslides than LSZ map produced using fuzzy gamma operator (Map III_D) in the study area. The major limitation in fuzzy set based approach was that a constant and unit weight was used for all the factors thereby giving equal importance to all the factors for LSZ mapping.

To overcome this limitation of constant and unit weight for all the causative factors in fuzzy set based approach, the ANN connection weight analysis was developed and implemented to determine the weights of different factors. It was observed that lithology was the most important and drainage buffer was the least important causative factors for landslide occurrences in the area. Lineament buffer showed the 2nd rank from the top. However, in conventional weighting approach, the drainage buffer was considered as the most important and aspect as the least important causative factors based on field inspection. Lithology was considered as 4th rank from top in order of importance, as these were not visible due to limited field of view. Drainage lines, which were obvious, were initially considered to be the most important factor for LSZ. However, the ANN derived weights showed that the real feature of importance was the lineaments. Further, the importance of lithology which was not so obvious from the field data was found to be the most important factor for LSZ mapping through weights derived from ANN connection weight approach. Hence, a fully objective approach (combined neural and fuzzy) of analysis could bring out the relative importance of different thematic data layers via weights and ratings in an unbiased manner.

Combination of the ANN derived weights and the ratings determined through fuzzy set based approach produced the LSZ Map IV, which showed systematic and a decreasing trend of variation in landslide density values from VHS to VLS zones. Also, an extremely high landslide density (i.e., 13.09) for VHS zone from this map

indicated a real physical situation on the ground. Hence, it was inferred from these results that LSZ Map IV generated through combined neural and fuzzy approach is the best.

A more rigorous comparison and evaluation of LSZ maps produced from different approaches in this study has been given in the next chapter.

Chapter 7

Comparative Evaluation of LSZ Maps

7.1 Introduction

In Chapters 5 and 6, the preparation of LSZ maps of the study area using four different approaches, namely conventional weighting approach (LSZ Map I_A), ANN black box approach (LSZ Map II), fuzzy set based approach (LSZ Map III) and combined neural and fuzzy approach (LSZ Map IV) was described. A comparative evaluation of these LSZ maps may throw invaluable light on their relative efficacy and mutual compatibility. This chapter deals with the comparative evaluation of the LSZ maps using three different analyses:

a) Landslide Density Analysis

b) Error Matrix Analysis

c) Difference Image Analysis

7.2 Landslide Density Analysis

Landslide density is defined as the ratio of the existing landslide area in percent to the area of each landslide susceptibility zone in percent, and is calculated here on the basis of the number of pixels. Landslide density values for each of the susceptibility zones (i.e., VHS, HS, MS, LS and VLS) for different LSZ maps have been calculated separately (Table 7.1). Usually, an ideal LSZ map should have the highest landslide density for VHS zone as compared to other zones and there ought to be a decreasing trend of landslide density values successively from VHS to VLS zone.

Table 7.1: Landslide densities of different susceptibility zones for various LSZ maps.

Landslide Susceptibility Zones	Landslide Density (computation based on pixel numbers)			
	Map I_A	Map II	Map III	Map IV
VHS	1.63	1.34	6.72	13.09
HS	1.79	1.50	1.11	1.58
MS	0.88	1.02	0.66	0.55
LS	0.41	0.49	0.26	0.40
VLS	0.19	0.12	0	0

It is observed from Table 7.1 that the landslide densities for VHS zone of LSZ Maps III and IV are much higher as compared to those obtained for other susceptibility zones. There is also a decreasing trend of landslide density values from VHS zone to VLS zone for Map III and IV. On the other hand, in case of LSZ Maps I_A and II, the landslide density is found to be only marginally higher for HS zone than VHS zone. It is also observed from Table 7.1 that the LSZ Maps I_A and II have a similar trend of landslide densities for various susceptibility zones, which is expected, since the LSZ Map I_A has been used as the reference map to generate the LSZ map II.

As far as the landslide density in VHS zone is concerned, it is observed that the LSZ Map IV has a much higher landslide density (>13) in this zone than that observed in same zone of other LSZ maps (e.g., 1.63 for Map I_A, 1.34 for Map II and 6.72 for Map III). Further, the Map IV also has a more systematic and reasonable trend of variation in landslide density values from VHS to VLS zones. Thus, the LSZ Map IV is interpreted to be the best LSZ map of the area (also see Section 7.4.3).

Thus, based on the landslide density values of different zones and their trend from VHS zone to VLS zone for all the LSZ maps, it is inferred that the combined neural and fuzzy approach developed and implemented for LSZ mapping appears to be significantly better than other approaches (i.e., conventional weighting, ANN black box and fuzzy set based approaches) used here.

7.3 Error Matrix Analysis

For a comparative evaluation of various LSZ maps, it is also important to investigate how pixels match or mismatch across LSZ maps based on different approaches. This aspect is discussed through error matrix analysis.

In the present context, the error matrix analysis of each LSZ map with respect to other LSZ maps is defined as the cross tabulation of distribution of pixels in various landslide susceptibility zones in a particular LSZ map and corresponding zones in another LSZ map. This analysis is based on cumulative number of pixels falling in each landslide susceptibility zone rather than on pixel by pixel basis (as discussed under another evaluation method in Section 7.4). The cells of error matrix provide a comparison on map-to-map basis to evaluate how pixels match or mismatch in different LSZ maps taken two at a time.

Here, three such error matrices have been generated to understand the distribution of number of pixels in various landslide susceptibility zones across the maps.

(a) Error matrix for LSZ Maps I_A and II

(b) Error matrix for LSZ Maps III and IV

(c) Error matrix for LSZ Maps I_A and IV

The reasons for selecting the above combinations are given below.

7.3.1 Error Matrix for LSZ Maps I_A and II

As has been described earlier, the LSZ Map I_A has been generated through conventional weighting approach and it has been used as a reference map for LSZ Map II, based on the ANN black box approach. Therefore, for comparing the two maps, an error matrix for these two LSZ maps has been generated (Tables 7.2).

Table 7.2: Error matrix of LSZ Maps I_A and II.

		LSZ Map I_A					Total
		VHS	HS	MS	LS	VLS	
LSZ Map II	VHS	23703	7306	94	10	0	31113
	HS	2686	90112	16237	362	0	109397
	MS	43	11448	98700	33831	5	144027
	LS	49	548	7498	96541	320	104956
	VLS	2	10	769	11338	6135	18254
Total		26483	109424	123298	142082	6460	407747

Table 7.2 shows that there is a high degree of agreement in the pixels of LSZ Maps I_A and II, particularly for VHS, HS, MS and LS zones, which is quite expected. However, there is some degree of mismatch in the VLS zone which is reflected by the

population of 6460 pixels in the VLS zone of LSZ Map I_A as against a population of 18254 pixels in this zone in LSZ Map II.

7.3.2 Error Matrix for LSZ Maps III and IV

The LSZ Maps III and IV have been prepared using two objective weight-rating approaches – the Map III based on fuzzy set based approach and the Map IV based on combined neural and fuzzy approach. Therefore, it is interesting to compare the two maps through error matrix (Tables 7.3).

Table 7.3: Error matrix of LSZ Maps III and IV.

		LSZ Map III					Total
		VHS	**HS**	**MS**	**LS**	**VLS**	
LSZ Map IV	**VHS**	4687	3511	1259	0	0	9457
	HS	15219	32495	31922	2766	0	82402
	MS	4817	54190	89885	47707	1023	197622
	LS	17	2392	37466	73475	3692	117042
	VLS	0	2	20	110	1092	1224
Total		24740	92590	160552	124058	5807	407747

The error matrix, given in Table 7.3 shows that there is a general agreement between these LSZ maps (Maps III and IV). About 50% pixels (201634 pixels out of total 407747 pixels), as indicated by the diagonal of the matrix, match in different zones. The VHS zone of LSZ Map IV has shown much focused population of 9457 pixels, whereas the LSZ Map III has a population of 24740 pixels in the VHS zone. This is responsible for some mismatch in the VHS zone.

The match or mismatch may be related to the type of weight/rating approaches adopted for the preparation of the two maps. The weights for the factors are considered as constant (i.e., unit weight for all the factors) in case of fuzzy set based

approach and only ratings for the categories obtained from fuzzy set based approach have been used for generating Map III. On the other hand, the ratings for the categories obtained from fuzzy set based approach have been integrated with the weights for the factors (obtained from ANN connection weight procedure) to prepare the LSZ Map IV. Therefore, in case of combined neural and fuzzy procedure, the factors have varied importance in terms of weights. This may be responsible for differences in the two maps and mismatch between LSZ Maps III and IV.

7.3.3 Error Matrix for LSZ Maps I_A and IV

It has been mentioned that the LSZ Map IV, prepared using fully objective combined neural and fuzzy approach, is seemingly the best LSZ map and shows a much focused distribution of pixels in VHS zone. The LSZ Map I_A has been generated using the most widely used conventional weighting approach. Therefore, a comparison of the two Maps I_A and IV has been carried out via error matrix (Table 7.4).

Table 7.4: Error matrix of LSZ Maps I_A and IV.

		LSZ Map I_A					Total
		VHS	**HS**	**MS**	**LS**	**VLS**	
LSZ Map IV	**VHS**	1114	2897	3433	2013	0	9457
	HS	11583	25584	27478	17150	607	82402
	MS	11344	56584	65401	61058	3235	197622
	LS	2442	24340	26952	61323	1985	117042
	VLS	0	19	34	538	633	1224
Total		26483	109424	123298	142082	6460	407747

The first point to be noted is that in LSZ Map IV, there is a big difference in the population of pixels in various landslide susceptibility zones. On the other hand, in

case of LSZ Map I_A, the population of pixels is quite the same in HS, MS and LS zones.

The error matrix shows that there is barely 37.8% match in number of pixels (154055 pixels out of total 407747 pixels), as indicated by the diagonal of the matrix. Thus, there is a lot of mismatch in number of pixels allocated to various zones in LSZ Maps I_A and IV. This mismatch may be due to the weight and rating assignment approaches of conventional weighting approach (resulting into LSZ Map I_A) versus combined neural and fuzzy approach (resulting into LSZ Map IV). As the weights and ratings have been assigned in a purely subjective manner in the former case and in a purely objective manner in the later case, the weights for the factors and the ratings for the categories significantly differ in both the approaches. This has resulted into two different LSZ maps (LSZ Maps I_A and IV), with a rather high mismatch. This point is further discussed in Section 7.4.3.

7.4 Difference Image Analysis

Difference image analysis elucidates how pixels shift from one landslide susceptibility zone to another zone across each LSZ map. Attributes of 1, 2, 3, 4 and 5 have been assigned respectively to the VLS, LS, MS, HS and VHS zones of all the four LSZ Maps. Difference images are generated by subtracting pixel attributes of one LSZ map from the other. Thus, a difference image can have a maximum of five different classes; viz, no difference, one-zone difference, two-zone difference, three-zone difference and four-zone difference. Fully matching pixels in the two LSZ maps would correspond to the no difference class in the difference image. Thus, the difference images show the spatial consistency between LSZ maps in terms of matching/mismatching of pixels in various landslide susceptibility zones.

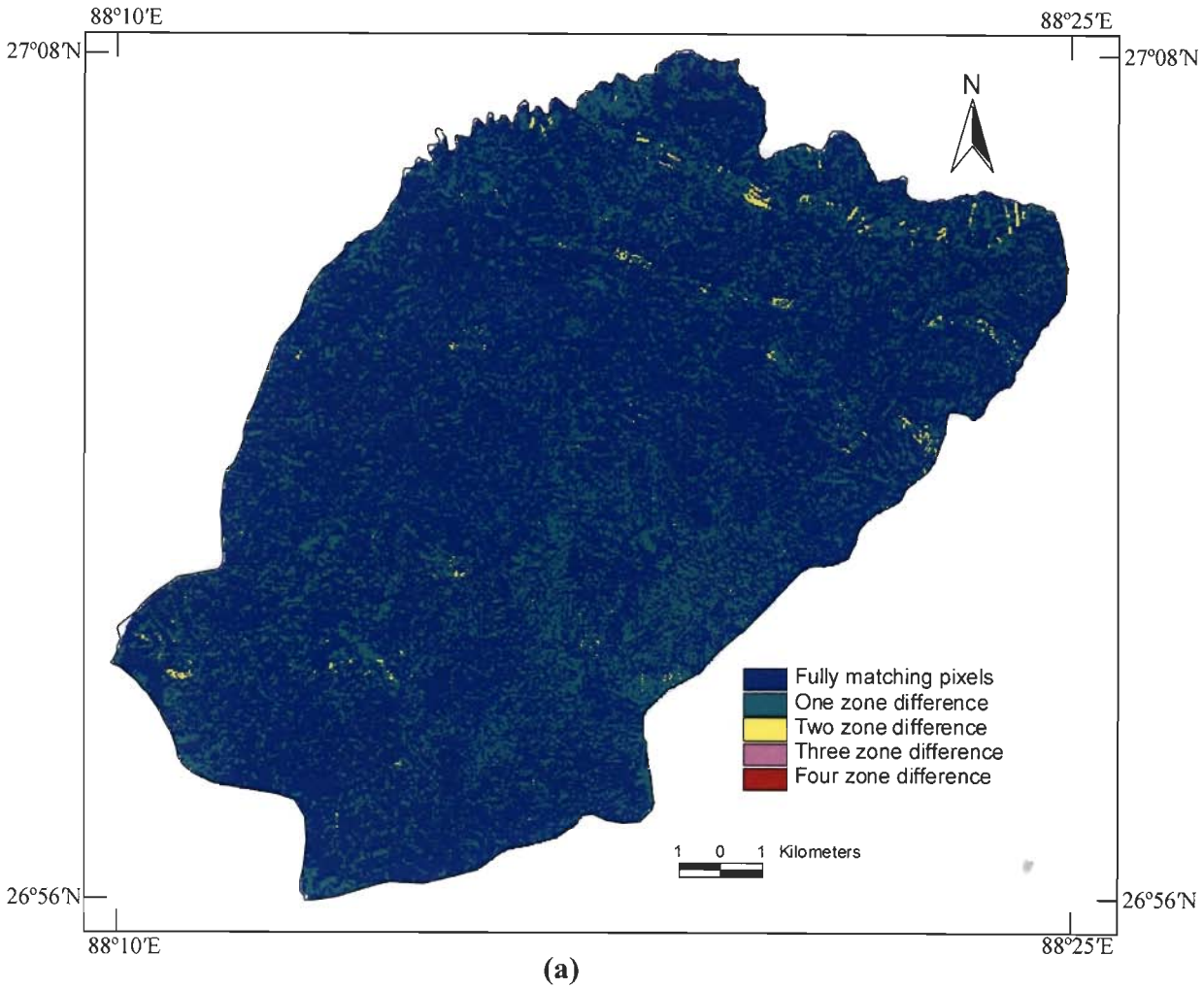
For this analysis, the same three different combinations of LSZ maps (i.e., Maps I_A & II, Maps III & IV and Maps I_A & IV) have been taken for mutual comparison, as has been done for the error matrix analysis, the logic of selecting the combination being the same. The results of these three difference images are presented in terms of cumulative number of pixels and percent areas covered in each of the difference classes (Table 7.5).

Table 7.5: Results of difference images of LSZ maps.

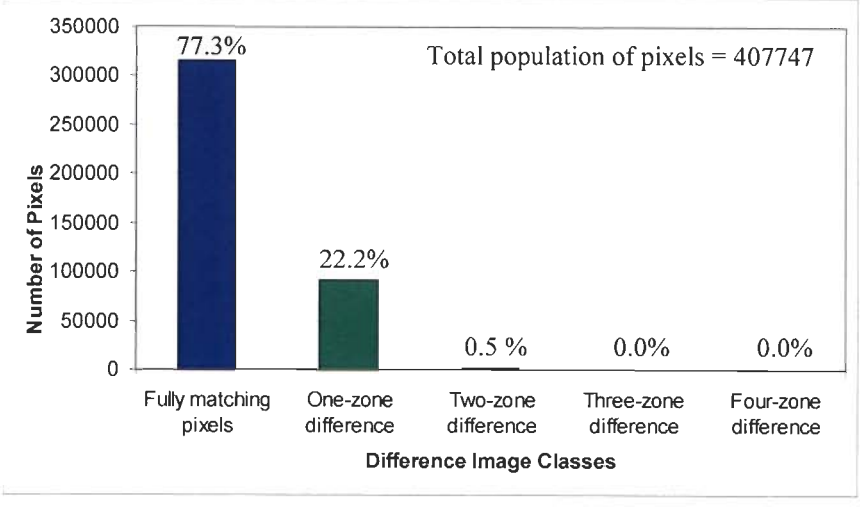
Figure	Difference images	No difference		One-zone difference		Two-zone difference		Three-zone difference		Four- zone difference	
		Number of pixels	% Area	Number of pixels	% Area	Number of pixels	% Area	Number of pixels	% Area	Number of pixels	% Area
7.1	I_A~II	315191	77.3	90664	22.2	1821	0.5	69	0.0	2	0
7.2	III~IV	201634	49.5	193817	47.5	12277	3.0	19	0.0	0	0
7.3	I_A~IV	154055	37.8	189075	46.4	59536	14.6	5081	1.2	0	0

7.4.1 Difference Image of LSZ Maps I_A and II

The LSZ Map I_A, prepared using conventional weighting approach, and Map II using the ANN black box approach, appear quite alike (see Figures. 5.3 and 6.6). A difference image of the two (Figure 7.1a) also shows a high degree of mutual correspondence and matching of landslide susceptibility zones throughout the area. This is in agreement with the results based on error matrix analysis (Table 7.2). About 77.3% pixels have full mutual matching and 22.2% pixels exhibit one-zone difference. Barely 0.5% pixels have two-zone difference (Figure 7.1b), and these appear to be related to a lithologic band in the northern part of the area. However, in general, there is a high degree of correspondence between LSZ Maps I_A and II.



(a)



(b)

Figure 7.1: (a) Difference image of LSZ Maps I_A~II and (b) Frequency distribution of pixels in difference image classes.

7.4.2 Difference Image of LSZ Maps III and IV

The LSZ Map III (prepared using fuzzy set based approach) (Figure 6.9) and the LSZ Map IV (prepared using combined neural and fuzzy approach) (Figure 6.15) have been generated based on objective weight-rating approaches. A difference image of the two (Figure 7.2a) shows a high degree of spatial matching. 49.5% pixels have full mutual matching and 47.5% pixels exhibit only one-zone difference (Figure 7.2b). 3.0% pixels have two-zone difference and these mainly appear to be related to a lithologic band in the northern part of the area, as discussed below.

It may be recalled that in case of LSZ Map III, prepared using fuzzy set based approach, all the causative factors have been given equal or unity weights whereas the importance of different categories of factors differ in terms of ratings. Ratings of categories vary between 0 and 1 for a causative factor (thematic layer) according to the fuzzy relation. In case of LSZ Map IV, the weights for factors derived from ANN and the ratings determined through fuzzy relation concept have been used. This has been done objectively without any bias. Therefore, the two LSZ maps differ slightly from each other.

In case of ANN derived weights, lithology has been assigned the highest weight of 4.8 (Rank 1) whereas weights for other causative factors (lineament buffer, slope, aspect, land use land cover and drainage buffer) vary from 2.1 to 0.2. As lithology has the highest and significantly higher weight than other factors, the importance of lithology has been brought out in the difference image (Figure 7.2a).

Further, for LSZ Map IV, the ANN derived weights for lineament buffer, slope, aspect, land use land cover and drainage buffer are 2.1, 1.3, 1.1, 0.5 and 0.2 respectively. On the other hand, for LSZ Map III, all these weights have been considered as equal and unity. Therefore, in difference image III~IV, the impact of

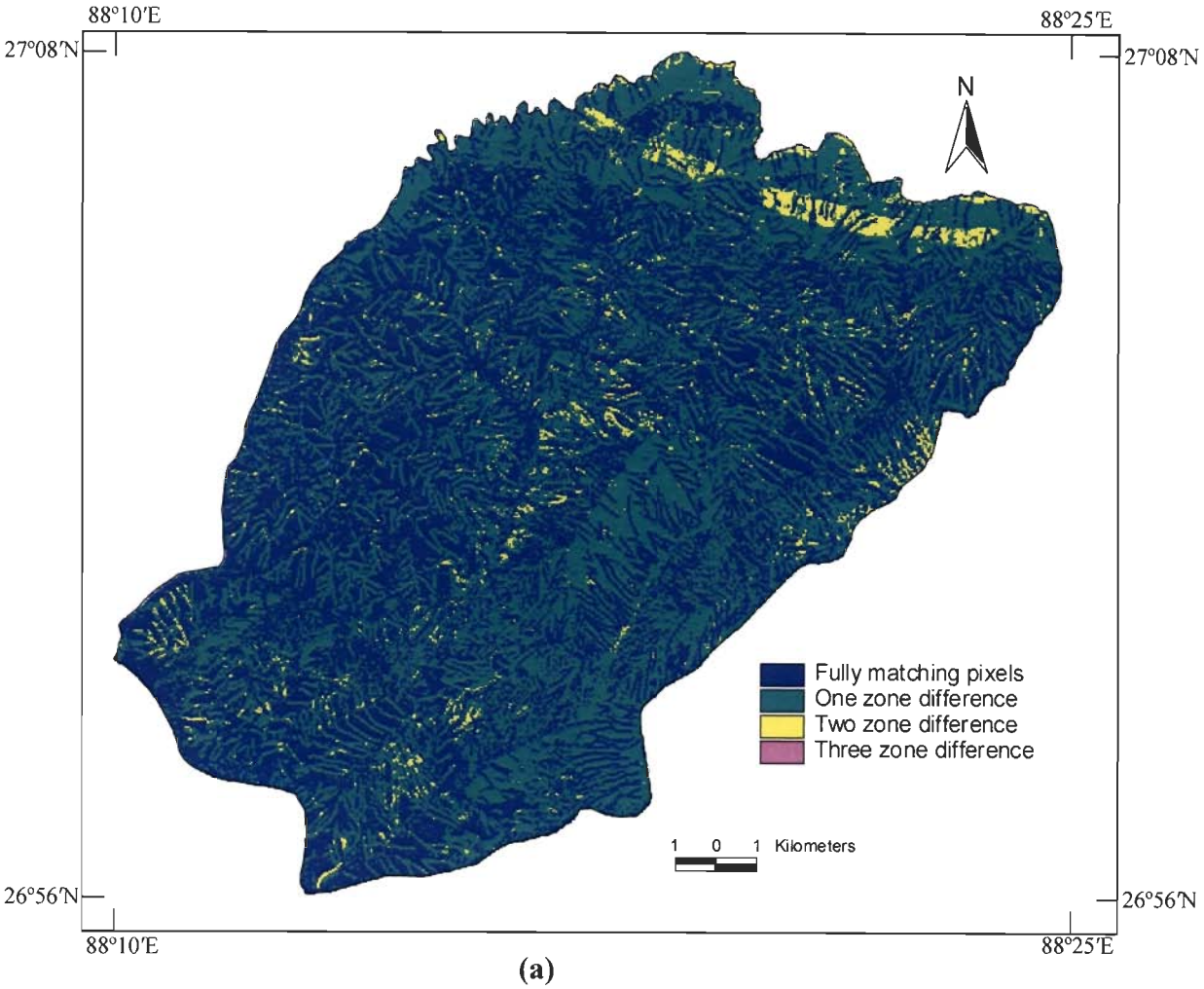
these factors is limited to fully matching and one-zone difference classes only (Figure 7.2a).

7.4.3 Difference Image of LSZ Maps I_A and IV

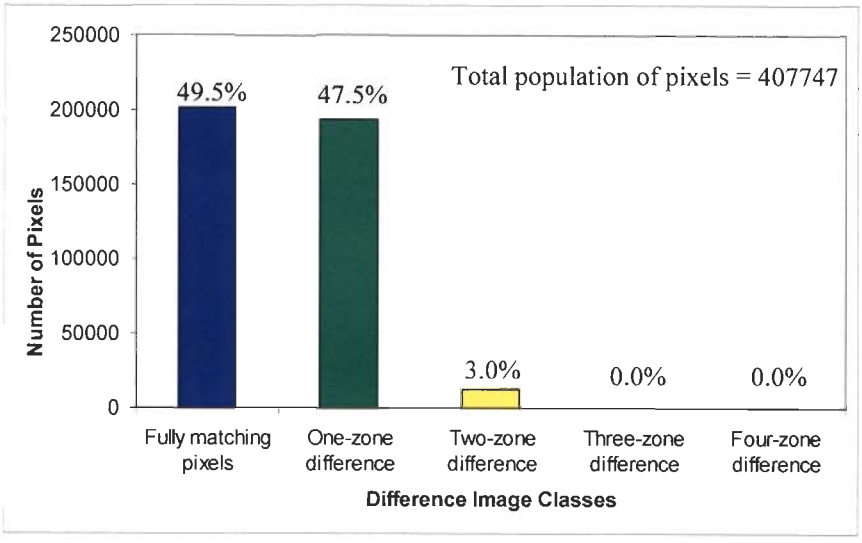
As mentioned earlier, the LSZ Map I_A has been prepared by using subjective weighting approach whereas the Map IV has been prepared using objective weight-rating approach based on combined neural and fuzzy. The two maps (Figures. 5.3 and 6.15) appear to exhibit the widest spatial difference, as was also seen through the error matrix analysis (Table 7.4). This is again confirmed in the difference image (Figure 7.3a), where only 37.8% pixels are found to be fully matching, 46.4% pixels exhibit one-zone difference, 14.6% pixels have two-zone difference and 1.2% pixels have three-zone difference (Figure 7.3b). This indicates a significant level of difference between the Maps I_A and IV.

A deeper look into the pattern of LSZ Maps I_A and IV reveals a wealth of interesting information. For preparing Map I_A, various important thematic layers considered (that we believed to be important through field observations) in that order, are: drainage buffer (weight = 9), lineament buffer (weight = 8), slope (weight = 7), lithology (weight = 6), land use land cover (weight = 4) and aspect (weight = 1) (Table 5.1). For this reason, the Map I_A shows a significant impact of drainage lines/buffers and a little impact of lithology.

On the other hand, by objective approach (combined neural and fuzzy) the weights of the factors obtained are: lithology (weight = 4.8), lineament buffer (weight = 2.1), slope (weight = 1.3), aspect (weight = 1.3), land use land cover (weight = 0.5) and drainage buffer (weight = 0.2) (Table 6.14). The Map IV (Figure 6.15)

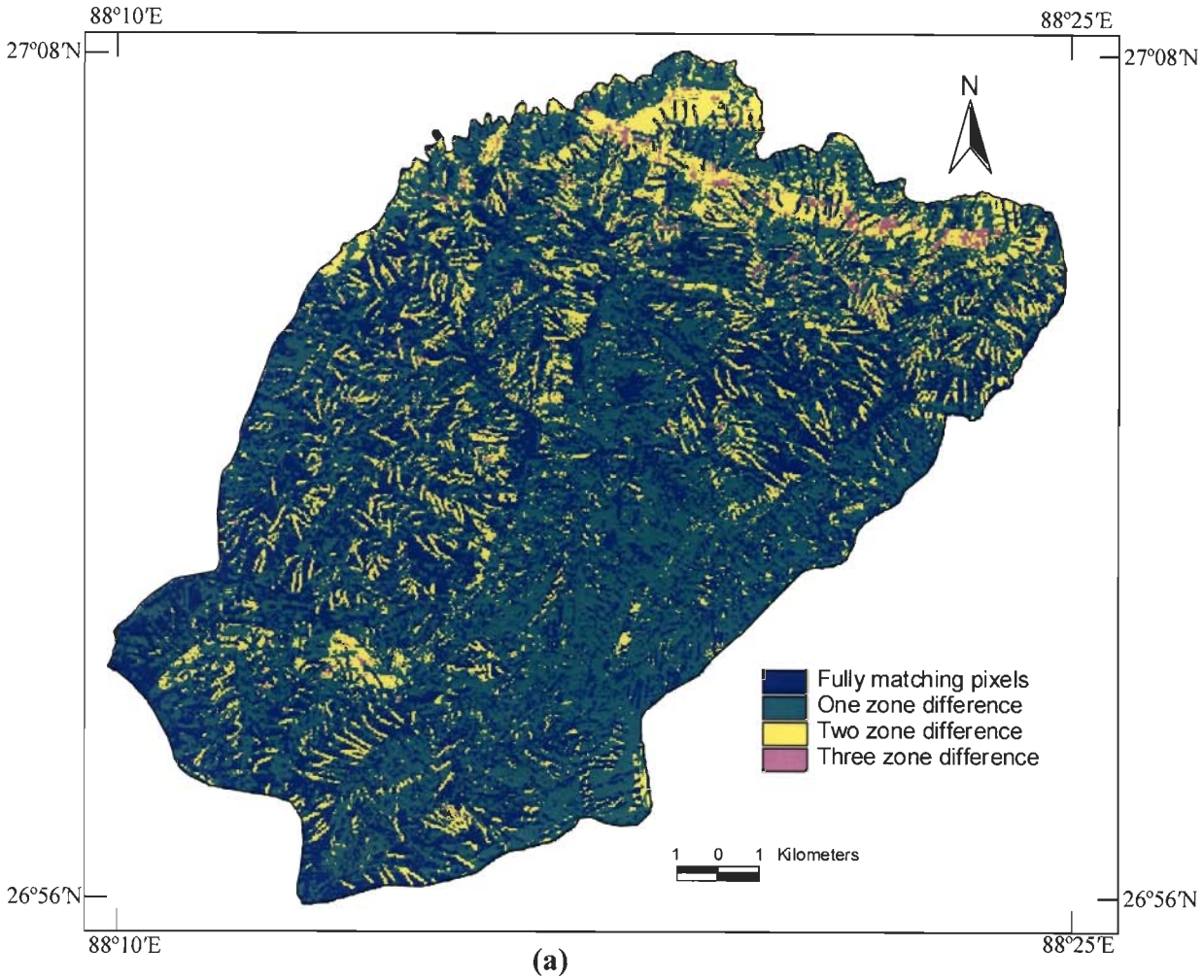


(a)

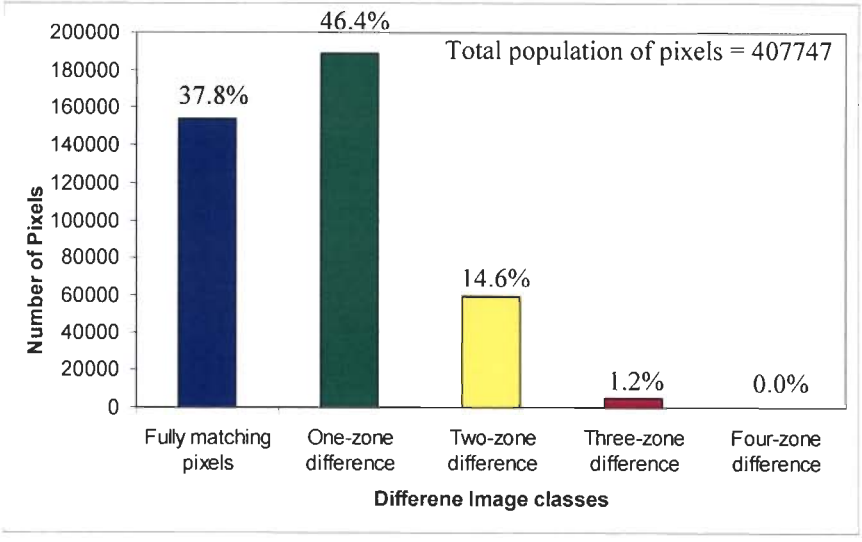


(b)

Figure 7.2: (a) Difference image of LSZ Maps III~IV and (b) Frequency distribution of pixels in difference image classes.



(a)



(b)

Figure 7.3: (a) Difference image of LSZ Maps I_A~IV and (b) Frequency distribution of pixels in difference image classes.

distinctly exhibits the impact of lithologic banding and lineament buffer, in accordance with the weights.

The image (Map I_A ~ Map IV) (Figure 7.3a) clearly shows the differences in the two LSZ maps, the most important being a band of two-zone difference lying in the northern part of the area which apparently relates to a lithologic band (as confirmed from the geologic map). Note that lithology has the highest rank (=1) in Map IV but a low importance (4^{th} from top) in Map I_A. This difference in importance is responsible for the prominent band in the difference image in Figure7.3a.

Two-zone differences are also seen at several places in the western and southern parts of the area (Figure 7.3a). These are apparently related to drainage buffer (highest importance in Map I_A and lowest importance in Map IV).

In the south-eastern part of Figure 7.3a, most pixels exhibit only no or one-zone difference. This is owing to the superimposition/coincidence of lineament buffer vis-à-vis drainage buffer here, i.e., the pixels being treated under drainage buffer in Map I_A and under lineament buffer in Map IV. This is quite possible in situations where drainage lines follow lineaments, i.e., lineaments are marked by (rectilinear/rectangular/angular) drainage. In essence, lineaments lead to fracturing of the terrain along which development of drainage takes place. However, in field due to limited field of view, drainage appears as a very conspicuous feature, and lineament is hardly observed. Considering the distribution of landslides in the field along the drainage lines, we presumably initially considered drainage line to be the most important input thematic layer for LSZ and produced the Map I_A. However, as a result of the objective spatial-domain regional analysis, drainage is reflected to the lowest rank (=6) in Map IV, and it is revealed that the real feature of importance is lineament, and drainage is only an apparent manifestation of the same at places.

Further, the importance of lithology which was not so obvious from field data (Map I_A) is brought to light by the fully objectively derived Map IV.

Thus, the comparative analysis shows the limitation of conventional weighting approach. In this approach, the weights are assigned based on the field observation, where field of view is limited. The fully objective approach (combined neural and fuzzy) of analysis on the other hand could bring out in an unbiased manner the relative importance (weights) of thematic data layers. Therefore, the above analysis brings out the fact of relative advantages of fully objective approach vis-à-vis conventional weighting approach.

Chapter 8

Landslide Risk Assessment

8.1 Introduction

Landslide Risk Assessment (LRA) is an extremely important aspect in practical applications of landslide studies. In general, risk can be defined as "the potential for adverse consequences, loss, harm or detriment by the hazard to human and the things that humans value" (Lee and Jones, 2004). In a more scientific way, risk has also been defined as a combination of the probability or frequency of occurrence of a particular hazard and the magnitude of the consequences of occurrence (Royal Society, 1992). This definition is quite useful as it identifies the importance of a phenomenon (landslides in the present case) in generating risk and the significance of consequences in the assessment of risk.

There is a range of risk assessment procedures varying from quantitative to qualitative estimations of risk, with the latter based more and more on expert judgement. Suitability of either qualitative or quantitative assessments depends on

both the desired accuracy of the outcome and the nature of the problem. Generally, for a large area, there is scarcity of available data for any quantitative analysis; a qualitative risk assessment may be more applicable. On the other hand, for specific sites, a detailed quantitative risk assessment may be required.

In this chapter, a brief review of the approaches for qualitative landslide risk assessment has been provided. This is followed with implementation of a novel landslide risk assessment approach developed in this study.

8.2 Landslide Risk Assessment (LRA) – A Brief Review

LRA approaches can be applied at different stages in the decision-making process, starting from developmental planning to specific site evaluation. Hence, risk assessment will also vary from a general indication of the threat across a region, to specific statements on levels of risk at a particular site. Thus, the LRA can be regional or site-specific in nature, depending upon the varied applications at different stages of the decision-making process. LRA can lead to the identification of areas with different levels of threat to human beings and the things that value to them. This information can be used for planning developmental activities in the area.

As mentioned earlier, LRA depends on two factors: (1) the probability of landslide hazard in a region and (2) the vulnerability of resources at risk.

The spatial distribution of landslide risk may be obtained by integrating landslide probability and vulnerability of population or property at spatial level in a GIS environment (Leone et al., 1996). The resulting map can be subdivided into areas of different risk zones.

The probability of landslide occurrences depends on both the inherent factors and triggering (external) factors. However, the triggering factors may change over a

very short time span and are thus very difficult to estimate. If triggering factors are not taken into account, the term "susceptibility" may be used to define the likelihood of occurrence of a landslide event as has been done in the present case (see Chapters 5 and 6).

