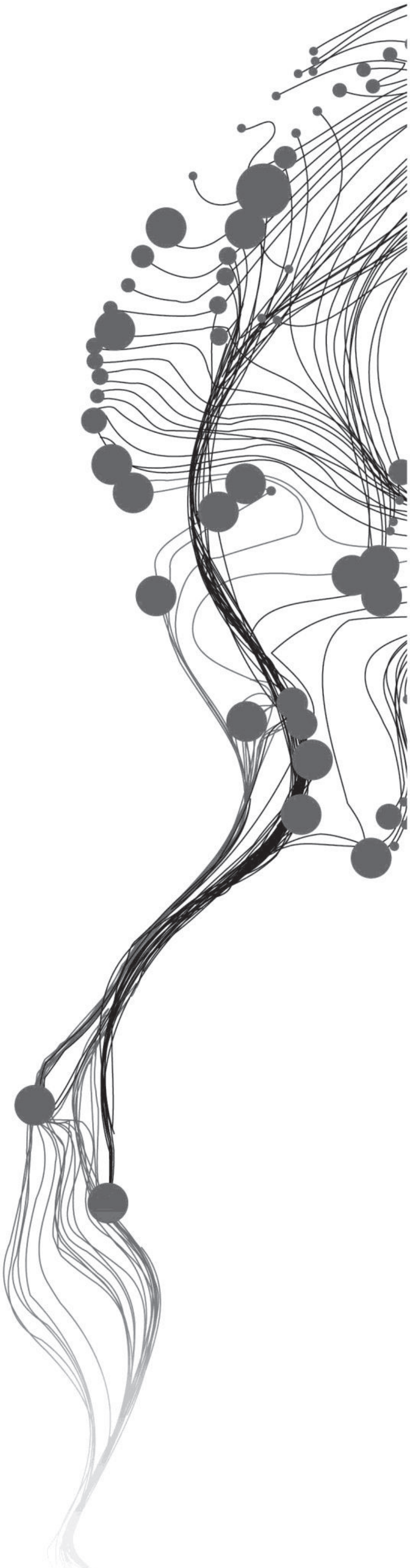
# EXTRACTION OF FOREST STRUCTURAL ATTRIBUTES AS INDICATORS OF LANDSLIDE-INDUCED DISTURBANCE USING LIDAR DATA

AMEYA GODE
FEBRUARY, 2012

SUPERVISORS:
Dr. H.A.M.J. van Gils
Dr. M.W. Straatsma



# EXTRACTION OF FOREST STRUCTURAL ATTRIBUTES AS INDICATORS OF LANDSLIDE-INDUCED DISTURBANCE USING LIDAR DATA

AMEYA GODE
Enschede, The Netherlands, February 2012

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: [Natural Resources Management]

SUPERVISORS:
Dr. H.A.M.J. van Gils
Dr. M.W. Straatsma

THESIS ASSESSMENT BOARD:
Dr. Y. A. Hussin (chair)
Dr. M. Gerke (external examiner, ITC)

# ABSTRACT

Europe's forests were hit hard by storms in 1990 and 1999, which caused 120 and 180 million m3 of damage respectively .Since forests are important in the context of carbon sequestration and biodiversity conservation, this triggered off a series of studies of forest disturbances/damage. Natural Resource scientists are researching many of these disturbance factors in order to understand them better and to develop control or mitigation methods. One of these natural disturbance factors is landslide activity. This study focuses on the effect of such slow moving translational, rotational and earth flow processes on forest in the Bois Noir landslide in the Alpes de Haute province of France. Previous researches in this area to study landslides under forests have been done with different methods like dendrogeomorphology and climate data analysis. This study makes use of high resolution airborne LiDAR data (with a mean point density of 180 points per $m^2$.). Since LiDAR has long proven to be the best known remote sensing application for studying forest structure, this study was also aimed at assessing the applicability and accuracy of using LiDAR data.

The prime focus of this study was to extract various forest structural attributes from the LiDAR point cloud and compare these attributes between 'stable' and 'unstable 'zones. From previous research and ecological knowledge, it was decided that the following structural attributes are to be taken into consideration while doing analysis: Tree Height, Diameter at Breast Height (DBH), tree inclination and orientation and canopy gaps.

The LiDAR point cloud was subjected to normalisation and gridding to form a canopy height model (CHM), which was used for most of the analysis. A region growing approach in ECognition was then used to segment tree crowns and canopy gaps. Also, any human induced features, edaphic features and 'obvious' areas where landslide had caused massive tree fall were excluded from analysis. This approach gave a 'gap map' of the area and also information about tree heights and tree density, which was then subjected to statistical analysis.

To validate the LiDAR data and assess its accuracy, field data was collected in the month of September 2011. From the analysis, it was found that tree heights can be estimated with an $R^2$ value of 0.72. The gap detection accuracy was 81%, in a canopy height model of 15cm grid size. Furthermore, it was proven that statistically there is a significant difference in the distributions of canopy gap area, gap shape and tree heights between stable and unstable zones. Thus, these can be considered as indicators of landslide activity, but it also needs some multi-temporal analysis to study the dynamics of these structural attributes.

# ACKNOWLEDGEMENTS

I would like to acknowledge all those people who have directly or indirectly supported this research.
Firstly, I would like to thank my first supervisor, Dr. Hein van Gils, for his support and guidance throughout the research, from field to office.
Also, many thanks to Dr. Menno Straatsma for his guidance and suggestions.
I would also like to acknowledge Khamarrul.A.Razak for his tireless efforts from fieldwork to post field data processing, including provision of a landslide delineation map.
Special thanks also go to Gregorijs Goldbergs, EOS department, ITC who orthorectified an aerial photograph for us in a crunch situation and which was used for guidance in fieldwork.
Last but not least, I would also like to thank my parents.

***"This research is dedicated to Café Rocks, Enschede, which made our weekends worthwhile"***

## TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

## 1.1. Forests in today's world:

On the world environment day (5th of June) 2011, the head of the United Nations Environment Program (UNEP), Achim Steiner stated that "*In fact we have learned through the understanding of the impacts of climate change, the loss of eco-system services, the prospects of water shortages, the impact of environmental disasters that our dependence on nature, on a healthy and functioning set of eco-systems remains critical to human well being, to the well being of our economy and indeed to the prospects of being able to sustain and feed and provide with the basic needs an ever growing population.*"

At about the time humankind discovered agriculture, there were approximately 6.2 billion hectares of forests covering the earth. Presently, approximately 4.0 billion hectares remain. Thus one third of the world's forests have already been lost. Much has been said about the advantages of forests and their vitality in maintaining the world climate stability.