Vulnerability may be defined as the level of potential damage, or degree of loss, of resources at risk, subjected to a landslide of a given intensity (Fell, 1994; Leone et al., 1996; Wong et al., 1997). Vulnerability assessment involves the understanding of the interaction between a given landslide and the affected resources. Generally, the vulnerability to landslide may depend on the volume and velocity of sliding, the distance of transportation of slided material, the resources at risk, their nature and proximity to the landslide (Finlay and Fell, 1995). The assessment of vulnerability is somewhat subjective and may be largely based on historic records. The appropriate vulnerability factor may be assessed systematically by expert judgement and can be expressed at a scale of 0 to 1.

In the present context of LRA, the probability of landslide hazard has been considered as landslide susceptibility, and vulnerability of resources at risk has been taken as resource damage potential (discussed later in Section 8.3). The LRA is aimed at providing a risk map that depicts the level of risk, in terms of different resources of a given region.

In the present study, LRA has been carried out at a regional scale. Therefore, a brief review of only qualitative landslide risk assessment approaches is given in this chapter. A number of qualitative LRA approaches can be found in the literature (e.g., Lee and Zones, 2004). These approaches are:

a) Risk registers

b) Relative risk scoring

c) Risk ranking matrices

d) Relative risk rating

e) Failure modes, effects and criticality analysis (FMECA)

8.2.1 Risk registers

A risk register is a document which keeps all the records of the known risks due to landslides in an area or at a particular site and also the decisions taken in monitoring and managing these risks. This register can serve the purpose of historical landslide data base. This is also useful in screening out the areas with minor or negligible landslide problems for planning developmental activities and also prioritizing landslide problems at an early stage of a project. Details on risk register concept and historic landslides have been discussed in several studies (e.g., Lee et al., 1998a; Lee, 1999; Lee and Clark, 2000 and Lee and Zones, 2004).

8.2.2 Relative Risk Scoring

In most cases, evaluation of risk in absolute terms is inappropriate due to the difficulties in assigning exact values for the hazard, for resources at risk and their possible consequences. Hence, it is expedient to assess the relative risk potential at different sites posed by particular hazards based on subjective appraisal.

The relative risk scoring approach uses the definition of risk as a function of hazard probability and adverse consequences. The landslide hazard probability and adverse consequence elements (resources at risk) are represented by relative scores or rank values and the risk is the product of these scores. The risk numbers thus produced are then used to classify each site within an arbitrarily defined scale of risk classes. This allows some comparison between different sites and provides a basis for

taking management decisions. The details of this approach along with examples are given in Lee and Zones (2004).

Boggett et al. (2000) used such approach to evaluate the problems of rockfalls and rockslides in South Shore Cliffs, Whitehaven, UK and identified the required remedial works. McDonnell (2002) developed and implemented an approach at a World Heritage site of basalt cliffs in Northern Ireland. The cliffline was divided into different sections and the risk within each section was calculated. The relative risk score was considered as a cumulative effect of the hazard score, visitor concentration score and visitor perception score. The hazard score was obtained by summation of four different components such as stability number (indicates hazard potential and obtained from slope stability analysis), slope angle, presence or absence of springs and water seeps, and presence or absence of dumped material. The resources at risk at this heritage site were visitors only. Thus, the concepts of visitor concentration score and visitor perception score were introduced. The perception score reflected visitor's awareness about the landslide hazard. The relative risk scores for all cliff sections were obtained and categorized into different risk classes.

Rautela and Lakhera (2000) prepared the vulnerability map of parts of Himalayas in Sirmur district, Himachal Pradesh, India, using 1991 census data and an LSZ map. The population of the area (1991 census) was categorized into five classes and was used to formulate the vulnerability coding of populations to devastation caused by landslides (Table 8.1). In this case, only population data was used for vulnerability studies, but not other resources/infrastructures were considered.

Table 8.1: Vulnerability coding of population classes to devastation caused by landslides (Rautela and Lakhera, 2000).

Landslide hazard classes	Population class				
	Sparse	**Low**	**Moderate**	**High**	**Very high**
Least	Low Vulnerability	Low Vulnerability	Low Vulnerability	Low Vulnerability	Moderate Vulnerability
Low	Low Vulnerability	Low Vulnerability	Low Vulnerability	Moderate Vulnerability	Moderate Vulnerability
Moderate	Low Vulnerability	Moderate Vulnerability	Moderate Vulnerability	High Vulnerability	High Vulnerability
High	Moderate Vulnerability	Moderate Vulnerability	High Vulnerability	Very High Vulnerability	Very High Vulnerability

Chau et al. (2004) presented landslide risk for Hong Kong as a function of hazard and exposure. Landslide hazard was represented by the LSZ map and the exposure was represented by population data (2001 census). The classification of the population was made based on the criterion that all the population classes have equal area (i.e., numbers of pixels). The class numbers of the susceptibility and population classes were normalized with respect to maximum class number in each category to obtain the index values. These normalized values represented the hazard index and exposure index. These indices of each pixel were multiplied to obtain the risk values of the area and the risk values were categorized to prepare the risk map of the area.

8.2.3 Risk Ranking Matrices

In this approach, risk is represented in the form of a risk matrix where subjective ranking of different risk levels is done as the likelihood of landslide hazard measured against the increasing severity of adverse consequences. This is fully based on expert judgements to make appropriate assessments of the likelihood of landslide events and adverse consequences.

Anbalagan and Singh (1996) implemented the LRA approach in mountainous terrain of Kumaun Himalayas, India, using a risk assessment matrix. Risk was considered as a function of hazard probability and damage potential. The damage potential was evaluated as very low (VLDP), low (LDP), moderate (MDP), high (HDP) and very high (VHDP) in terms of loss of life and/or injuries as well as loss of land and property (Table 8.2). For example, damage potential of resources damage to >50 dwellings or damage of very thick vegetated area or damage of >2000m of road, is treated as VHDP.

Table 8.2: Damage potential of different resources at risk (Anbalagan and Singh, 1996).

Damage potential (DP)	Number of dwellings likely to be damaged	Land use land cover categories	Length of road damage (m)
VLDP	<2	Barren	<100
LDP	2-5	Sparsely vegetated	101-500
MDP	5-10	Mod. Vegetated/ agricultural land	501-1000
HDP	10-50	Thickly vegetated	1001-2000
VHDP	>50	Very thickly vegetated	>2000

The hazard probability of slope facets such as very low (VLHP), low (LHP), moderate (MHP), high (HHP) and very high (VHHP) was obtained from the LSZ map. These datasets on damage potential and hazard potential were integrated manually based on a slope facet concept and a risk assessment matrix was formed with a five fold classifications such as very low risk (VLR), low risk (LR), moderate risk (MR), high risk (HR) and very high risk (VHR).

van Dine et al. (2002) used the concept of risk matrices for a qualitative risk assessment of a forest land at Perry Ridge under British Columbia Ministry of Forests. The probability of landslide hazard was rated as high, moderate, low, very low or none based on the past occurrence of landslides, independent of their sizes. The consequences were rated as high, moderate or low based on the resources at risk (people, property and water supply). Three different risk matrices were developed, one for each of the resources at risk. The very high, high, moderate, low and very low risk zones for different combinations of hazard and consequences to resources were assessed based on the risk matrices and landslide management recommendations were made for different locations.

Cardinali et al. (2002) used both relative risk scoring and risk ranking matrices approaches to describe the landslide risk assessment of parts of Umbria, Central Italy. The study area was divided into a series of landslide susceptibility zones and the risk within each susceptibility zone was determined as a function of susceptibility and vulnerability. Landslide hazard for each susceptibility zone was defined in terms of landslide frequency and intensity. Levels of hazard were defined using a two-digit coding, one each for landslide frequency and intensity. Such coding system was used to determine whether the hazard was due to high frequency of landslides or intensity or both. Estimates of vulnerability of each type of resource at risk were based on the relationship between intensity and type of landslide and the likely damage due to the landslides. Three different levels of damage such as aesthetic, functional and structural were envisaged. A risk matrix was prepared using the coding system of landslide hazard and different levels of damage. Here also, a unique coding value termed as specific risk index was used in the risk matrix instead of qualitative terms such as low, medium and high. In order to provide a measure of total risk, the specific

landslide risk indices for each susceptibility zone were categorized into one of the landslide risk zones such as very high, high, moderate and low.

8.2.4 Relative Risk Rating

Relative risk rating approach adopts method similar to those used in the relative risk scoring and risk ranking matrices approaches. It is a descriptive approach in which a range of risk categories are defined, each with a certain degree of hazard and level of consequence. According to Palmer et al. (2002), this approach has proved its usefulness in situations where the resources at risk are uniform or broadly similar throughout an area, but have spatial variation in the degree of hazard. This technique also provides a means of identifying the relative risk throughout the area.

In this approach, the area is divided into different units based on the ground characteristics such as geology, landform, soil, topography, etc. Information on the distribution, nature and frequency of landslides, various resources at risk and the expected levels of consequences within each unit are then gathered. Risk categories are assigned to each unit based on the hazard and consequence conditions within it.

8.2.5 Failure Modes, Effects and Criticality Analysis (FMECA)

The FMECA approach provides a structured framework for the qualitative analysis of various components of an engineered slope using engineering judgement to generate scores or rankings. The details of this approach are discussed in Lee and Zones (2004). However, this technique is applicable only for the risk assessment of structural failures in an engineered slope. The FMECA approach has been used as a risk assessment tool in the dam industry (Sandilands et al., 1998; Hughes et al., 2000), coastal slopes (Lee, 2003), etc.

8.3 Landslide Risk Assessment Approaches Developed

The above literature review on qualitative landslide risk assessment reveals that the risk register approach is a heuristic approach that utilizes field records of risks related to landslide occurrences in this study. Other approaches namely relative risk scoring, risk ranking matrices and relative risk rating are quite similar to each other. In these approaches, linguistic coding of various resources at risk with respect to their damage potential is done. In most cases, only the population as resource element has been taken into consideration for risk assessment. The other resources (e.g., land use land cover) that humans value have not been generally considered. The risk assessment matrix has also been generated in terms of linguistic coding only.

In the present study, landslide risk has been considered as a function of landslide potential or susceptibility and the resource damage potential and is given as,

$$LR = f(LP, RDP) \tag{8.1}$$

where LR is the landslide risk, LP is the landslide potential and RDP is the resource damage potential. Two different approaches for landslide risk assessment have been developed and implemented to prepare the LRA maps of the study area.

(a) LRA using danger pixels

(b) LRA using Fuzzy Concept

8.3.1 LRA Using Danger Pixels

Danger pixels can be defined as pixels those appear to be under real risk due to landslides. The steps involved in this approach, LRA using danger pixels, are given in Figure 8.1. Danger pixels have been considered as those pixels which lie in VHS and

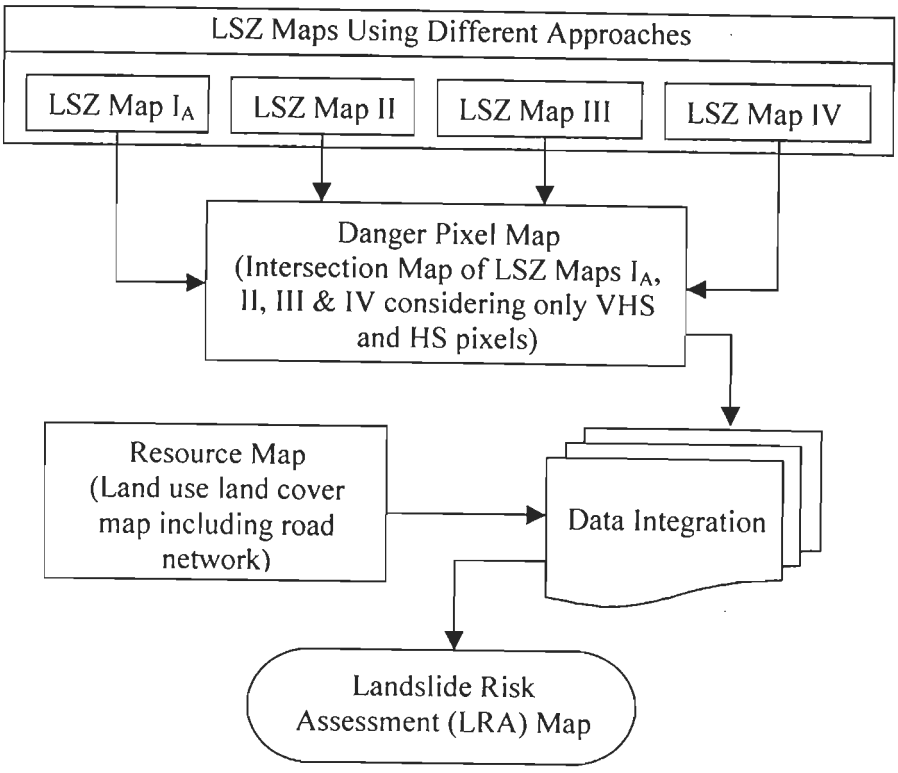


Figure 8.1: Steps for landslide risk assessment (LRA) using danger pixels.

HS zones in all the four LSZ maps – Maps I_A (conventional weighting approach), Map II (ANN black box approach), Map III (fuzzy set based approach) and Map IV (combined neural and fuzzy approach). For generating a danger pixel map, the VHS and HS zones in each LSZ map were merged together and the remaining landslide susceptibility zones (MS, LS and VLS) were masked out. Various LSZ maps were then integrated to generate the danger pixel map (Figure 8.2). Hence, the danger pixel map is an intersection map of all the four LSZ maps with (VHS + HS) zones combined. In other words, these pixels do not lie in MS/LS/VLS zone in any of the four LSZ maps (I_A, II, III and IV). The danger pixel map represents pixels that appear to be under real danger from landslide point of view. Hence, the danger pixel map is a binary image with a value of 1 for danger pixels and a value of 0 for masked areas.

The next step consists of generating a resource map that has included all the existing land use land cover categories and the road network of the area. The road network has been digitized from the topographic maps of the area in a vector layer and then rasterized. The land use land cover map (refer Figure 4.19) described earlier in Chapter 4, has been integrated with the road network layer to prepare the resource map of the area (Figure 8.3). This includes different resource categories namely thick forest, sparse forest, tea plantation, agricultural land, barren land, habitation, road, water body and river sand. Hence, the resource map is an image with different numerical attributes for different resource categories. In the danger pixel map (Figure 8.2), out of a total number of 33935 danger pixels, 2585 pixels fall under barren land category. As barren land is not an important resource category from damage point of view, these pixels are ignored for landslide risk assessment. Hence, the remaining 31350 pixels covering different resource categories (habitation, road, agriculture, tea plantation, thick forest and sparse forest) have been considered under risk due to landslides. The danger pixel map and the resource map have been multiplied to generate the LRA map (referred as LRA Map I) of the study area. The LRA Map I (Figure 8.4) shows spatial distribution of different resource categories that appear to be under real danger due to landslides. The distribution of danger pixels in different resource categories is given in Table 8.3.

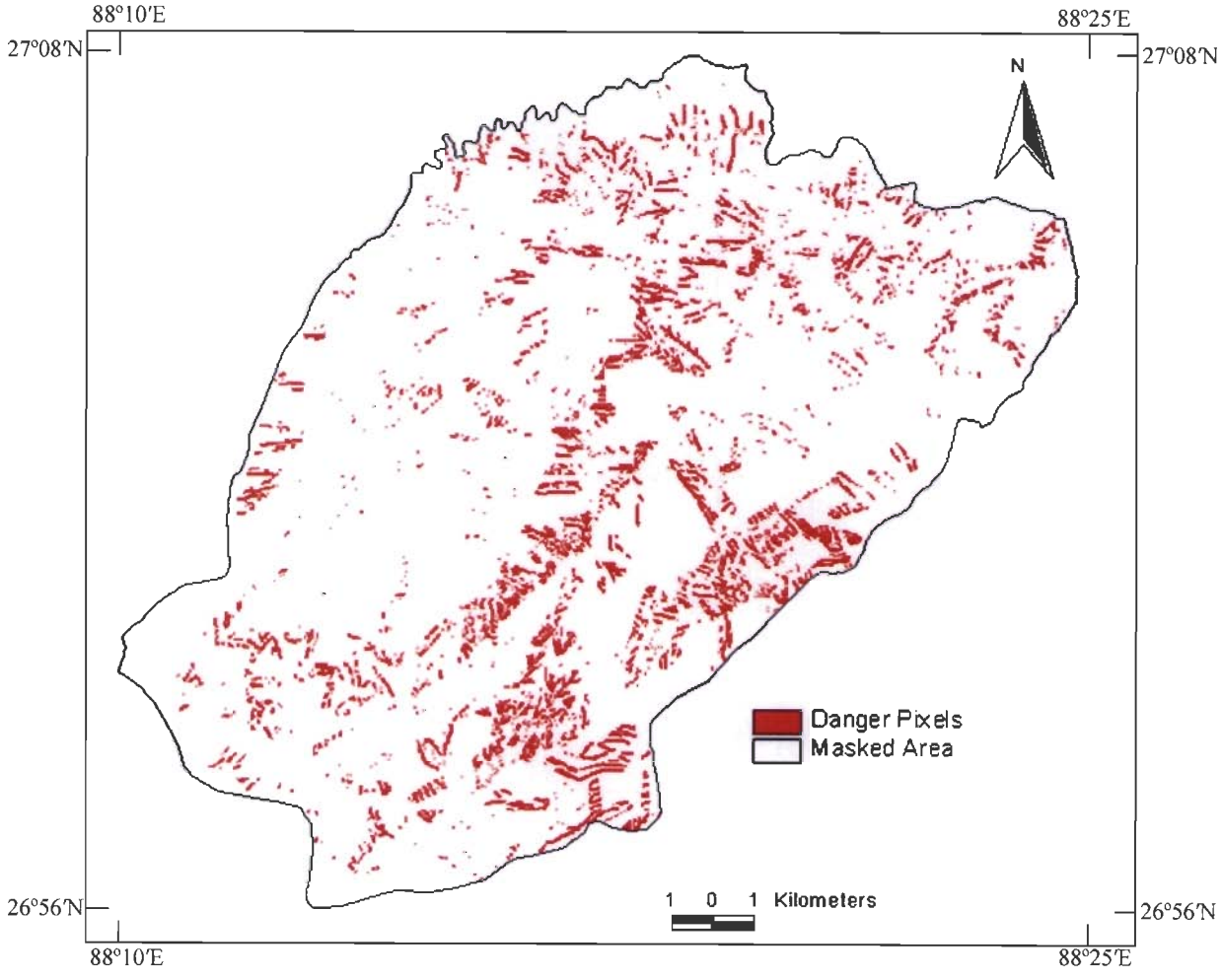


Figure 8.2: Danger pixel map of the study area.

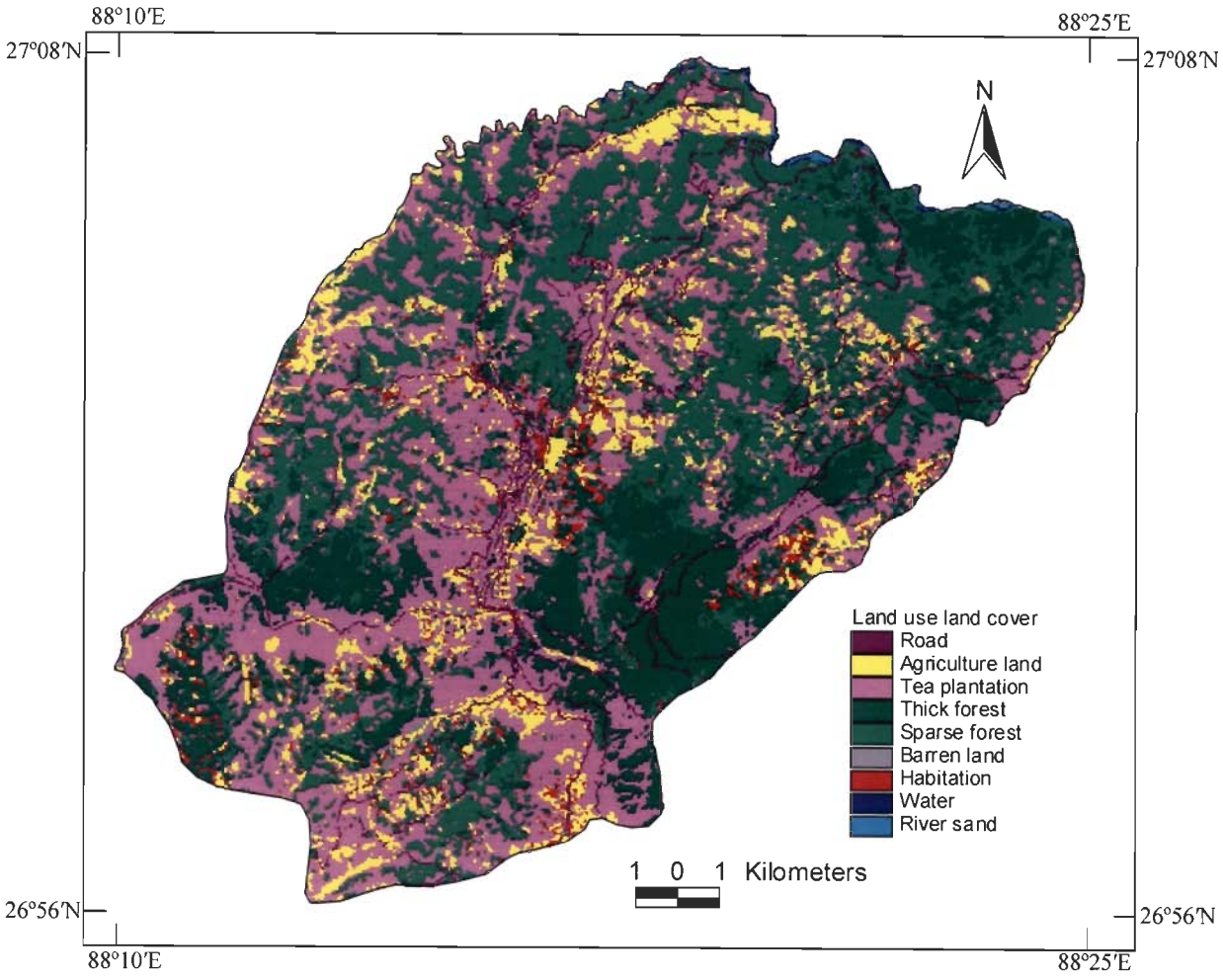


Figure 8.3: Resource map of the study area.

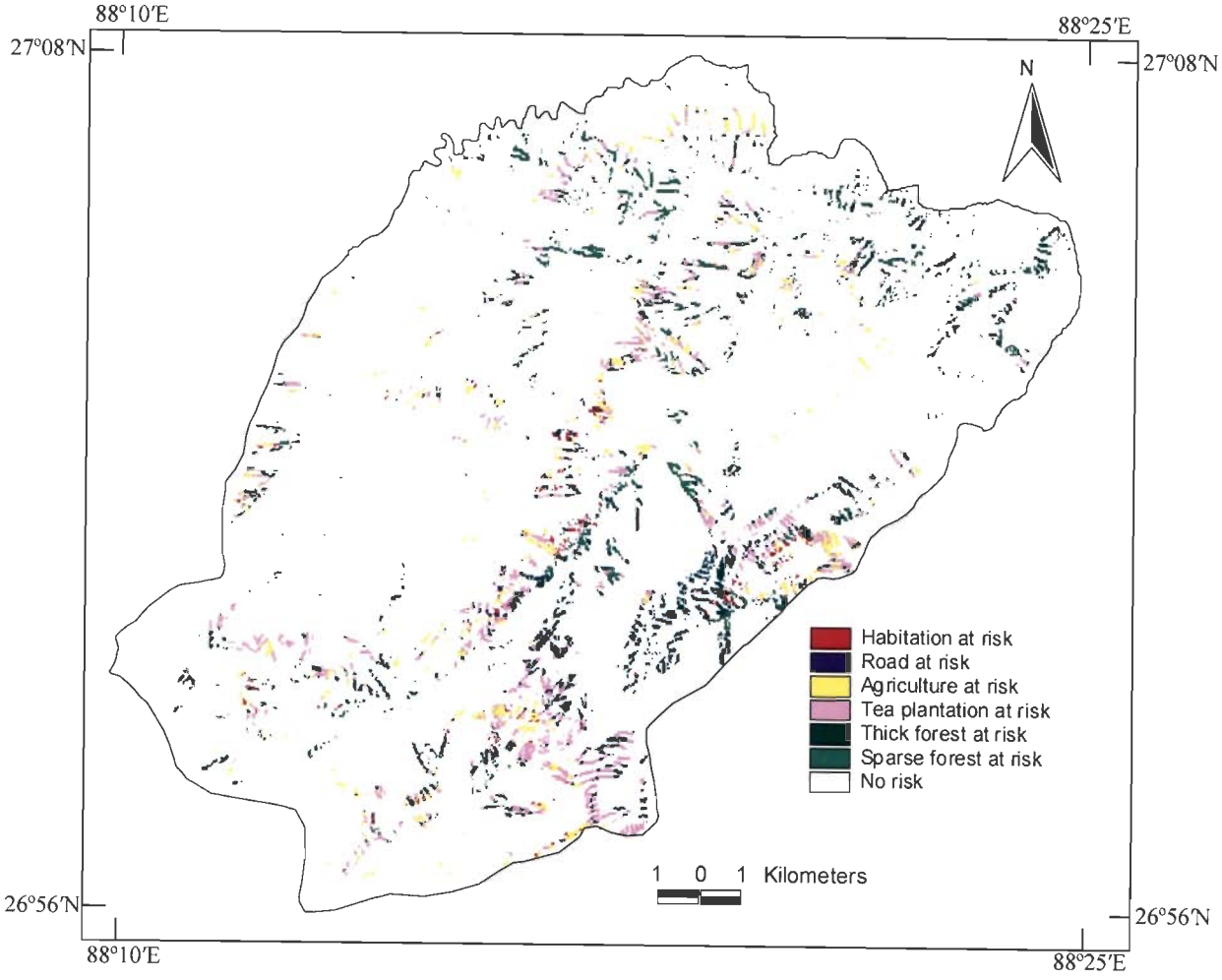


Figure 8.4: Landslide risk assessment map (LRA Map I) using danger pixels.

Table 8.3: Distribution of pixels of various resource categories under risk using danger pixels.

Resource Categories under Risk	Number of Pixels
Habitation	1015
Road	921
Agriculture	4517
Tea Plantation	10517
Thick Forest	5760
Sparse Forest	8620
Total	31350

It is observed from LRA Map I (Figure 8.4) that the habitation around Darjeeling and Ghum are under risk due to landslides. A portion of road from Sonada to Ghum is also under risk due to landslides. Mostly the tea plantation in the southern part and thick forests in the southeastern part of the study area are under risk due to landslides.

8.3.2 LRA Using Fuzzy Concept

This approach is an extension of risk ranking matrices approach of Anbalagan and Singh (1996). According to them, the landslide potential and the damage potential of various resource elements have been categorized into qualitative terms such as very low, low, moderate, high and very high. Also, the risk ranking matrices have been developed in qualitative terms. However, in this approach, the landslide potential and the damage potential of various resource elements can be quantified in terms of fuzzy membership values as per their relative importance to risk assessment. Thus, the risk assessment matrix can be generated with numerical values, which can be classified into different risk zones. In the present study, a qualitative approach based on fuzzy

linguistic rules has been developed and implemented for the generation of LRA map of parts of the study area in a raster based GIS environment.

This approach is a combination of risk scoring and risk matrix. The best LSZ map of the area namely Map IV, prepared using the combined neural and fuzzy approach, has been used as an input layer to provide landslide potential. Further, land use land cover map including the road network has been used as the input layer to correspond to the resource map, which has been used to derive information on resource damage potential.

The fuzzy membership values for landslide susceptibility zones and different land use land cover categories have been assigned on the basis of a linguistic scale derived from expert knowledge. These two layers in terms of their fuzzy membership values have been multiplied in a raster GIS environment to generate a LRA map which has been classified into different risk zones.

The implementation of this approach has been carried out using three different steps; as shown in Figure 8.5:

a) Risk scoring of LSZ Map IV to yield landslide potential raster layer
b) Risk scoring of Resource Map to yield resource damage potential raster layer
c) Generation of LRA Map

a) Risk scoring of LSZ Map IV to yield landslide potential raster layer

The LSZ Map IV (Figure 6.15) has five different landslide susceptibility zones namely VHS, HS, MS, LS and VLS. As per definition, the VHS zone has the highest landslide potential as compared to other susceptibility zones and the VLS zone has the

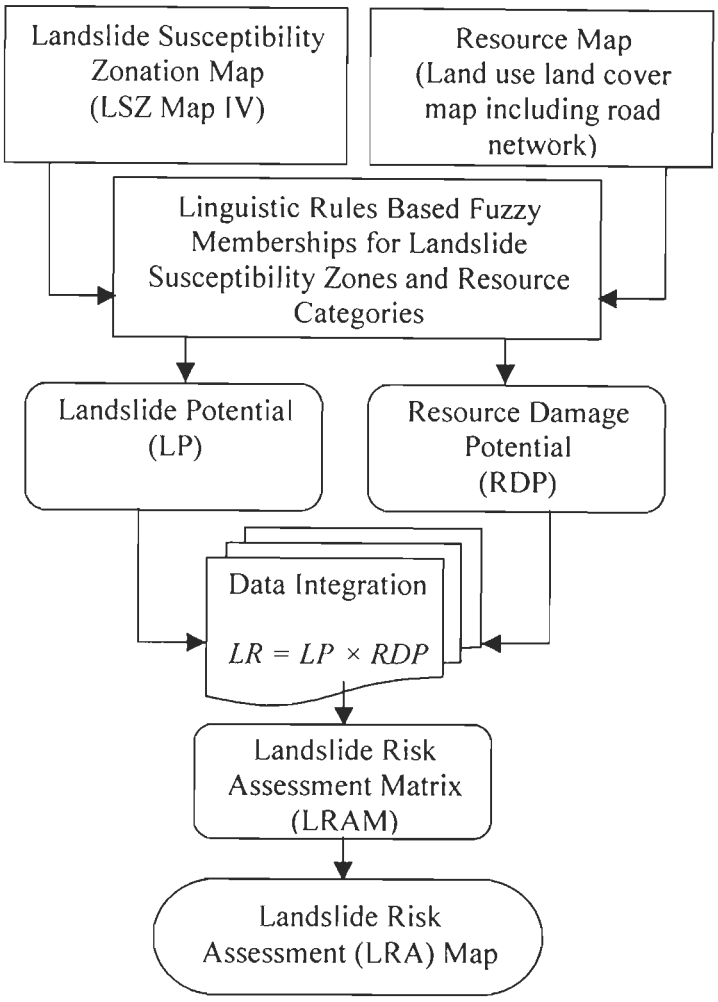


Figure 8.5: Steps for landslide risk assessment (LRA) using fuzzy concept

least landslide potential. Accordingly, linguistic rules have been designed for risk scoring of these susceptibility zones. Fuzzy membership values representing the landslide potential (LP) based on these linguistic rules have been assigned to each susceptibility zone for landslide risk assessment (Table 8.4).

Table 8.4: Linguistic rules for risk scoring of various landslide susceptibility zones.

Landslide Susceptibility Zone	Linguistic Rules for Risk Scoring	Fuzzy Membership value for Landslide Potential
VHS	On-going severe landslides widespread. Landslide almost certain to occur.	1.0
HS	On-going landslide activities evident at many places. Most likely occurrence of landslides under adverse conditions.	0.8
MS	Landslides have occurred in the past locally. Possible occurrence of landslides under adverse conditions.	0.55
LS	Landslides unlikely to occur. Slopes are generally stable.	0.3
VLS	Very rare or no occurrence of landslides. Inherently stable slopes naturally.	0.1

b) Risk scoring of resource map to yield resource damage potential raster layer

The resource map (Figure 8.3) has nine different categories namely habitation, road, thick forest, sparse forest, tea plantation, agricultural land, barren land, water body and river sand. These resource categories will be subjected to landslides resulting in resource damages. The damage potential ought to be related to the importance of these categories to the society. Keeping this in view, the habitation area (buildings and property etc.) has been assigned the highest damage potential, since the damage in this case could be in the form of direct deaths, injuries and property loss. The next important resource category is the road network, as it will have a direct

impact on the essential infrastructure/services (i.e., road network) of the society. The damage in this case will result in lack of connectivity in the area and also will affect the rescue and rehabilitation process during post-disaster management stage. The agricultural land and tea plantation areas have been treated at par, given the same damage potential due to landslides. In this case, there is not likely to be any direct impact on the population, but the effect will be on the economy. The resource categories like thick forest and sparse forest have been considered to be of quite similar importance, as far as damage potential is considered, as these will not have any direct impact on the society. The rest of the categories like barren land, water body and river sand have been considered as having little landslide damage potential.

Accordingly, the linguistic rules have been developed for risk scoring of these resource categories and the fuzzy membership values representing the resource damage potential (RDP) based on these linguistic rules have been assigned to each resource category for landslide risk assessment (Table 8.5).

c) Generation of LRA map

It has been mentioned earlier that the landslide risk is a combination of landslide potential and resource damage potential at a particular site. In the earlier paragraphs, landslide potential and resource damage potential have been given fuzzy membership values. Thus, there are two input layers: one corresponding to landslide potential and another corresponding to resource damage potential. In both layers, each pixel has fuzzy membership values. Hence, landslide risk can be obtained by the following equation:

$$LR = LP \times RDP \tag{8.2}$$

where landslide risk, landslide potential and resource damage potential are represented by *LR*, *LP* and *RDP* respectively. Thus, landslide risk values for different combinations of landslide potential and resource damage potential can be represented in the form of an LRA Matrix (Table 8.6).

Table 8.5: Linguistic rules for risk scoring of various resource categories for damage potential.

Resource Category	Linguistic Rules for Risk Scoring	Fuzzy Membership Value for Damage Potential
Habitation	Direct impact on population and assets such as buildings and property etc. Damages in the form of deaths, injuries and property loss.	1.00
Road	Impact on essential infrastructure/services (i.e., road network). Damage in the form of lack of connectivity in the area, that could also affect the rescue and rehabilitation process during post-disaster management stage.	0.60
Agriculture	Direct impact on economy (earnings) and essential food items for survival.	0.35
Tea plantation	Direct impact on economy (earnings).	0.35
Thick forest	Loss of forest resource of the nation, though no direct impact on individual economy.	0.15
Sparse forest	Loss of some forest resource of the nation, though no direct impact on individual economy.	0.15
Barren land	Little damage	0.05
Water	Little damage	0.05
River sand	Little damage	0.05

Table 8.6: LRA matrix for different combinations of landslide potential and resource damage potential. (red - very high risk, pink – high risk, yellow – moderate risk, blue – low risk and green – very low risk).

		Landslide Potential (LP)				
		VHS (1.0)	HS (0.8)	MS (0.55)	LS (0.3)	VLS (0.1)
Resource Damage Potential (RDP)	Habitation (1.0)	1.0	0.8	0.55	0.3	0.1
	Road (0.6)	0.6	0.48	0.33	0.18	0.06
	Agriculture (0.35)	0.35	0.28	0.19	0.11	0.04
	Tea Plantation (0.35)	0.35	0.28	0.19	0.11	0.04
	Thick Forest (0.15)	0.15	0.12	0.08	0.05	0.02
	Sparse Forest (0.15)	0.15	0.12	0.08	0.05	0.02
	Barren (0.05)	0.05	0.04	0.03	0.02	0.01
	River Sand (0.05)	0.05	0.04	0.03	0.02	0.01
	Water (0.05)	0.05	0.04	0.03	0.02	0.01

It is observed from the above table that the LRA values have a range from 1.00 to 0.01. The value is 1.00 for very high landslide potential in habitated areas, and the value is 0.01 for very low landslide potential in resource categories such as barren land, water bodies and river sand. The range of landslide risk values from 0.01 to 1.00 has been segmented into five different landslide risk zones as per the scheme in Table 8.7 and the LRA map (referred as LRA Map II) (Figure 8.6) of the area has been prepared.

Table 8.7: Scheme of segmentation of landslide risk values into various landslide risk zones.

Landslide Risk Values	Landslide Risk Zones	Colour
$0.0 < LR \leq 0.1$	Very Low Risk Zone (VLR)	Green
$0.1 < LR \leq 0.2$	Low Risk Zone (LR)	Blue
$0.2 < LR \leq 0.4$	Moderate Risk Zone (MR)	Yellow
$0.4 < LR \leq 0.6$	High Risk Zone (HR)	Pink
$LR > 0.6$	Very High Risk Zone (VHR)	Red

The LRA Map II has been superimposed on the resource map to determine the spatial distribution of different risk zones in various resource categories (Table. 8.9).

Table 8.8: Spatial distribution of risk zones and resource categories

Risk Zones	Number of pixels in different resource categories (% of total area)									Total number of pixels (% of total area)
	Habita -tion	Road	Agricul -ture	Tea Planta -tion	Thick Forest	Sparse Forest	Barren Land	Water	River Sand	
VHR	2496 (0.61)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	2496 (0.61)
HR	4422 (1.09)	2782 (0.68)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	7204 (1.77)
MR	2820 (0.69)	7140 (1.75)	13040 (3.20)	27415 (6.72)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	50415 (12.36)
LR	0 (0.00)	4797 (1.18)	20246 (4.96)	108097 (26.51)	16246 (3.98)	23627 (5.80)	0 (0.00)	0 (0.00)	0 (0.00)	173013 (42.43)
VLR	8 (0.00)	316 (0.08)	121 (0.03)	527 (0.13)	54570 (13.38)	103735 (25.44)	13319 (3.27)	892 (0.22)	935 (0.23)	174619 (42.83)
Total	9746 (2.39)	15035 (3.69)	33407 (8.19)	136039 (33.36)	70816 (17.36)	127362 (31.24)	13319 (3.27)	892 (0.22)	935 (0.23)	407747 (100)

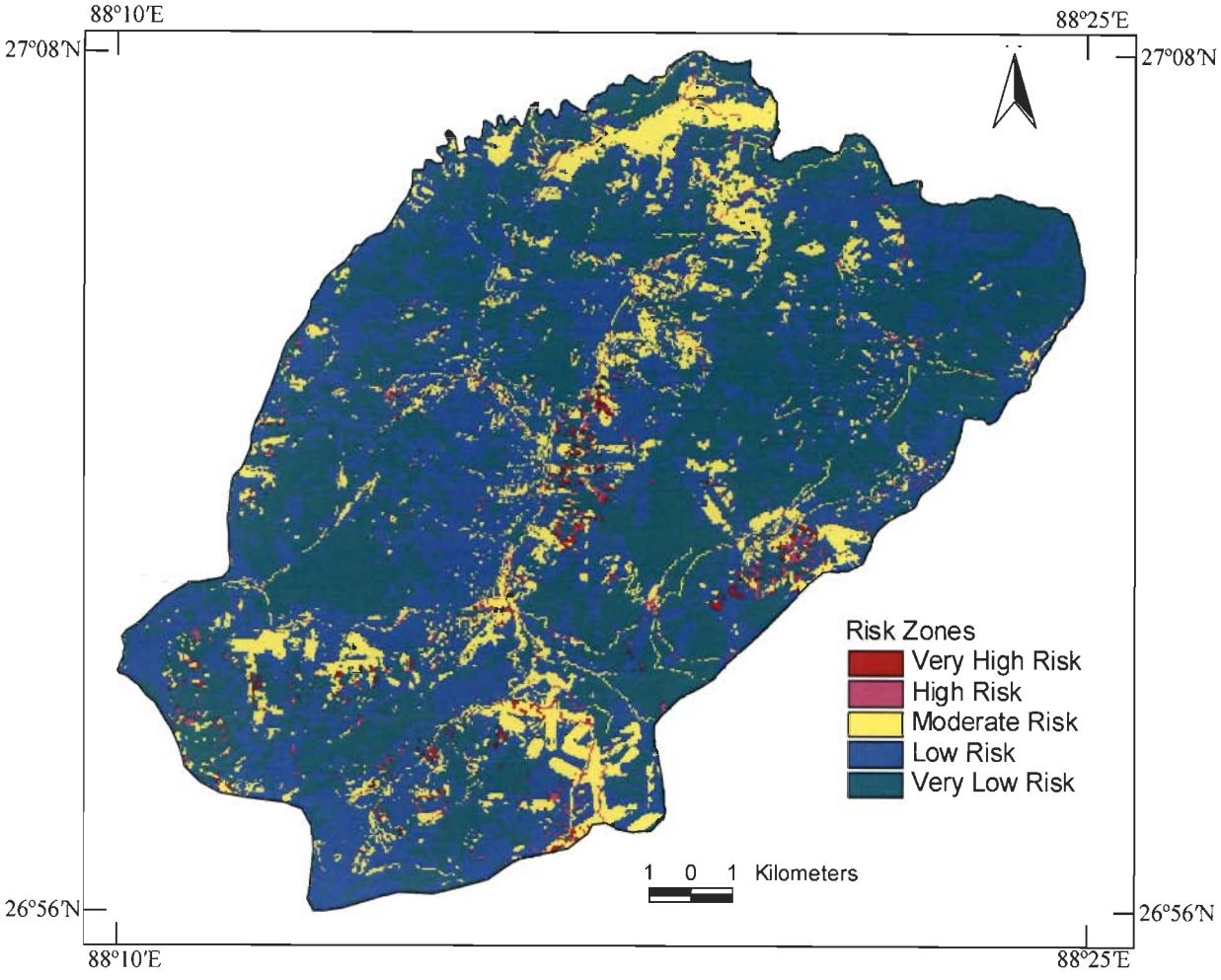


Figure 8.6: Landslide risk assessment map (LRA Map II) of the study area using fuzzy concept.

It can be observed from Table 8.8 that 2496 pixels (0.61% of total area) are under VHR zone. This is due to a combination of habitation resources being located under very high landslide susceptibility zone. Further, 7204 pixels (1.77% of total area) are under HR zone and this comprises partly habitation (4422 pixels) and partly road (2782 pixels). The MR zone occupies 12.36% of the total area (50415 pixels) that is mainly covered by tea plantation (6.72%), agriculture (3.20%), and road (1.75%). An area of 42.43% (173013 pixels) is represented by LR zone that is mainly covered by tea plantation (26.51%), sparse forest (5.80%), agriculture (4.96%) and thick forest (3.98%). Further, a large area of 42.83% (174619 pixels) is represented by VLR zone that is covered mainly by sparse forest (25.44%) and thick forest (13.38%).

Finally, a closer look at the LRA Map II (Fig.8.6) reveals that landslides pose very high/high risk to the following resources:

- Habitation in and around Sonada, Darjeeling and northeastern part of Tiger hill is under VHR zone.
- A section of road from Sonada to Ghum is under HR zone.

8.4 Summary

The LRA Map I produced from danger pixel concept does not infer the degree of severity of risk to different resource categories due to landslides. However, the LRA Map II produced from fuzzy concept depicts different degrees of severity of risk from VHR to VLR for various resource categories due to landslides.

It can be observed that in case of LRA Map I, the habitations in and around Sonada, Darjeeling and northeastern part of Tiger hill is under VHR zone and a section of road from Sonada to Ghum is under HR zone, whereas in case of LRA Map I, these habitations and road section are inferred to be under risk due to landslides.

Thus, broadly both the LRA Maps I and II agree. However, other finer details such as roads in some other sections of Darjeeling-Lamahata, Darjeeling-Limbu Basti and Darjeeling-Rishihat have been inferred under HR zone.

Further, four different LSZ maps produced from conventional weighting, ANN black box, fuzzy set based and combined neural and fuzzy approaches have been used as an input to generate LRA Map I using danger pixel concept whereas only the LSZ map produced from combined neural and fuzzy approach (as has been found to be the best LSZ map of the area) has been used as an input to generate the LRA Map II using fuzzy concept. Hence, both these LRA Maps I and II can not be compared spatially.

Chapter 9

Summary and Conclusions

Landslides happen to be the most common natural hazards in the Himalayan regions that cause damage to both property and life every year. The study of landslides has drawn global attention mainly due to increasing awareness of its socio-economic impacts and also due to increasing pressure of urbanization on the mountain environment. Hence, identification of landslide-prone areas is essential for safer strategic planning of future developmental activities. Therefore, landslide susceptibility zonation (LSZ) becomes important whereby an area is divided into different potential zones due to landslides and mass movements. The relative importance of factors (weights) and their categories (ratings) plays a vital role in LSZ studies. These weights and ratings can be determined by implementing different approaches, which at times are very subjective in nature. Therefore, developing a suitable approach for determination of weights and ratings objectively and their

implementation in a geographic information system (GIS) environment for LSZ mapping is highly important and is an active research area even today.

Several approaches for LSZ mapping have been developed which can be grouped into two broad categories such as qualitative and quantitative. In qualitative approaches, subjective decision rules are applied to define weights and ratings based on the experience of experts (Nagarajan et al. 1998; Gupta et al. 1999; Rautela and Lakhera, 2000; Saha et al., 2002; Sarkar and Kanungo, 2004; etc.).The qualitative approaches can be sub-grouped into three major categories namely distribution analysis, geomorphic analysis and map combination approaches. To remove subjectivity in qualitative analysis, some quantitative approaches have been employed to determine the weights and ratings of factors and their categories respectively for LSZ studies. These quantitative approaches can be sub-grouped into five major categories namely statistical analysis, deterministic analysis, probabilistic models, distribution free approaches and landslide frequency analysis. During the last five years, bivariate statistical models (Lin and Tung, 2003; He et al., 2003; Suzen and Doyuran, 2004; Saha et al., 2004; etc.), multivariate statistical models (Dhakal et al., 2000; Clerici et al., 2002; etc.) and probabilistic prediction models (Lee et al., 2002a; 2002b; Chi et al., 2002a; Lan et al., 2004; etc.) have been implemented for LSZ studies. Apart from these approaches, distribution-free approaches such as fuzzy set based models (Chi et al., 2002b; Gorsevski et al., 2003; Tangestani, 2003; Ercanoglu and Gokceoglu, 2004; Metternicht and Gonzalez, 2005), artificial neural network models (Arora et al., 2004; Gomez and Kavzoglu, 2005; Yesilnacar and Topal, 2005) and combined neural and fuzzy models (Elias and Bandis, 2000; Lee et al., 2004) have also been proposed for LSZ studies.

Research Objectives and Scope of the Work

The literatures on LSZ mapping indicates that different approaches have been adopted in different parts of the world. In view of the gaps identified based on the available literatures, the main objective of the present research was to develop an approach for LSZ mapping leading to risk assessment through the use of advanced approaches and their implementation within the domain of remote sensing and GIS. Specific objectives of the present study were therefore enumerated as:

i) Implementation of conventional weighting approach, ANN black box approach, fuzzy set based approach and combined neural and fuzzy approach for LSZ mapping.

ii) A rigorous evaluation of LSZ maps generated from different approaches to examine their relative efficacy.

iii) Development and implementation of fuzzy set based approach for landslide risk assessment.

Study Area, Data Used and Software Tools

The study area is located in Darjeeling Himalayas and is bounded by latitude 26°56′ - 27°8′ N and longitude 88°10′ – 88°25′ E covering an area of about 254 km^2. The Darjeeling Himalayas lie within the Lesser- and Sub-Himalayan belts. The tectonic units in the area occur in inverted order of stratigraphic superposition. The Daling Group of rocks comprises of low-grade metamorphic rocks, which includes slate, phyllite, schist, quartzite, greywacke and epidiorite. The Darjeeling Group consists primarily of foliated gneisses. Rocks of the Paro Sub-group, which have similar lithologic characteristics as that of the Darjeeling Group, are present at lower

elevations. The main land use practice in the area is tea plantation. The area in the eastern part is dominated by thick forest.

Data from following sources was collected to generate various thematic data layers:

a) Remote sensing images from IRS-1C LISS-III multispectral sensor (acquired on 22nd March, 2000) and IRS-1D-PAN sensor (acquired on 3rd April, 2000)

b) Survey of India (SOI) topographic maps at 1:50,000 scale (Sheet Nos. 78 A/4, A/8, B/1) and 1:25,000 scale (Sheet Nos. 78 A/8/2, A/8/3, A/8/6, B/5/1)

c) Published Geological map (Geological Survey of India)

d) Extensive field data on landslides and land use/land cover

The remote sensing images were processed in ERDAS imagine software. The GIS analysis was carried out in ArcView software. The fuzzy set based approach was implemented through MS Excel spreadsheet and the ANN approaches were implemented in MATLAB software.

Thematic Data Layer Preparation

Various thematic data layers pertaining to causative factors were prepared using the above mentioned data for LSZ mapping. These were generated using remote sensing and GIS tools.

Digital elevation model and its derivatives

The Digital Elevation Model (DEM) is an excellent source to derive topographic attributes responsible for landslide activity in a region. Therefore, a DEM was generated by digitization of contours at 20 m and 40 m interval from the topographic maps. The DEM was subsequently used to derive slope and aspect data layers at a spatial resolution corresponding to the nominal pixel size of 25 m (approximately equivalent to IRS LISS III image pixel resolution) using standard processes in raster GIS.

Lithology

Different rock types have varied composition and structure, which contribute to the strength of the material. The stronger rocks give more resistance to the driving forces as compared to the weaker rocks, and hence are less prone to landslides and vice versa. The lithology data layer was prepared by digitizing the polygons from the geo-referenced geological map of Sikkim-Darjeeling area. Minor modifications in lithological boundaries at some places were also incorporated in this vector layer after field verification. This lithology data layer was later rasterized at 25 m spatial resolution to match the nominal resolution of IRS LISS-III image for further processing.

Lineament buffer

Lineaments are the structural features which describe the zone/plane of weakness, fractures and faults along which landslide susceptibility is higher. It has generally been observed that the probability of landslide occurrence increases at sites close to lineaments. Lineaments were interpreted from the PAN and LISS III images.

There was no major thrust/fault reported in the study area, but mega lineaments were identified. A distance buffer map was generated with four buffer zones at 125 m intervals up to 500 m and another buffer zone beyond 500 m to establish the influence of lineaments on landslide occurrence.

Drainage buffer

Many landslides in hilly areas occur due to the erosional activity associated with drainage. Therefore, a drainage data layer was prepared by digitizing the stream lines from the topographic maps. This layer was updated by overlaying it on LISS-III image. The ordering of streams was performed on the basis of Strahler's classification scheme. A distance buffer map with 25 m buffer zone around 1^{st} and 2^{nd} order drainages only was considered for further analysis.

Land use land cover

Land use land cover is also a key factor responsible for landslide occurrences. The incidence of landslide is inversely related to the vegetation density. Eight dominant land use land cover classes namely thick forest, sparse forest, tea plantation, agriculture, barren, built up, water bodies and river sand in the area have been deciphered. The four spectral bands of LISS III image, DEM and NDVI images were integrated to prepare a land use land cover map by a multi-source classification process using the most widely adopted maximum likelihood classifier. The accuracy of the map was 94.7%. A very small portion of the study area was covered by cloud and its shadow in the LISS III image. Initially, this portion was masked out. Subsequently, masked portion of the land use land cover map thus prepared was filled

with the land use land cover information obtained from topographic maps and field surveys.

Existing landslide distribution

The mapping of existing landslides is essential to study the relationship between the actual landslide distribution in the area and the causative factors. High spatial resolution IRS-1C-PAN and PAN-sharpened LISS-III images were used to produce a landslide distribution map, which was verified from field surveys. A total of 101 landslides showing areas occupied by sliding activity were identified. The majority of landslides have areal extent of 500 m^2 to 2000 m^2.

Landslide Susceptibility Zonation (LSZ) Maps

Four different approaches (Conventional weighting, ANN black box, fuzzy set based and combined neural and fuzzy approaches) were implemented to generate LSZ maps.

Conventional Weighting Approach

The most commonly used conventional weighting approach involved assignment of weights and ratings to the thematic data layers and their categories based on the knowledge of the study area and the opinions of experts on the subject. The weighted data layers were generated by multiplying the weight of the layer with the ratings of the corresponding categories of each layer. A Landslide Susceptibility Index (LSI) map was generated by arithmetically integrating the weighted layers. The LSI values ranged from 21 to 310, which were categorized into five susceptible zones with class boundaries at 68, 137, 176 and 236 using segmentation by natural breaks.

Further, success rate curves method for segmentation was also adopted to fix the boundaries of landslide susceptibility zones statistically and to minimise subjectivity in arbitrarily selecting the natural boundaries of different zones and accordingly an LSZ map was prepared. A comparison of the outputs of both the segmentation methods was carried out.

Both the LSZ maps did not show any definite pattern for the distribution of susceptible zones. It was found that VHS and HS zones together occupied 33.3% of total area and contained 58.7% of existing landslide area in case of segmentation using natural breaks, whereas in case of success rate curves method of segmentation, VHS and HS zones together occupied 28.1% of total area and contained 45.1% of existing landslide area. Finally, the LSZ map prepared through conventional weighting approach using segmentation by natural breaks was used for comparative evaluation with other LSZ maps.

ANN Black Box Approach

A feed forward multi-layer ANN with one input layer, two hidden layers and one output layer was designed. The input layer contained 6 neurons each representing a causative factor. The output layer contained a single neuron corresponding to existing landslide locations. But, due to rather less number of existing landslides as identified from field and remote sensing images, the neural network accuracies were very low. Therefore, the LSZ Map obtained from conventional weighting approach was considered as reference map, to derive representative sample of pixels with known LSZ class. Two independent training and testing datasets were formed. Each dataset consisted of 2500 mutually exclusive pixels corresponding to 500 pixels per landslide susceptibility zone of the reference map. The Levenberg-Marquardt back-

propagation algorithm was used to train the neural networks. A total of 39 neural network architectures were designed, trained and tested. The network architecture 6/13/7/1 (i.e., 6 neurons in input layer, 13 neurons in 1st hidden layer, 7 neurons in 2nd hidden layer and one neuron in output layer) with high training data accuracy (correlation coefficient of 0.918, RMSE 0.112 and 74.4% correct) and high testing data accuracy (correlation coefficient of 0.896, RMSE 0.126 and 72.6% correct) was found to be the most appropriate one. This network was used to prepare the LSZ map of the entire study area.

It is found that VHS and HS zones together occupied 34.6% of total area and contained 50.7% existing landslide area. There is lot of similarity between LSZ maps prepared using conventional weighting approach and ANN black box approach. This may be due to the fact that the conventional one was used as the reference map for generating the ANN black box based LSZ map.

Fuzzy Set Based Approach

In the fuzzy set based approach, ratings of each category of a given thematic layer were determined using the concept of fuzzy relation. The cosine amplitude similarity method was used to determine the membership degrees of categories by establishing the strength of relationship (r_{ij}) between the existing landslides and the categories.

By assigning the ratings of the 35 categories, 35 images of r_{ij} were generated. The corresponding r_{ij} images for various categories of a thematic layer were composited together to generate an R_l image for that thematic layer, where l varies from 1 to t thematic layers (e.g., 6 thematic layers in the present case). The integration of these 6 thematic layers (R_l images) was performed to obtain landslide susceptibility

index (LSI) using arithmetic overlay operation. The range of LSI values obtained varied from 0.014 to 0.252 and were divided into five landslide susceptibility zones using success rate curves method. Accordingly, the most appropriate landslide susceptibility zone boundaries were fixed at LSI values of 0.111, 0.137, 0.163 and 0.189. The spatial distribution of existing landslides in the LSZ map showed that VHS zone occupied 6.1% of total area and contained 41.0% of existing landslide area. Further, HS and VHS zones together occupied 28.8% of the total area and contained 66.1% of existing landslide area.

To bring fuzziness in the integration process also, fuzzy gamma operator was also used. However, it yielded inferior results and was not pursued further. Hence, the LSZ map prepared through fuzzy set based approach and integrated using arithmetic overlay operation was used for further analysis and comparison.

Combined Neural and Fuzzy Approach

This approach involved three main stages:

(a) Determination of weights of thematic layers through ANN connection-weight analysis.

(b) Determination of ratings for categories using cosine amplitude similarity concept.

(c) Integration of ratings and weights in GIS to generate an LSZ map.

In this approach, the ANN connection weights are used to characterize the input data sources (e.g., the thematic layers) in terms of ranks or weights. The connection weight matrices for input-hidden, hidden-hidden and hidden-output layers are obtained for a two-hidden layer network. Simple matrix multiplications of these weight matrices give rise to the final weight matrix corresponding to the causative

factors. The absolute values of these weights were considered to rank the factors meaning thereby that the factor with maximum absolute weight is assigned as rank 1 and the factor with the minimum absolute weight as rank 6.

A feed forward back-propagation ANN with one input layer, two hidden layers and one output layer was considered. The data for the input neurons corresponded to the normalized ratings (r_{ij}) of the categories. The output corresponded to the presence or absence of landslide at the pixel. 100 neural network architectures were designed, trained and tested. The adjusted weights of input-hidden, hidden-hidden and hidden-output connections for each network were captured and analyzed to obtain the weights for thematic layers corresponding to six causative factors. These factors were ranked according to their absolute weights for each network. Considering all the 100 networks, the rank of a factor was decided based on the rank observed by the maximum number of networks (majority rule). Subsequently, the normalized average of the weights of these networks at a scale of 0-10 for a particular factor was calculated and assigned as the weight of that factor. The arithmetic integration of six thematic data layers representing the ratings of the categories (obtained from fuzzy set based approach) and weights for the layers was done to obtain the LSI for each pixel. The LSI values ranged from 0.030 to 0.408. Here also, the success rate curve method was used to classify the LSI values into five different susceptibility zones to produce the LSZ map. Accordingly, the most appropriate boundaries of landslide susceptibility zones were fixed at LSI values of 0.208, 0.253, 0.299 and 0.344.

In this LSZ map, VHS zone occupied only 2.3% of the total study area and contained 30.1% of landslide area. Further, VHS and HS zones together occupied 22.5% of the total area and contained 62% of existing landslide area. This LSZ map showed preferential distribution of higher landslide susceptibility zones along

structural discontinuities (lineaments), which should indeed be the case. Also, the Darjeeling gneiss rock type in south-eastern part, feldspathic greywacke and Reyang quartzite in the northern part of the study area clearly indicated moderate to very high susceptibility zones.

Comparative Evaluation of LSZ Maps

The LSZ maps produced using four different LSZ mapping approaches namely conventional weighting approach, ANN black box approach, fuzzy set based approach and combined neural and fuzzy approach were rigorously evaluated using three different approaches:

a) Landslide density analysis

b) Error matrix analysis

c) Difference image analysis.

Landslide Density Analysis

Landslide density is defined as the ratio of the percent existing landslide area to percent area of each landslide susceptibility zone, and can be calculated on the basis of the number of pixels. It was found that the LSZ Maps produced from conventional and ANN black box approaches had a similar trend of landslide densities for various susceptibility zones. This result was on expected lines, as the conventional weighting based LSZ map was used as the reference map to train ANN. Landslide densities of VHS zone for fuzzy and combined neural and fuzzy based LSZ Maps were found to be much higher as compared to those of other susceptibility zones. There was also a decreasing trend of landslide density values from VHS zone to VLS zone for these maps. The LSZ Map produced from combined neural and fuzzy approach had a much

higher landslide density (>13) of VHS zone as compared to other LSZ maps. Based on the landslide density analysis, it is inferred that the LSZ map produced from combined neural and fuzzy approach appears to be significantly better than those produced from other approaches (fuzzy, conventional and ANN black box approaches).

Error Matrix Analysis

Three different error matrices for different LSZ map combinations were generated to understand the distribution of number of pixels in different LSZ maps. There was a high degree of matching in the pixels of LSZ Maps produced from conventional and ANN black box approaches, particularly for VHS, HS, MS and LS zones, which was expected. There was also a general correspondence between the two LSZ Maps produced from fuzzy and combined neural and fuzzy approaches. The VHS zone of LSZ Map produced from combined neural and fuzzy approach showed much focused population in comparison to LSZ Map produced from fuzzy approach. This is responsible for some mismatches in the VHS zone. The match or mismatch in number of pixels between these two LSZ Maps may be due to the fact that the weights of the factors are considered as constant (i.e., equal importance for all the factors) in case of fuzzy set based approach, whereas in case of combined neural and fuzzy approach the factors have varied importance in terms of weights. There was a lot of mismatch in number of pixels between LSZ Maps produced from conventional and combined neural and fuzzy approaches. This mismatch may be attributed to the differences in weight and rating assignment procedures of conventional weighting approach versus combined neural and fuzzy approach.

Difference Image Analysis

Difference image analysis elucidates how pixels shift from one landslide susceptibility zone to another zone, based on the LSZ mapping approach adopted. Three different combinations of LSZ maps were taken for mutual comparison. A difference image of LSZ maps produced from conventional and ANN black box approaches showed a high degree of mutual correspondence and matching of landslide susceptibility zones throughout the area. This was in agreement with the error matrix result. About 77.3% pixels had full mutual matching and 22.2% pixels exhibited one-zone difference. Barely 0.5% pixels had two-zone difference, and these appeared to be related to a lithologic band in the northern part of the area. A difference image of maps produced from fuzzy and combined neural and fuzzy approaches showed a high degree of spatial matching. About 49.5% pixels had full mutual matching and 47.5% pixels exhibited only one-zone difference. About 3.0% pixels had two-zone difference and these mainly appeared to be related to a lithologic band in the northern part of the area. The two LSZ maps slightly differed from each other because of the differences in the weights of the factors in these two approaches. As lithology has the highest and significantly higher weight than other factors in combined neural and fuzzy approach, the importance of lithology has been brought out in the difference image. Since, there was not much difference between ANN derived weights for lineament buffer, slope, aspect, land use land cover and drainage buffer and unity weights in fuzzy set based approach, the impact of these factors was limited to fully matching and one-zone difference classes only in this difference image. The difference image of maps produced from conventional and combined neural and fuzzy approaches appeared to exhibit the widest spatial difference, where only 37.8% pixels were found to be fully matching, 46.4% pixels exhibited one-zone

difference, 14.6% pixels had two-zone difference and 1.2% pixels had three-zone difference. The most important was a two-zone difference band in the northern part of the difference image marking a lithologic band in the study area. This is because of the fact that lithology has the highest weight in combined neural and fuzzy approach compared to a lower weight in conventional approach based LSZ map. Further, a large number of two-zone differences are also seen in the western and southern parts of the area, which are apparently related to drainage buffer. In the south-eastern part, most pixels exhibit only no difference or one-zone difference. This is interpreted to be related to the fact that drainage lines follow lineaments. In brief, the LSZ mapping and comparative analysis has brought out the relative advantages of fully objective approach vis-à-vis other approaches.

The following broad conclusions may be drawn based on the comparative evaluation of different LSZ maps produced from conventional weighting, ANN black box, fuzzy and combined neural and fuzzy approaches:

a) The LSZ maps produced from conventional and ANN black box approaches are quite similar to each other, as indicated by landslide density values, error matrix and difference image. This is also expected as the LSZ map of conventional approach was used as the reference map for producing the LSZ map using ANN black box approach.
b) The LSZ maps produced from fuzzy set based and combined neural and fuzzy approaches exhibit similarity to each other in terms of landslide density values, error matrix and difference image analysis. This is because both the maps have been generated using objective data processing techniques.

c) The LSZ maps produced from conventional weighting and combined neural and fuzzy approaches are found to exhibit the widest mutual spatial differences. This is again considered to be related to the manner in which the two maps have been generated – the previous one being based on highly subjective conventional weighting approach and the later one being derived from fully objective combined neural and fuzzy approach.

d) The LSZ map produced from combined neural and fuzzy approach is considered to be the best LSZ map of the area. This is because of the fact that it has a much higher landslide density value for VHS zone (> 13), as compared to other maps (1.63 for Map I, 1.34 for Map II, and 6.72 for Map III), and has a more systematic and reasonable trend of decreasing landslide values from VHS through VLS zones.

e) As far as the effects of various thematic layers on the spatial patterns of LSZ maps are concerned, it is observed that each one of the LSZ maps genuinely reflects the relative weights of the input thematic layer. For example, in the conventional weighting approach, drainage buffer has been given the highest weight, and this is well seen on LSZ map produced from conventional weighting approach as well as the LSZ map produced from ANN black box approach. Further, in the combined neural and fuzzy approach, the highest weight pertains to lithology, followed by lineament buffer, and the effects of these causative factors are well observed on the LSZ map.

f) A difference image analysis of LSZ maps produced from conventional weighting and combined neural and fuzzy approaches is highly revealing. Firstly, the importance of lithology, which was not so obvious from field data (in LSZ map using conventional approach), is brought to light by the fully objectively derived LSZ map using combined neural and fuzzy approach. Further, it is seen that there

is a coincidence of lineament buffer vis-à-vis drainage buffer at places, i.e., the same pixels are treated under drainage buffer in the LSZ map produced from conventional approach and under lineament buffer in the LSZ map produced from combined neural and fuzzy approach. This implies that drainage at places follows lineaments, i.e. lineaments have led to fracturing of the terrain along which drainage has developed. Although in the field, drainage lines appear to control the distribution of landslides, the objective spatial domain regional analysis reveals that the real feature of importance is the lineament.

Thus, the comparative analysis of LSZ maps shows the limitation of conventional weighting approach, where weights are assigned based on field observations, which has its obvious limitations of limited perspective views (e.g. drainage being given highest weight etc.). The fully objective approach (combined neural-fuzzy) on the other hand could bring out in an unbiased manner the relative importance (weights) of thematic data layers (lithology and lineament). Therefore, the above analysis elucidates the relative advantages of fully objective approach vis-à-vis conventional weighting approach for LSZ mapping

Landslide Risk Assessment

In the present study, landslide risk is considered to be a function of landslide potential or susceptibility and the resource damage potential. Two different approaches namely (1) LRA using danger pixels and (2) LRA using Fuzzy Concept, for landslide risk assessment were developed and implemented to prepare the LRA maps of the study area.

LRA using danger pixels

A novel concept of danger pixel was introduced for landslide risk assessment. *Danger pixels* are considered as those pixels which lie in VHS and HS zones in all four LSZ maps produced from different approaches. The danger pixel map is an intersection map of all the four LSZ maps with (VHS + HS) zones combined. A resource map including all the existing land use land cover types (thick forest, sparse forest, tea plantation, agricultural land, barren land, habitation, water body and river sand) and the road network of the area was also prepared. In this case, danger pixel map contributed to landslide potential and the resource map contributed to resource damage potential. The danger pixel map and the resource map were integrated to generate the LRA map of the study area. The LRA map showed spatial distribution of different resource categories that appeared to be under real danger due to landslides. It can be observed from the LRA map that the habitation around Darjeeling and Ghum are under risk due to landslides. A portion of road from Sonada to Ghum is also under risk due to landslides. Mostly the tea plantation in the southern part and thick forests in the southeastern part of the study area are under risk due to landslides. The LRA map produced using danger pixels does not infer the degree of severity of risk to different resource categories due to landslides.

LRA using Fuzzy Concept

The landslide potential and the damage potential of various resource elements can be quantified in terms of fuzzy membership values as per their relative importance to risk assessment. Thus, the risk assessment matrix can be generated with numerical values, which can be classified into different risk zones. In the present study, a qualitative approach based on fuzzy linguistic rules has been developed and

implemented for the generation of LRA map of parts of the study area in a raster based GIS environment.

In this approach, LSZ map, prepared using the fully objective combined neural and fuzzy approach and which was found to be the best LSZ map of the area, was used as an input to provide landslide potential. Further, the resource map was used as another input layer to derive information on resource damage potential. Linguistic rules were developed for risk scoring of landslide susceptibility zones and resource categories and the fuzzy membership values representing the landslide potential (LP) and resource damage potential (RDP) based on these linguistic rules were assigned to each susceptibility zone and resource category for landslide risk assessment. Landslide risk values for different combinations of landslide potential and resource damage potential were obtained by integrating LP and RDP layers and were represented in the form of a LRA matrix. The range of landslide risk values was segmented into five different landslide risk zones and the Landslide Risk Assessment (LRA) map of the area was prepared. It was observed that 2496 pixels (0.61% of total area) is under very high risk zone. This is due to a combination of habitation resources being located under very high landslide susceptibility zone at selected places. Further, 7204 pixels (1.77% of total area) is under high risk zone and this comprises partly habitation (4422 pixels) and partly road (2782 pixels). The LRA Map revealed that landslides pose very high risk to selected sites of habitation in Sonada, Darjeeling and northeastern part of Tiger hill, and HR to a section of road from Sonada to Ghum. The LRA Map produced from fuzzy concept depicts different degrees of severity of risk from VHR to VLR for various resource categories due to landslides.

Thus, it can be concluded that the LRA map produced from danger pixel concept does not infer the degree of severity of risk to different resource categories

due to landslides whereas the LRA map produced from fuzzy concept depicts different degrees of severity of risk from very high risk to very low risk. Broadly, both the LRA maps agree in identifying some habitations and road section under risk due to landslides. However, finer details with different degrees of severity of risk can be inferred from the LRA map produced using fuzzy concept.

Scope for Future Research

Briefly, as far as methodology is concerned, the use of fuzzy and combined neural and fuzzy approaches has been the novelty of this research, as this has increased the objectivity in the weight assignment process for LSZ mapping. Further, image based comparative analysis of LSZ maps and approaches involving 'danger pixel' concept and fuzzy concept for risk assessment are other important contributions in this work. Nevertheless, this study has led to some further research areas which can be targeted in future. These can be enumerated as,

a) Increase in the computational efficiency of neural and fuzzy set based approaches for LSZ mapping. Further refinements can be made so that these can be used at operational level.
b) Development of technique for detail mapping of site specific risk elements with the help of high resolution satellite stereo data aimed at quantitative risk assessment.

References

Acharya, S.K., (1989) The Daling Group, its nomenclature, tectono-stratigraphy and structural grain: with notes on their possible equivalents, *Geological Survey of India*, SPN-22, 5-13.

Aleotti, P., Baldelli, P. and Polloni, G., (1996a) Landsliding and Flooding Event Triggered by Heavy Rains in the Tanaro Basin (Italy). In: *Proceedings International Congress Interpraevent*, Garmisch-PartenKirchen, 1, 435-446.

Aleotti, P. and Chowdhury, R., (1999) Landslide Hazard Assessment: Summary, Review and New Perspectives, *Bulletin of Engineering Geology & Environment*, 58, 21-44.

Anbalagan, R., (1992) Landslide Hazard Evaluation and Zonation Mapping in Mountainous Terrain, *Engineering Geology*, 32, 269-277.

Anbalagan, R. and Singh, B., (1996) Landslide Hazard and Risk Assessment Mapping of Mountainous Terrains – A case study from Kumaon Himalaya, India, *Engineering Geology*, 43, 237-246.

Anderson, J. M., Hardy, E. E., Roach, J. T. and Witmert, R. E., (1976) A Land Use Classification System for use with Remote Sensing Data, *U. S. Geological Survey Professional Paper*, No. 964, Washington DC: Government Printing Office.

Apan, A. A., (1997) Land Cover Mapping for Tropical Forest Rehabilitation Planning using Remotely-Sensed Data, *International Journal of Remote Sensing*, 18(5), 1029-1049.

Arora, M.K. and Agarwal, K., (2002) A Computer Program for Sampling Design to Assess Image Classification Accuracy, *Photogrammetry Journal of Finland*, 18(1), 33-43.

Arora, M.K., Das Gupta, A.S. and Gupta, R.P., (2004) An Artificial Neural Network Approach for Landslide Hazard Zonation in the Bhagirathi (Ganga) Valley, Himalayas, *International Journal of Remote Sensing*, 25(3), 559-572.

Arora, M.K. and Mathur, S., (2001) Multi-source Classification using Artificial Neural Network in a Rugged Terrain, *GeoCarto International*, 16(3), 37-44.

Atkinson, P.M. and Massari, R., (1998) Generalised Linear Modelling of Susceptibility to Landsliding in the Central Apennines, Italy, *Computers & Geosciences*, 24(4), 373-385.

Atkinson, P.M. and Tatnall, A.R.L., (1997) Neural networks in remote sensing, *International Journal of Remote Sensing*, 18, 699-709.

Bernknopf, R. L., Campbell, R. H., Brookshire, D. S. and Shapiro, C. D., (1988) A Probabilistic Approach to Landslide Hazard Mapping in Cincinnati, Ohio, with Application for Economical Evaluation, *Bulletin American Association of Engineering Geologists*, 25, 39-56.

Binaghi, E., Luzi, L., Madella, P., Pergalani, F. and Rampini, A., (1998) Slope Instability Zonation: A comparison between Certainty Factor and Fuzzy Dempster-Shafer Approaches, *Natural Hazards*, 17, 77-97.

Blanc, R. P. and Cleveland, G. B., (1968) Natural Slope Stability as related to Geology, San Clemente Area, Orange and San Diego Counties, California, *California Division of Mines and Geology Special Report*, 98, 19p.