Although most of the forest damage is anthropological, the damages caused to forests by natural hazards cannot be overlooked. Also, many of the causal agents of natural disasters are due to a direct or indirect effect of human beings.

Europe's forests were hit hard by storms in 1990 and 1999, which caused 120 and 180 million m3 of damage respectively (UNECE/FAO2000b). Since forests are important in the context of carbon sequestration and biodiversity conservation, this triggered off a series of studies of forest disturbances/damage. Storms, wild fires, insects, biological invasions and landslides were established to be some of the major factors causing forest damage. These can be on a continental scale or a regional to local scale. Also the total forest area and total stand volume in Europe has increased during the last century (Holmsgard 1982, UN-ECE/ FAO 2000c), which can also be perceived as a greater resource vulnerable to damage.

Thus, it becomes quintessential to detect disturbances in forest ecosystem, especially in areas which are prone to natural hazards. However, accessibility and costs of fieldwork also remain an issue of concern. Remote sensing techniques have been applied by many researchers to analyze these disturbances(Thomas et al., 2011) However, the applicability of using low to medium resolution satellite data is limited, especially for detecting growth anomalies and understory vegetation. (Hais, Jonásová, Langhammer, & Kucera, 2009)

Management of forest ecosystems to sustain desired benefits requires knowledge of how forests change over time in response to natural disturbances and management activities. Disturbances include both stresses and destructive agents; these include invasive species (diseases as well as plant and animal pests); fire; changes in climate and serious weather events such as hurricanes and ice storms; pollution of the air, water, and soil; real estate development of forest lands; and timber harvest. Some of these are caused by humans, in part or entirely, others are not. Some problems may not be obvious, others are painfully so—such as when gypsy moth populations are so large that when GM frass (excrement) falls, it can be heard.

Many of these changes can seriously affect the ability of particular fish and wildlife to inhabit wildland ecosystems. Natural Resource scientists are researching many of these disturbance factors in order to understand them better and to develop control or mitigation methods.

## 1.2. Landslides and forest disturbances:

The term "landslide" describes many types of downhill earth movements ranging from rapidly moving catastrophic rock avalanches and debris flows in mountainous regions to more slowly moving earth slides (www.redcross.org). Landslides can be slowly moving, or can be a sudden event, rapidly destroying the features in its path. Gravity is the prime driving force for the landslide once the initiation is done. The factors that can cause this initiation are heavy rainfall, erosion, poor construction practices, freezing and thawing, earthquake, and volcanic eruptions. Landslides are typically associated with periods of heavy rainfall or rapid snowmelt and tend to worsen the effects of flooding. Areas burned by forest and brush fires are particularly susceptible to landslides. (Flageollet, Maquaire, Martin, & Weber, 1999)

Another geomorphological process which is prominently seen in mountainous regions is debris flow. Debris flows—sometimes referred to as mudslides, mudflows, lahars, or debris avalanches—are common types of fast-moving landslides. These flows generally occur during periods of heavy rainfall or rapid snowmelt. They usually start on steep hillsides as shallow landslides that liquefy and accelerate to speeds that are typically about 10 miles (16 kilometres) per hour, but can exceed 35 miles (56 kilometres) per hour.

Landslides constitute a common mass movement process and a widespread hazard in mountain and hill slope environments where they repeatedly cause damage and destruction to settlements, transportation corridors, or even lead to the loss of life (John F, 1978). To avoid damage or fatalities, data are needed on the frequency and magnitude of past events to perform an appropriate hazard assessment. The damages that landslides cause to life and property are well known, and have been studied extensively for disaster management. Much has been written on the impacts of landslides on the total environment, including effects on people, their homes and possessions, farms and livestock, industrial establishments and other structures, and lifelines. However, few authors have discussed the effects of landslides on the natural environment, i.e., on (1) the morphology of the Earth's surface, particularly that of mountain and valley systems, both on the continents and beneath the oceans; (2) the forests and grasslands that cover much of the continents, and (3) the native wildlife that exist on the Earth's surface and in its rivers, lakes, and seas.(Schuster & Highland, 2003). In this study, we will deal with the effect of landslides on forest structure, namely tree structural parameters which are affected by the process. The word 'effect' has been used here instead of 'damage' because the effect of landslides on forest ecosystem is not necessarily negative.

Landslides or other natural disturbances are some of the key factors which lead to regeneration in a forest, and also help maintain the health of the ecosystem. This process could be even compared to a human body, where a disease is supposed to "cleanse" the insides and produce more immune cells. Moreover, recent studies indicate that natural disturbances are important for forest regeneration and the maintenance of species diversity in riparian vegetation (Baker, 1990). Morphological alterations that landslides cause on forests, like creating gaps in canopies, transporting layers of soil, water to low-lying regions have a great impact on the ecological balance of forests. The heterogeneous topography of riparian forests caused by these disturbances has a great influence on the distribution and regeneration of trees (River, 2005).

### 1.3. LiDAR (Light Detection and Ranging):

LiDAR (Light Detection and Ranging, also LADAR) is an optical remote sensing technology that can measure the distance to, or other properties of a target by illuminating the target with light, often using pulses from a laser. Lidar systems are active remote sensing devices that measure the time of travel needed for a pulse of laser energy sent from the airborne system to reach the ground and reflect back to the sensor. The time measurement is converted into a distance measurement that is used to derive a precise three dimensional characterization of reflecting ground surfaces, including forest vegetation. In areas with dense vegetation cover, lidar pulses will mainly reflect from the top and from within the vegetation canopy, with some laser pulses penetrating to the ground and therefore providing an accurate ground elevation (Sorin C. Popescu, Wynne, & Nelson, 2002).

Research on LiDAR change detection has only begun with a few studies using topographic change mapping to monitor coastal erosion(Flageollet, et al., 1999; http://www.csc.noaa.gov/crs/tcm) . Airborne LiDAR has demonstrated its capability to measure useful canopy properties, such as height and cover, using commercial airborne laser scanning systems(Næsset, 2002). Unlike two-dimensional imagery, the vertical component of three-dimensional LiDAR data allows the analyst/user to separate ground vs. vegetation information, which is a prerequisite to most LiDAR applications, many of which focus exclusively on one or the other component(Hudak, Evans, & Stuart Smith, 2009)

The most common use of LiDAR is to build a Digital Elevation Model (DEM), which is the primary product of LiDAR survey, and used for various applications (namely geomorphological, hydrological).The DEM is essentially an interpolated surface that represents topographic variation in three dimensions. The 'object' above the ground in a DEM might be buildings in developed areas or vegetation in forested areas. Majority of LiDAR data users however look no further beyond the DEM, ignoring the residual variation which represents the features above the earth's surface. The ground vs non ground features are not independent of each other, which can be easily proven by studying a LiDAR derived DEM in a vegetated and a non vegetated area. The DEM of a non vegetated area is much 'smoother', as compared to that of a forested environment, where the DEM appears much 'rougher', which is because lesser points penetrate through the canopy of trees to hit the ground. The points hitting the canopy (or above ground features) which are considered 'noise' by LiDAR users dealing with the geomorphological features, are in fact an ace for natural resource managers or wildlife ecologists. Most obvious is the canopy height information, while another measure having a direct physical basis is percent canopy cover, calculated as the percentage of LiDAR returns intercepted by the vegetation canopy, within a bin size (cell resolution) specified by the user(Hudak, et al., 2009).