Boggett, A. D., Mapplebeck, N. J. and Cullen, R. J., (2000) South Shore Cliffs, Whitehaven – Geomorphological Survey and Emergency cliff Stabilisation Works, *Quarterly Journal of Engineering Geology and Hydrogeology*, 33, 213-226.

Bonham-Carter, G. F., (1994) *Geographic Information Systems for Geoscientists: Modelling with GIS*, Pergamon, Ottawa, 398p.

Bowman, H. N., (1972) Natural Slope Stability in the city of Greater Woolongong, *N.S.W. Geological Survey Records*, 14(2), 159-222.

Brabb, E.E., (1984) Innovative Approaches to Landslide Hazard and Risk Mapping. In: *Proceedings 4th International Symposium on Landslides*, Toronto, Canada, 1, 307-324.

Brabb, E. E., (1991) The World Landslide Problem, *Episodes*, 14(1), 52-61.

Brabb, E. E., Pampeyan, E. H. and Bonilla, M. G., (1972) Landslide Susceptibility in San Mateo County, California, U.S. *Geological Survey Miscellaneous Field Studies Maps*, MF-360.

Bruzzone, L., Conese, C., Maselli, F. And Roli, F., (1997) Multi-source Classification of Complex Rural areas by Statistical and Neural-Network Approaches, *Photogrammetric Engineering and Remote Sensing*, 63(5), 523-533.

Canuti, P., Frascati, F., Garzonio, C. A. and Rodolfi, C., (1979) Dinamica Morfologica di un Ambiente Sogetto a Fenomeni Franosi e ad Intensa Attiva Agricola, *Consiglio Nazionale delle Ricerche*, Perugia, Italy, 142, 81-102.

Cardinali, M., Reichenbach, P., Guzzetti, F., Ardizzone, F., Antonini, G., Galli, M., Cacciano, M., Castellani, M. and Salvati, P. (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.

Carmassi, F., Liberati, G., Ricciardi, C. and Sciotti, M., (1992) Stability Evaluation for unified Power Plant siting in Geothermal Areas. In: *Proceeding of 6th International Symposium on Landslides*, Christchurch, 1, 893-898.

Carrara, A., (1983) Multivariate Models for Landslide Hazard Evaluation, *Mathematical Geology*, 15(3), 403-426.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V. and Reichenbach, P., (1990) Geographical Information System and Multivariate Models in Landslide Hazard Evaluation. In: *Proceeding 6th International Conference and Field Workshop on Landslides*, Switzerland-Austria-Italy, 1, 17-28.

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 Process & Landforms*, 16, 427-445.

Carrara, A. and Guzzetti, F. (Eds.), (1995) *Geographical Information Systems in Assessing Natural Hazards,* Kluwer Academic Publisher, Dordrecht, The Netherlands, 353p.

Carrara, A. and Merenda, L., (1976) Landslide Inventory in Northern Calabria, Southern Italy, *Bulletin Geological Society of America*, 87, 1153-1162.

Chau, K. T., Sze, Y. L., Fung, M. K., Wong, W. Y., Fong, E. L. and Chan, L. C. P., (2004) Landslide Hazard Analysis for Hong Kong using Landslide Inventory and GIS, *Computers and Geosciences*, 30, 429-443.

Chavez, P. S. Jr., (1988) An Improved Dark Object Subtraction Technique for Atmospheric Correction of Multispectral Data, *Remote Sensing of Environment*, 24, 459-479.

Chi, K-H., Park, N-W. and Lee, K., (2002a) Identification of Landslide Area using Remote Sensing Data and Quantitative Assessment of Landslide Hazard. In: *Proceedings of IEEE International Geosciences and Remote Sensing Symposium*, 19 July, Toronto, Canada.

Chi, K-H., Park, N-W. and Chung, C-J., (2002b) Fuzzy Logic Integration for Landslide Hazard Mapping using Spatial Data from Boeun, Korea. In: *Proceedings of Symposium on Geospatial Theory, Processing and Applications*, Ottawa.

Choubey, V. D. and Litoria, P. K., (1990) Terrain Classification and Land Hazard Mapping in Kalsi-Chakrata Area (Garhwal Himalaya), India, *ITC Journal*, 1, 58-66.

Chung, C-J. F. and Fabbri, A. G., (1993) The representation of Geoscience Information for data integration, *Nonrenewable Resources*, 2(2), 122-139.

Chung, C-J. F. and Fabbri, A. G., (1998) Three Bayesian Prediction Models for Landslide Hazard. In: Buccianti, A. (ed.) *Proceeding of International Association for Mathematical Geology Annual Meeting (IAMG'98)*, Ischia, Italy, 204-211.

Chung, C-J. F. and Fabbri, A. G., (1999) Probabilistic Prediction Models for Landslide Hazard Mapping, *Photogrammetric Engineering & Remote Sensing*, 65(12), 1389-1399.

Chung, C-J. F. and Leclerc, Y., (1994) A Quantitative Technique for Zoning Landslide Hazard. *Proceeding of International Association for Mathematical Geology Annual Conference*, Mont Tremblant, Quebec, 87-93.

Civco, D. L., (1989) Topographic Normalization of Landsat Thematic Mapper Digital Imagery, *Photogrammetric Engineering & Remote Sensing*, 55(9), 1303-1309.

Clerici, A., Perego, S., Tellini, C. and Vescovi, P., (2002) A Procedure for Landslide Susceptibility Zonation by the Conditional Analysis Method, *Geomorphology*, 48, 349-364.

Coates, D. R., (1977) Landslide Perspectives. In: Coates, D. R. (ed.) *Landslides: Reviews in Engineering Geology 3*, Geological Society of America, Boulder, Colorado.

Colby, J. D., (1991) Topographic Normalization in Rugged Terrain, *Photogrammetric Engineering & Remote Sensing*, 57(5), 531-537.

Congalton, R. G., (1991) A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37, 35-46.

Cooke, R. U. and Doornkamp, J. C., (1990) *Geomorphology in Environmental Management*, Clarendon Press, Oxford, 410p.

Coppin, N. and Richards, I., (1990) *Use of Vegetation in Civil Engineering.* Butterworths, London.

Cotecchia, V., (1978) Systematic Reconnaissance Mapping and Registration of Slope Movements, *Bulletin of International Association of Engineering Geology*, 17, 5-37.

Crozier, M. J., (1986) *Landslides: Causes, Consequences and Environment*, Croom Helm, Australia, 252p.

Cruden, D. M. and Varnes, D. J., (1996) Landslide Types and Processes (Chapter 3). In: Turner, A. K. and Schuster, R. L. (eds.) *Landslides-Investigation and Mitigation*, Transportation Research Board Special Report 247, National Research Council, USA, 36-75.

Csaplovics, E., (1998) High Resolution Space Imagery for Regional Environmental Monitoring – Status quo and Future Trends, *International Archives of Photogrammetry and Remote Sensing*, 32(7), 211-216.

Curran, P. J. and Foody, G. M., (1994) The use of Remote Sensing to Characterize the Regenerative states of Tropical Forests. In: Foody, G. and Curran, P. (eds.) *Environmental Remote Sensing from Regional to Global Scales*, John Wiley and Sons, Chichester, 44-83.

Davis, J.C., (1986) *Statistics and Data Analysis in Geology*, John Wiley & Sons, New York, 646p.

Demers, M. N., (2000) *Fundamentals of Geographic Information Systems*, 2nd Edition, John Wiley & Sons, New York, US, 498p.

Dikau, R., Brunsden, D., Schrott, L. and Ibsen, M.L. (Eds.), (1996) *Landslide Recognition: Identification, Movement and Causes*, John Wiley & Sons, Chichester, UK, 251p.

Dhakal, A. S., Amada, T. and Aniya, M., (2000) Landslide Hazard Mapping and its Evaluation using GIS: An Investigation of Sampling Schemes for a Grid-cell based Quantitative Method, *Photogrammetric Engineering & Remote Sensing*, 66(8), 981-989.

Dietrich, E. W., Reiss, R., Hsu, M. L. and Montgomery, D. R., (1995) A Process-based Model for Colluvial Soil Depth and Shallow Landsliding using Digital Elevation Data, *Hydrological Processes*, 9, 383-400.

Dobrovolny, E. and Schmoll, H. R., (1974) Slope Stability Map of Anchorage and Vicinity, Alaska, *U.S. Geological Survey Miscellaneous Inventory Maps*, I-787-E.

Donati, L. and Turrini, M. C., (2002) An Objective Method to Rank the Importance of the Factors Predisposing to Landslides with the GIS Methodology:

Application to an Area of the Apennines (Valnerina: Perugia, Italy), *Engineering Geology*, 63, 277-289.

Dubois, D. and Prade, H., (1980) *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York.

Eiumnoh, A. and Shrestha, P., (2000) Application of DEM Data to Landsat Image Classification: Evaluation in a Tropical Wet-Dry Landscape of Thailand, *Photogrammetric Engineering & Remote Sensing*, 66(3), 297-304.

Elias, P.B. and Bandis, S.C., (2000) Neurofuzzy Systems in Landslide Hazard Assessment, In: *Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, July 2000, 199-202.

EOSAT, (1994) *Merge Process Enhances Value of Data Sets*, EOSAT Notes, 9, 5.

Ercanoglu, M. and Gokceoglu, C., (2004) Use of Fuzzy Relations to Produce Landslide Susceptibility Map of a Landslide Pron Area (West Black Sea Region, Turkey), *Engineering Geology*, 75(3&4), 229-250.

Espizua, L. E. and Bengochea, J. D., (2002) Landslide Hazard and Risk Zonation Mapping in the Rio Grande Basin, Central Andes of Mendoza, Argentina, *Mountain Research and Development*, 22(2), 177-185.

Eyles, G. O., (1983) The Distribution and Severity of Present Soil Erosion in New Zealand, *New Zealand Geographer*, 39(1), 12-28.

Fell, R., (1994) Landslide Risk Assessment and Acceptable Risk, *Canadian Geotechnical Journal*, 31, 261-272.

Fenti, V., Silvano, S. and Spagna, V., (1979) Methodological Proposal for an Engineering Geomorphological Map, Forecasting Rockfalls in the Alps, *Bulletin of International Association of Engineering Geology*, 19, 134-138.

Finlay, P. and Fell, R., (1995) *A Study of Landslide Risk Assessment in Hong Kong*. Report for the geotechnical Engineering Office, Hong Kong.

Fisher, P. F. and Pathirana, S., (1990) The Evaluation of Fuzzy Membership of Land Cover Classes in the Suburban Zone, *Remote Sensing of Environment*, 34(2), 121-132.

Foody, G. M., (2002) Status of Land Cover Classification Accuracy Assessment, *Remote Sensing of Environment*, 80, 185-201.

Foody, G. M. and Arora, M. K., (1996) Incorporating Mixed Pixels in the Training, Allocation and Testing Stages of Supervised Classifications, *Pattern Recognition Letters*, 17, 1389-1398.

Foody, G. M., Campbell, N. A., Trodd, N. M. and Wood, T. F., (1992) Derivation and Applications of Probabilistic Measures of Class Membership from Maximum Likelihood Classification, *Photogrammetric Engineering & Remote Sensing*, 58, 1335-1341.

Frank, T. D., (1988) Mapping Dominant Vegetation Communities in the Colorado Rocky Mountain Front Range with LANDSAT Thematic Mapper and Digital Terrain Data, *Photogrammetric Engineering & Remote Sensing*, 54(12), 1727-1734.

Freund, J. E., (1992) *Mathematical Statistics*, Fifth Edition, Printice-Hall of India Private Limited, New Delhi, India.

Gomez, H. and Kavzoglu, T., (2005) Assessment of Shallow Landslide Susceptibility using Artificial Neural Networks in Jabonosa River Basin, Venezuela, *Engineering Geology*, 78(1-2), 11-27.

Gong, P., (1996) Integrated Analysis of Spatial Data for Multiple Sources: using Evidential Reasoning and Artificial Neural Network Techniques for Geological Mapping, *Photogrammetric Engineering & Remote Sensing*, 62, 513-523.

Gorsevski, P. V., Gessler, P. E. and Jankowski, P., (2003) Integrating a Fuzzy k-means Classification and a Bayesian Approach for Spatial Prediction of Landslide Hazard, *Journal of Geographical Systems*, 5, 223-251.

Greenbaum, D., Tutton, M., Bowker, M., Browne, T., Buleka, J., Greally, K., Kuna, G., McDonald, A., Marsh, S., O'Connor, E. and Tragheim, D., (1995) Rapid Methods of Landslide Hazard Mapping: Papua New Guinea Case Study, *British Geological Survey*, Technical Report WC/95/27.

Gupta, R.P., (2003) *Remote Sensing Geology*, 2nd Edition, Springer-Verlag, Berlin Heidelberg , Germany, 655p.

Gupta, R.P. and Joshi, B.C., (1990) Landslide Hazard Zonation using the GIS Approach – A case Study from the Ramganga Catchment, Himalayas, *Engineering Geology*, 28, 119-131.

Gupta, V., Sah, M.P., Virdi, N.S. and Bartarya, S.K., (1993) Landslide Hazard Zonation in the Upper Satlej Valley, District Kinnaur, Himachal Pradesh, *Journal of Himalayan Geology*, 4, 81-93.

Gupta, R.P., Saha, A.K., Arora, M.K. and Kumar, A., (1999) Landslide Hazard Zonation in a part of Bhagirathy Valley, Garhwal Himalayas, using integrated Remote Sensing – GIS, *Journal of Himalayan Geology*, 20(2), 71-85.

Guzzetti, F., Carrara, A., Cardinali, M. and Reichenbach, P., (1999) Landslide Hazard Evaluation: A Review of Current Techniques and their Application in a multi-scale Study, Central Italy, *Geomorphology*, 31, 181-216.

Hagan, M. T., Demuth, H. B. and Beale, M. H., (1996) *Neural Network Design*, PWS Publishing, Boston, MA.

Hagan, M. T. and Mehnaj, M., (1994) Training Feedforward Networks with the Marquardt Algorithm, *IEEE Transactions on Neural Networks*, 5(6), 989-993.

Hansen, A., (1984) Landslide Hazard Analysis. In: Brundsen, D. and Prior, D. B. (eds.) *Slope Instability*, Wiley, New York, 523-602.

Hansen, A., Franks, C. A. M., Kirk, P. A., Brimicombe, A. J. and Tung, F., (1995) Application of GIS to Hazard Assessment, with particular reference to Landslides in Hong Kong. In: Carrara, A. and Guzzetti, F. (eds.) *Geographical Information Systems in Assessing Natural Hazards*, Kluwer Academic Publisher, Dordrecht, The Netherlands, 135-175.

Hayashi, C., (1952) On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of View. *Ann. Inst. Statist. Math.*, 4, 69-98.

Haykin, S., (1999) *Neural Networks: A Comprehensive Foundation*, Second Edition. Prentice Hall, New Jersey.

He, Y. P., Xie, H., Cui, P., Wei, F. Q., Zhong, D. L. and Gardner, J. S., (2003) GIS-based Hazard Mapping and Zonation of debris flows in Xiaojiang Basin, Southwestern China, *Environmental Geology*, 45, 286-293.

Hearn, G. J., (1995) Landslide and Erosion Hazard Mapping at Ok Tedi Copper Mine, Papua New Guinea, *Quarterly Journal of Engineering Geology*, 28, 47-60.

Heckerman, D., (1986) Probabilistic Interpretation of MYCIN's Certainty Factors. In: Kanal, L. N. and Lemmer, J. F. (eds.) *Uncertainty in Artificial Intelligence*, Elsevier, New York, 298-311.

Hines, J. W., (1997) *Fuzzy and Neural Approaches in Engineering*. Wiley, New York, 210p.

Holben, B. and Justice, C., (1981) An Examination of Spectral Band Ratioing to Reduce the Topographic Effect on Remotely Sensed Data, *International Journal of Remote Sensing*, 2(2), 115-133.

Hsli, K., (1975) Catastrophic Debris Stream (Sturzstorm) generated by Rockfalls, *Geological Society of America Bulletin*, 86, 129-140.

Hughes, A., Hewlett, H., Samuels, P. G., Morris, M., Sayers, P., Moffat, I., Harding, A. and Tedd, P., (2000) Risk Management for UK Reservoirs. *Construction Industry Research and Information Association (CIRIA) C542*, London.

Humbert, M., (1977) Risk Mapping of areas exposed to Movements of Soil and Sub-soil: French "Zermos" maps, *Bulletin of International Association of Engineering Geologists*, 16, 80-82.

Hutchinson, J. N., (1988) General Report: Morphological and Geotechnical Parameters of Landslides in relation to Geology and Hydrology. In: *Proceeding of Fifth International Symposium on Landslides*, Lausanne, 1, 3-35.

Hutchinson, J. N., (1996) Keynote Paper: Landslide hazard Assessment. In: Bell, D. H. (ed.) *Landslides*, Balkema, Rotterdam, The Netherlands, 3, 1805-1841.

Hutchinson, J. N. and Chandler, M. P., (1991) A Preliminary Landslide Hazard Zonation of the undercliff of the Isle of Wight. In: Chandler, R. J. (ed.) *Slope Stability Engineering, Development and Applications*, Thomas Telford, 197-206.

Ives, J. D. and Messerli, B., (1981) Mountain Hazard Mapping in Nepal: Introduction to an Applied Mountain Research Project, *Mountain Research and Development*, 1, 223-230.

Jade, S. and Sarkar, S., (1993) Statistical Models for Slope Instability Classification, *Engineering Geology*, 36, 91-98.

Janssen, L. F., Jaarsma, J. and Linder, E. van der., (1990) Integrating Topographic Data with Remote Sensing for Land-Cover Classification, *Photogrammetric Engineering & Remote Sensing*, 48(1), 123-130.

Jensen, J. R., (1996) *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd Edition, Prentice-Hall, New Jersey, US, 318p.

Jones, A. R., Settle, J. J. and Wyatt, B. K., (1988) Use of Digital Terrain Data in the Interpretation of SPOT-1 HRV Multispectral Imagery, *International Journal of Remote Sensing*, 9(4), 669-682.

Juyal, G. P., (2002) Landslides in the Uttaranchal Himalaya and their Rehabilitation by Bio-engineering Measures. In: Nainwal, H. C. and Prasad, C. (eds.) *Geodynamics and Environment Management of Himalaya*, HNB Garhwal University, Srinagar Garhwal, India, 182-190.

Kawakami, H. and Saito, Y., (1984) Landslide Risk Mapping by a Quantification Method. In: *Proceedings of 4th International Symposium on Landslides*, Toronto, 2, 535-540.

Kawata, Y., Ueno, S. and Kusaka, T., (1988) Radiometric Correction for Atmospheric and Topographic Effects on Landsat MSS Images, *International Journal of Remote Sensing*, 9, 729-748.

Kienholz, H., (1978) Maps of Geomorphology and Natural Hazard of Griendelwald, Switzerland, scale 1:10,000, *Artic and Alpine Research*, 10, 169-184.

Kienholz, H., Hafner, H., Schneider, G. and Tamrakar, R., (1983) Mountain Hazards Mapping in Nepal's Middle Mountains: maps of Land Use and Geomorphic Damages (Kathmandu-Kakani Area), *Mountain Research and Development*, 3(3), 195-220.

Kienholz, H., Schneider, G., Bichsel, M., Grunder, M. and Mool, P., (1984) Mapping of Mountain Hazards and Slope Stability, *Mountain Research and Development*, 4, 247-266.

Lakhera, R.C. and Champatiray, P. K., (1996) Satellite Census of Landslide prone areas in Himalaya, *Himalayan Geology*, 17, 209-216.

Landry, J., (1979) carte ZERMOS. Zones exposes a des risques lies aux mouvements du sol et du sous-sol, region de Lons-le-Saunier a Poligny (Jura), Orleans, *Bur. de Rech. Geol. et Miniere*, 1 Map, 14.

Lan, H. X., Zhou, C. H., Wang, L. J., Zhang, H. Y. and Li, R. H., (2004) Landslide Hazard Spatial Analysis and Prediction using GIS in the Xiaojiang Watershed, Yunnan, China, *Engineering Geology*, 76, 109-128.

Lee, E. M., (1999) Coastal Planning and Management: the impact of the 1993 Holbeck Hall Landslide, Scarborough, East Midlands Geographer, 21, 78-91.

Lee, E. M., (2003) Coastal Change and Cliff Instability: Development of a Framework for Risk Assessment and Management. Unpublished PhD thesis, University of Newcastle upon Tyne.

Lee, S., Choi, J., Chwae, U. and Chang, B., (2002a) Landslide Susceptibility Analysis using Weight of Evidence. In: *Proceedings of IEEE International Geosciences and Remote Sensing Symposium*, 19 July, Toronto, Canada.

Lee, S., Choi, J. and Min, K. D., (2002b) Landslide Susceptibility Analysis and Verification using the Bayesian Probability Model, *Environmental Geology*, 43, 120-131.

Lee, E. M. and Clark, A. R., (2000) The Use of Archive Records in Landslide Risk Assessment: Historical Landslide events on the Scarborough Coast, UK. In: Bromhead, E. N., Dixon, N. and Ibsen, M. L. (eds.) *Landslides: In Research, Theory and Practice*, Thomas Telford, London, 904-910.

Lee, E. M., Clark, A. R. and Guest, S., (1998a) An Assessment of Coastal Landslide Risk, Scarborough, UK. In: Moore, D. and Hungr, O. (eds.) *Engineering Geology: The View from the Pacific Rim*, 1787-1794.

Lee, S. and Min, K. D., (2001) Statistical Analysis of Landslide Susceptibility at Yongin, Korea, *Environmental Geology*, 40(9), 1095-1113.

Lee, S., Ryu, J., Won, J. and Park, H., (2004) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network, *Engineering Geology*, 71, 289-302.

Lee, E. M. and Zones, D. K. C., (2004) *Landslide Risk Assessment*. Thomas Telford, London, 454 p.

Leone, F., Aste, J. P. And Leroi, E., (1996) Vulnerability Assessment of Elements Exposed to Mass-movement: Working toward a Better Risk Perception. In: Senneset, R. (ed.) *Landslides*, Balkema, Rotterdam, The Netherlands, 263-269.

Leroi, E., (1996) Landslide Hazard-Risk Maps at different scales: Objectives, Tools and Developments. In: Senneset, R. (ed.) *Landslides*, Balkema, Rotterdam, The Netherlands, 35-51.

Lin, M-L. and Tung, C-C., (2003) A GIS-based Potential Analysis of the Landslides induced by the Chi-Chi Earthquake, *Engineering Geology*, 71, 63-77.

Lu, P. F. and An, P., (1999) A Metric for Spatial Data Layers in Favorability Mapping for Geological Events, *IEEE Transaction on Geosciences & Remote Sensing*, 37, 1194-1198.

Luzi, L. and Pergalani, F., (1999) Slope Instability in Static and Dynamic Conditions for Urban Planning: the Oltre Po Pavese Case History (Regione Lombardia-Italy), *Natural Hazards*, 20, 57-82.

Maharaj, R., (1993) Landslide Processes and Landslide Susceptibility Analysis from an upland Watershed: A case study from St. Andrew, Jamaica, West Indies, *Engineering Geology*, 34, 53-79.

Mantovani, F., Soeters, R. and Van Westen, C. J., (1996) Remote Sensing Techniques for Landslide Studies and Hazard Zonation in Europe, *Geomorphology*, 15, 213-225.

Mark, R. K. and Ellen, S. D., (1995) Statistical and Simulation Models for Mapping Debris-flows Hazard. In: Carrara, A. and Guzzetti, F. (eds.) *Geographical Information Systems in Assessing Natural Hazards*, Kluwer Academic Publisher, Dordrecht, The Netherlands, 93-106.

Marston, R., Miller, M. and Devkota, L., (1998) Geoecology and Mass Movements in the Manaslu Ganesh and Langtang-Jural Himals, Nepal, *Geomorphology*, 26, 139-150.

Mather, P. M., (1999) *Computer Processing of Remotely-Sensed Images: An Introduction*, Second Edition. John Wiley and Sons, Chichester.

McDermid, G. and Franklin, S., (1995) Remote Sensing and Geo-morphometric Discrimination of Slope Processes, *Zeitschrift fur Geomorphologie*, 101, 165-185.

McDonnell, B. A., (2002) Hazard Identification and Visitor Risk Assessment at the Giant's Causeway World Heritage Site, Ireland. In: McInnes, R. G. and Jakeways, J. (eds.) *Instability – Planning and Management*, Thomas Telford, London, 527-534.

McKean, J., Buechel, S. and Gaydos, L., (1991) Remote Sensing and Landslide Hazard Assessment, *Photogrammetric Engineering & Remote Sensing*, 57(9), 1185-1193.

Mehrotra, G. S., Sarkar, S. and Dharmaraju, R., (1991) Landslide Hazard Assessment in Rishikesh-Tehri Area, Garhwal Himalaya, India. In: Bell, F. G. (ed.) *Landslides*, Balkema, Rotterdam, The Netherlands, 1001-1007.

Mehrotra, G. S., Sarkar, S., Kanungo, D. P., Mahadevaiah, K., (1996) Terrain Analysis and Spatial Assessment of Landslide Hazards in parts of Sikkim Himalaya, *Geological Society of India*, 47, 491-498.

Mejia-Navarro, M., Wohl, E. E. And Oaks, S. D., (1994) Geological Hazards, Vulnerability and Risk Assessment using GIS: Model for Glenwood Springs, Colorado, *Geomorphology*, 10, 331-354.

Meneroud, J. P. and Calvino, A., (1976) carte ZERMOS. Zones exposes a des risques lies aux mouvements du sol et du sous-sol a 1:25000, region de la Moyenne Vesubie (Alps-Matitimes), Orleans, *Bur. de Rech. Geol. et Minieres*, 1 Map, 11.

Metternicht, G. and Gonzalez, S., (2005) FUERO: Foundations of a Fuzzy Exploratory Model for Soil Erosion Hazard Prediction, *Environmental Modelling & Software*, 20(6), 715-728.

Michelson, D. B., Liljeberg, B. M. and Pilesjo, P., (2000) Comparison of Algorithms for Classifying Swedish Landcover Using Landsat TM and ERS-1 SAR Data, *Remote Sensing of Environment*, 71(1), 1-15.

Nagarajan, R., Mukherjee, A., Roy, A., Khire, M. V., (1998) Temporal remote Sensing Data and GIS Application in Landslide Hazard Zonation of part of Western Ghat, India, *International Journal of Remote Sensing*, 19(4), 573-585.

Naithani, A. K., (1999) The Himalayan Landslides, *Employment News*, 23(47), 20-26 February, 1-2.

Navalgund, R. R. and Kasturirangan, K., (1983) The Remote Sensing Satellite – A Program Overview. In: *Proceedings of Indian Academy of Sciences; Engineering – Sciences – Remote Sensing – III*, 6, 313-336.

Obermier, S. F., (1979) Slope Stability Map of Fairfax County, Virginia, *U.S. Geological Survey Miscellaneous Field Studies Map*, MF-1072.

O'Leary, D., Friedman, J. and Pohn, H., (1976) Lineament, Linear, Lineation: Some Proposed New Standards for Old Terms, *Geological Society of America Bulletin*, 87, 1463-1469.

Pachauri, A. K., Gupta, P. V. and Chander, R., (1998) Landslide Zoning in a part of the Garhwal Himalayas, *Environmental Geology*, 36(3-4), 325-334.

Pachauri, A. K. and Pant, M., (1992) Landslide Hazard Mapping based on Geological Attributes, *Engineering Geology*, 32, 81-100.

Palmer, J. S., Clark, A. R., Cliffe, D. and Eade, M., (2002) The Management of Risk on the Chalk Cliffs at Brighton, UK. In: McInnes, R. G. and Jakeways, J. (eds.) *Instability – Planning and Management*, Thomas Telford, London, 355-362.

Paola, J. D. and Schowengerdt, R. A., (1995) A Review and Analysis of Back-propagation Neural Networks for Classification of Remotely Sensed Multi-spectral Imagery, *International Journal of Remote Sensing*, 16, 3033-3058.

Pike, R. J., (1988) The Geometric Signature: Quantifying Landslide-terrain Types from Digital Elevation Models, *Mathematical Geology*, 20(5), 491-511.

Prakash, A., Fielding, E. J., Gens, R., Genderen, J. L. van and Evans, D. L., (2001) Data Fusion for Investigating Land Subsidence and Coalfire Hazards in a Coal Mining Area, *International Journal of Remote Sensing*, 22(6), 921-932.

Radbruch-Hall, D. H. and Crowther, K. C., (1976) Landslides: Causes and Effect, *International Association Engineering Geologist Bulletin*, 14, 205-216.

Radbruch-Hall, D. H. and Varnes, D. J., (1973) Map showing Areas of Estimated Relative Amounts of Landslides in California, *U.S. Geological Survey Miscellaneous Inventory Map*, I-747.

Rautela, P. and Lakhera, R. C., (2000) Landslide Risk Analysis between Giri and Tons Rivers in Himachal Himalaya (India), *International Journal of Applied Earth Observation & Geoinformation (JAG)*, 2(3-4), 153-160.

Remondo, J., Gonzalez, A., Diaz de Teran, J. R., Cendrero, A., Fabbri, A. and Chung, C-J. F., (2003) Validation of Landslide Susceptibility Maps; Examples and Applications from a case study in Northern Spain, *Natural Hazards*, 30, 437-449.

Rib, H. T. and Liang, T., (1978) Recognition and Identification. In: Schuster, R.L. and Krizek, R. J. (eds.) *Landslides Analysis and Control*, Washington Transportation Research Board, Special Report, National Academic of Sciences , WA, 176, 34-80.

Richards, J. A. and Jia, X., (1999) *Remote Sensing Digital Image Analysis: An Introduction*, 3rd Edition, Springer-Verlag, Heidelberg, Germany, 363p.

Ripley, B., (1996) *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge.

Rodriguez Ortiz, J. M., Hinojosa, J. A. And Prieto, C., (1978) Regional Studies on Mass Movements in Spain. In: *Proceeding of 3rd International Congress, International Association Engineering Geologist,* Section I, 1, 267-277.

Ross, T. J., (1995) *Fuzzy Logic with Engineering Applications*, McGraw-Hill, New York.

Roy, K. K. (1976) Some problems on stratigraphy and tectonics of Darjeeling and Sikkim Himalayas. In: *Recent Geological Studies in the Himalaya, Seminar, 1971*, Geol. Surv. Ind. Misc. Pub., 24(2), 279-394.

Royal Society, (1992) *Risk: Analysis, Perception and Management*, Report of a Royal Society Study Group. Royal Society, London.

Rupke, J., Cammeraat, E., Seijmonsbergen, A. C. and van Westen, C. J., (1988) Engineering Geomorphology of Widentobel Catchment, Appenzell and Sankt Gallen, Switzerland: A Geomorphological Inventory System applied to Geotechnical Appraisal of Slope Stability, *Engineering Geology*, 26, 33-68.

Sabins, F. F., (1996) *Remote Sensing Principles and Interpretations*, W. H. Freeman & Co., New York, US, 494p.

Saha, A. K., Arora, M. K., Csaplovics, E. and Gupta, R. P., (2005) Land Cover Classification Using IRS LISS III Image and DEM in a Rugged Terrain: A Case Study in Himalayas, *Geocarto International*, 20(2), 33-40.

Saha, A. K., Gupta, R. P. and Arora, M. K., (2002) GIS-based Landslide Hazard Zonation in a part of the Himalayas. *International Journal of Remote Sensing*, 23, 357-369.

Saha, A. K., Gupta, R. P., Sarkar, I., Arora, M. K. and Csaplovics, E., (2005) An Approach for GIS-based Statistical Landslide Susceptibility Zonation – with a case Study in the Himalayas, *Landslides*, 2, 61-69.

Sandilands, N. M., Noble, M. and Findlay, J. W., (1998) Risk Assessment Strategies for Dam based Hydro Schemes. In: *The Prospect for Reservoirs in the 21st Century*, Thomas Telford, London.

Sanjeevi, S., Vani, K. and Lakshmi, K., (2001) Comparison of Conventional and Wavelet Transform Techniques for Fusion of IRS-1C LISS-III and PAN Images. In: Proceedings of ACRS 2001 – 22nd Asian Conference on Remote Sensing, 5-9 November, Singapore, 1, 140-145.

Saraf, A. K., (2000) IRS-1C-PAN Depicts Chamoli Earthquake Induced Landslides in Garhwal Himalayas, India, *International Journal of Remote Sensing*, 21(12), 2345-2352.

Sarkar, S. and Kanungo, D. P., (2004) An Integrated Approach for Landslide Susceptibility Mapping using Remote Sensing and GIS. *Photogrammetric Engineering & Remote Sensing*, 70(5), 617-625.

Sarkar, S., Kanungo, D. P. and Mehrotra, G. S., (1995) Landslide Hazard Zonation: A case Study in Garhwal Himalaya, India, *Mountain Research and Development*, 15(4), 301-309.

Sassa, K., (1988) Special Lecture: Geotechnical Model for the Motion of Landslides. In: *Proceeding 5th International Symposium on Landslides*, Lausanne, 1, 37-55.

Schalkoff, R. J., (1997) *Artificial Neural Networks*, Wiley, New York.

Schuster, R., (1996) Socioeconomic significance of landslides. In: Turner, A.K., Schuster, R.L. (eds.) *Landslides: Investigation and Mitigation*, Transportation Research Board, National Research Council, Special Report, National Academic Press, Washington, DC, 247, 12-36.

Seeley, M. W. and West, D. O., (1990) Approach to Geologic Hazard Zoning for Regional Planning, Inyo National Forest, California and Nevada, *Bulletin American Association of Engineering Geologist*, 27(1), 23-35.

Selby, M., (1993) *Hillslope Materials and Processes*. Oxford University Press, Oxford.

Sesagiri, D. N. and Badrinarayan, S., (1982) The Nilgiri Landslides, *Geological Survey of India Miscellaneous Publication*, 57, 32p.

Shalan, M. A., Arora, M. K. and Ghosh, S. K., (2003) An Evaluation of Fuzzy Classifications from IRS 1C LISS III Imagery: a Case Study, *International Journal of Remote Sensing*, 24(15), 3179-3186.

Shanmugam, P. and Sanjeevi, S., (2001) Analysis and Evaluation of Fusion Techniques using IRS-1C LISS-3 and PAN Data for Monitoring Coastal Wetlands of Vedaranniyam, Tamil Nadu. In: *Proceedings of the National Symposium on Advances in Remote Sensing Technology with special emphasis on High Resolution Imagery & Annual Convention of Indian Society of Remote Sensing (ISRS)*, 11-13 December, Ahmedabad, India.

Sharma, P. K., Chopra, R., Verma, V. K. And Thomas, A., (1996) Flood Management using Remote Sensing Technology: The Punjab (India) Experience, *International Journal of Remote Sensing*, 17(17), 3511-3521.

Shortliffe, E. H. and Buchanan, G. G., (1975) A Model of inexact Reasoning in Medicine, *Mathematical Biosciences*, 23, 351-379.

Siddle, H. J., Jones, D. B. and Payne, H. R., (1991) Development of a Methodology for Landslip Potential Mapping in the Rhonda Valley. In: Chandler, R. J. (ed.) *Slope Stability Engineering, Development and Applications*, Thomas Telford, 137-148.

Sinha, B. N., Varma, R. S. and Paul, D. K., (1975) Landslides in Darjeeling District (West Bengal) and Adjacent Areas, *Bulletins of the Geological Survey of India*, Series B, No.36, 1-45.

Skempton, A. W. and Hutchinson, J. N., (1969) Stability of Natural Slopes and Embankment Foundations. In: *Proceeding of Seventh International Conference on Soil Mechanics and Foundation Engineering*, Mexico City, State of the Art Volume, 291-340.

Soeters, R. and van Westen, C. J., (1996) Slope Instability Recognition, Analysis and Zonation. In: Turner, A. K. and Schuster, R. L. (eds.) *Landslides, Investigation and Mitigation*, Transportation Research Board, National Research Council, Special Report 247, National Academy Press, Washington, DC, U.S.A., 129-177.

Strahler, A. N., (1964) Quantitative Geomorphology of Basins and Channel Networks. In: Chow, V.T. (ed.) *Handbook of Applied Hydrology*, Mcgraw Hill, New York.

Strahler, A. H., Logan, T. L. And Bryant, N. A., (1978) Improving Forest Cover Classification Accuracy from Landsat by Incorporating Topographic Information. In: *Proceedings of 12th Symposium Remote Sensing Environment*, Ann Arbor, Michigan, US, 2, 927-942.

Suzen, M. L. and Doyuran, V., (2004) Data driven Bivariate Landslide Susceptibility Assessment using Geographical Information Systems: A Method and Application to Asarsuyu Catchment, Turkey, *Engineering Geology*, 71, 303-321.

Suzen, M. L. and Toprak, V., (1998) Filtering of Satellite Images in Geological Lineament Analyses: An Application to a Fault Zone in Central Turkey, *International Journal of Remote Sensing*, 19(6), 1101-1114.

Takei, A., (1982) Limitation Methods of Hazard Zones in Japan. In: Takei and Aulitzky (eds.) *Report of Japanese-Austrian Joint Research, Forecast of Disaster Zone in Mountainous Regions*, 1980-1981, Kyoto University Laboratory of Erosion Control Research Bulletin, 1, 7-25.

Tangestani, M. H., (2003) Landslide Susceptibility Mapping using the Fuzzy Gamma Operation in a GIS, Kakan Catchment Area, Iran. *Proceeding of Map India Conference 2003.*

Turner, A.K. and Schuster, R.L., (1996) Landslides: Investigation and Mitigation, Transportation Research Board, National Research Council, Special Report 247, *National Academy Press*, Washington, DC, 675p.

Turrini, M. C. and Visintainer, P., (1998) Proposal of a Method to define areas of Landslide Hazard and an Application to an area of the Dolomites, Italy, *Engineering Geology*, 50, 255-265.

van Dine, D. F., Jordan, P. and Boyer, D.C., (2002) An Example of Risk Assessment from British Columbia, Canada. In: McInnes, R. G. and Jakeways, J. (eds.) *Instability – Planning and Management*, Thomas Telford, London, 399-406.

van Westen, C. J., (1993) Application of Geographic Information Systems to Landslide Hazard Zonation, Ph. D. thesis, Technical University of Delft, International Institute for Aerospace Survey and Earth Sciences, The Netherlands, *ITC Publication,* 15(1), 245p.

van Westen, C. J., (1994) GIS in Landslide Hazard Zonation: A Review, with Examples from the Andes of Colombia. In: Price, M. and Heywood, I. (Eds.) *Mountain Environments and Geographic Information System,* Taylor & Francis, Basingstoke, UK, 135-165.

van Westen, C. J., (1997) Statistical Landslide Hazard Analysis. In: *Application Guide, ILWIS 2.1 for Windows*, ITC, Enschede, The Netherlands, 73-84.

van Westen, C. J. and Bonilla, J. B., (1990) Mountain Hazard Analysis Using a PC Based GIS. In: Price, D. G. (ed) *Proceeding Sixth International Congress, International Association of Engineering Geology, Amsterdam*, A. A. Balkema, Rotterdam, The Netherlands, 265-271.

van Westen, C. J., Rengers, N. and Soeters, R., (2003) Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment, *Natural Hazards*, 30, 399-419.

van Westen, C. J., Rengers, N., Terlien, M. T. J., and Soeters, R., (2000) Prediction of the occurrence of Slope Instability Phenomena through GIS-based Hazard Zonation, *Geol Rundsch*, 86, 404-414.

van Westen, C. J., Soeters, R. and Sijmons, K., (2000) Digital Geomorphological Landslide Hazard Mapping of the Alpago Area, Italy, JAG, 2(1), 51-60.

Varnes, D. J., (1978) Slope Movement Types and Processes, Landslides Analysis and Control. *Special Report 176, Transportation Research Board*, Washington, DC, 11-80.

Varnes, D. J., (1984) Landslide Hazard Zonation: A Review of Principles and Practice. *UNESCO*, Paris, 1-63.

Virdi, N. S., Sah, M. P. and Bartarya, S. K., (1997) Mass Wasting, its Manifestations, Causes and Control: some case Histories from Himachal Himalaya. In: Agarwal, D. K., Krishna, A. P., Joshi, V., Kumar, K. and Palni, M. S. (Eds.) *Perspectives of Mountain Risk Engineering in the Himalayan Region*, Gyanodaya Prakashan, Nainital, 111-130.

Welch, R. and Ehlers, M., (1987) Merging Multiresolution SPOT HRV and Landsat TM Data, *Photogrammetric Engineering & Remote Sensing*, 53(3), 301-303.

Wieczorek, G. F., (1984) Preparing a Detailed landslide-Inventory map for Hazard Evaluation and Reduction, *Bulletin of International Association of Engineering Geologists*, 21, 337-342.

Wieczorek, G. F., Gori, P. L., Jager, S., Kappel, W. M. and Negussey, D., (1996) Assessment and Management of Landslide Hazards near Tully Valley Landslide, Syracuse, New York, USA. In: *Proceeding Seventh International Symposium on Landslides*, Trondheim, 1, 411-416.

Wong, H. N., Ho, K. K. S. and Chan, Y. C., (1997) Assessment of Consequence of Landslides. In: Cruden, D. and Fell, R. (eds.) *Landslide Risk Assessment*, Balkema, Rotterdam, 111-149.

Wright, R. H. and Nilsen, T. H., (1974) Isopleth Map of Landslide Deposits, Southern San Francisco Bay Region, California, *US Geological Survey Miscellaneous Field Studies Map*, MF-550 (Scale 1:250,000).

www.em-dat.net EM-DAT: The OFDA/CRED International Disaster Database. Universite Catholique de Louvain, Brussels, Belgium.

Yesilnacar, E. and Topal, T., (2005) Landslide Susceptibility Mapping: A Comparison of Logistic Regression and Neural Networks Methods in a Medium Scale Study, Hendek region (Turkey), *Engineering Geology*, 79, 251-266.

Yin, K. L. and Yan, T. Z., (1988) Statistical Prediction Model for Slope Instability of Metamorphosed rocks. In: Bonnard, C. (ed.) *Proceeding Fifth International Symposium on Landslides*, Lausanne, Balkema, Rotterdam, The Netherlands, 2, 1269-1272.

Zadeh, L. A., (1965) Fuzzy Sets, *IEEE Information and Control*, 8, 338-353.

Zadeh, L. A., (1973) Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Transaction on System Management Cybern.*, SMC-3, 1, 28-46.

Zhou, W., (1999) Verification of the Nonparametric Characteristics of Backpropagation Neural Networks for Image Classification. *IEEE Transaction on Geoscience and Remote Sensing*, 37, 771-779.

Zimmermann, M., Bichsel, M. and Kienholz, H., (1986) Mountain Hazards Mapping in the Khumbu Himal, Nepal, with Prototype Map, Scale 1:50,000, *Mountain Research and Development*, 6(1), 29-40.

Appendix-I

Absolute weights of the factors for 100 networks and their ranks (in brackets) in Combined Neural and Fuzzy Approach (I – Input layer, HA – 1st hidden layer, HB – 2nd hidden layer and O – Output layer)

Network architecture (Number of neurons)				Weights/Ranks					
I	HA	HB	O	Landuse	Lithology	Slope	Slope Aspect	Drainage	Lineament
6	3	1	1	777.28 (5)	2079.90 (2)	2411.26 (1)	1800.36 (3)	223.01 (6)	1566.76 (4)
6	3	2	1	6179.52 (2)	73.08 (6)	9656.00 (1)	4392.46 (3)	392.65 (5)	3429.80 (4)
6	3	3	1	667.32 (2)	273.60 (6)	1275.30 (1)	520.10 (4)	380.92 (5)	592.75 (3)
6	3	6	1	586.66 (1)	528.97 (2)	504.45 (3)	14.38 (6)	259.72 (4)	174.22 (5)
6	3	9	1	1888.67 (2)	509.38 (6)	1414.31 (5)	1657.50 (3)	1656.19 (4)	1963.04 (1)
6	3	12	1	864.19 (3)	997.12 (2)	3803.06 (1)	284.80 (4)	270.54 (5)	256.54 (6)
6	3	15	1	957.52 (3)	161.69 (6)	3231.73 (1)	818.30 (4)	245.78 (5)	2818.30 (2)
6	4	2	1	137.83 (6)	4892.39 (1)	3016.78 (3)	402.67 (5)	538.84 (4)	4356.30 (2)
6	4	3	1	2.52 (6)	1855.51 (1)	734.36 (3)	824.28 (2)	8.97 (5)	409.01 (4)
6	4	4	1	514.19 (4)	2428.08 (1)	1379.22 (2)	58.79 (6)	149.19 (5)	970.97 (3)
6	4	8	1	496.85 (5)	5241.70 (1)	944.50 (2)	926.01 (3)	481.89 (6)	655.68 (4)
6	4	12	1	352.60 (6)	1860.84 (2)	1854.54 (3)	1039.32 (4)	482.77 (5)	2529.14 (1)
6	4	15	1	56.26 (5)	1478.20 (3)	2186.46 (2)	18.18 (6)	88.42 (4)	4454.33 (1)
6	5	4	1	692.31 (4)	4677.18 (1)	936.76 (3)	2439.74 (2)	283.65 (6)	325.64 (5)
6	5	5	1	174.67 (5)	2216.54 (1)	860.39 (3)	106.35 (6)	270.65 (4)	1070.07 (2)
6	5	8	1	80.145 (5)	4486.20 (1)	308.36 (3)	1033.94 (2)	277.32 (4)	12.27 (6)
6	5	10	1	880.42 (2)	1364.38 (1)	330.73 (5)	119.19 (6)	372.14 (4)	727.92 (3)
6	5	11	1	311.73 (4)	1449.69 (1)	390.59 (3)	41.56 (6)	82.32 (5)	681.28 (2)
6	6	2	1	3595.15 (5)	10595.48 (3)	14304.04 (1)	3804.32 (4)	637.57 (6)	12049.52 (2)
6	6	3	1	319.26 (4)	1690.69 (1)	115.99 (6)	1375.69 (2)	269.99 (5)	813.06 (3)
6	6	8	1	35.46 (6)	295.15 (4)	641.91 (2)	352.56 (3)	90.78 (5)	848.14 (1)
6	6	15	1	220.44 (5)	2626.90 (1)	2364.05 (2)	315.42 (4)	200.13 (6)	1579.98 (3)

6	7	15	1	473.81 (5)	5488.43 (1)	7.92 (6)	665.88 (3)	646.88 (4)	1505.49 (2)
6	7	11	1	2250.56 (3)	7353.68 (1)	1704.30 (4)	980.09 (5)	485.81 (6)	4894.85 (2)
6	7	10	1	1524.44 (2)	249.56 (5)	578.35 (4)	740.41 (3)	76.49 (6)	2112.65 (1)
6	7	8	1	1546.08 (3)	1372.87 (4)	1159.06 (5)	3274.82 (1)	31.00 (6)	1760.37 (2)
6	7	7	1	4095.99 (4)	19934.87 (1)	4441.44 (3)	2336.50 (6)	3173.95 (5)	9108.41 (2)
6	7	6	1	451.95 (5)	957.00 (3)	1089.00 (1)	508.33 (4)	187.49 (6)	979.97 (2)
6	7	5	1	266.49 (5)	353.18 (4)	582.74 (3)	933.54 (1)	134.67 (6)	896.44 (2)
6	7	4	1	1099.73 (3)	149.46 (5)	136.19 (6)	3016.52 (1)	791.59 (4)	2807.38 (2)
6	8	4	1	361.18 (4)	198.46 (5)	3517.52 (1)	362.51 (3)	111.93 (6)	1544.56 (2)
6	8	6	1	853.70 (3)	505.24 (4)	1786.32 (1)	243.47 (5)	138.59 (6)	968.34 (2)
6	8	8	1	552.82 (4)	770.79 (3)	7074.27 (1)	978.35 (2)	319.76 (6)	545.31 (5)
6	8	10	1	324.42 (5)	2474.24 (1)	452.20 (4)	1089.35 (3)	238.49 (6)	1412.21 (2)
6	8	13	1	157.73 (6)	1582.24 (2)	1385.46 (3)	3444.34 (1)	267.97 (5)	1271.30 (4)
6	8	15	1	451.96 (5)	1905.89 (2)	3503.26 (1)	679.70 (4)	226.35 (6)	728.56 (3)
6	9	4	1	1065.73 (5)	15498.44 (1)	8155.96 (2)	1.39 (6)	2689.41 (4)	2834.79 (3)
6	9	5	1	893.35 (4)	1361.85 (2)	2646.25 (1)	624.76 (5)	406.81 (6)	1232.99 (3)
6	9	6	1	499.22 (5)	10688.35 (1)	3621.90 (3)	4024.76 (2)	363.46 (6)	1996.34 (4)
6	9	9	1	47611.92 (2)	6563.81 (5)	75822.35 (1)	316.92 (6)	20861.53 (3)	18647.96 (4)
6	9	11	1	111.03 (5)	2200.99 (1)	929.69 (2)	633.12 (4)	102.36 (6)	759.94 (3)
6	9	12	1	254.01 (6)	6710.73 (1)	2384.98 (2)	2203.75 (3)	808.75 (5)	887.17 (4)
6	9	15	1	184.32 (5)	6626.27 (1)	2556.05 (2)	1917.29 (4)	53.11 (6)	2310.42 (3)
6	10	7	1	1111.93 (4)	4130.81 (1)	517.87 (5)	1310.90 (3)	204.16 (6)	1456.27 (2)
6	10	8	1	1442.70 (5)	17247.35 (1)	7239.24 (2)	2805.78 (3)	1237.61 (6)	2038.46 (4)
6	10	9	1	329.64 (4)	3981.79 (1)	1947.38 (2)	4.34 (6)	183.47 (5)	581.05 (3)
6	10	10	1	154.35 (6)	381.64 (5)	2355.6 (1)	1846.99 (2)	573.02 (4)	995.78 (3)
6	10	11	1	115.11 (6)	2769.11 (1)	404.17 (4)	442.36 (3)	345.11 (5)	1204.41 (2)
6	10	13	1	296.16 (5)	1456.71 (2)	697.81 (4)	1098.74 (3)	226.65 (6)	1841.11 (1)
6	10	15	1	222.49 (6)	3424.24 (1)	529.38 (3)	303.56 (5)	308.61 (4)	1072.01 (2)
6	11	6	1	230.12 (6)	1817.27 (3)	2530.74 (2)	2589.36 (1)	232.93 (5)	638.15 (4)

6	11	9	1	175.06 (6)	5159.88 (1)	2626.81 (2)	2430.67 (3)	362.24 (5)	2082.62 (4)
6	11	11	1	304.92 (6)	470.10 (4)	624.13 (3)	342.83 (5)	1130.48 (2)	2501.71 (1)
6	11	12	1	121.63 (6)	1127.53 (3)	1204.2 (2)	949.84 (4)	314.16 (5)	1653.06 (1)
6	11	14	1	334.18 (6)	17369.92 (2)	20142.8 (1)	2858.97 (4)	3012.91 (3)	1536.55 (5)
6	11	15	1	10.98 (6)	102.95 (5)	883.29 (3)	1644.43 (1)	271.20 (4)	1494.54 (2)
6	12	5	1	31.68 (6)	190.06 (4)	498.57 (3)	1014.51 (1)	48.46 (5)	857.55 (2)
6	12	8	1	92.86 (6)	375.67 (4)	1217.53 (2)	1308.38 (1)	281.12 (5)	1155.51 (3)
6	12	11	1	1135.35 (5)	1836.02 (3)	3381.29 (1)	1229.44 (4)	49.94 (6)	3109.85 (2)
6	12	12	1	197.97 (6)	588.07 (3)	335.32 (4)	1519.91 (1)	239.30 (5)	1125.91 (2)
6	12	13	1	65.08 (5)	25.83 (6)	701.58 (3)	1342.88 (1)	208.01 (4)	1339.70 (2)
6	12	14	1	749.65 (5)	1300.74 (3)	2208.40 (2)	2291.31 (1)	465.71 (6)	1203.24 (4)
6	12	15	1	813.40 (4)	290.44 (5)	5612.11 (1)	4850.02 (2)	261.01 (6)	3788.27 (3)
6	12	16	1	186.09 (5)	686.50 (2)	305.85 (4)	563.86 (3)	156.38 (6)	1507.12 (1)
6	13	5	1	402.39 (5)	438.27 (4)	858.77 (1)	770.98 (2)	178.95 (6)	668.86 (3)
6	13	8	1	708.64 (4)	278.39 (5)	879.03 (3)	982.14 (2)	40.04 (6)	1093.78 (1)
6	13	10	1	536.98 (2)	360.85 (3)	221.54 (5)	254.59 (4)	1.45 (6)	643.68 (1)
6	13	11	1	974.54 (2)	198.60 (5)	1044.4 (1)	772.58 (3)	21.06 (6)	596.81 (4)
6	13	12	1	45.82 (6)	2655.71 (1)	1733.42 (2)	1271.56 (4)	152.15 (5)	1399.50 (3)
6	13	13	1	188.01 (6)	1259.83 (1)	198.13 (5)	1214.39 (2)	241.01 (4)	908.10 (3)
6	13	14	1	666.84 (4)	1339.05 (1)	1106.0 (2)	506.28 (5)	227.23 (6)	901.74 (3)
6	13	15	1	1882.72 (4)	12553.39 (2)	14454.7 (1)	203.93 (6)	883.18 (5)	4654.30 (3)
6	14	5	1	1872.10 (5)	3616.66 (2)	2899.08 (3)	1934.07 (4)	764.00 (6)	4022.12 (1)
6	14	8	1	474.74 (5)	4373.27 (1)	1082.22 (3)	814.52 (4)	244.70 (6)	1742.11 (2)
6	14	10	1	49.01 (6)	5887.91 (1)	511.92 (3)	285.19 (4)	180.10 (5)	2660.06 (2)
6	14	11	1	316.80 (4)	1321.94 (2)	1205.20 (3)	247.71 (5)	108.61 (6)	2629.64 (1)
6	14	12	1	1452.91 (4)	2569.58 (2)	2004.63 (3)	1013.99 (5)	276.52 (6)	2999.28 (1)
6	14	13	1	3906.12 (4)	36871.42 (1)	17613.5 (2)	3337.9 (5)	1161.31 (6)	12864.49 (3)
6	15	5	1	205.23 (5)	475.24 (4)	1274.36 (2)	788.27 (3)	170.93 (6)	1682.22 (1)
6	15	8	1	1388.94 (6)	5683.37 (1)	4588.69 (2)	2352.95 (4)	2248.74 (5)	3496.00 (3)

6	15	10	1	3716.74 (4)	11945.31 (2)	22117.0 (1)	7655.26 (3)	568.63 (6)	1416.85 (5)
6	15	12	1	153.07 (5)	1778.14 (2)	1773.22 (3)	1022.02 (4)	26.75 (6)	2882.96 (1)
6	15	13	1	776.58 (3)	2786.48 (1)	730.74 (4)	694.65 (5)	98.51 (6)	1581.14 (2)
6	15	15	1	178.29 (4)	448.56 (2)	5.23 (6)	1121.98 (1)	6.66 (5)	350.99 (3)
6	15	18	1	245.83 (4)	163.48 (5)	3170.20 (1)	1696.55 (3)	54.75 (6)	1746.83 (2)
6	16	5	1	973.58 (3)	3385.76 (1)	389.42 (5)	837.86 (4)	172.15 (6)	2498.29 (2)
6	16	15	1	77.90 (6)	335.18 (3)	267.22 (4)	889.98 (2)	98.83 (5)	1436.89 (1)
6	17	4	1	531.58 (4)	100.43 (6)	625.08 (3)	843.10 (2)	246.94 (5)	2022.85 (1)
6	17	14	1	43.47 (6)	19713.23 (1)	8582.08 (3)	5656.08 (4)	1005.93 (5)	8858.36 (2)
6	17	19	1	517.75 (3)	511.03 (4)	155.15 (5)	1009.21 (2)	91.01 (6)	1738.15 (1)
6	18	5	1	962.64 (5)	2504.72 (1)	1310.22 (2)	1260.30 (3)	237.48 (6)	1092.08 (4)
6	18	16	1	15.49 (6)	2813.82 (1)	971.74 (3)	892.54 (4)	109.64 (5)	1490.17 (2)
6	18	20	1	372.92 (5)	2887.00 (1)	809.40 (3)	2254.60 (2)	17.36 (6)	626.54 (4)
6	19	5	1	565.31 (5)	3700.04 (2)	4368.5 (1)	2534.22 (4)	357.54 (6)	2702.33 (3)
6	19	8	1	79.75 (5)	983.89 (3)	1066.4 (2)	622.66 (4)	67.13 (6)	1292.38 (1)
6	19	12	1	1573.75 (5)	4500.5 (2)	3022.5 (3)	4666.32 (1)	1581.74 (4)	939.47 (6)
6	20	7	1	368.44 (4)	3408.37 (1)	294.68 (5)	770.58 (3)	17.19 (6)	1888.81 (2)
6	20	13	1	711.61 (6)	3739.11 (2)	2641.21 (3)	1232.29 (5)	1244.54 (4)	4177.32 (1)
6	20	17	1	1320.04 (3)	1427.87 (2)	403.19 (6)	1080.75 (4)	492.68 (5)	1815.77 (1)
6	20	18	1	415.22 (5)	1446.52 (1)	1116.5 (2)	423.25 (4)	1.88 (6)	887.14 (3)

Appendix – II

List of Publications Based on this Research Work

1. **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, 347-366.

2. **Kanungo, D.P., Arora, M.K., Gupta, R.P., Sarkar, S., (2005)** GIS based Landslide Hazard Zonation using Neuro-fuzzy Weighting. In: *Proceeding 2nd Indian International Conference on Artificial Intelligence (IICAI-05)*, 20-22 Dec., 2005, Pune, India, 1222-1237.

3. **Kanungo, D.P., Arora, M.K., Sarkar, S., Gupta, R.P., (2006)** Fuzzy Gamma Operator for Integration of Thematic Maps. In: *Indo-Australian Conference on Information Technology in Civil Engineering (IAC-ITCE)*, 20-21 Feb., 2006, Roorkee, India (Abstract published).

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Engineering Geology 85 (2006) 347–366
ENGINEERING GEOLOGY
www.elsevier.com/locate/enggeo

A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas

D.P. Kanungo [a,*], M.K. Arora [b], S. Sarkar [a], R.P. Gupta [c]

[a] *Geotechnical Engineering Division, Central Building Research Institute, Roorkee 247 667, India*
[b] *Department of Civil Engineering, Indian Institute of Technology Roorkee, Roorkee 247 667, India*
[c] *Department of Earth Sciences, Indian Institute of Technology Roorkee, Roorkee 247 667, India*

Received 12 September 2005; received in revised form 29 January 2006; accepted 13 March 2006
Available online 22 May 2006

Abstract

Landslides are one of the most destructive phenomena of nature that cause damage to both property and life every year, and therefore, landslide susceptibility zonation (LSZ) is necessary for planning future developmental activities. In this paper, apart from conventional weighting system, objective weight assignment procedures based on techniques such as artificial neural network (ANN), fuzzy set theory and combined neural and fuzzy set theory have been assessed for preparation of LSZ maps in a part of the Darjeeling Himalayas. Relevant thematic layers pertaining to the causative factors have been generated using remote sensing data, field surveys and Geographic Information System (GIS) tools. In conventional weighting system, weights and ratings to the causative factors and their categories are assigned based on the experience and knowledge of experts about the subject and the study area to prepare the LSZ map (designated here as Map I). In the context of objective weight assignments, initially the ANN as the black box approach has been used to directly produce an LSZ map (Map II). In this approach, however, the weights assigned are hidden to the analyst. Next, the fuzzy set theory has then been implemented to determine the membership values for each category of the thematic layer using the cosine amplitude method (similarity method). These memberships are used as ratings for each category of the thematic layer. Assuming weights of each thematic layer as one (or constant), these ratings of the categories are used for the generation of another LSZ map (Map III). Subsequently, a novel weight assignment procedure based on ANN is implemented to assign the weights to each thematic layer objectively. Finally, weights of each thematic layer are combined with fuzzy set derived ratings to produce another LSZ map (Map IV). The maps I–IV have been evaluated statistically based on field data of existing landslides. Amongst all the procedures, the LSZ map based on combined neural and fuzzy weighting (i.e., Map IV) has been found to be significantly better than others, as in this case only 2.3% of the total area is found to be categorized as very high susceptibility zone and contains 30.1% of the existing landslide area.

Keywords: Landslide susceptibility zonation; GIS; Remote sensing; ANN; Fuzzy; Combined neural and fuzzy

1. Introduction

Landslides are amongst the most damaging natural hazards in the hilly regions. The study of landslides has drawn global attention mainly due to increasing

* Corresponding author. Tel.: +91 1332 283418; fax: +91 1332 272272.
E-mail addresses: dpk_cbri@yahoo.com (D.P. Kanungo), manojfce@iitr.ernet.in (M.K. Arora), shantanu_cbri@yahoo.co.in (S. Sarkar), rpgesfes@iitr.ernet.in (R.P. Gupta).

0013-7952/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.enggeo.2006.03.004

awareness of its socio-economic impacts and also increasing pressure of urbanization on the mountain environment (Aleotti and Chowdhury, 1999). Landslides have represented 4.89% of the natural disasters that occurred worldwide during the years 1990 to 2005 (www.em-dat.net). According to Schuster (1996) and Ercanoglu and Gokceoglu (2004), this trend is expected to continue in future due to increased unplanned urbanization and development, continued deforestation and increased regional precipitation in landslide prone areas due to changing climatic patterns. Hence, the identification of landslide-prone areas is essential for safer strategic planning of future developmental activities in the region. Therefore, Landslide Susceptibility Zonation (LSZ) of an area becomes important whereby the area may be divided into near-homogeneous domains and ranked according to degrees of potential hazard due to mass movements (Varnes, 1984). The area may thus be categorized as very high susceptibility (VHS), high susceptibility (HS), moderate susceptibility (MS), low susceptibility (LS) and very low susceptibility (VLS) zones to produce an LSZ map.

Landslide susceptibility zonation studies in the Himalayas have conventionally been carried out based on manual interpretation of a variety of thematic maps and their superimposition (Anbalagan, 1992; Pachauri and Pant, 1992; Gupta et al., 1993; Sarkar et al., 1995; Mehrotra et al., 1996; Virdi et al., 1997). However, this approach is time consuming, laborious and uneconomical with data collected over long time intervals. In recent times, due to the availability of a wide range of remote sensing data together with data from other sources in digital form and their analysis using GIS, it has now become possible to prepare different thematic layers corresponding to the causative factors that are responsible for the occurrence of landslides in a region (Gupta and Joshi, 1990; van Westen, 1994; Nagarajan et al., 1998; Gupta, 2003). The integration of these thematic layers with weights assigned according to their relative importance in a GIS environment leads to the generation of an LSZ map (Gupta et al., 1999; Saha et al., 2002; Sarkar and Kanungo, 2004; Saha et al., 2005). However, in the studies cited above, the weights were assigned on the basis of the experience of the experts about the subject and the area. The weighting system was thus highly subjective and might therefore contain some implicit biases towards the assumptions made.

For minimizing the subjectivity and bias in the weight assignment process, quantitative methods, namely, statistical analysis, deterministic analysis, probabilistic models, distribution-free approaches and landslide frequency analysis may be utilized. During the last 5 years, bivariate statistical models (Lin and Tung, 2003; He et al., 2003; Suzen and Doyuran, 2004; Saha et al., 2005; etc.), multivariate methods (Dhakal et al., 2000; Clerici et al., 2002; etc.) and probabilistic prediction models (Chi et al., 2002a; Lee et al., 2002a,b; Lan et al., 2004; etc.) have been implemented for LSZ studies. Apart from these methods, some work on distribution-free approaches such as fuzzy set based methods (Chi et al., 2002b; Gorsevski et al., 2003; Tangestani, 2003; Metternicht and Gonzalez, 2005; Ercanoglu and Gokceoglu, 2004), artificial neural network (ANN) models (Arora et al., 2004; Gomez and Kavzoglu, 2005; Yesilnacar and Topal, 2005) and neuro-fuzzy models (Elias and Bandis, 2000; Lee et al., 2004; Kanungo et al., 2005) have recently been attempted for LSZ studies. Due to some success of neural networks and fuzzy set theories in these studies, an attempt has been made here to develop an objective procedure that takes into account the advantages of both neural networks and fuzzy set theory for landslide susceptibility zonation in a part of Darjeeling Himalayas.

2. Weight assignment procedures

The basic pre-requisite for landslide susceptibility zonation studies is the determination of weight and rating values representing the relative importance of factors and their categories respectively for landslide occurrence. These weights and ratings can be determined based on the subjective expert opinions as well as based on an objective analysis. Four different procedures have been implemented here to determine the weights and ratings in order to generate LSZ maps. This section highlights the salient features of these procedures.

2.1. Conventional weighting procedure

In this weighting scheme, factors and their categories are assigned numerical values based on the experience of experts on the subject and about the study area. The numerical values (generally at an ordinal scale from 0 to 9) assigned to factors are termed as weights and those assigned to categories of factors are termed as ratings (Lee et al., 2004). Higher is the numerical value of weight or rating, greater is its influence on the occurrence of landslide.

2.2. Artificial neural network procedure

ANN, which is a useful technique for regression and classification problems, has been successfully applied in

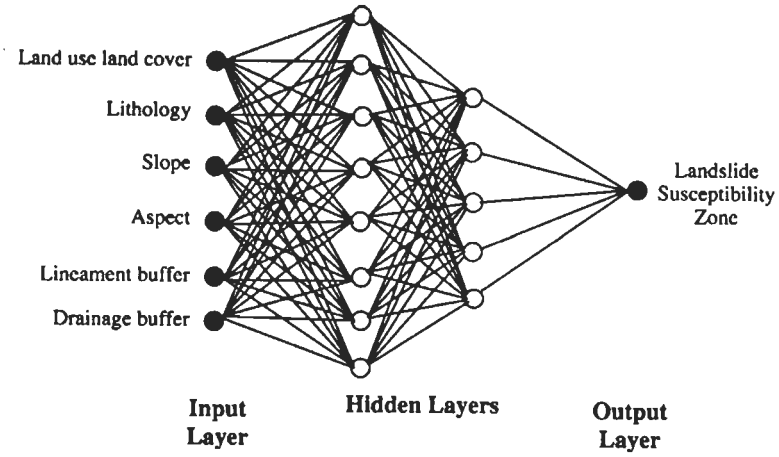


Fig. 1. A schematic diagram of artificial neural network for LSZ using ANN black box procedure.

other fields, and promises to be suitable for the delineation of areas prone to landslide activity. It has been found that ANNs have several advantages for LSZ mapping, as these are non-linear and thus have the capability to analyse complex data patterns. Also, ANN can process data at varied measurement scales such as continuous, ordinal and categorical data, a scenario which is often encountered in LSZ mapping.

An ANN comprises of a number of neurons that work in parallel to transform input data into output classes. A feed-forward multilayer network is generally adopted, which consists of three layers namely input, output and hidden layers in between these two (Paola and Schowengerdt, 1995). Each layer in a network contains sufficient number of neurons depending on the specific application. The neurons in a layer are connected to the neurons in the next successive layer and each connection carries a weight (Atkinson and Tatnall, 1997). The input layer receives the data from different sources (e.g., thematic layers). Hence, the number of neurons in the input layer depends on the number of input data sources. The hidden and output layers actively process the data. The number of hidden layers and their neurons are often determined by trial and error (Gong, 1996). The number of neurons in output layers is fixed by the application and is represented by the class being mapped (e.g., LSZ classes in the present case). Each hidden neuron responds to the weighted inputs it receives from the connected neurons from the preceding input layer (Lee et al., 2004). Once the combined effect on each hidden neuron is determined, the activation at this neuron is determined via a transfer function (Yesilnacar and Topal, 2005). Any differentiable nonlinear function can be used as a transfer function, but a sigmoid function is generally used though there are many other functions (Schalkoff, 1997). The sigmoid function constrains the outputs of a network between 0 and 1.

An important characteristic of a neural network is its capability to learn from the data being processed. The network weights are adjusted in the training process, which can be executed through a number of learning algorithms based on backpropagation learning (Ripley, 1996; Haykin, 1999; Zhou, 1999; Lee et al., 2004; Gomez and Kavzoglu, 2005; Yesilnacar and Topal, 2005). The most widely used backpropagation algorithms are gradient descent and gradient descent with momentum. These are often too slow for the solution of practical problems. The faster algorithms use standard numerical optimizers such as conjugate gradient, quasi-Newton and Levenberg–Marquardt approach. In this study, Levenberg–Marquardt algorithm (implemented as TRAINLM in MATLAB software) has been used for training the neural network. The details of this algorithm can be found in Hagan and Menhaj (1994) and Hagan et al. (1996). Unlike gradient descent algorithms, it does not consider learning rate and momentum factor as its parameters. However, the main scalar parameter involved in this algorithm is mu (μ), which is modified in an adaptive fashion after giving an initial random value.

In back propagation learning, the difference (i.e., error) between the neural network outputs and target outputs is back propagated through the neural network and is minimized by updating interconnection weights between the layers (Arora et al., 2004; Lee et al., 2004). The process of back propagating the error is repeated iteratively until the error is minimized to an acceptable value and the adjusted weights are obtained, which are then used to determine the network outputs. The performance of the network depends on the accuracy obtained over a set of testing dataset. If the network is

trained and tested to an acceptable accuracy, then the adjusted weights are used to determine the outputs of each unknown data sample. This approach has been called as ANN black box approach.

In this study, a feed forward multi-layer ANN with one input layer, two hidden layers and one output layer has been considered to produce an LSZ map. The input layer contains 6 neurons each representing a causative factor that contributes to the occurrence of the landslide. The output layer contains a single neuron representing one of the five LSZ classes (VLS, LS, MS, HS and VHS) for a given set of input values for a pixel.

By varying the number of neurons in both the hidden layers, the neural networks are run several times to identify the most appropriate neural network architecture based on training and testing accuracies. A schematic diagram of the best neural network architecture is given in Fig. 1. The whole dataset of the study area is then processed with the most accurately trained and tested network to generate the LSZ map of the study area. The neural network processing has been implemented in Neural Network Tool Box of MATLAB Software.

2.3. Fuzzy set based procedure

Fuzzy relation concept defined by Zadeh (1973) is based on the theory of fuzzy sets. A fuzzy set can be explained as a set containing elements that have varying degrees of membership in the set (Ross, 1995). Fuzzy relations play an important part in fuzzy modeling and are based on the philosophy that everything is related to some extent (Dubois and Prade, 1980). In this paper, one of the well known similarity methods, cosine amplitude method (Ross, 1995; Ercanoglu and Gokceoglu, 2004), has been used to determine the relationship between the landslide occurrence and the categories responsible for such activity. The membership degrees of categories of each factor are determined by establishing the strength of the relationship (r_{ij}) between the existing landslides and the categories.

Let n be the number of categories of a thematic layer represented as an array $X=\{x_1, x_2, ..., x_n\}$, each of its elements, x_i, is a vector of pixels p (i.e., the size of the image in the present context) and can be expressed as,

$$x_i = \{x_{i1}, x_{i2}, \ldots, x_{ip}\} \tag{1}$$

Each element of a relation, r_{ij}, results from a pairwise comparison of a category of a thematic layer i (i.e., layer corresponding to a causative factor) with a category of thematic layer j (i.e., landslide distribution layer), say x_i and x_j containing elements x_{ik} and x_{jk} respectively. In the cosine amplitude method, for example, r_{ij} (membership grades) between categories of a thematic layer and that of the landslide distribution layer are computed by the following equation with its values ranging from 0 to 1 ($0 \le r_{ij} \le 1$).

$$r_{ij} = \frac{\left[\sum_{k=1}^{p} x_{ik}x_{jk}\right]}{\sqrt{\left(\sum_{k=1}^{p} x_{ik}^2\right)\left(\sum_{k=1}^{p} x_{jk}^2\right)}} \tag{2}$$

The r_{ij} value for a category can be defined as the ratio of total number of landslide pixels in the category to the square root of the multiplication of total number of pixels in that category and the total number of landslide pixels in the area. Values of r_{ij} close to 0 indicate dissimilarity, whereas values close to 1 indicate the similarity between the two datasets. Eq. (2) leads to $(n-1)$ r_{ij} images corresponding to each category of the thematic layers under consideration. These images show r_{ij} values at the pixels belonging to the category in question, whereas rest of the pixels indicate 0 values. The corresponding r_{ij} images for various categories of a thematic layer are composited together to generate an r_{ij} image for that thematic layer and is represented as R_l, where l varies from 1 to t thematic layers belonging to each causative factor (e.g., 6 thematic layers in the present case). The fuzzy processing has been implemented in MS Excel spreadsheet and ArcView GIS software.