Thus, Nelson (Nelson, Krabill, & Tonelli, 1988) recommended the use of the laser-derived stand profiles for the retrieval of stand characteristics of a forest. Various studies have shown that height, volume, biomass, crown diameter, stem density, or diameter at breast height estimates can be produced using LiDAR data (Næsset, 2002). Moreover, LiDAR can also be used to detect forest disturbances, either evident from data (such as treefall) or derived from tree structural attributes. Mackey (Mackey, Roering, &

McKean, 2009) manually recognized trees on the single LiDAR-derived image and five historical aerial photos and tracked tree displacement over 42 years in order to quantify decadal-scale slide deformation and observed the long-term sediment flux in earth flow-prone terrain. The LiDAR processing techniques are becoming so advanced that a study in 2004 reliably delineated canopy gaps less than 1 m2 (Vepakomma, St-Onge, & Kneeshaw, 2011).

### 1.4. Problem Statement:

Hudak (2009) groups LiDAR applications for forestry into 3 categories:

(1) Characterization of forest structure which includes canopy surface, canopy interior and individual trees;
(2) Natural resource applications which encompass forest inventory, fire and fuels, ecology and wildlife, geology, geomorphology, and surface hydrology; and
(3) Sensor integration.

A lot of study has been done with LiDAR to characterize forest structure and make a forest inventory for calculating biomass with high accuracy, because carbon is currently the 'hot topic' in the world (S.C. Popescu, Wynne, & Nelson, 2003).

Also, LiDAR has been used for change detection, because it can facilitate the process at a very minute level. Change detection for individual trees was indeed a huge breakthrough in forest management, when Persson (2001) concluded that individual trees can be detected with up to 70% accuracy and tree height can be measured with an accuracy better than 1 m. Laser based crown delineation is a better technique in comparison to single tree-based estimation with orthoimages, since it measures the geometrical properties of trees directly(Hyyppä & Hyyppä, 1999).

So far, many researchers have attempted to characterize forest stand/ tree structural parameters as indicators of forest disturbance (J.R. Runkle, 1982). Comparisons also have been made between forest responses to natural versus human induced disturbances.

An alternative method of studying forest disturbances induced by landslides has been dendrogeomorphology. The study of tree rings tell us about the age of trees and the time when the landslide occurred. The type of deformation in trees tells us about the type of landslide affecting the tree. A study done in 2011 concludes that "comparison of tree ring data with historical records and aerial photographs clearly demonstrates the spatiotemporal accuracy of the reconstruction of landslides"(Lopez Saez et al., 2011). However, these studies have been more for *reconstructing past landslides*.

Thus, for researchers attempting to study forest disturbances induced by landslides, the two biggest problems were:

1) Resolution of data being low
2) Lack of enough field data for validation.

The researches that have been done to study forest disturbances mostly focus on canopy gaps as indicators of disturbance. In 1770 Finnish botanist Pehr Kalm, in his book Travels in North America, was probably the first researcher to describe the general occurrence and important ecological role of tree-fall disturbance in primeval temperate forests. Since then, canopy gaps have been of prime importance when it came to studying forest disturbances. Koukoulas and Blackburn (2004) combined LiDAR imagery and GIS to quantify the spatial properties of canopy gaps. Most of the studies have been dealing with gap dynamics, as a function of the regeneration capacity of the forest stand. A study concludes that multi-temporal medium density LiDAR enables the detection of new gaps with a very high accuracy, and can potentially be used to measure growth on an individual crown, or window basis(St-Onge & Vepakomma, 2004).

Thus, large-area forest inventories using sample plots are perhaps the best applications for laser scanning at the individual tree level, concludes a study(Yu, Hyyppa, Kaartinen, & Maltamo, 2004).

So, we see that forest stand characteristics have been studied for biomass and ecological study purposes. Very few studies have been done to characterize them as indicators of forest disturbance. In fact, a research paper states its follow up as "By exploring long-term and large-scale mortality and recruitment processes, we should be able to validate or improve our understanding of forest succession processes developed from earlier small spatial and temporal studies. Such new insights may have direct implications for forest managers who seek silvicultural and management strategies with a natural disturbance based underpinning"(Vepakomma, St-Onge, & Kneeshaw, 2008).

## 1.5. Objectives:

There is a serious lacking of studies which use LiDAR data to extract tree structural parameters as indicators of disturbance. The latest effort to do so was by Razak (2011), who characterized tree inclination and tree height dissimilarities as indicators of landslide activity.
Thus, the aim of this study is to come up with more indicators of landslide activity. The main objective of the study can be phrased as "to extract tree structure anomalies from high point density LiDAR data and establish them as indicators of landslide-induced forest disturbances".
Since this single objective is too broad, and unspecific about which tree structure anomalies are to be studied, some research specific objectives were also framed, which are as follows:

1. To collect tree structure data from the field, namely tree height, diameter at breast height, inclination angle, type of deformation, canopy area, tree orientation, tree density and canopy gaps from both disturbed and undisturbed forest for validation and comparison.
2. To extract canopy gap area, shape and location.
3. To extract other tree structural parameters (tree height, DBH, canopy area, inclination angle) from the high density LiDAR data and classify them by tree species, tree age and type of landslide.
4. To compare these parameters of a landslide affected forest to those of a forest unaffected by landslides to study which parameters show maximum variability.

These research objectives lead to research questions, namely:
1.1. Which tree structural parameters differ significantly in affected and unaffected forests?

1.2. What is the relation between structural anomaly and the type of the landslide affecting the tree?

2.1. To what extent can we derive tree parameters from the high density LiDAR data?

2.2. How reliable is the LiDAR data for analysing the tree structure?

So, to proceed with answering these research questions, some research hypotheses had to be put forward.

Research Hypotheses:
1. Parameters like tree height, DBH and Canopy Gaps differ highly in 'stable' and 'unstable' forest.

2. There is a strong relationship between the type of landslide (rotational, transitional, rock fall) and the type of deformation observed.