2.4. Combined Neural and fuzzy procedure

In another procedure, the ANN connection weights may be used to characterize the input data sources (e.g., the causative factors) in terms of ranks or weights. In this process, the connection weight matrices for input–hidden, hidden–hidden and hidden–output layers are obtained for a two-hidden layer network. Simple matrix multiplications of these weight matrices give rise to the final weight matrix corresponding to the factors (Olden et al., 2004). For example, if a network of 6/14/8/1 architecture (representing 6 neurons in the input layer, 14 neurons in the 1st hidden layer, 8 neurons in the 2nd hidden layer and one neuron in the output layer) is considered, connection weight matrices of 6×14, 14×8 and 8×1 are obtained. The product of 6×14 and 14×8 matrices gives a resultant matrix of 6×8. Subsequently, the product of 6×8 and 8×1 matrices gives an output

[1] Input(I) – Hidden A (HA) Connection Weights

	HA_1	HA_2	HA_3	HA_4	HA_5	HA_6	HA_7	HA_8	HA_9	HA_{10}	HA_{11}	HA_{12}	HA_{13}	HA_{14}
I_1	20.45	-12.98	3.52	-0.98	23.70	9.68	-1.92	-22.60	12.37	-3.26	5.35	12.36	16.72	9.68
I_2	-49.57	-16.84	51.54	87.70	-20.19	52.71	-79.29	-24.83	43.90	-91.70	-61.46	-46.21	-43.25	15.49
I_3	-31.63	-48.57	-52.19	47.39	-11.27	52.44	-5.31	-12.58	44.33	3.84	-46.30	24.22	44.39	-66.34
I_4	12.83	-9.89	18.44	0.15	21.10	-3.93	-16.33	-24.43	-20.15	14.86	-15.81	3.08	-23.33	13.79
I_5	-2.28	9.72	6.48	3.14	-0.06	8.78	8.53	3.95	5.98	1.89	-3.46	-4.09	-4.07	-0.30
I_6	13.46	-12.42	14.70	5.39	-11.54	9.72	2.81	-3.41	10.30	6.19	17.86	25.66	1.28	20.66

[2] Hidden A (HA) – Hidden B (HB) Connection Weights

	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
HA_1	-1.41	2.15	-3.63	-1.70	2.96	1.06	-1.71	-2.60
HA_2	3.10	-1.31	0.69	-1.21	-0.90	0.60	-1.92	2.65
HA_3	1.42	-0.88	-0.07	-1.85	-0.90	-0.20	0.19	2.53
HA_4	-1.17	-0.25	-0.66	1.76	2.04	1.55	-0.72	0.60
HA_5	2.56	-2.74	0.08	-5.36	-0.76	-1.32	1.86	-2.10
HA_6	1.98	1.50	-0.45	3.35	-1.65	1.91	-3.33	1.71
HA_7	-1.54	2.76	2.04	-3.07	0.60	-1.84	-2.96	2.43
HA_8	1.84	-1.46	1.27	0.90	-2.89	-1.83	0.76	-0.08
HA_9	-2.55	-0.45	-0.96	-1.43	-3.94	-0.21	-3.50	-0.20
HA_{10}	1.42	-1.84	-2.44	0.59	0.42	-4.59	-3.45	-1.41
HA_{11}	0.01	-2.68	0.48	1.37	1.55	-0.52	-0.13	-0.71
HA_{12}	-0.30	-1.71	-1.84	2.01	1.34	0.27	-1.09	-0.04
HA_{13}	-2.62	-1.19	1.75	2.32	-1.97	-2.60	1.65	-2.55
HA_{14}	4.22	0.96	-1.32	-2.72	-2.99	-2.59	0.61	-2.55

[3] Connection Weight Products of [1] & [2]

	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
I_1	-64.53	-10.19	-125.47	-111.32	10.13	-11.87	-21.14	-184.95
I_2	66.70	259.49	29.64	234.11	-377.06	884.93	225.86	446.47
I_3	-686.94	81.93	48.73	741.85	-52.38	229.35	-242.08	24.89
I_4	193.99	11.19	-204.81	-211.31	151.55	13.25	70.24	-83.25
I_5	48.23	21.67	19.55	-36.32	-60.27	3.15	-96.88	95.38
I_6	-4.84	0.50	-154.42	69.02	13.00	-42.62	-138.08	-59.90

[4] Hidden B – Output Connection Weights

	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
Output	3.43	-1.61	-3.87	4.83	2.71	-2.14	-3.67	-3.25

[5] Connection Weight Products of [3] & [4]

X \ Y	HB_1	HB_2	HB_3	HB_4	HB_5	HB_6	HB_7	HB_8
I_1	-221.34	16.41	485.57	-537.68	27.45	25.40	77.58	601.09
I_2	228.78	-417.78	-114.71	1130.75	-1021.83	-1893.75	-828.91	-1451.03
I_3	-2356.2	-131.91	-188.59	3583.14	-141.95	-490.81	888.43	-80.89
I_4	665.39	-18.02	792.61	-1020.63	410.70	-28.36	-257.78	270.56
I_5	165.43	-34.89	-75.66	-175.43	-163.33	-6.74	355.55	-309.99
I_6	-16.60	-0.81	597.61	333.37	35.23	91.21	506.75	194.68

$$I_X = \left| \sum_{Y=1}^{8} HB_{XY} \right|$$

	Weight	Rank	Normalized Weight
I_1	474.48	5	0.544
I_2	4368.48	1	5.007
I_3	1081.22	3	1.239
I_4	814.47	4	0.933
I_5	245.06	6	0.281
I_6	1741.44	2	1.996

Fig. 2. Steps for computations of connection weight matrices of ANN to characterize input data layers in terms of ranks and weights (in the present study, I_1—Land use land cover, I_2—Lithology, I_3—Slope, I_4—Aspect, I_5—Drainage buffer and I_6—Lineament buffer); Note: In step [5], X: 1–6; Y: 1–8.

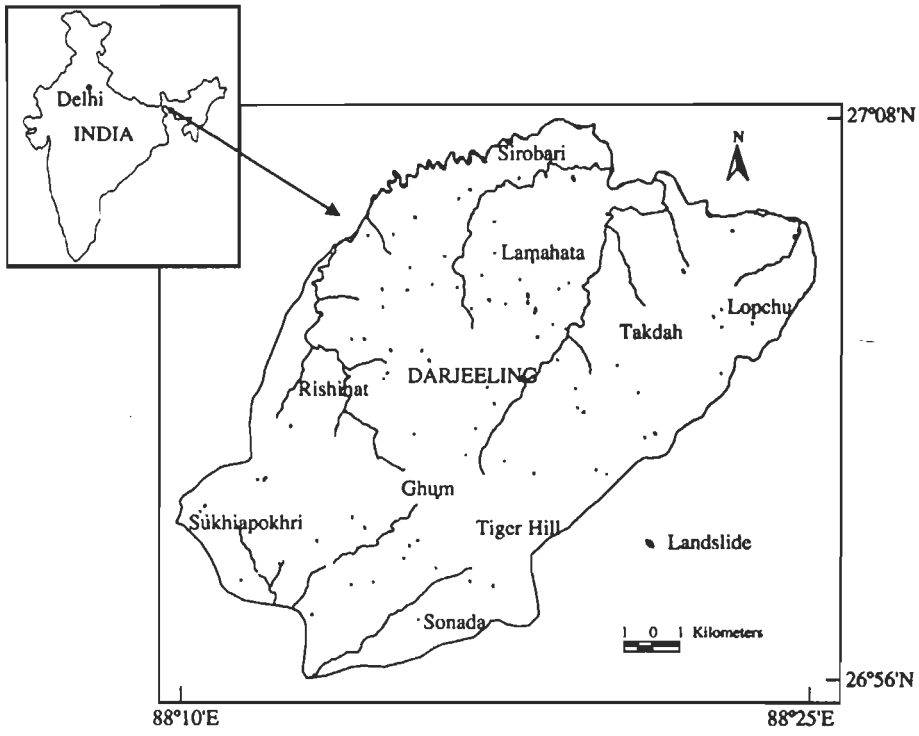


Fig. 3. The study area with landslide distribution in Darjeeling Himalayas.

matrix of 6×1 which corresponds to the weights of 6 factors. The absolute values of these weights are considered in the present work to rank the factors meaning thereby that the factor with maximum absolute weight is assigned as rank 1 and the factor with the minimum absolute weight as rank 6. This has been illustrated in Fig. 2.

3. Study area

The Darjeeling Himalayas, encompassing a total area of 3000 km^2 rise abruptly from the alluvial plains of West Bengal and attain a maximum elevation of about 2600 m. The area lies between Sikkim in the North, Bhutan in the east and Nepal in the west. The southern foot hill region is characterized by East–West trending highly dissected platform of terrace deposits. The southerly flowing river Tista approximately divides the Darjeeling Himalayas into two parts, the eastern and the western parts occupied by Kalimpong and Darjeeling hills respectively. The Tista River however, does not fall within the study area. The study area encompasses Darjeeling hill which lies between latitude 26°56′–27°8′N and longitude 88°10′–88°25′E and covers an area of about 254 km^2 (Fig. 3). The main habitat areas are Darjeeling, Sonada and Sukhiapokhri.

The study area is highly dissected by many ridges and valleys. The maximum elevation of 2584 m occurs at the Tiger hill. The area is dominated by slopes ranging between 15° and 35° while steep slopes of >35° occupy smaller area. In general, the gentle slopes of 0–15° were found on the ridges and at places in the region of lower relief also.

The Darjeeling Himalayas lie within the Lesser and Sub Himalayan belts. The tectonic units in the area occur in inverted order of stratigraphic superposition. Various rock groups have been named locally (Acharya, 1989). The Daling group of rocks comprises of low-grade metamorphic rocks and includes slate, phyllite, schist, quartzite, greywacke and epidiorite. The Darjeeling Group consists primarily of foliated gneisses. Rocks of the Paro Sub-group, which have similar characteristics to the Darjeeling Group, are present at lower elevations.

The annual rainfall in the area is of the order of 3000 mm to 6000 mm. The rainfall pattern is highly seasonal with a maximum rainfall during the monsoon season from June to October. The main land use practice in the study area is tea plantation. The agriculture land is mostly developed around the habitat areas. In general, the area is dominated by thick forest particularly in the eastern part.

4. Thematic data layer preparation

Various thematic data layers corresponding to causative factors namely lithology, slope, aspect, lineaments, land use land cover and drainage have been prepared. These factors fall under the category of preparatory factors, which make the slope susceptible to movement without actually initiating it and thus, are considered responsible for the occurrence of landslides in the region for which pertinent data can be collected from available resources as well as from the field. The triggering factors such as rainfall and earthquakes, set off the movement by shifting the slope from a marginally stable to an actively unstable state. Further, the attributes of the ground (internal factors) in terms of landslide susceptibility have been considered here. Rainfall and earthquakes are external factors and temporal phenomena. Also, past data on these external factors in relation to landslide occurrence are not available. Therefore, these factors have not been included in this study. A thematic layer corresponding to the landslide distribution map has also been prepared to establish a spatial correlation between existing landslides and the causative factors, which will be helpful for the preparation and evaluation of LSZ maps of the area using different weighting procedures.

The IRS-1C-LISS-III (acquired on 22nd March, 2000) and 1D-PAN (acquired on 3rd April, 2000) data along with Survey of India topographic maps at 1:25,000 and 1:50,000 scale, and the geological map at 1:250,000 scale published by Acharya (1989) are the main data sources used to generate these thematic data layers. Extensive field data have been collected during the years 2001 to 2003 to collect information on existing landslide distribution to assist in creation of training and testing datasets, finding out fuzzy membership values and validation of LSZ maps. The months of March and April were preferred for field data collection as these coincided with the date of remote sensing data acquisition. All the thematic data layers have been resampled to match the nominal spatial resolution (i.e., 25 m) of IRS-LISS-III multispectral image.

4.1. Landslide distribution map

The mapping of existing landslides is essential to study the relationship between the landslide distribution and the causative factors. As, it is not possible to map each and every landslide through field surveys in such a terrain, a comprehensive mapping of landslide has been undertaken through remote sensing image interpretation. The identification of landslides on remote sensing image is based on the spectral characteristics, shape, contrast and the morphological expression. In general, there is a distinct spectral contrast between landslides and the background area. High spatial resolution IRS-1C-PAN and PAN-sharpened LISS-III images have been used for landslide mapping. On the PAN image, landslides appear as features of very light tones due to rock debris without any vegetation on the slope. After enhancing the contrast of the PAN image, landslides occurring in barren areas can also be identified. A few old landslides are identified on the basis of their shape, landform and drainage. On the PAN-sharpened LISS-III image, the landslides appear as bright-white features (due to high reflectance) that are easily distinguished from other features. Further, landslides are characterized by fan shape, sharp lines of break in topography and sometimes a local drainage anomaly. Often, the toe part of the slide gives rise to a debris flow channel. Many of the landslides identified on remote sensing images have also been verified in the field.

A total of 101 landslides of varying dimensions (180 m^2 to 27400 m^2) have been identified from remote sensing images and field surveys. A majority of landslides have areal extent of 500 m^2–2000 m^2. Most of the observed landslides are rock slides. However, in some cases, complex types of failure are also possible.

The landslides thus identified have been digitized as polygons in separate vector layers one each for remote sensing derived and field mapped, which are then merged into single landslide layer. This layer has been converted to a rasterized landslide distribution map (Fig. 3) for further analysis.

4.2. Digital elevation model and its derivatives

The Digital Elevation Model (DEM) is an excellent source to derive topographic attributes responsible for landslide activity in a region. Therefore, a DEM at a spatial resolution corresponding to pixel size 25 m × 25 m has been generated by digitization of contours on the topographic maps. The DEM is subsequently used to derive the slope and aspect data layers. Slope angle is one of the key factors in inducing slope instability. The slope data layer consists of 5 classes with 10° interval as per the slope classification used in other studies (Anbalagan, 1992; Gupta et al., 1999; Dhakal et al., 2000). Aspect is defined as the direction of maximum slope of the terrain surface and has an indirect influence on slope instability. In general, the south facing slopes have lesser vegetation density as compared to the north facing slopes and hence, the erosional activity is relatively more in the former case (Sinha et al., 1975). The aspect data layer derived here

represents nine classes, namely, N, NE, E, SE, S, SW, W, NW and flat as per the classification given in other studies (Sarkar and Kanungo, 2004; Saha et al., 2005).

4.3. Lithology

Different rock types (or lithology) have varied composition and structure, which contribute to the strength of the material. The stronger rocks give more resistance to the driving forces as compared to the weaker rocks, and hence are less prone to landslides and vice versa. The lithology data layer has been prepared by digitizing the polygons from the co-registered geological map of Sikkim–Darjeeling area (Acharya, 1989) in a vector layer. Necessary modifications have also been incorporated in this vector layer after field verification. This lithology data layer is later rasterized at 25 m spatial resolution. The six rock types present in this data layer are Darjeeling gneiss, Paro gneiss, Lingtse granite gneiss, feldspathic greywacke, and quartzites of the Paro sub-group and the Reyang group.

4.4. Lineaments

Lineaments are the structural features which describe the zone/plane of weakness, fractures and faults along which landslide susceptibility is higher. It has generally been observed that the probability of landslide occurrence increases at sites close to lineaments, which not only affect the surface material structures but also make contribution to terrain permeability causing slope instability. Lineaments have been interpreted from the PAN and LISS-III images. The individual bands of LISS-III image are enhanced using linear contrast stretching followed with a 3×3 edge filters to highlight high frequency features. Subsequently, all the four bands are layer stacked to produce the edge-enhanced image which has been used for visual interpretation of lineaments. The lineaments have been interpreted based on the tonal contrast, structural alignments and rectilinear trends of morphological features and linear stream courses that are conspicuous by their abrupt changes in the course. There is no major thrust/fault reported in the study area, but mega lineaments have been identified. The interpreted lineaments have been digitized on-screen and subsequently rasterized to produce the lineament data layer. Initially, buffer zones at 250 m intervals were created. These buffer zones were spatially cross-correlated with the landslide pixels in the area and it was observed that 98% of landslide pixels occurred in 1st two buffer zones (up to 500 m). Hence, it was decided to consider four buffer zones at 125 m intervals up to 500 m and another buffer zone beyond 500 m to establish the influence of lineaments on landslide occurrence. Thus, a lineament buffer layer consisting of five classes such as 0–125 m, 125–250 m, 250–375 m, 375–500 m and >500 m has been prepared.

4.5. Drainage

Many of the landslides in hilly areas occur due to the erosional activity associated with drainage. Therefore, a drainage data layer has been prepared by digitizing the drainages from the topographic maps in a vector layer. Later, this layer has been overlaid on IRS-LISS-III image for updating the drainages. This was felt necessary as most of the 1st order drainages, which were not present on the topographic maps, could be interpreted from the LISS-III image, which also showed change in the course of the river and other major drainages at some places. The ordering of the drainage has been performed on the basis of Strahler's classification scheme (Strahler, 1964). Drainages up to 6th order have been observed in the study area. Subsequently, the vector layer has been rasterized at 25 m spatial resolution. Initially, 25 m buffer zones on either side of the drainages for all the drainage orders were created. It was observed from the spatial correlation of landslide distribution in these buffer zones that majority of landslide pixels occurred in the 1st and 2nd order drainage buffers only. Therefore, 25 m buffer zones around these drainages only have been considered to create a drainage buffer layer for further analysis.

4.6. Land use land cover

Land use land cover is also a key factor responsible for landslide occurrences. The incidence of landslide is inversely related to the vegetation density. Hence, barren slopes are more prone to landslide activity as compared to the forest area. Eight dominant land use land cover classes namely thick forest, sparse forest, tea plantation, agriculture, barren, built up, water bodies and river sand have been considered similar to other studies (Sarkar and Kanungo, 2004; Saha et al., 2005). A very small portion of the study area is covered by cloud and its shadow in the LISS-III image. Initially, this portion was masked. The four spectral bands of LISS-III image, Digital Elevation Model (DEM) and Normalized Difference Vegetation Index (NDVI) image have been considered to prepare a land use land cover map in a multi-source classification process using the most widely adopted maximum likelihood classifier. The map has been prepared at an overall classification accuracy of 94.7%. Subsequently, masked portion of the land use land cover

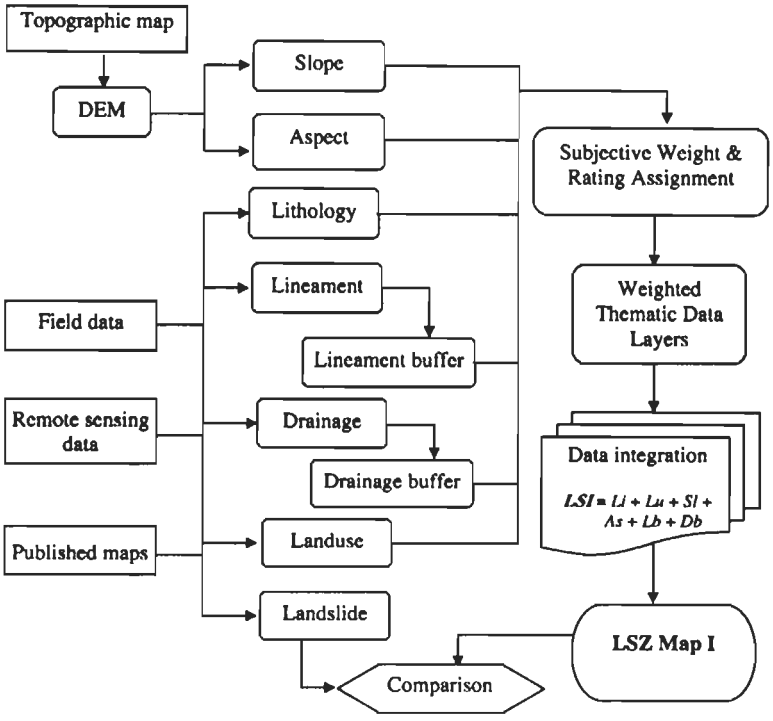


Fig. 4. Flow diagram showing different stages of preparing LSZ map using conventional weighting procedure.

map thus prepared has been filled with the land use land cover information obtained from field surveys to generate the final land use land cover layer.

5. Implementation of various procedures for LSZ

The LSZ mapping was performed in GIS environment to categorize each and every pixel of the dataset to one of the landslide susceptibility zones. GIS tool allows for the storage and manipulation of information concerning different factors as distinct layers and thus provide an excellent tool for LSZ mapping. The thematic data layers were prepared in GIS platform and the data were stored as attributes for further analysis. The weights and ratings of the thematic layers and their categories respectively were determined using four different weighting procedures as mentioned earlier. All the four procedures were implemented in ArcView GIS software to generate LSZ maps.

5.1. LSZ using conventional weighting procedure

5.1.1. Implementation

The conventional weighting procedure involves assignment of weights and ratings to the thematic layers and their categories respectively based on the knowledge of the study area and the experience on the subject. Different steps of this procedure for LSZ are given in Fig. 4. Such weighting scheme was used by Sarkar and Kanungo (2004) with a different combination of thematic layers and their categories.

In the study area, it was observed that most of the landslides were associated with drainage channels and hence the maximum weight was assigned to the drainage layer. Also, maximum rating of 9 was assigned to 1st order drainage buffer category as most of the landslides initiate from the lower order drainages. The next important factor considered was the lineament. Here, the maximum rating was given to the 0–125 m lineament buffer category as the nearness to the lineaments controls the occurrence of the landslide. Since the steeper slopes are more prone to landslide, the slope classes were given ratings in the descending order. The competent rocks such as quartzite, greywacke are less susceptible to landslides than the gneisses as more number of landslides are observed in gneissic rock in the field. Hence, the ratings to lithology categories were assigned accordingly. Occurrence of landslides also depends on the type of the land use land cover. Barren slopes are more susceptible to erosion as compared to areas with thick forest and hence, maximum rating was assigned to the barren slopes and minimum to the thick forest. The slope aspect has an indirect influence on slope instability. Based upon the landslide distribution, south and east

Table 1
Weights and ratings for thematic layers and their categories (conventional weighting procedure)

Thematic layers	Categories	Weights	Ratings
Drainage buffer	1. 1st order	9	9
	2. 2nd order		5
Lineament buffer	1. 0–125 m	8	9
	2. 125–250 m		7
	3. 250–375 m		5
	4. 375–500 m		3
	5. >500 m		1
Slope	1. 0–15°	7	1
	2. 15–25°		3
	3. 25–35°		5
	4. 35–45°		7
	5. >45°		9
Lithology	1. Darjeeling gneiss	6	7
	2. Feldspathic greywacke		3
	3. Paro gneiss		5
	4. Lingse granite gneiss		9
	5. Paro quartzite		1
	6. Reyang quartzite		1
Land use land cover	1. Agriculture land	4	5
	2. Tea plantation		3
	3. Thick forest		1
	4. Sparse forest		7
	5. Barren land		9
	6. Habitation		2
	7. Water body		0
	8. River sand		0
Aspect	1. Flat	1	0
	2. North		1
	3. Northeast		4
	4. East		7
	5. Southeast		8
	5. South		9
	7. Southwest		6
	8. West		3
	9. Northwest		2

facing slopes were considered more prone to landslides than the other slopes (Dhakal et al., 2000). Considering these facts and field observations, ratings for slope aspect categories were assigned accordingly. The weights and ratings thus assigned to each thematic layer and their categories are given in Table 1.

The weighted thematic data layers were generated by algebraically multiplying the weight of the layer with the ratings of the corresponding categories of each layer. In the present case, six weighted thematic data layers namely lithology, land use land cover, slope, aspect, lineament buffer and drainage buffer were produced. These layers were laid over one another and arithmetically added according to the following equation to generate a Landslide Susceptibility Index (LSI) map in GIS,

$$LSI = Li + Lu + Sl + As + Lb + Db \quad (3)$$

where Li, Lu, Sl, As, Lb and Db are abbreviations for the weighted thematic layers for lithology, land use land cover, slope, aspect, lineament buffer and drainage buffer respectively.

The LSI values range from 21 to 310, which were then categorized to produce landslide susceptibility classes. A judicious way for such classification is to search for abrupt changes in values (Davis, 1986). The classification procedure reported by Sarkar and Kanungo (2004) was followed. The class boundaries were drawn at LSI values of 68, 137, 176 and 236 to obtain five susceptibility zones. The LSZ map (referred here as Map I), thus obtained is shown in Fig. 5a. The area covered by five different landslide susceptibility classes and the area of landslides occupied per class have also been determined (Table 2).

5.1.2. Analysis of the output LSZ Map I

The five susceptibility zones in LSZ Map I were distributed all over the study area. The map did not show any definite pattern for the distribution of susceptibility zones. It can be inferred from Table 2 that 33.3% of HS and VHS areas together could predict 58.7% of existing landslide area. It was again observed from this map that the VHS and HS zones represented mostly the 1st and 2nd order drainage buffer areas. This happened as the drainage layer was assigned the maximum weight in this conventional procedure.

5.2. LSZ using ANN Black Box procedure

5.2.1. Implementation

The flow diagram indicating the different steps followed in this procedure is shown in Fig. 6. Initially, the field data on landslides (i.e., existing landslide distribution) were used to process the neural network. But, due to less number of landslide pixels, the neural network accuracies were found to be very low. Therefore, similar to earlier studies (Arora et al., 2004), the map obtained from conventional approach was considered as reference map, to derive representative sample sizes. Two independent sets of training and testing data were randomly selected with their land slide susceptibility class known from the LSZ map I. Each dataset consisted of 500 mutually exclusive pixels corresponding to each landslide susceptibility zone (Foody and Arora, 1997). The training dataset was used to train various network architectures while the testing dataset was used to control the overtraining of the network and to evaluate the accuracy of the networks. The input values for neural network processing correspond to the attributes of the category of a thematic layer (as mentioned in 2nd

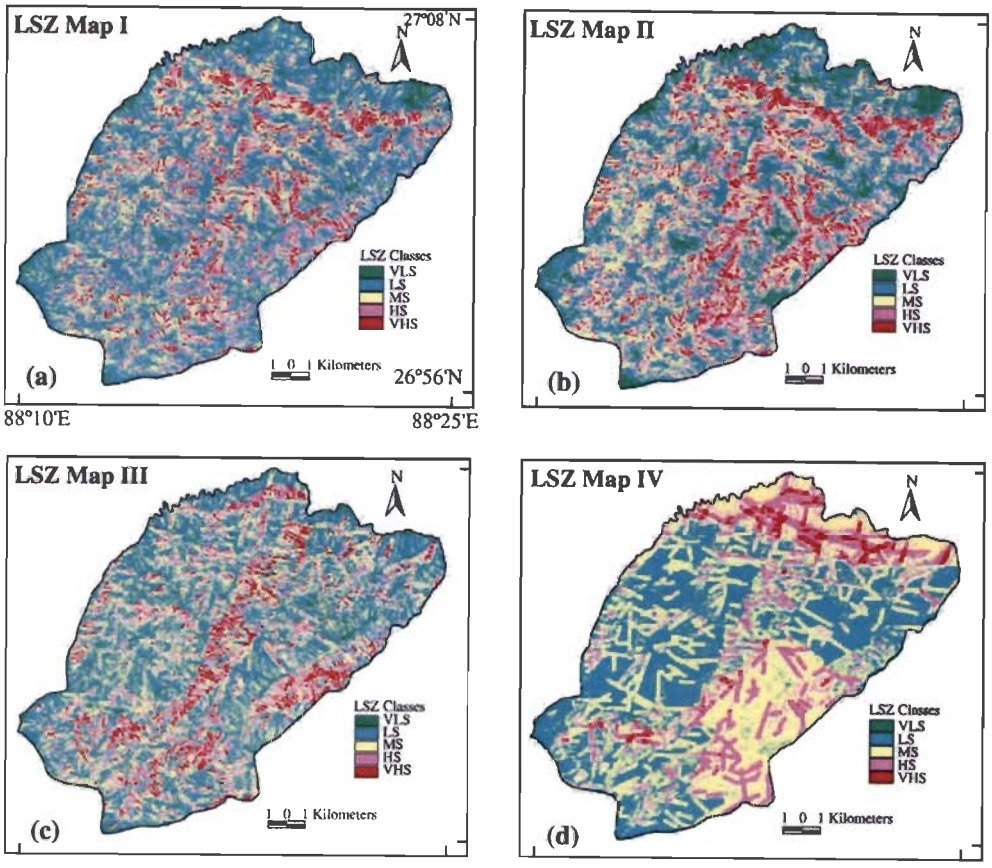


Fig. 5. LSZ maps using four different procedures. (a) Conventional weighting procedure. (b) ANN black box procedure. (c) Fuzzy set based procedure. (d) Combined neural and fuzzy procedure.

column of Table 2), which were normalized with respect to the highest value within each thematic layer. The normalized attribute values of different landslide susceptibility zones (VLS, LS, MS, HS and VHS) in the desired output in order of increasing susceptibility correspond to 0.2, 0.4, 0.6, 0.8 and 1.0 respectively.

A total of 39 neural network architectures were created by varying the number of neurons in both the hidden layers. The training process was initiated by assigning arbitrary initial connection weights, which were constantly updated until an acceptable accuracy over training data was reached. These adjusted weights obtained from the trained network were subsequently used to process the testing dataset to examine the generalization capability of the network.

The performance of the networks was evaluated by determining both the training and testing data accuracies

Table 2
Distribution of landslide susceptibility zones and landslides in different LSZ maps

Landslide susceptibility zones	LSZ Map I		LSZ Map II		LSZ Map III		LSZ Map IV	
	Area (%)	Landslide area occupied per class (%)	Area (%)	Landslide area occupied per class (%)	Area (%)	Landslide area occupied per class (%)	Area (%)	Landslide area occupied per class (%)
Very high susceptibility zone	6.5	10.6	7.7	10.3	6.1	41.0	2.3	30.1
High susceptibility zone	26.8	48.1	26.9	40.4	22.7	25.1	20.2	31.9
Moderate susceptibility zone	30.2	26.5	35.4	36.1	39.4	25.9	48.4	26.5
Low susceptibility zone	34.9	14.5	25.8	12.7	30.4	8.0	28.8	11.5
Very low susceptibility zone	1.6	0.3	4.2	0.5	1.4	0.0	0.3	0.0

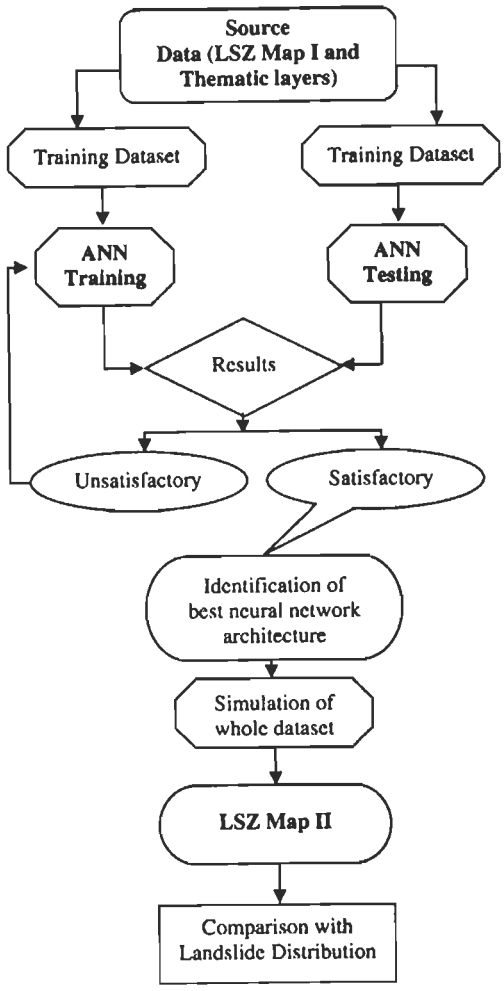


Fig. 6. Flow diagram showing different steps of ANN black box procedure.

in terms of correlation coefficient, root mean squared error (RMSE) (Freund, 1992) and also by the percent correct or overall classification accuracy (Congalton, 1991). The training and testing accuracies for some networks are given in Table 3. A variation in both training and testing data accuracies can be seen as the neural network architectures change. This infers that there exists an optimal network architecture for a given dataset. In the present case, the network architecture 6/13/7/1 with training data accuracy (correlation coefficient of 0.918, root mean square error 0.112 and 74.4% correct) and the testing data accuracy (correlation coefficient of 0.896, root mean square error 0.126 and 72.6% correct) was the most appropriate one. Therefore, weights obtained from this network were subsequently used to obtain the network output of each pixel. The output values ranged from 0.062 to 0.993, which were categorized into one of the five landslide susceptibility zones (Table 4) to produce the LSZ map in GIS. The LSZ map (referred here as Map II) thus produced is shown in Fig. 5b. The area covered by different landslide susceptibility zones and the area of landslides occupied per class were also determined (Table 2).

5.2.2. Analysis of the output LSZ Map II

It can be inferred from Table 2 that in case of LSZ Map II, an area of 34.6% belonging to VHS and HS zones put together could predict only 50.7% landslide area. Further, the map does not depict any definite pattern, as the susceptibility zones are distributed all over the area. Moreover, the drainage buffer layer has a major control on this LSZ map also. This may be due to the fact that the LSZ Map I was used as the reference map for training the network to produce an ANN black box based LSZ map (Map II). Hence, there is lot of similarity between LSZ Maps I and II.

Table 3
Training and testing data accuracies of ANN black box procedure (bold indicates the best acceptable architecture in this study)

NN architecture	Correlation coefficient		RMSE		Accuracy (%)		
	Training	Testing	Training	Testing	Training	Testing	Diff.
6–4–2–1	0.877	0.870	0.136	0.139	64.9	63.2	1.7
6–5–2–1	0.886	0.875	0.131	0.137	67.4	65.0	2.4
6–6–4–1	0.892	0.881	0.128	0.134	68.0	65.3	2.7
6–7–5–1	0.906	0.893	0.120	0.127	71.0	69.3	1.7
6–8–5–1	0.908	0.891	0.119	0.128	70.8	68.7	2.1
6–9–5–1	0.911	0.893	0.117	0.127	72.1	68.6	3.5
6–10–4–1	0.912	0.893	0.116	0.128	71.4	69.8	1.6
6–11–3–1	0.915	0.897	0.114	0.125	73.9	70.0	3.9
6–12–4–1	0.912	0.892	0.116	0.128	72.9	70.9	2.0
6–13–5–1	0.915	0.899	0.114	0.124	73.8	70.3	3.5
6–13–7–1	**0.918**	**0.896**	**0.112**	**0.126**	**74.4**	**72.6**	**1.8**
6–13–9–1	0.918	0.893	0.112	0.128	73.2	68.0	5.2
6–14–3–1	0.912	0.892	0.116	0.128	71.9	70.1	1.8
6–15–4–1	0.919	0.891	0.112	0.128	74.2	69.2	5.0
6–15–6–1	0.915	0.894	0.114	0.127	73.8	69.5	4.3
6–15–8–1	0.914	0.894	0.115	0.127	72.9	69.6	3.3
6–16–3–1	0.917	0.890	0.113	0.129	72.7	69.0	3.7
6–16–7–1	0.920	0.889	0.111	0.130	73.8	67.2	6.6

5.3. LSZ using fuzzy set based procedure

5.3.1. Implementation

In the fuzzy set based procedure, ratings of each category of a given thematic layer were determined on the basis of cosine amplitude similarity method, which were then integrated in GIS by considering the weight of each thematic layer as one (or constant) to generate the LSZ map.

The cosine amplitude method as described earlier was adopted to determine the ratings of the categories of factors. The landslide distribution map and different categories of thematic layers taken one at a time were considered as two datasets for the computation of rating or strength of relationship (r_{ij}). In the landslide distribution layer, pixels belonging to landslides were assigned as 1 and the rest were assigned as 0. Similarly, a value of 1 was assigned to a pixel belonging to a particular category of a thematic layer and a value of 0 to the rest. Hence, in total 36 data layers in binary form were generated, which contained 35 layers of categories of thematic layers (Table 5) and one layer of landslide distribution. These layers were used for the determination of r_{ij} in GIS so as to generate 35 images denoting r_{ij} values. The r_{ij} values thus obtained are listed in Table 5.