3. Tree parameters can be derived from LiDAR data with (more than) 70% accuracy.

Thus, the main tree structural parameters which will be focussed on are tree height, DBH, crown projection area, inclination and orientation of tree, and gaps in the canopy. These were chosen because of two reasons:

i) Damage to a forest can be of two types: ecological and mechanical. Ecological damages should be studied over a certain period of time, which involves extensive field work, which requires a lot of time and money. Thus the mechanical damages can be studied from the most obvious of tree structural attributes, as above.

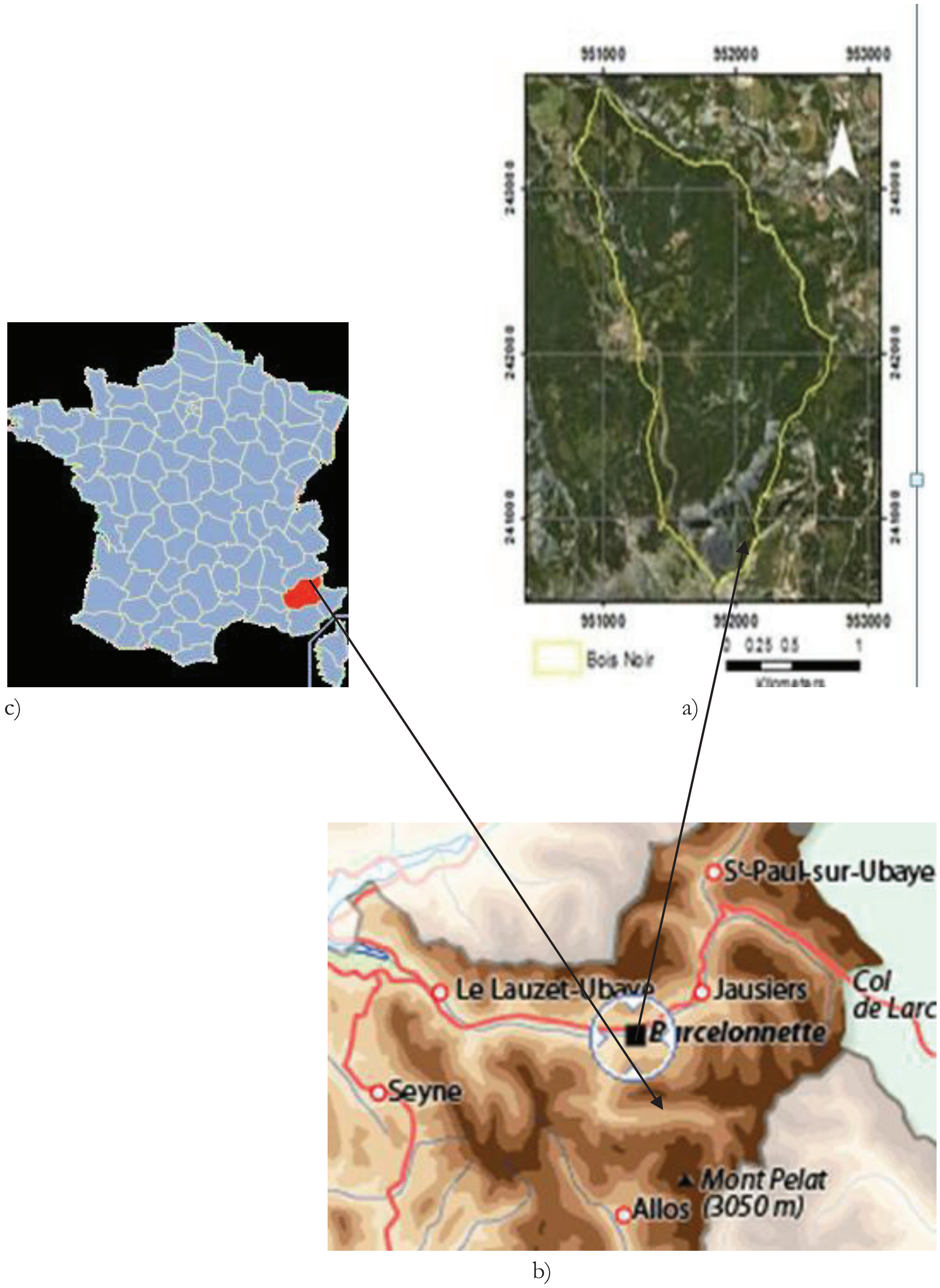
ii) Previous researches (irrespective of which method was used to study disturbances) indicate that the above listed parameters were the best indicators of forest disturbance. As an example, in 1982, James Runkle studied canopy gaps for the geography of forest disturbance(J.R. Runkle, 1982). Extensive field work was done in this study, and the author also wrote a manual with guideline protocols for sampling canopy gaps, which was followed while sampling in field for this study too.

# 2. MATERIALS AND METHODS

## 2.1. Study area:

The Bois Noir landslide(44°23'27 N, 6°45'27 E; is situated in Barcelonnette basin in the South-western French Alps, a tributary of the Ubaye river, Aples de Haute Provence, France. The landslide body is 300 m long, 500 m wide, and ranges from 1,600 to 1,680 m a.s.l. in elevation. The Bois Noir slope segment is characterized by an irregular topography with slope gradients ranging between 10° and 35°(Thiery, Malet, Sterlacchini, Puissant, & Maquaire, 2007), and the site is covered by *Pinus uncinata* and grasslands(K. A. Razak, et al., 2011). Lying on an East West axis, the broad Barcelonnette basin slopes up from 1100 to 3000 m altitude(Flageollet, et al., 1999). The Bois Noir catchment is characterized by a dry and mountainous climate with strong inter-annual rainfall variability (e.g. annual rainfall may vary between 400 and 1400 mm). These prime geomorphic and climatic factors explain the development of the slope by rotational or translational shallow landslides which usually affect the uppermost 2 to 6 m (Thiery, et al., 2007). Forest covers 92% of the total surface area and consists mainly of black pine with some deciduous trees (Thiery et al., 2004). All the landslides composing the landslide complex are typically shallow and occur at the interface between the bedrock and the surface deposits. In a study, this area has been identified as highly susceptible to landslides although most of the area is covered by vegetation(Thiery, et al., 2007).A landslide inventory map at 1:10,000 scales, based on aerial-photo interpretation, field surveys and historical records was created. Tilted and deformed trees, as well as recent scarps and open cracks, clearly indicate that the Bois Noir landslide has been subject to multiple reactivation in the recent past. All the landslides composing the landslide complex are typically shallow and occur at the interface between the bedrock and the surface deposits. These deformed trees, or *drunken trees* as they are more popularly called, will be the main interest in this study.