The corresponding r_{ij} images for various categories of a thematic layer were combined together to generate an R_l image for that thematic layer. The integration of these R_l images for various thematic layers can be performed in several ways to compute LSI values. The simplest approach is to add this arithmetically which is similar to any other conventional GIS integration process. Alternatively, to bring fuzziness in the integration process also, the use of fuzzy algebraic sum, fuzzy algebraic product and fuzzy gamma operator can be put forth. In view of this, the LSI values were computed in two different ways: (a) using arithmetic integration and (b) using fuzzy gamma operator. The performance of the two methods was also examined. It was found that the arithmetic overlay approach of thematic layer integration yielded better results than the fuzzy gamma operator for this dataset. Therefore, this approach was considered further.

In the arithmetic overlay approach, the LSI for each pixel of the study area was obtained using the following equation,

$$\mathrm{LSI} = \sum_{l=1}^{t} (R_l) \tag{4}$$

The LSI values were found to lie in the range from 0.014 to 0.252. The observed mean (μ_o) and standard

Table 4
Classification scheme for neural network output values for LSZ mapping (ANN black box procedure)

Range of values	Landslide susceptibility zone
0–0.3	Very low susceptibility zone
0.3–0.5	Low susceptibility zone
0.5–0.7	Moderate susceptibility zone
0.7–0.9	High susceptibility zone
>0.9	Very high susceptibility zone

Table 5
Fuzzy ratings for different categories of factors

Thematic layers corresponding to factors	Categories	Number of pixels	Number of landslide pixels	Fuzzy rating (r_{ij})
Land use land cover	Agriculture land	35,692	85	0.0488
	Tea plantation	142,541	84	0.0243
	Thick forest	72,685	38	0.0229
	Sparse forest	131,088	65	0.0223
	Barren land	14,237	58	0.0638
	Habitation	10,341	9	0.0295
	Water	970	0	0
	River sand	1005	0	0
Lithology	Darjeeling gneiss	73,371	77	0.0324
	Feldspathic greywacke	45,938	61	0.0364
	Paro gneiss	247,242	158	0.0253
	Lingtse granite gneiss	20,926	15	0.0268
	Paro quartzite	12,154	14	0.0339
	Reyang quartzite	8089	14	0.0416
Slope	0–15°	51,380	23	0.0212
	15–25°	146,974	117	0.0282
	25–35°	144,495	131	0.0301
	35–45°	50,246	58	0.0340
	>45°	14,329	10	0.0264
Aspect	Flat	2072	0	0
	N	59,880	22	0.0192
	NE	45,077	32	0.0266
	E	52,868	73	0.0372
	SE	45,689	77	0.0411
	S	37,630	49	0.0361
	SW	29,860	20	0.0259
	W	55,132	26	0.0217
	NW	79,148	40	0.0225
Drainage buffer	25 m along 1st order drainage	116,168	102	0.0296
	25 m along 2nd order drainage	27,690	44	0.0399
Lineament buffer	0–125 m	146,761	243	0.0407
	125–250 m	108,929	35	0.0179
	250–375 m	72,380	36	0.0223
	375–500 m	41,360	17	0.0203
	>500 m	38,317	8	0.0144

deviation (σ_o) from the probability distribution curve of these LSI values are 0.150 and 0.024 respectively. The LSI values were divided into five distinct classes (susceptibility zones) with boundaries at ($\mu_o - 1.5\ m\sigma_o$), ($\mu_o - 0.5\ m\sigma_o$), ($\mu_o + 0.5\ m\sigma_o$) and ($\mu_o + 1.5\ m\sigma_o$) where m is a positive, non-zero value (Saha et al., 2005) which controls in fixing the most appropriate boundaries within the LSI range for landslide susceptibility classes. This classification scheme was adopted to fix the boundaries of classes statistically and also to avoid the subjectivity in arbitrarily selecting the boundaries of classes as was done in the conventional procedure.

Several LSZ maps of the study area were prepared for different values of m. The cumulative percentage of landslide occurrences in various susceptibility zones ordered from VHS to VLS were plotted against the cumulative percentage of area of the susceptibility zones for these LSZ maps. These curves, defined as success rate curves (Chung and Fabbri, 1999; Lu and An, 1999; Lee et al., 2002b), were used to select the appropriate value of m to decide the suitability of a LSZ map. Five representative success rate curves corresponding to $m=0.8$, 1.0, 1.1, 1.2 and 1.4 are shown in Fig. 7a. The suitability of any LSZ map can be judged by the fact that more percentage of landslides must occur in VHS zone as compared to other zones. It can be seen from Fig. 7a that for 10% of the area in VHS zone, the curves corresponding to $m=0.8$, 1.0, 1.1, 1.2 and 1.4 show the landslide occurrences of 42.7%, 47.7%, 48.6%, 47.3% and 41.8% respectively. Hence, for the first 10% area, the curve corresponding to $m=1.1$ has the highest success rate. Based on this analysis, the LSZ map corresponding to $m=1.1$ appears to be the most appropriate one for the study area. Accordingly, the landslide susceptibility class

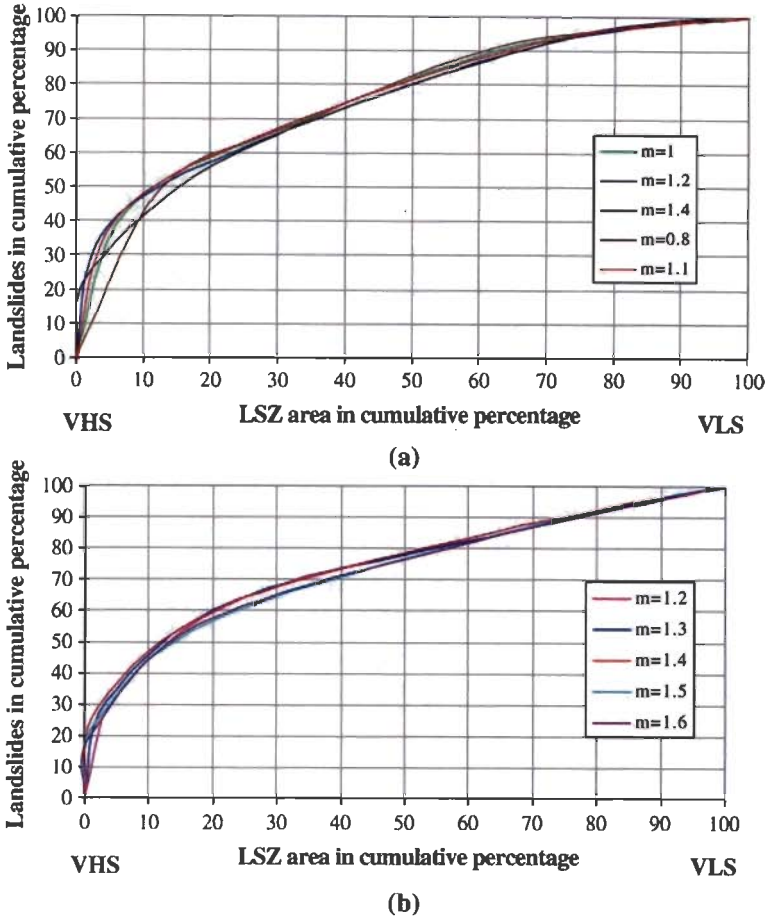


Fig. 7. Success rate curves to select the appropriate LSZ map. (a) Fuzzy set based procedure. (b) Combined neural and fuzzy procedure.

boundaries were fixed at LSI values of 0.110, 0.136, 0.163 and 0.189. The LSZ map (referred here as Map III) thus produced is shown in Fig. 5c. The area covered by different landslide susceptibility zones and the area of landslides occupied per class are also given in Table 2.

5.3.2. Analysis of output LSZ Map III

The visual inspection of the LSZ Map III depicts an overall NNE–SSW zonation trend in the area. It has been observed that the southeast and east facing slopes are more prone to landslides than other slopes. Hence, it can be stated that there is a topographic control over this LSZ map. Further, the spatial correlation between the landslide distribution and the LSZ map shows that 41.0% of landslide area has predicted over 6.1% area of VHS zone. It can also be stated that 28.8% of the total area occupied by HS and VHS zones are able to predict 66.1% of the total landslide area (Table 2).

5.4. LSZ using combined neural and fuzzy procedure

5.4.1. Implementation

The combined neural and fuzzy approach involves three main stages:

1) determination of weights of thematic layers through ANN connection-weight approach
2) determination of ratings for categories of thematic layers using cosine amplitude method
3) integration of ratings and weights using GIS to arrive at the LSZ map.

The methodology for LSZ using this procedure is shown in Fig. 8.

A feed forward back-propagation multi-layer ANN with one input layer, two hidden layers and one output layer was considered to determine the weights of the causative factors. The number of neurons in the input layer equals the number of input thematic layers. The

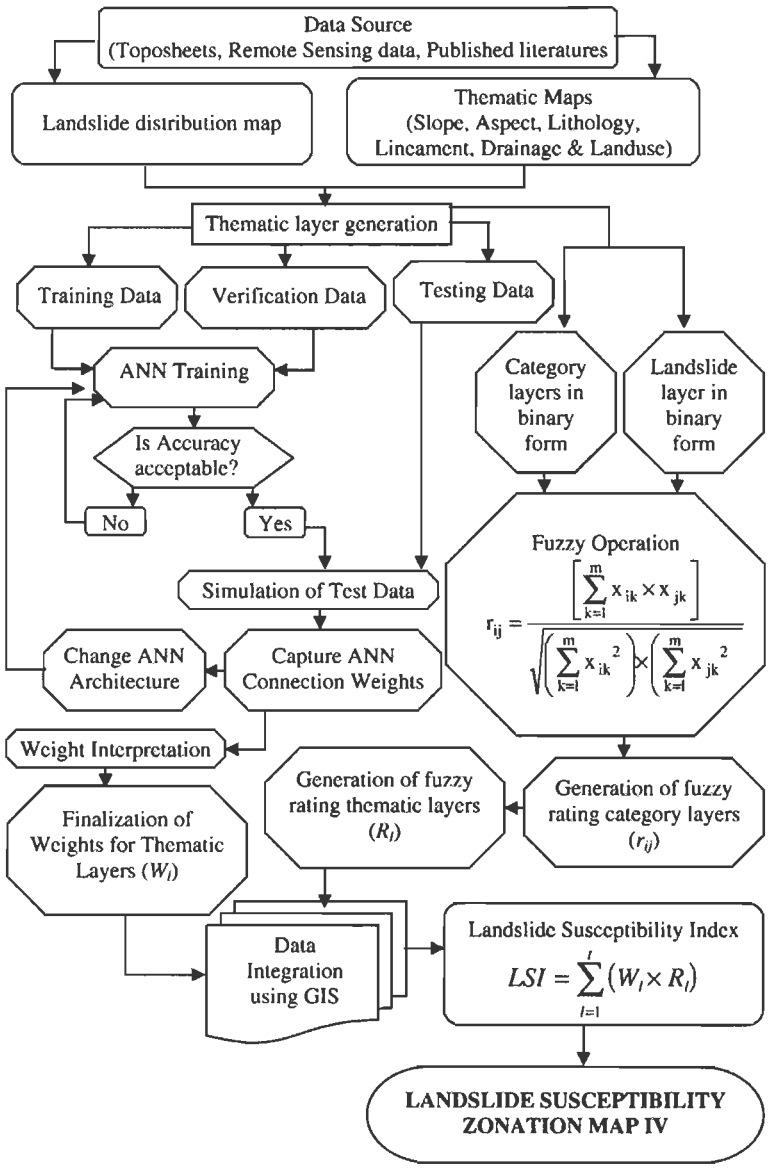


Fig. 8. Flow diagram showing the combined neural and fuzzy procedure for LSZ mapping.

data at each neuron of the input layer correspond to the weighted normalized rating or r_{ij} of the corresponding category (i.e., last column of Table 5). The output layer consists of a single neuron representing presence or absence of landslide for a pixel. Hence, the output value is either 0 or 1. The number of neurons in the hidden layers is varied by running the networks several times to achieve the desired training and testing data accuracies.

One set each of training, verification and testing data were randomly generated from the study area. The datasets consist of 226 pixels each, out of which 113 pixels were landslide pixels and rest 113 pixels were no landslide pixels. All the pixels in the datasets were mutually exclusive (Foody and Arora, 1997). The training dataset was used to train different network architectures while the verification dataset was used simultaneously with the training dataset to control the overtraining of the network. The testing dataset was used to evaluate the accuracy of the networks. Similar to ANN black box approach, the well known back-propagation learning algorithm was used to train the neural networks. 100 neural network architectures were designed, trained and tested. The training process was

initiated with arbitrary initial connection weights, which were constantly updated until an acceptable accuracy was reached. The training accuracy observed for the networks was of the order of 75% to 90%.

The final adjusted weights of the trained network were used to derive outputs of the testing data to evaluate the performance of the network. The testing accuracy observed for the networks was of the order of 60% to 70%. The adjusted weights of input–hidden, hidden–hidden and hidden–output connections for each network were captured for further analysis. Simple matrix multiplication was performed on these weight matrices to obtain a 6×1 weight matrix for each network which represents the weights of six causative factors (thematic layers) in this study. These causative factors were ranked according to the corresponding absolute weights for each network which means the higher the value of absolute weight, the more crucial the factor is for the occurrence of landslide. Considering all the 100 networks, the rank of a factor was decided based on the rank observed by the maximum number of networks (majority rule). Out of the 100 networks, 41 networks categorized lithology as rank 1 (most important), 31 networks categorized lineament as rank 2, 30 networks categorized slope as rank 3, 27 networks categorized aspect as rank 4, 33 networks categorized land use land cover as rank 5 and 49 networks categorized drainage as rank 6 (least important). These results are summarized in Table 6. Subsequently, the weighted normalized average of the weights of these networks at a scale of 0–10 for a particular factor was calculated and assigned as the weight of that factor (W_l) for the preparation of LSZ map. The weights thus obtained through ANN for all the factors are listed in Table 7.

It has been observed that the network with architecture 6/14/8/1 has been found to be the best for this dataset as it produced the same ranking pattern as mentioned above. The normalized weights obtained through this network at a scale from 0 to 10 for lithology, lineament buffer, slope, aspect, land use land cover and drainage buffer are 5.007, 1.996, 1.239, 0.933, 0.544 and 0.281 respectively (Fig. 2). These weights are almost at par with the weights obtained through majority rule (Table 7) which have been finally considered for producing the LSZ map.

The integration of 6 thematic layers representing the ratings for the categories (R_l) of the layers (obtained from fuzzy logic) and weights for the layers (W_l) (obtained from ANN) was performed by using simple arithmetic overlay operation in GIS. Hence, this procedure has been named here as combined neural and fuzzy weighting procedure. The LSI for each pixel of the study area was thus obtained by using the following equation.

$$\mathrm{LSI} = \sum_{l=1}^{t} (W_l \times R_l) \tag{5}$$

The LSI values were found to lie in the range from 0.030 to 0.408. The success rate curve approach was used to classify the LSI values into five different susceptibility zones to produce the LSZ map. Five representative success rate curves corresponding to $m = 1.2$, 1.3, 1.4, 1.5 and 1.6 are shown in Fig. 7b. It can be observed that for 10% of the area in VHS zone the curves corresponding to $m = 1.2$, 1.3, 1.4, 1.5 and 1.6 show the landslide occurrences of 43.9%, 45.6%, 46.7%, 43.3% and 43.9% respectively. Hence, for the first 10% area, the curve corresponding to $m = 1.4$ has the highest success rate. Based on this analysis, the LSZ map corresponding to $m = 1.4$ appears to be the most appropriate one for the study area. Accordingly, the boundaries of landslide susceptibility zones were fixed at LSI values of 0.208, 0.253, 0.299 and 0.344. The LSZ map (referred here as Map IV) thus produced is given in Fig. 5d. The area covered by different landslide susceptibility zones and the area of landslides occupied per class are also given in Table 2.

5.4.2. Analysis of output LSZ Map IV

In LSZ Map IV, only 2.3% of the total study area was occupied by the VHS zone and 30.1% of landslide area was predicted over this zone. It was also inferred that 22.5% of the total area occupied by VHS and HS zones could predict 62% of landslide area (Table 2). Further, this LSZ map has shown preferential distribution of higher landslide susceptibility zones along structural discontinuities (lineaments), which should indeed be the case. The buffer zones of lineaments have clearly indicated the VHS and HS zones in the north and southeast parts of the area. Therefore, it indicates the "ghost-effect" of lineaments on LSZ map as stated by Saha et al. (2005). Also, the Darjeeling gneiss rock type in southeastern part, feldspathic greywacke and Reyang quartzite in the northern part of the study area have clearly indicated moderate to very high susceptibility zones. Most of the lineaments up to 125 m buffer zone in these rock types have indicated high and very high susceptibility zones. Hence, it depicts the importance of lithology (i.e., rock types) as well as lineaments on the LSZ.

6. Comparative analysis and discussion

The LSZ maps were prepared using four different weighting procedures in a GIS-based approach. The

Table 6
Ranks of factors based on majority rule in combined neural and fuzzy weighting procedure (the number represents the number of artificial neural networks out of 100 networks categorizing a factor with respect to a particular rank and the rank corresponding to the maximum number of neural networks for a factor represents the final rank of that factor)

Factors	Number of networks						Final rank (majority rule)
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	
Land use land cover	1	8	10	22	33	26	5
Lithology	41	21	12	10	10	6	1
Slope	23	24	30	9	9	5	3
Aspect	13	15	22	27	12	11	4
Drainage buffer	0	1	2	17	31	49	6
Lineament buffer	22	31	24	15	5	3	2

comparative analysis of different LSZ maps has been described below.

The pattern of percent area distribution of susceptibility classes in different LSZ maps prepared in this study appears to be quite similar to that obtained in other LSZ studies in the Himalayan regions (Gupta et al., 1999; Arora et al., 2004; Sarkar and Kanungo, 2004; Saha et al., 2005). However, in the LSZ Map I prepared using conventional weighting procedure, the LS zone occupied the maximum percent area (34.9%) in comparison to the MS zone which occupied 30.2% area.

Further, the VHS zone in the LSZ Map IV occupied 2.3% of the total study area, whereas in all other LSZ maps the area occupied by VHS is more than 6% of total area. Subsequently, the landslide distribution map was spatially cross-checked with all the four LSZ maps. The landslide distribution in the VHS and HS zones of LSZ maps (Table 2) indicate that the LSZ maps produced by fuzzy and combined neural and fuzzy procedures could predict more landslides in these zones as compared to other two LSZ maps.

Moreover, it can also be observed that the LSZ map produced by combined neural and fuzzy procedure shows preferential distribution of higher landslide susceptibility zones along structural discontinuities (lineaments) as compared to other LSZ maps, which may indeed be the case. Overall, the buffers of lineaments have left traces on the LSZ map. Owing to the landslide susceptibility of the terrain, the lineaments ought to leave some traces (termed as "ghost effect" in Saha et al., 2005) on the LSZ map.

Furthermore, the Darjeeling gneiss rock type in southeastern part, feldspathic greywacke and Reyang quartzite in the northern part of the study area have clearly indicated MS to VHS zones. Most of the lineaments up to 125 m buffer zone in these rock types have indicated HS and VHS zones. Hence, it depicts the importance of lithology (i.e., rock types) as well as lineaments on LSZ.

On the basis of these results, it can be concluded that the LSZ map derived from combined neural and fuzzy weighting procedure appears to be the best amongst all the weighting procedures and may thus be a useful way of assigning weights to the factors in an objective manner thereby minimizing the subjectivity.

7. Conclusions

In this study, four different weighting procedures viz. conventional based on subjective weighting, ANN black box, fuzzy logic and combined neural and fuzzy were applied for LSZ mapping in part of Darjeeling Himalayas and a comparative evaluation was carried out. The combined neural and fuzzy weighting integration produced the most accurate LSZ map. This may be attributed to the following reasons:

1) It represents an objective approach where weights for factors are determined through ANN connection weight approach and ratings of the categories of factors are determined through cosine amplitude similarity method based on fuzzy relation concept.
2) The LSZ map reflects preferential distribution of higher landslide susceptibility zones along lineaments which may indeed be the case.

Table 7
Weights of thematic layers derived through ANN (combined neural and fuzzy weighting procedure)

Thematic layers	ANN derived weights
Lithology	4.807
Lineament buffer	2.113
Slope	1.318
Aspect	1.065
Land use land cover	0.495
Drainage buffer	0.202

3) It delineates a relatively small area (only 2.3% of total area) for VHS zone, which can be more meaningful for practical applications.

Therefore, the integration of different factors in GIS environment using the combined neural and fuzzy weighting procedure may serve as one of the key objective approaches in this direction because of the fact that it can narrow down the potential susceptibility zones in a meaningful way for planning future developmental activities and implementation of disaster management programmes in hilly terrains.

References

Acharya, S.K., 1989. The Daling Group, its nomenclature, tectonostratigraphy and structural grain: with notes on their possible equivalents. Geol. Surv. Indones., Spec. Publ. 22, 5–13.

Aleotti, P., Chowdhury, R., 1999. Landslide Hazard Assessment: Summary, Review and New Perspectives. Bull. Eng. Geol. Environ. 58, 21–44.

Anbalagan, R., 1992. Landslide Hazard Evaluation and Zonation Mapping in Mountainous Terrain. Eng. Geol. 32, 269–277.

Arora, M.K., Das Gupta, A.S., Gupta, R.P., 2004. An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. Int. J. Remote Sens. 25 (3), 559–572.

Atkinson, P.M., Tatnall, A.R.L., 1997. Neural networks in remote sensing. Int. J. Remote Sens. 18, 699–709.

Chi, K.-H., Park, N.-W., Lee, K., 2002a. Identification of Landslide Area using Remote Sensing Data and Quantitative Assessment of Landslide Hazard. Proc. IEEE Int. Geosciences and Remote Sensing Symp., 19 July, Toronto, Canada.

Chi, K.-H., Park, N.-W., Chung, C.-J., 2002b. Fuzzy logic integration for landslide hazard mapping using spatial data from Boeun, Korea. Proc. of Symp. on Geospatial Theory, Processing and Applications, Ottawa.

Chung, C.-J.F., Fabbri, A.G., 1999. Probabilistic prediction models for landslide hazard mapping. Photogramm. Eng. Remote Sensing 65, 1389–1399.

Clerici, A., Perego, S., Tellini, C., Vescovi, P., 2002. A procedure for landslide susceptibility zonation by the conditional analysis method. Geomorphology 48, 349–364.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46.

Davis, J.C., 1986. Statistics and Data Analysis in Geology. John Wiley & Sons, New York. 646 pp.

Dhakal, A.S., Amada, T., Aniya, M., 2000. Landslide hazard mapping and its evaluation using GIS: an investigation of sampling schemes for a grid-cell based quantitative method. Photogramm. Eng. Remote Sensing 66 (8), 981–989.

Dubois, D., Prade, H., 1980. Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York.

Elias, P.B., Bandis, S.C., 2000. Neurofuzzy systems in landslide hazard assessment. Proc. 4th Int. Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Science, pp. 199–202.

Ercanoglu, M., Gokceoglu, C., 2004. Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). Eng. Geol. 75 (3&4), 229–250.

Foody, G.M., Arora, M.K., 1997. An evaluation of some factors affecting the accuracy of classification by an artificial neural network. Int. J. Remote Sens. 18, 799–810.

Freund, J.E., 1992. Mathematical Statistics, Fifth Edition. Printice-Hall of India Pvt. Ltd., New Delhi, India.

Gomez, H., Kavzoglu, T., 2005. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Eng. Geol. 78 (1–2), 11–27.

Gong, P., 1996. Integrated analysis of spatial data for multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping. Photogramm. Eng. Remote Sensing 62, 513–523.

Gorsevski, P.V., Gessler, P.E., Jankowski, P., 2003. Integrating a fuzzy *k*-means classification and a bayesian approach for spatial prediction of landslide hazard. J. Geogr. Syst. 5, 223–251.

Gupta, R.P., 2003. Remote Sensing Geology, 2nd Edition. Springer-Verlag, Berlin Heidelberg, Germany.

Gupta, R.P., Joshi, B.C., 1990. Landslide Hazard Zonation using the GIS Approach—A case Study from the Ramganga Catchment, Himalayas. Eng. Geol. 28, 119–131.

Gupta, V., Sah, M.P., Virdi, N.S., Bartarya, S.K., 1993. Landslide Hazard Zonation in the Upper Satlej Valley, District Kinnaur, Himachal Pradesh. J. Himal. Geol. 4, 81–93.

Gupta, R.P., Saha, A.K., Arora, M.K., Kumar, A., 1999. Landslide Hazard Zonation in a part of Bhagirathy Valley, Garhwal Himalayas, using integrated Remote Sensing—GIS. J. Himal. Geol. 20 (2), 71–85.

Hagan, M.T., Menhaj, M., 1994. Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Netw. 5 (6), 989–993.

Hagan, M.T., Demuth, H.B., Beale, M.H., 1996. Neural Network Design. PWS Publishing, Boston, MA.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation, Second edition. Prentice Hall, New Jersey.

He, Y.P., Xie, H., Cui, P., Wei, F.Q., Zhong, D.L., Gardner, J.S., 2003. GIS-based hazard mapping and zonation of debris flows in Xiaojiang Basin, Southwestern China. Environ. Geol. 45, 286–293.

Kanungo, D.P., Arora, M.K., Gupta, R.P., Sarkar, S., 2005. GIS-based landslide hazard zonation using neuro-fuzzy weighting. Proc. 2nd Ind. Int. Conf. on Artificial Intelligence (IICAI-05), Pune, pp. 1222–1237.

Lan, H.X., Zhou, C.H., Wang, L.J., Zhang, H.Y., Li, R.H., 2004. Landslide hazard spatial analysis and prediction using GIS in the Xiaojiang Watershed, Yunnan, China. Eng. Geol. 76, 109–128.

Lee, S., Choi, J., Chwae, U., Chang, B., 2002a. Landslide susceptibility analysis using weight of evidence. Proc. of IEEE Int. Geosciences and Remote Sensing Symposium, 19 July, Toronto, Canada.

Lee, S., Choi, J., Min, K., 2002b. Landslide susceptibility analysis and verification using the bayesian probability model. Environ. Geol. 43, 120–131.

Lee, S., Ryu, J., Won, J., Park, H., 2004. Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Eng. Geol. 71, 289–302.

Lin, M.-L., Tung, C.-C., 2003. A GIS-based Potential Analysis of the Landslides induced by the Chi-Chi Earthquake. Eng. Geol. 71, 63–77.

Lu, P.F., An, P., 1999. A metric for spatial data layers in favorability mapping for geological events. IEEE Trans. Geosci. Remote Sens. 37, 1194–1198.

Mehrotra, G.S., Sarkar, S., Kanungo, D.P., Mahadevaiah, K., 1996. Terrain analysis and spatial assessment of landslide hazards in parts of Sikkim Himalaya. Geol. Soc. India 47, 491–498.

Metternicht, G., Gonzalez, S., 2005. FUERO: foundations of a fuzzy exploratory model for soil erosion hazard prediction. Environ. Model. Softw. 20 (6), 715–728.

Nagarajan, R., Mukherjee, A., Roy, A., Khire, M.V., 1998. Temporal remote sensing data and GIS application in landslide hazard zonation of part of Western Ghat, India. Int. J. Remote Sens. 19, 573–585.

Olden, J.D., Joy, M.K., Death, R.G., 2004. An accyrate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecol. Model. 178, 389–397.

Pachauri, A.K., Pant, M., 1992. Landslide hazard mapping based on geological attributes. Eng. Geol. 32, 81–100.

Paola, J.D., Schowengerdt, R.A., 1995. A review and analysis of backpropagation neural networks for classification of remotely sensed multi-spectral imagery. Int. J. Remote Sens. 16, 3033–3058.

Ripley, B., 1996. Pattern Recognition and Neural Networks. Cambridge Univ. Press, Cambridge.

Ross, T.J., 1995. Fuzzy Logic with Engineering Applications. McGraw-Hill, New York.

Saha, A.K., Gupta, R.P., Arora, M.K., 2002. GIS-based landslide hazard zonation in a part of the Himalayas. Int. J. Remote Sens. 23, 357–369.

Saha, A.K., Gupta, R.P., Sarkar, I., Arora, M.K., Csaplovics, E., 2005. An approach for GIS-based statistical landslide susceptibility zonation—with a case study in the Himalayas. Landslides 2, 61–69.

Sarkar, S., Kanungo, D.P., 2004. An integrated approach for landslide susceptibility mapping using remote sensing and GIS. Photogramm. Eng. Remote Sensing 70 (5), 617–625.

Sarkar, S., Kanungo, D.P., Mehrotra, G.S., 1995. Landslide hazard zonation: a case study in Garhwal Himalaya, India Moun. Res. Dev. 15 (4), 301–309.

Schalkoff, R.J., 1997. Artificial Neural Networks. Wiley, New York.

Schuster, R., 1996. Socioeconomic significance of landslides. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation: Special Report, vol. 247. National Academic Press, Washington, DC, pp. 12–36.

Sinha, B.N., Varma, R.S., Paul, D.K., 1975. Landslides in Darjeeling district (West Bengal) and adjacent areas. Bull. Geol. Surv. India, Ser. B 36, 1–45.

Strahler, A.N., 1964. Quantitative geomorphology of basins and channel networks. In: Chow, V.T. (Ed.), Handbook of Applied Hydrology. Mcgraw Hill, New York.

Suzen, M.L., Doyuran, V., 2004. Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu Catchment, Turkey. Eng. Geol. 71, 303–321.

Tangestani, M.H., 2003. Landslide susceptibility mapping using the fuzzy gamma operation in a GIS, Kakan Catchment Area, Iran. Proc. of Map India Conference.

van Westen, C.J., 1994. GIS in landslide hazard zonation: a review, with examples from the Andes of Colombia. In: Price, M., Heywood, I. (Eds.), Mountain Environments and Geographic Information System. Taylor & Francis, Basingstoke, pp. 135–165.

Varnes, D.J., 1984. Landslide Hazard Zonation: A Review of Principles and Practice. UNESCO, Paris, pp. 1–63.

Virdi, N.S., Sah, M.P., Bartarya, S.K., 1997. Mass wasting, its manifestations, causes and control: some case histories from Himachal Himalaya. In: Agarwal, D.K., Krishna, A.P., Joshi, V., Kumar, K., Palni, M.S. (Eds.), Perspectives of Mountain Risk Engineering in the Himalayan Region. Gyanodaya Prakashan, Nainital, pp. 111–130.

www.em-dat.net EM-DAT: The OFDA/CRED International Disaster Database. Universite Catholique de Louvain, Brussels, Belgium.

Yesilnacar, E., Topal, T., 2005. Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Eng. Geol. 79, 251–266.

Zadeh, L.A., 1973. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Trans. Syst. Man Cybern SMC-3 (1), 28–46.

Zhou, W., 1999. Verification of the nonparametric characteristics of backpropagation neural networks for image classification. IEEE Trans. Geosci. Remote Sens. 37, 771–779.

GIS based Landslide Hazard Zonation using Neuro-Fuzzy Weighting

D. P. Kanungo[1], M. K. Arora[2], R. P. Gupta[3], S. Sarkar[1]

[1] Geotechnical Engineering Division, Central Building Research Institute, Roorkee 247 667, India
dpk_cbri@yahoo.com, shantanu_cbri@yahoo.co.in
[2] Department of Civil Engineering, Indian Institute of Technology Roorkee, Roorkee 247 667, India
manojfce@iitr.ernet.in
[3] Department of Earth Sciences, Indian Institute of Technology Roorkee, Roorkee 247 667, India
rpgesfes@iitr.ernet.in

Abstract. Landslides are one of the most frightening and destructive phenomena of nature that cause damage to both property and life every year, especially in the Himalayas. Landslide Hazard Zonation (LHZ) is therefore necessary for planning future developmental activities and implementation of disaster management programmes in mountainous terrains. A number of qualitative and quantitative methods have been used for the LHZ studies. The conventional approaches for LHZ suffer from the subjective weighting system where weights are assigned to causative factors responsible for landslide activity based on the experience of experts. In this study, an objective weighting approach based on neuro-fuzzy technique has been implemented for LHZ in a part of the Darjeeling Himalayas. This is considered to be a distribution free approach. Relevant thematic layers pertaining to the causative factors related to landslide activity have been generated using remote sensing and Geographic Information System (GIS) techniques. The membership values for each category of the thematic layers have been determined using the cosine amplitude method (similarity method) and are used as ratings for each category of the thematic layer. Artificial neural networks (ANNs) have been applied to determine the weights for each causative factor represented by a thematic layer. The integration of these weights and ratings, determined through a combined neuro-fuzzy approach in GIS environment led to the generation of LHZ map of acceptable accuracy in terms of areas estimated for each landslide hazard zone.

1 Introduction

Landslides are amongst the most damaging natural hazards in the hilly regions. The study of landslides has drawn global attention mainly due to increasing awareness of its socio-economic impacts and also increasing pressure of urbanization on the mountain environment [1]. Hence, the identification of landslide-prone areas is essential for safer strategic planning of future developmental activities in the region. Therefore, Landslide Hazard Zonation (LHZ) becomes important whereby an area

Bhanu Prasad (Editor): IICAI-05, pp. 1922-1937, 2005.

may be divided into near-homogeneous domains and ranked according to degrees of potential hazard due to mass movements [2].

Landslide hazard zonation studies in the Himalayas have conventionally been carried out based on manual interpretation of a variety of thematic maps and their superimposition ([3], [4], [5], [6], [7], [8]). However, this approach is time consuming, laborious, uneconomical with data collected over long time intervals and fraught with errors. In recent times, due to the availability of a wide range of remote sensing data and their analysis using GIS, it has now become possible to prepare different thematic layers representing causative factors that are responsible for the occurrence of landslide in a region ([9], [10], [11], [12]). The integration of these thematic layers with weights assigned according to their relative importance in GIS environment leads to an LHZ map ([13], [14], [15], [16]). In these studies, the weights were assigned on the basis of the experience of the experts about the subject and the area. The weighting system was thus subjective and might therefore contain some implicit bias towards the assumptions made.

In order to minimize the subjectivity and bias in the weight assignment process, the Artificial Neural Network (ANN) and fuzzy concepts may be applied. These are distribution free approaches and have been applied in various studies related to landslides ([17], [18], [19], [20]). Due to the successful implementation of neural networks and fuzzy set theories in these studies, an attempt has been made here to develop a weight assignment process that combines neural network and a fuzzy set based techniques for landslide hazard zonation in a part of Darjeeling Himalayas.

2 Neuro-Fuzzy Approach

The fuzzy relation concept has been applied to determine the ratings of the categories of thematic layers whereas ANN has been applied to determine the weights of each thematic layer. In the present context, the rating of a category of a thematic layer represents the contribution of that category for the occurrence of landslides whereas the weight of a thematic layer (factor) defines the relative importance of that factor with respect to other factors responsible for landslide activities in the terrain. This section highlights the salient features of fuzzy relation concept and the ANN.

2.1 Fuzzy relation concept

Fuzzy relation concept defined by [21] is based on the theory of fuzzy sets. In the fuzzy set theory, membership degrees of elements have varying degrees of support or confidence in (0,1) interval. Thus, a fuzzy set can be explained as a set containing elements that have varying degrees of membership in the set [22]. Fuzzy relations play an important part in fuzzy modeling. Fuzzy relations are based on the philosophy that everything is related to some extent or unrelated [23]. There are different ways such as Cartesian product, closed-form expression, linguistic rules of knowledge, similarity methods in data manipulation, etc. to develop numerical values characterizing a fuzzy relation. All these methods attempt to determine some sort of similarity in data [22].

In this paper, one of the well known similarity methods, cosine amplitude method, has been used to determine the relationship between the landslide occurrence and the factors responsible for such activity. The membership degrees of the categories of each factor are calculated by the strength of the relationship (r_{ij}) between the existing landslides and the factors.

Let n be the number of categories of the thematic layers represented as an array $X = \{x_1, x_2, \ldots, x_n\}$, each of its elements, x_i, is a vector of pixels p (i.e., the size of the image in the present context) and can be expressed as,

$$x_i = \{x_{i1}, x_{i2}, \ldots, x_{ip}\} \tag{1}$$

Each element of a relation, r_{ij}, results from a pairwise comparison of a category of a thematic layer i (i.e., layer corresponding to a causative factor) with a category of thematic layer j (i.e., landslide distribution layer), say x_i and x_j containing elements x_{ik} and x_{jk} respectively. In the cosine amplitude method, r_{ij} (membership grades) between categories of a thematic layer and that of landslide distribution layer are computed by the following equation with its values ranging from 0 to 1 ($0 \leq r_{ij} \leq 1$).

$$r_{ij} = \frac{\left[\sum_{k=1}^{p} x_{ik} \times x_{jk}\right]}{\sqrt{\left(\sum_{k=1}^{p} x_{ik}^2\right) \times \left(\sum_{k=1}^{p} x_{jk}^2\right)}} \tag{2}$$

Values of r_{ij} close to 0 indicate dissimilarity whereas values close to 1 indicate the similarity between two data sets. The equation (2) leads to n-1 r_{ij} images corresponding to each category of the thematic layers under consideration. These images show r_{ij} values at the pixels belonging to the category in question whereas rest of the pixels indicate 0 values, and thus may be treated as binary images. The corresponding r_{ij} images for various categories of a thematic layer are composited together to generate an r_{ij} image for that thematic layer and is represented as R_l, where l varies from 1 to t thematic layers belonging to each causative factor (e.g., 6 thematic layers in the present case).

2.2 Artificial Neural Networks

An ANN comprises of a number of neurons that work in parallel to process the input data into output classes. A feed-forward multilayer network is generally used, which consists of three layers such as input, output and a hidden layer in between these two. Each layer in a network contains sufficient number of neurons depending on the specific problem. The neurons in a layer are connected to the neurons of the next successive layer and each connection carries a weight. The input layer receives the data from different sources (e.g., thematic layers). Hence, the number of neurons in the input layer depends on the number of data sources. The hidden and output layers

actively process the data. Generally, the number of neurons in input and output layers is fixed by the application. The number of hidden layers and their neurons are often determined by trial and error [24]. The hidden layers introduce non-linearity to the network to produce an output through an activation function. The most common activation function is a sigmoid function [25].

The important characteristic of a neural network is its capability to learn or train. The network weights are adjusted in the training process that depends on the type of learning algorithm used. The most widely used learning algorithm is the back-propagation algorithm. This algorithm computes and minimizes the error based on the desired and network-derived outputs iteratively. This process stops when the error converges to some minimum value and the network weights are adjusted. The error function, E, can be expressed as follows.

$$E = 0.5\sum_{i=1}^{s}(T_i - O_i)^2 \tag{3}$$

where T_i is the target output vector, O_i is the network output vector and s is the number of training samples [18]. The adjusted weights of a trained network are used to determine the network outputs of a testing dataset with known desired outputs. The performance of the network depends on the accuracy of the testing dataset. If the network is trained and tested to an acceptable accuracy, then the network connection weights may be used to characterize the input data sources (e.g., the causative factors) in terms of ranks or weights. In this process, the connection weight matrices for input-hidden, hidden-hidden and hidden-output layers are obtained for a two-hidden layer network. Simple matrix multiplications of these weight matrices give rise to the final weight matrix corresponding to the factors [26]. For example, if a network of 6-5-4-1 architecture is considered, connection weight matrices of 6×5, 5×4 and 4×1 are obtained. The product of 6×5 and 5×4 matrices gives a resultant matrix of 6×4. Subsequently, the product of 6×4 and 4×1 matrices gives an output matrix of 6×1 which corresponds to the weights of 6 factors. The absolute values of these weights are considered in the present work to rank the factors meaning thereby that the factor with maximum absolute weight is assigned as rank 1 and the factor with the minimum absolute weight as rank 6.

Several neural network architectures were experimented with to finalize the rank of each factor based on the rankings observed by these networks by applying the majority rule. Finally, the absolute weights of corresponding networks for a particular factor were averaged. The average weights of each factor were then normalized at a scale of 0 to 10. The weight thus obtained was treated as the final weight value assigned to that factor.

3 Study Area

The Darjeeling Himalayas, encompassing a total area of 3000 km^2 rises abruptly from the alluvial plains of North Bengal and attains a maximum elevation of about 2600 meters. The area lies between Sikkim in the North, Bhutan in the east and Nepal in the west. The study area encompasses Darjeeling hill which lies within the latitude 26° 56′ - 27° 8′ N and longitude 88°10′ – 88°25′ E and covers an area of about 250

km^2 (Fig.1). The area is traversed by many ridges and valleys. The major ridges are at the elevation of 2000 m to 2600 m whereas valleys are at elevation range of 500 m to 1000 m. The maximum elevation in the area occurs at Tiger hill at 2584 m.

The annual rainfall in the area is of the order of 3000 mm to 6000 mm. The rainfall pattern is highly seasonal with maximum rainfall falling during the monsoon months of June to October. The maximum monthly rainfall of 580 mm was observed in July 2001 at Darjeeling [15].

The Darjeeling Himalaya lies within Lesser and Sub Himalayan belt. The tectonic units in the area occur in a reverse order of stratigraphic superposition. In the study area, the Daling group is comprised of low-grade metamorphic rocks, while the Darjeeling Group consists primarily of gneisses. Rocks of the Paro Sub-group, which have similar characteristics to the Darjeeling Group, are present at lower elevations [27]. The presence of lineaments, hot springs, lakes, water falls, abrupt changes in the river gradient and erraticity in the terrace distribution suggested that the area has been undergoing differential uplift along a number of regional and local sub-vertical faults [28].

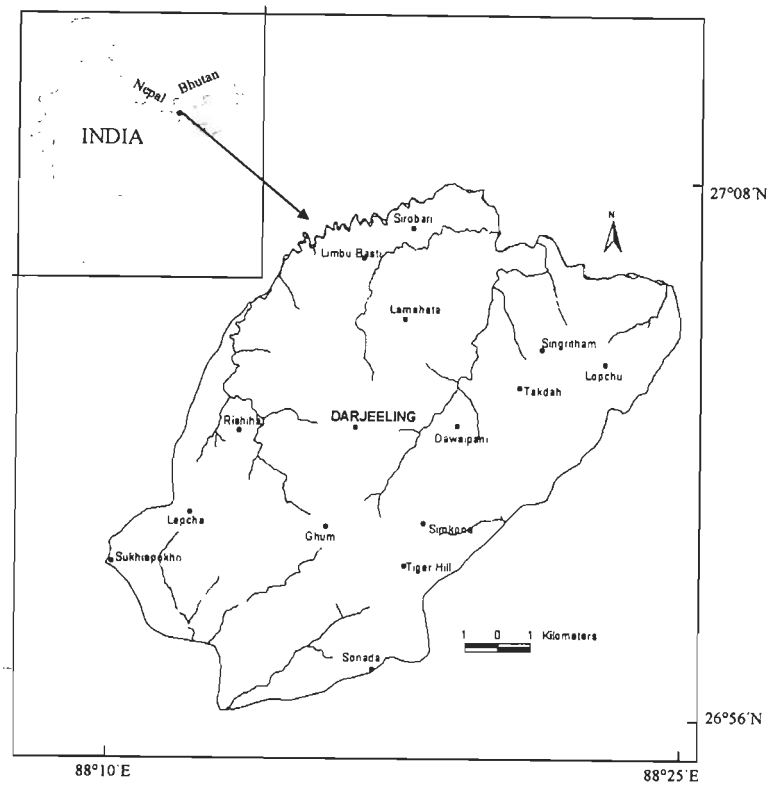


Fig.1. Study area

4 Thematic Data Layer Preparation

Various thematic data layers corresponding to causative factors such as slope, aspect, lithology, lineament, landcover and drainage responsible for the occurrence of landslides in the region were prepared. A thematic layer corresponding to the landslide distribution map was also prepared. The data sources utilized to generate these thematic data layers were Survey of India toposheets (1:25000 and 1:50000 scale), published geological maps, satellite IRS-1C-LISS III multispectral image (March, 2000) and IRS-1D-PAN image (April, 2000). For verification and accuracy assessment of landslide, land use land cover and geological features, field-surveys were also carried out. All the thematic data layers were resampled to 25m pixel size to match the nominal spatial resolution of IRS-LISS III multispectral image.

4.1 Landslide Distribution Map

The mapping of existing landslides is essential to study the relationship between the landslide distribution and the causative factors. A landslide distribution map was prepared using field data and the interpretation of satellite remote sensing data. The detection of landslides in remote sensing data is based on the spectral characteristics, size, shape, contrast and the morphological expression. The interpretability is influenced by the contrast that results from the spectral difference between the landslide and its surroundings.

In this study, PAN and PAN-sharpened LISS III images were used for landslide detection. As the landslides were mostly devoid of vegetation, they showed high reflectance. They were found to be mostly circular to elliptical in shape. Majority of the landslides observed were debris slide. However, in some cases, rotational and complex types of failure were also present. Few rock falls of small magnitude were also identified in the field. The landslide distribution map thus prepared contained 101 landslides showing the areas occupied by sliding activity. It includes source area and debris cover. The majority of landslides have aerial extent of 500 m^2 - 2000 m^2. The landslides identified on the satellite images were digitized as polygons in a vector layer. The field identified landslides were digitized in a separate vector layer and later merged into the landslide layer generated from PAN sharpened image. This layer was rasterized to produce the landslide distribution map for further analysis.

4.2 Digital Elevation Model and its derivatives

The Digital Elevation Model (DEM) has been used to derive information on topographic variations in the area which are responsible for landslide activity. DEM was generated from digitized contours of 10m and 20m from toposheets. From the DEM, slope and aspect data layers were derived. Slope angle is one of the key factors in inducing slope instability. The slope map represents the spatial distribution of slope values in the area. These were classified into 5 classes with 10° interval as per slope classification given in other studies ([3], [13], [29]). Aspect is defined as the direction of maximum slope of the terrain surface. The aspect map derived here represents nine classes, namely, N, NE, E, SE, S, SW, W, NW and flat.

4.3 Lithology

Different rock types (or lithology) have varied inherent characteristics such as composition and structure, which contribute to the strength of the material. The stronger rocks give more resistance to the driving forces as compared to weaker rocks, and hence are less prone to landslides and vice versa. The lithology map was derived from the geological map (scale 1:250000) of Sikkim-Darjeeling area [27] which was scaled to 1:25000 for field investigation. Necessary modifications were incorporated after field verification to prepare the lithology map. The six rock types present in the area were Darjeeling gneiss, Paro gneiss, Lingtse granite gneiss, feldspathic greywacke, and quartzites of the Paro sub-group and the Reyang group. The Lingtse granite gneiss rocks were strongly foliated and sheared showing very high weathering at most locations. The Paro gneisses were stronger and coarse grained than the Darjeeling gneiss of higher metamorphic grade. The quartzites were stronger than the other rocks in the area. However, all the rocks were folded, faulted and sheared to varying degrees and they have been subjected to high levels of weathering along the drainage channels.

4.4 Lineament

It has been generally observed that the probability of landslide occurrence is increased at sites which are close to lineaments. The lineaments, not only affect surface material structures but also make contribution to terrain permeability, favoring slope instability. The lineaments showing fractures, discontinuities and shear zones were interpreted from the PAN and PAN-sharpened LISS III images. There was no major thrust/fault reported in the study area, but mega lineaments were identified from the satellite data. These lineaments were mapped from the sharpened image by visual interpretation on the basis of tonal contrast, structural alignments, rectilinear trends of morphological features and linear stream courses that were conspicuous by their abrupt changes in the course. These were digitized and subsequently rasterized to generate a lineament data layer. Five buffer zones with 125m intervals were generated to establish the influence of lineaments on landslide occurrence.

4.5 Drainage

Most of the landslides in hilly areas are generally initiated due to the erosional activity associated with drainage, and thus it is an important factor. The drainage map was prepared from the topographic maps by digitizing the streams in a vector layer. Later, this layer was overlaid on IRS LISS III image for cross-checking and modification of the map. The ordering of the drainages was performed on the basis of Strahler classification [30]. Subsequently, this map was rasterized. As the 1st and 2nd order drainages play an important role in initiating landslides, a 25m buffer zone for both these drainage orders was created and considered for further analysis.

4.6 Land use land cover

Land use land cover map in hilly areas generally shows distribution of forest cover, water bodies and types of land use practices. This is one of the key factors for landslide occurrences. The incidence of landslide is inversely related to the vegetation density. Hence, barren slopes are more prone to landslide activity as compared to the forest area. In this study, eight classes of land use land cover were considered. These classes are thick forest, sparse forest, tea plantation, agriculture, barren, built up, water bodies and river sand. This land use land cover layer was generated by employing supervised digital image classification approach using four spectral bands of LISS III image, Digital Elevation Model (DEM) and NDVI image. Maximum likelihood classifier was used to classify these data to produce a land use land cover classification with an overall accuracy of 94.39%. Subsequently, the land use land cover information in a very small portion of the classified image, which was masked by the cloud and its shadow in the original LISS III image, was obtained from field surveys. The remotely sensed derived land use land cover classification was updated with the field derived information for the masked region.

5 Methodology and Implementation

A neuro-fuzzy approach has been adopted in this study to carry out landslide hazard zonation mapping. This involves three main stages; i) determination of ratings for categories of thematic layers using cosine amplitude method, ii) determination of weights of thematic layers through ANN implementation and iii) integration of ratings and weights using GIS to arrive at the final LHZ map. The sequence of stages of methodology for LHZ is given in Fig. 2.

5.1 Implementation of fuzzy relation concept *via* cosine amplitude method

The cosine amplitude method as described earlier has been adopted to determine the ratings of the categories of factors. The landslide distribution and the different categories of thematic layers taken one at a time were considered as two data sets for the computation of rating or strength of relationship (r_{ij}). The pixels in the landslide areas were assigned as 1 and rest of the pixels were assigned as 0 in the landslide distribution layer. Similarly, a value of 1 is assigned to a particular category of a thematic layer and a value of 0 to rest of the pixels. Hence, in total 36 data layers in binary form were generated, which contain 35 layers of categories of thematic layers (Table 1) and one layer for landslide distribution. These layers were used for determination of r_{ij} in GIS. The results are provided in Table 1.

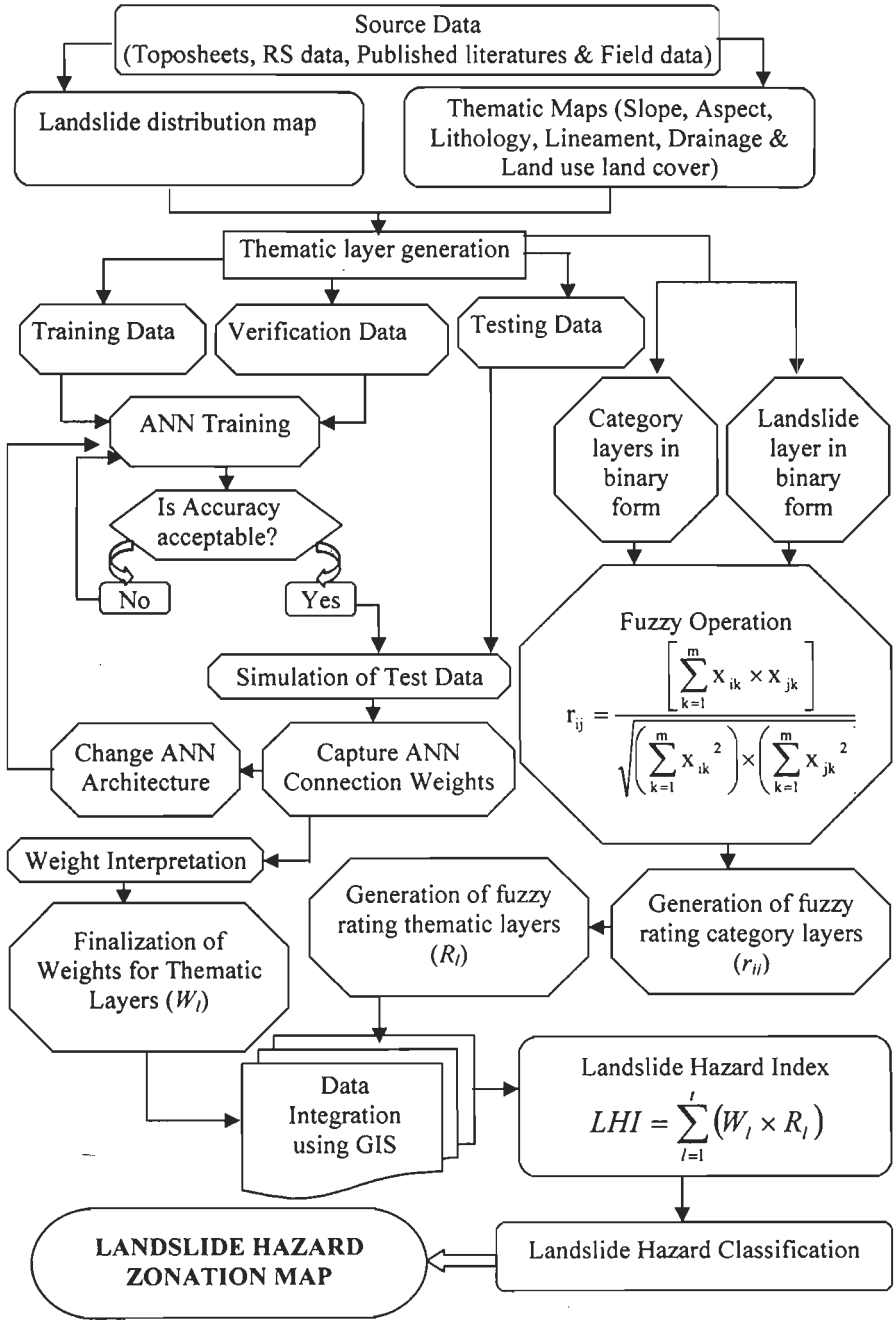


Fig.2. Flow diagram showing the methodology followed in the present study

Table 1. Fuzzy ratings for different categories of factors

Thematic layers	Categories	Number of pixels	Number of Landslide pixels	Fuzzy Rating (r_{ij})
Land use land cover	Agriculture Land	35692	85	0.0488
	Tea Plantation	142541	84	0.0243
	Thick Forest	72685	38	0.0229
	Sparse Forest	131088	65	0.0223
	Barren Land	14237	58	0.0638
	Habitation	10341	9	0.0295
	Water	970	0	0
	River Sand	1005	0	0
Lithology	Darjeeling Gneiss	73371	77	0.0324
	Feldspathic Graywacke	45938	61	0.0364
	Paro Gneiss	247242	158	0.0253
	Lingtse Granite Gneiss	20926	15	0.0268
	Paro Quartzite	12154	14	0.0339
	Reyang quartzite	8089	14	0.0416
Slope	0°-15°	51380	23	0.0212
	15°-25°	146974	117	0.0282
	25°-35°	144495	131	0.0301
	35°-45°	50246	58	0.0340
	>45°	14329	10	0.0264
Aspect	Flat	2072	0	0
	N	59880	22	0.0192
	NE	45077	32	0.0266
	E	52868	73	0.0372
	SE	45689	77	0.0411
	S	37630	49	0.0361
	SW	29860	20	0.0259
	W	55132	26	0.0217
	NW	79148	40	0.0225
Drainage Buffer	1st Order	116168	102	0.0296
	2nd Order	27690	44	0.0399
Lineament Buffer	0-125m	146761	243	0.0407
	125-250m	108929	35	0.0179
	250-375m	72380	36	0.0223
	375-500m	41360	17	0.0203
	>500m	38317	8	0.0144

5.2 ANN Implementation

A feed forward back-propagation multi-layer artificial neural network with one input layer, two hidden layers and one output layer was considered to determine the weights

of the causative factors in order to produce a LHZ map. The neurons in the input layer represent different thematic layers corresponding to causative factors. A pixel corresponds to a set of input values to ANN. The output corresponds to the presence or absence of landslide in the pixel. Hence, the output value is either 0 or 1. The number of neurons in the hidden layers was varied by running the networks several times to achieve the desired training and testing data accuracies.

One set each of training, verification and testing data were randomly generated from the study area. The data sets consist of 226 pixels each, out of which 113 pixels were landslide pixels and rest 113 pixels were no landslide pixels. All the pixels in the datasets were mutually exclusive [31]. The training dataset was used to train different network architectures while the verification dataset was used simultaneously with the training dataset to control the overtraining of the network. The testing dataset was used to evaluate the accuracy of the trained networks. The training, verification and testing data files for the input layer are two-dimensional matrices with rows representing values of the categories of each thematic layer and columns representing the number of pixels. The values of the categories in the input layer correspond to the weighted normalized rating or r_{ij} of the corresponding category (last column of Table 1). Similarly the training, verification and testing data files for the desired output are two-dimensional matrices with rows representing the presence or absence of landslide with values 1 or 0 respectively and columns representing the number of pixels.

The well known back-propagation learning algorithm was used to train the neural networks. About 100 neural network architectures were designed by varying the number of neurons in both the hidden layers. In this study, Levenberg-Marquardt algorithm (TRAINLM in MATLAB 6.1) for training function, gradient descent with momentum as learning function, mean squared error as performance function and sigmoid transfer function were used to train the networks. The maximum number of epochs was set at 10000 for each network. The training process was initiated and the arbitrary initial connection weights were constantly updated as the training progressed, until an acceptable accuracy was reached. The training accuracy observed for the networks was of the order of 75% to 90%.

The final weights of the trained network were used to simulate the testing data to evaluate the performance of the network. The testing accuracy observed for the networks was of the order of 60% to 70%. The adjusted weights of input-hidden, hidden-hidden and hidden-output connections for each network were finally captured for further calculations. Simple matrix multiplication was performed on these weight matrices to obtain a final 6×1 weight matrix for each network which represents the weights of six causative factors in this study. These causative factors were ranked according to the corresponding absolute weights for each network which means higher the value of absolute weight, more crucial the factor is for the occurrence of landslide. Considering all the 100 networks, the rank of a factor was decided based on the rank observed by maximum number of networks. Out of 100 networks, 41 networks categorized lithology as rank 1 (most important), 31 networks for lineament as rank 2, 30 networks for slope as rank 3, 27 networks for aspect as rank 4, 33 networks for land use land cover as rank 5 and 49 networks for drainage as rank 6 (least important).

Subsequently, the weighted normalized average of the weights of these networks for a particular factor was calculated and assigned as the weight of that factor for

preparation of LHZ map. The weights thus obtained through ANN for all the factors are listed in Table 2.

Table 2. Weights of thematic layers derived through ANN approach

Thematic layers	ANN derived weights
Lithology	4.807
Lineament Buffer	2.113
Slope	1.318
Aspect	1.065
Land use land cover	0.495
Drainage buffer	0.202

5.3 GIS based LHZ

The integration of 6 thematic layers representing the ratings for the categories (R_l) of the factors and weights for the factors (W_l) was performed by using simple arithmetic overlay operation in GIS. The Landslide Hazard Index (LHI) for each pixel of the study area was obtained by using the following equation.

$$LHI = \sum_{l=1}^{l} (W_l \times R_l) \tag{4}$$

The LHI values were found to lie in the range from 0.03042 to 0.40796. The observed mean (μ_o) and standard error (σ_o) from the probability distribution curve of these LHI values are 0.27605 and 0.03259 respectively. The LHI values were divided into five distinct classes (hazard zones) with boundaries at ($\mu_o - 1.5m\sigma_o$), ($\mu_o - 0.5m\sigma_o$), ($\mu_o + 0.5m\sigma_o$) and ($\mu_o + 1.5m\sigma_o$) where m is a positive, non-zero value (Saha et al., 2005). Several LHZ maps of the study area were prepared for different values of m. The cumulative percentage of landslide occurrences in various hazard zones ordered from very high hazard (VHH) to very low hazard (VLH) were plotted against the cumulative percentage of area of the hazard zones for these LHZ maps. These curves, defined as success rate curves by [32], [33], [34], were used to select the appropriate value of m to decide the suitability of a LHZ map. Five representative success rate curves corresponding to m = 1.2, 1.3, 1.4, 1.5 and 1.6 are shown in Fig.3. The suitability of any LHZ map can be judged by the fact that more percentage of landslides should occur in very high hazard zone as compared to other zones. It was observed from Fig.3 that for 10% of the area in very high hazard zone the curves corresponding to m = 1.2, 1.3, 1.4, 1.5 and 1.6 show the landslide occurrences of 43.9%, 45.6%, 46.7%, 43.3% and 43.9% respectively. Hence, for the first 10% area, the curve corresponding to m=1.4 has the highest success rate. Based on this analysis, the LHZ map corresponding to m = 1.4 appears to be the most appropriate one for the study area. Accordingly, the landslide hazard class boundaries were fixed at LHI values of 0.20761, 0.25324, 0.29886 and 0.34449.

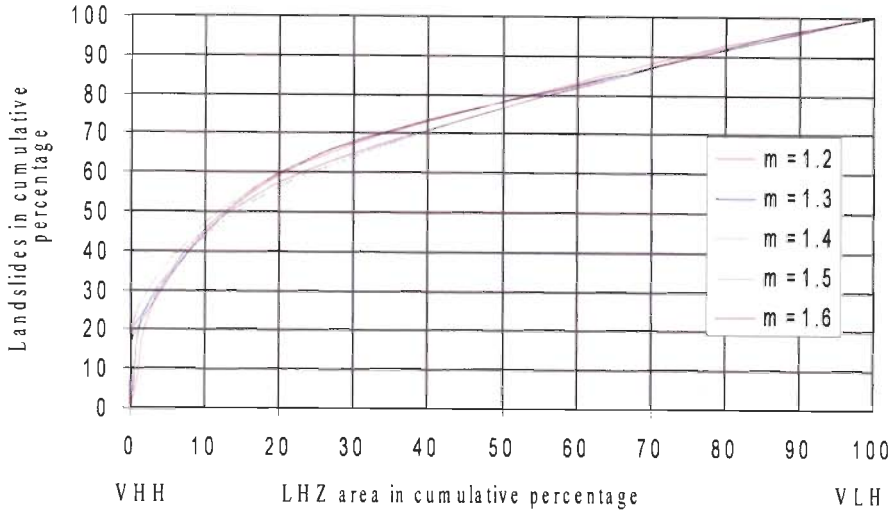


Fig.3. Success rate curves to select the best LHZ map

6 Results and Discussion

The LHZ map (Fig.4) was prepared using the above GIS-based neuro-fuzzy approach. The map shows that 2.3% of total area falls in very high hazard zone, 20.2% in high hazard zone, 48.4% in moderate hazard zone, 28.8% in low hazard zone and 0.3% in very low hazard zone. The percent area distribution of these hazard classes appears to be quite similar to that obtained in other LHZ studies in the Himalayan regions ([13], [15], [16], [18]). This confirms the usefulness of the approach developed in this study. Subsequently, the landslide distribution map was spatially cross-checked with the neuro-fuzzy derived LHZ map. The results show that about 30.1% of landslide area falls in very high hazard zone, 31.9% in high hazard zone, 26.5% in moderate hazard zone, 11.5% in low hazard zone and the very low hazard zone is devoid of landslides. From this LHZ map, it can also be observed that the buffers of lineaments have clearly indicated the very high and high hazard zones in the northern and south-eastern parts of the area. Overall, the buffers of lineaments have left traces on the LHZ map. Owing to the landslide susceptibility of the terrain, the lineaments ought to leave some traces on the LHZ map. Therefore, it indicates major influence of lineaments on landslide hazard zonation. Also, the Darjeeling gneiss rock type in south-eastern part, feldspathic greywacke and Reyang quartzite in the northern part of the study area have clearly indicated the hazard zones. Hence, it depicts the importance of lithology (i.e., rock types) on landslide hazard zonation. On the basis of these results, it can be concluded that the LHZ map derived from neuro-fuzzy weighting approach appears to conform with the real situations of landslide occurrences in the region and may thus be a useful way of integrating several factors

in an objective manner thereby minimizing the subjectivity in the weight assignment process.

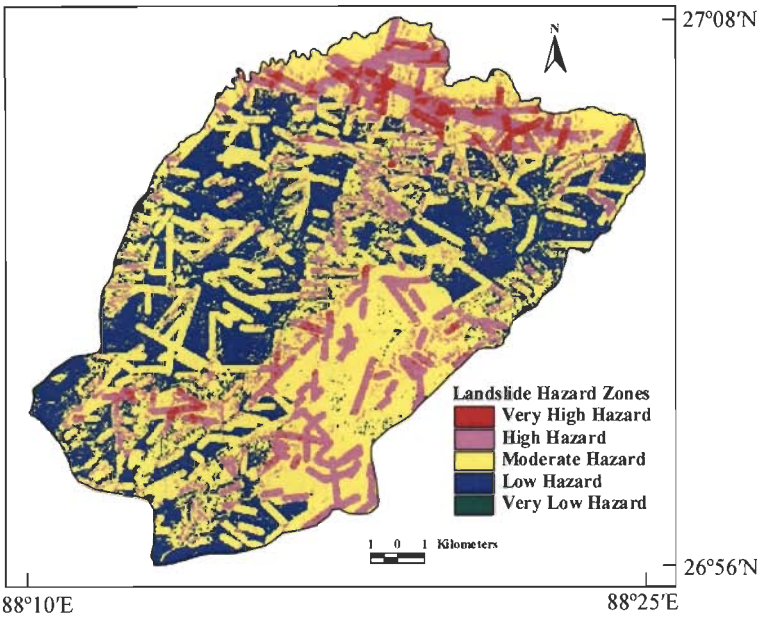


Fig.4. LHZ map using neuro-fuzzy weighting approach

7 Conclusions

Landslides are one of the major concerns worldwide because of loss of life and property every year in mountainous terrain. Hence, landslide hazard zonation mapping is essential to delineate potential disaster prone areas. In this study, a GIS-based neuro-fuzzy approach has been applied for the preparation of LHZ map. This methodology may serve as one of the successful objective approaches in this direction because of the fact that it can narrow down the potential hazard zones in a meaningful way for practical applications.

References

1. Aleotti, P., Chowdhury, R.: Landslide Hazard Assessment: Summary, Review and New Perspectives. Bull. Eng. Geol. Environ. **58** (1999) 21-44
2. Varnes, D.J.: Landslide Hazard Zonation: A Review of Principles and Practice. UNESCO, Paris (1984) 1-63
3. Anbalagan, R.: Landslide Hazard Evaluation and Zonation Mapping in Mountainous Terrain. Eng. Geol. **32** (1992) 269-277

4. Pachauri, A.K., Pant, M.: Landslide Hazard Mapping based on Geological Attributes. Eng. Geol. **32** (1992) 81-100
5. Gupta, V., Sah, M.P., Virdi, N.S., Bartarya, S.K.: Landslide Hazard Zonation in the Upper Satlej Valley, District Kinnaur, Himachal Pradesh. J. Him. Geol. **4** (1993) 81-93
6. Sarkar, S., Kanungo, D.P., Mehrotra, G.S.: Landslide Hazard Zonation: A case Study in Garhwal Himalaya, India. Moun. Res. Dev. **15**, 4 (1995) 301-309
7. Mehrotra, G.S., Sarkar, S., Kanungo, D.P., Mahadevaiah, K.: Terrain Analysis and Spatial Assessment of Landslide Hazards in parts of Sikkim Himalaya. Geol. Soc. Ind. 47 (1996) 491-498
8. Virdi, N.S., Sah, M.P., Bartarya, S.K.: Mass Wasting, its Manifestations, Causes and Control: some case Histories from Himachal Himalaya. In: Agarwal, D.K., Krishna, A.P., Joshi, V., Kumar, K., Palni, M.S. (eds.): Perspectives of Mountain Risk Engineering in the Himalayan Region. Gyanodaya Prakashan, Nainital (1997) 111-130
9. van Westen, C.J.: GIS in landslide Hazard Zonation: A Review, with Examples from the Andes of Colombia. In: Price, M., Heywood, I.: Mountain Environments and Geographic Information System. Taylor & Francis, Basingstoke U.K. (1994) 135-165
10. Gupta, R.P., Joshi, B.C.: Landslide Hazard Zonation using the GIS Approach – A case Study from the Ramganga Catchment, Himalayas. Eng. Geol. **28** (1990) 119-131
11. Gupta, R.P.: Remote Sensing Geology. Springer-Verlag, Berlin Heidelberg (1991)
12. Nagarajan, R., Mukherjee, A., Roy, A., Khire, M.V.: Temporal remote Sensing Data and GIS Application in Landslide Hazard Zonation of part of Western Ghat, India. Int. J. Rem. Sens. **19** (1998) 573-585
13. Gupta, R.P., Saha, A.K., Arora, M.K., Kumar, A.: Landslide Hazard Zonation in a part of Bhagirathy Valley, Garhwal Himalayas, using integrated Remote Sensing – GIS. Him. Geol. **20**, 2 (1999) 71-85
14. Saha, A.K., Gupta, R.P., Arora, M.K.: GIS-based Landslide Hazard Zonation in a part of the Himalayas. Int. J. Rem. Sens. **23** (2002) 357-369
15. Sarkar, S., Kanungo, D.P.: An Integrated Approach for Landslide Susceptibility Mapping using Remote Sensing and GIS. Photogram. Eng. & Rem. Sens. **70**, 5 (2004) 617-625
16. Saha, A.K., Gupta, R.P., Sarkar, I., Arora, M.K., Csaplovics, E.: An approach for GIS-based Statistical Landslide Susceptibility Zonation – with a case Study in the Himalayas. Landslides **2** (2005) 61-69
17. Elias, P.B., Bandis, S.C.: Neurofuzzy Systems in Landslide Hazard Assessment. Proc. 4th Int. Symp. Spatial Accuracy Assessment in Natural Resources and Environmental Science (2000) 199-202
18. Arora, M.K., Das Gupta, A.S., Gupta, R.P.: An Artificial Neural Network Approach for Landslide Hazard Zonation in the Bhagirathy (Ganga) Valley, Himalayas. Int. J. Rem. Sens. **25**, 3 (2004) 559-572
19. Ercanoglu, M., Gokceoglu, C.: Use of Fuzzy Relations to Produce Landslide Susceptibility Map of a Landslide Pron Area (West Black Sea Region, Turkey). Eng. Geol. 75, 3&4 (2004) 229-250

20. Gomez, H., Kavzoglu, T.: Assessment of Shallow Landslide Susceptibility using Artificial Neural Networks in Jabonosa River Basin, Venezuela. Eng. Geol. **78**, 1-2 (2005) 11-27
21. Zadeh, L.A.: Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Trans. Syst. Man. Cybern. SMC-**3**, 1 (1973) 28-46
22. Ross, T.J.: Fuzzy Logic with Engineering Applications. McGraw-Hill, New York (1995)
23. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
24. Gong, P.: Integrated Analysis of Spatial Data for Multiple Sources: using Evidential Reasoning and Artificial Neural Network Techniques for Geological Mapping. Photogram. Eng. & Rem. Sens. **62** (1996) 513-523
25. Schalkoff, R.J.: Artificial Neural Networks. Wiley, New York (1997)
26. Olden, J.D., Joy, M.K., Death, R.G.: An Accyrate Comparison of Methods for Quantifying Variable Importance in Artificial Neural Networks using Simulated Data. Ecological Modelling **178** (2004) 389-397
27. Acharya, S.K., Ray, K.K.: Daling Group and Related Rocks. Geol. Surv. Ind. SPN-**22** (1989)
28. Banerji, P.K., Guha, P.K., Dhiman, L.C.: Inverted Metamorphism in the Sikkim-Darjeeling Himalaya. Geol. Soc. Ind. **21** (1980) 330-342
29. Dhakal, A.S., Amada, T., Aniya, M.: Landslide Hazard Mapping and its Evaluation using GIS: An Investigation of Sampling Schemes for a Grid-cell based Quantitative Method. Photogram. Eng. & Rem. Sens. **66**, 8 (2000) 981-989
30. Strahler, A.N.: Quantitative Geomorphology of Basins and Channel Networks. In: Chow, V.T. (ed.): Handbook of Applied Hydrology, Mcgraw Hill, New York (1964)
31. Foody, G.M., Arora, M.K.: An Evaluation of some Factors Affecting the Accuracy of Classification by an Artificial Neural Network. Int. J. Rem. Sens. **18** (1997) 799-810
32. Chung, C-J.F., Fabbri, A.G.: Probabilistic Prediction Models for Landslide Hazard Mapping. Photogram. Eng. & Rem. Sens. **65** (1999) 1389-1399
33. Lu, P.F., An, P.: A Metric for Spatial Data Layers in Favorability Mapping for Geological Events. IEEE Tran. Geosci. Rem. Sems. **37** (1999) 1194-1198
34. Lee, S., Choi, J., Min, K.: Landslide Susceptibility Analysis and Verification using the Bayesian Probability Model. Environ. Geol. **43** (2002b) 120-131