*Fig.2.1. A field photograph showing 'drunken trees' in a landslide affected zone in the study area.*




a) Bois Noir landslide
b) Study area: Barcellonnette basin.
c) France map: Google maps

*Fig2.2. Study area location.*

## 2.2. Materials:

### 2.2.1. Data:

a) Landslide map:
A map showing the whole study area delineated into stable and unstable zones was prepared by Khamarrul.A.Razak (2011). This was to be the basis of comparison.
b) Ortho photo:
An aerial photograph of the study area captured at the same time as the LiDAR data acquisition was available for use, especially for guidance in field. Resolution of the photograph was 30cm.
c) LiDAR data:
An HDAL dataset was acquired in July 2009 using a hand-held laser scanning system. This system consists of a RIEGL VQ-480 laser scanner, a Topcon Legacy GGD GPS and an iMAR FSAS inertial measurement unit (IMU). Specifications are given in Tab. 1. An airborne LiDAR campaign was carried out under snow-free conditions using a helicopter flying about 300 m above the ground. Several flight lines were acquired over the same area to increase the point density over the forested terrain. Here we used about eight million points with a mean point density of 180 points $m^{-2}$.

| | |
|---|---|
| Acquisition (month/year) | July 2009 |
| Laser scanner | RIEGL VQ480i |
| IMU system | IMAR FSAS (record up to 500 Hz) |
| Positional system | Topcon legacy (record up to 5Hz) |
| Laser pulse repetition rate | 300 kHz |
| Beam divergence | 0.3 mrad |
| Laser beam footprint | 75 mm at 250 m |
| | 60° |
| Scanning method | Rotating multi-facet mirror |

Table 2.1: Metadata for the Airborne LiDAR data acquisition

### 2.2.2. Fieldwork Materials:

Measurement of tree structural attributes was a task which had to be completed with the highest level of accuracy possible, as the field data was to be used for validation of attributes extracted from LiDAR data. The following instruments/materials were used for fieldwork:

1. Caliper (60cm): For measuring Diameter at Breast Height (DBH) of trees.
2. Nikon laser rangefinder: For measuring height of trees.
3. Suunto clinometer PM5 : For measuring orientation of trees.
4. Compass: For measuring orientation of trees.
5. Diameter tape: For measuring canopy width and axes of canopy gaps.
6. Leica differential GPS system 1200: For getting coordinates of the observed points with millimeter level accuracy.

### 2.2.3. Software used:

1. LAStools.
2. ECognition.
3. ArcGis 10.

4. Quick Terrain Modeler (trial version)
5. SPSS 16.

## 2.3. Methodology:

### 2.3.1. Fieldwork:

A field campaign was carried out from in September 2011.

Sampling was carried out on a purely purposive basis, because of the following reasons:

1)Accessibility: Since the area has been subjected to landslides and earth/debris flows over the last few centuries, the whole landscape has a very rugged topography, with slopes having an uneven gradient and ridges and cracks running through the area. Thus, plots and transects had to be chosen in areas which are accessible, and make it possible for the researcher to stand near the tree for measurement!

2). Objectives: Three different researchers, three different objectives (namely, gap analysis, carbon estimation and forest inventory). Thus, sampling had to be done in:

i) Unstable areas with drunken trees.

ii) Stable areas with straight trees(with uniform species)

3) GPS: A Leica differential GPS system 1200 combined with a total station was used to get millimeter level accuracy for tree position. Thus, from the previously established GPS stations a transfer had to be done to establish the 'center' of the plot. This was a fairly time-consuming job, and thus the stable and unstable plots had to be within visible range of each other to facilitate the positioning of the GPS.

The sampling was carried out in two phases, transects and plots.

1. **Transect based sampling**: This sampling was done in form of a line transect of 25 m, to sample for canopy gaps. The length of transect was first measured using a Nikon laser rangefinder, and gaps were located and their coordinates taken with the Leica GPS. The thresholds set for an area to be defined as a 'gap' were:

i) The understory vegetation should be no more than 5 meters in height.

ii) The surrounding vegetation should be a minimum 7 meters in height.

The serial numbers of surrounding trees, the length of major and minor axis of gap (measured using diameter tape) were noted down and 2 photographs of each gap were taken using a Nikon camera combined with an SLR lens, keeping the camera on a flat surface at 1 meter height from the ground. Transects were sampled in both stable and unstable areas, and one in a rockfall affected area, making the total number of transects 10, and the number of gaps 43. The trees in transects were exclusively coniferous.

2. **Plot based sampling**:

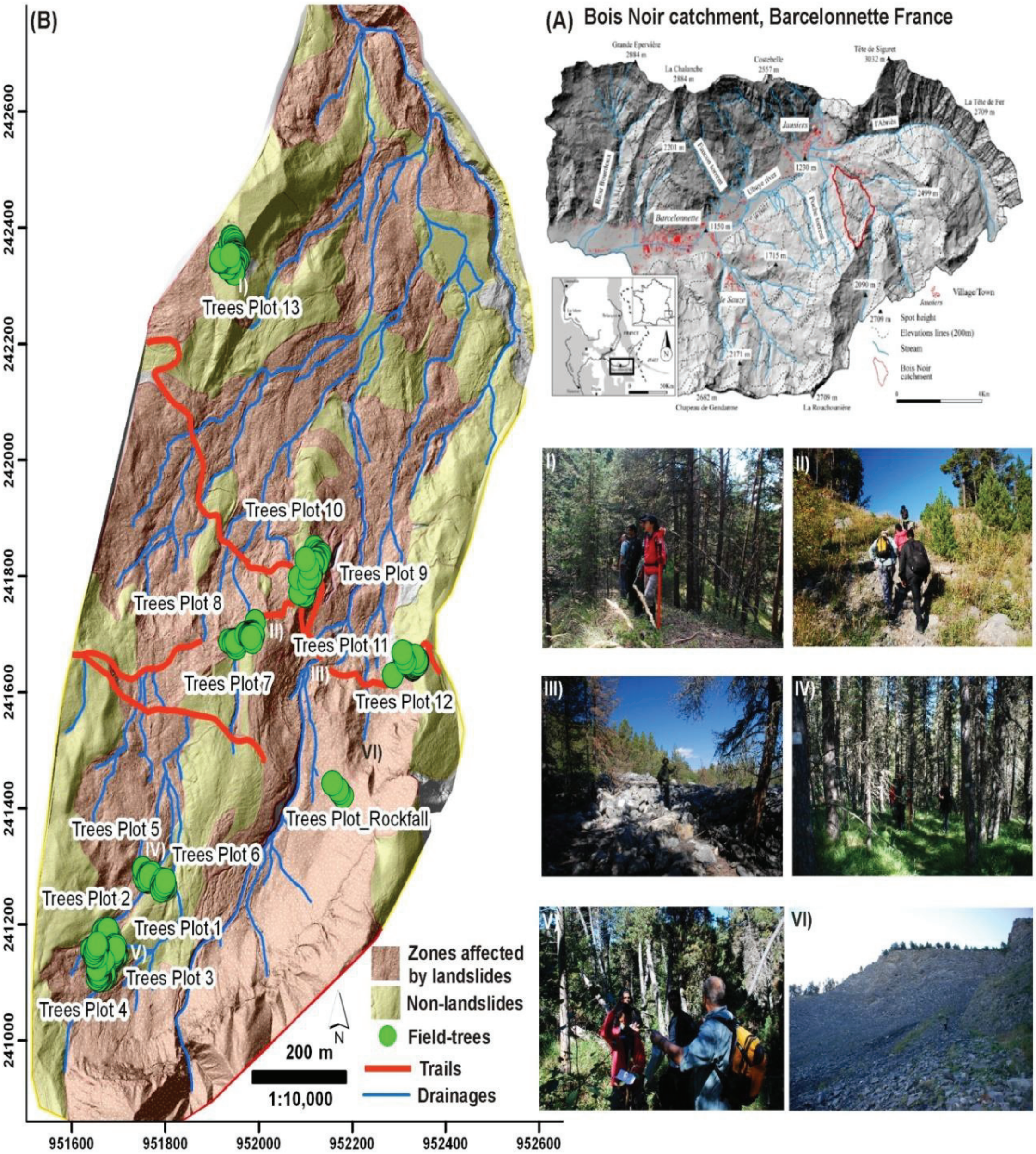
This sampling was done to measure tree structural attributes, namely height, diameter at breast height (DBH), canopy width, inclination, orientation and type of deformation if any. Plots were selected on a purposive basis, and in both stable and unstable areas. Based on the range of visibility from the total station, trees were selected in the plots and the attributes were measured. Canopy width was measured in North – South direction and East to West direction with a measuring tape. Tree height was measured with a Nikon laser rangefinder. A caliper was used to measure DBH. For the inclination and orientation, a Suunto PM5 and were used respectively. A total of 13 plots were sampled, 6 in the unstable areas, 7 in stable areas. The plots were spread out throughout the study area as evenly as possible.

3. **Age sampling**:

   As an additional approach, increment boring was done on 6 trees in every plot to get the age of the respective trees.

4. **Additional information:**

   The slope of the area where a plot or transect was located was noted down. Also, some other disturbances in the forest like disease, wind throw were detected. Moreover, across all the plots in the south, the main problem detected was that of lack of thinning, which could be a management flaw; and the presence of mistletoe parasite on most of the trees.




*A) An overview of the Bois Noir catchment; B) Study area delineated into stable and unstable zones showing the locations of sampling, prepared by Khamarrul Razak ; I to VI) fieldwork photos*

*Fig.2.3. Overview of field work.*

### 2.3.2. Primary Steps using LAStools:

Primarily, the dense LiDAR point cloud had to be filtered as a part of steps towards building a Canopy Height Model, or any other kind of analysis. Different filtering algorithms are available in the literature. Methods include TIN densification, iterative residual analysis, region growing, multiscale curvature analysis, or segment based classification. A comparison can be found in Sithole and Vosselman (Sithole & Vosselman, 2004).The first very step towards filtering the data was to extract ground points from the data, as the ground points come in use while calculating height, or inclination and orientation. For this purpose the software LAStools was used.

LAStools is a freeware which provides an easy-to-use platform that implements reading and writing of LIDAR points from and to the ASPRS LAS format (version 1.0-1.3). It comes with various other utilities, like LASGround, LASHeight, LASBoundary and many more, which can be run from LAStools, which is a simple GUI for running the utilities, or from command prompt, whichever convenient to the user. This software was used because of two major reasons:

i) It is a freeware, thus reducing the total costs of the research.
ii) It can deal with a very large number of points, at high densities, quiet efficiently.

1. **Bare Earth Extraction:**

Thus, for the ground-non ground classification of LiDAR points, the LASGround utility of LAStools was used. This tool is purely for bare-earth extraction, and classifies the LiDAR points into ground (class 2) and non-ground (class 1). By default, the tool has a setting called "forests and hills" , which uses a step size of 5 metres. The step size determines the search radius in which low points are selected. The tool considers only the last returns of LiDAR, which can be also changed to all returns. However, since the study area is hilly, forested and not extremely steep slope, the default settings were used to run LASGround on the LiDAR data of the whole study area (which was merged 17 files into one for convenience). The result was a LAS file (default format of LiDAR data) with ground-non ground classification of points. (for command line, see Appendix 2)

2. **Normalisation of point cloud:**

Next, it was important to find the precise height of each non-ground point above the ground, as the height was to be used in several operations (like classification, tree height calculation, canopy gap delineation) later on in the study. For this purpose, LASHeight is the most elligible utility in LAStools. This tool computes the height of each LAS point above the ground. This assumes that grounds points have already been classified (classification == 2) so they can be identified and used to construct a ground TIN. The tool "reads LIDAR in LAS/LAZ/ASCII format, triangulates the ground points into a TIN (or whatever other point class was selected with '-class 4' or '-classification 3'), and then calculates the elevation of each point with respect to this TIN" (http://www.cs.unc.edu/~isenburg/lastools/download/lasheight_README.txt). By default the resulting heights are quantized, scaled with a factor of 10, clamped into an unsigned character between 0 and 255, and stored in the "user data" field of each point.

This tool also has an option to 'drop below', which excludes all points below a specified height. This option ideally, would have facilitated the identification of canopy gaps. However, this option led to a data loss when gridded into a raster, and was not used for the final analysis.

Also, to avoid quantizing and clamping,the 'replace z' option in LASHeight was used. This option 'replaces z', or the elevation value of each point with the computed height. That means that afterwards all ground points will have an elevation of zero and all other points will have an elevation that equals their height above (or below) the ground TIN at their x and y location. In a sense this would "normalize" the elevations of all points in respect to their surrounding ground truth. This step was important because it was the height aspects of trees that were the prime focus, thus the height of each point above the ground with the height of ground points being zero were essentially required.

The output of this procedure was a LAS file with points classified as ground and non-ground, with each non-ground point having a height value above zero (ground level).(For command line, see appendix 2 ).

**3. Gridding:**

A Canopy Height Model is the representation of the difference between the top of the canopy surface and the underlying ground topography. It is derived by filtering LiDAR point clouds to separate ground and canopy hits.

The most established method to derive a canopy height model is to generate a Digital Terrain Model (DTM) by considering only last returns of LiDAR, and a Digital Surface Model (DSM) considering only first returns (Corresponding & Blackburn, 2004; St-Onge & Vepakomma, 2004). However, an accuracy assessment of the DTM is required, using ground points collected in field. Since no such ground points were used in field, an alternative approach towards building a canopy height model was used.

Since there already existed a LAS file which was filtered, classified into ground and non-ground, and the heights of each point calculated, that was in itself a 3D version of a canopy height model. However, to analyse tree heights and gaps, and also to visualise the whole study area without confusion, a 2D representation was required. The LAS file resulting from the above methods (LASGround and LASHeight) was then subjected to the LASGrid operation.

The option chosen to grid in this study was 'elevation' and the 'highest' of elevation was selected for gridding.

The other things needed to specify is the grid size. The grid size depends on what purpose will the Canopy Height Model be used for. In this study, grid sizes of 1m, 70 cm, 60cm, 50cm and 15cm were used, out of which size 50cm and 15cm were used in two different canopy height models.

Furthermore, this tool can efficiently deal with billions of LiDAR points. The default usage of memory (500 MB) was changed to 1500 mb while gridding because of the high density point cloud, which occupies a lot of memory.

The output was specified as .tif format and the LAS file of the whole study area was gridded to generate a Canopy Height Model. (For command line, see appendix 2 )

### 2.3.3. Gap Detection and Delineation:

One important aspect of the study was to delineate gaps from the LiDAR data. This attempt has been done in a few previous studies, and different methods were used to delineate gaps in 2D and 3D(Vehmas, Packalén, Maltamo, & Eerikäinen, 2011). However, to begin with, it was necessary to visualize the Canopy Height Model for detecting gaps sampled in field. Firstly, raster calculator of the Map Algebra toolset of ArcGis was used to assign a 'nodata value' to everything below 5 metres of height, as the threshold specified, so that made the view more clear about which pixels should be considered a gap pixels. Another approach was to produce a binary raster, with a '0'value to all pixels with a highest pixel value below 5 metres, and a '1'value to all those above metres. Both the methods proved equally effective, however the binary raster method was more useful when it came to make a separate layer involving only gaps. The projection of the Canopy Height Model was set to Lambert Conformal Conic, in the system of NTF Lambert Zone III, and the field observation points were added to the Canopy Height Model. Visually it was assessed whether or not the gap coordinates were falling within the 'nodata'or the '0'zone, and that particular gap was marked as 'correctly detected'. In some cases, the field observation point fell at the edge of a gap and canopy. For such cases, a threshold of one metre was set. Any point which was within one metre of a 'non-vegetation' zone was to be considered as a correctly detected gap. The detected gap was then added as a feature class, and delineated by editing with a polygon tool with the maximum accuracy possible. It was also taken into consideration that the gap should be delineated *as it was measured in the field.* This factor had to be considered because at some places the gap was not closed, it followed a narrow passage of pixels of 'non-vegetation' into another gap, or a zone of earthflow where the vegetation had completely disappeared. There were two major motives behind this manual delineation:

i) To assess the accuracy with which gaps can be detected and their area calculated as compared to the values sampled in field. This was, by no way to be taken into consideration in the final gap analysis. However, several different methods have been tried to delineate gaps perfectly, but since no universally accepted definition of canopy gaps exists, manual delineation was done as an alternative analysis.

ii) Although studying gap dynamics was not an objective in this study, it was an interesting issue that *why* a gap was not detected, or in delineation, the area turned out to be more/less than the value sampled in the field.

### 2.3.4. Automatic Method for delineating gaps:

Digitizing gaps was one way to delineate them and assess their area and distribution. However, the very basic flaw in this method was the bias due to manual work. An automatic method of delineating gaps had to be undertaken so as to get a zero bias set of values for gaps around the study area. Moreover, it was not possible to digitize each and every single gap in the study area.

Researchers have tried many different methods to delineate gaps automatically from LiDAR or a LiDAR-induced CHM. A successful attempt was made in 2010 to extract canopy gaps directly from LiDAR, avoiding any interpolation method(Gaulton & Malthus, 2010). However, this study was done over a continuous cover forest, assuming that the tree heights and ages were fairly uniform. However, the study area in this research was a landslide affected forest, so by no means could it be considered a continuous cover forest. Moreover, software which could efficiently handle such a large point cloud and locate local maximas was not available. Thus, this method was not tried out in this study.

Another approach was carried out by Zhang (Zhang, 2008) to delineate gaps using a black top hat mathematical morphological function. However, this research was in a mangrove forest to detect large gaps created by lightning, so this method was not tested in this study either.

Blackburn and Koukoulas (Corresponding & Blackburn, 2004) used a fixed height threshold method, followed by shrinking operations in ArcGis, to delineate gaps in a forest. This method was tried out in this study; a brief overview of this method is as follows:

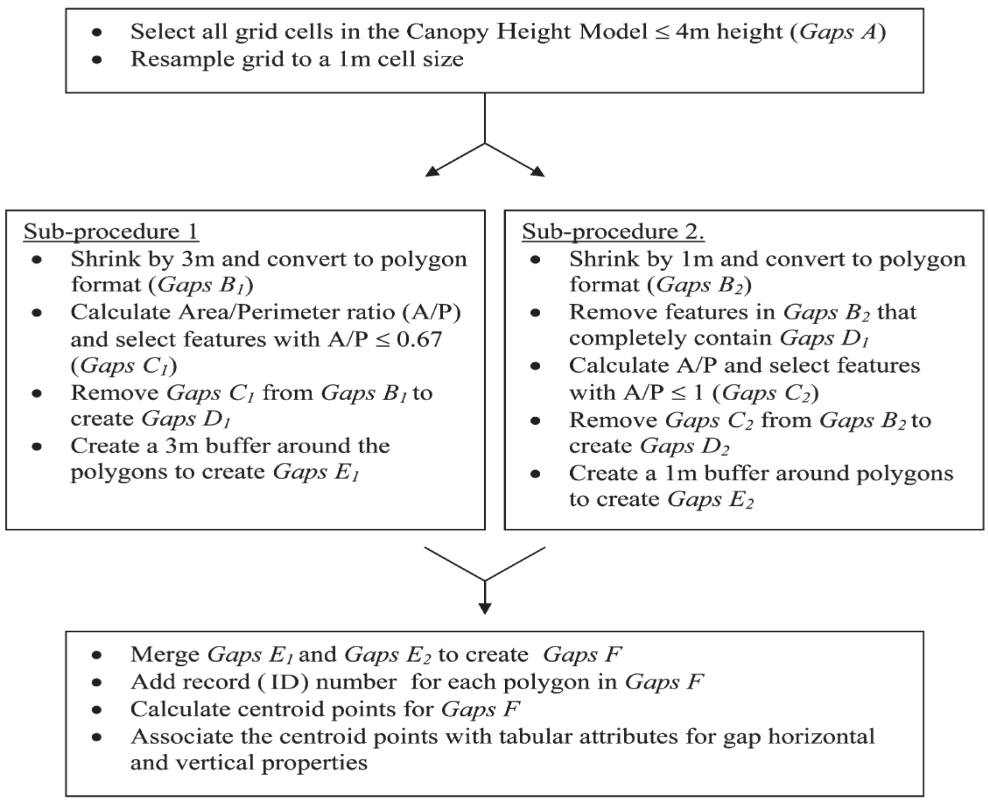



Fig 2.4. *Summary Procedure for gap delineation, (Corresponding & Blackburn, 2004).*

Another approach which seemed most convincing was employed by Vepakomma (2008). This was basically a region growing algorithm, with a predefined threshold of 5 meters. Since this kind of delineation would prove useful for other purposes like tree height extraction, a delineation procedure in ECognition was undertaken.

Usually, a region-growing based approach starts with applying a Gaussian filter of appropriate size. However, in this case, no Gaussian filter was used because a CHM generated from LiDAR data of such a high point density, facilitates individual tree crown visualization quite efficiently. To be on the safe side, a Gaussian low pass filter of 3x3 was applied, but with an insignificant change in the final output.

The delineation process started by applying chessboard segmentation. This is the most basic of segmentation processes, which divides the whole image into equal squares of a specified size. In this case, a size of 2 by 2 pixel-sized objects was selected. This would facilitate two processes later on:

i) Identifying valleys, roads and non-vegetation areas efficiently.
ii) Refining very small objects, which could be spurious gaps or intra canopy gaps (created due to penetration of laser beams through the canopy)

Following the chessboard segmentation, it was necessary to identify gaps, or rather, areas of no or low-lying vegetation as a start. A shadow masking process is considered to be the most ideal for this purpose. Under the shadow masking parent process, an 'assign class' algorithm was applied to the unclassified layer, setting a threshold of 5 meters(Vepakomma, et al., 2008) for the mean pixel value. All the pixels having a value of less than 5 meters were subsequently classified as 'gaps'. All the gaps were merged into one class using the merge region algorithm.

Next process was that of distinguishing between gaps. Since, gaps caused due to natural disturbances were the prime focus of this study, the following type of objects classified as 'gaps' had to be eliminated:

1) The mountainous area in the south-east of the study area, which was completely bare, thus classified as a gap.
2) The open area right in the middle of the study area, which had been created due to farming practices centuries ago. (This was concluded from personal communication with Mr.Jean-phillippe Malet, Ecole et Observatoire des Sciences de la Terre [EOST] and an old house uninhabited for decades in that area told the same story).
3) Trails, roads, and huge patches of non-vegetation created by landslide processes. The huge patches had to be eliminated because firstly, such huge gaps are obviously due to a disturbance, so no detailed study of them is really required.

For this, a 'remove objects' algorithm was applied, giving a threshold of 1000 pixels for area of gap object. Thus, every 'gap' having an area of more than 1000 pixels was removed (the value was determined solely by trial-and-error). This produced a class of gaps which were not to be considered in analysis. Furthermore, another morphological algorithm was applied to perform closing of canopy gaps, with an elliptic fit of greater than 0.2. Elliptic fit is the portion of area of the selected polygon (in this case, the canopy gap), that is covered by the largest eclipse that fits inside the polygon without exceeding its borders and passes through the centroid of the polygon. Such a low value was chosen because not necessarily all gaps can be accounted as an ellipse even roughly.

It was also necessary to identify trees, especially to extract tree heights. Thus, another assign class algorithm was applied to assign all pixels (unclassified) with a mean pixel value of more than 5 as trees. This was followed by the algorithm of find local extrema and find enclosed by image algorithms to detect tree tops. To remove false local maxima (tree top) and local minima (seeds or tree crown edge) grown tree tops and seeds which were in the proximity of one another were merged before starting the region growing process. Then region growing from tree tops was started until significant boundaries of tree crowns found. Grown region was stopped based on visual examination.

A summary of the approach is presented in appendix (3).

An example of the delineated polygons, with their maximum height points exported as a shapefile, is shown in fig 2.5.

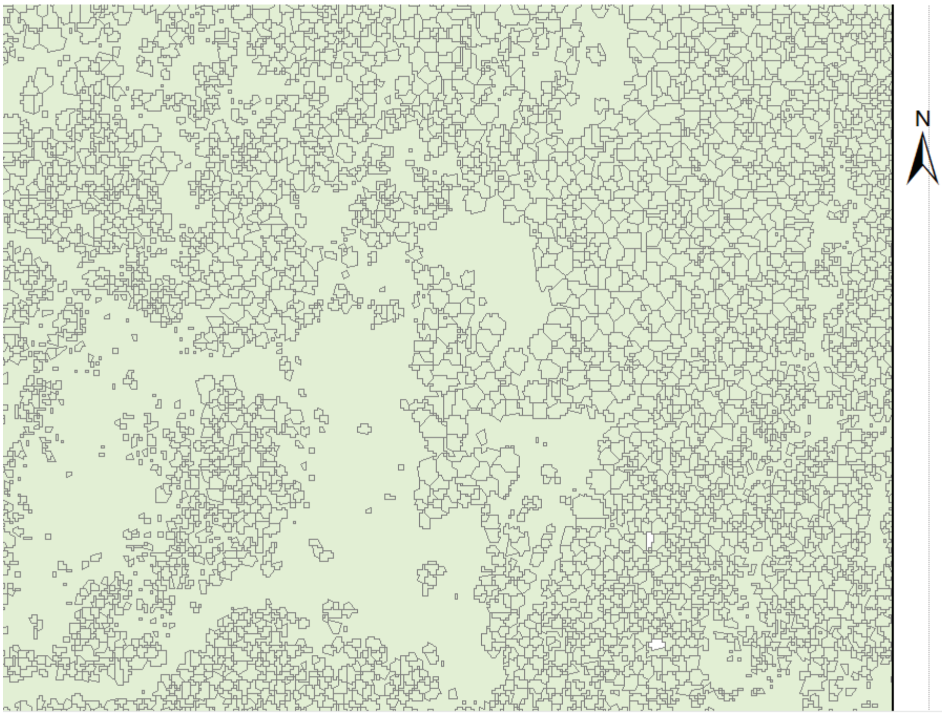



Figure 2.5. Output of ECognition, shapefile showing tree polygons and their maximum height points.

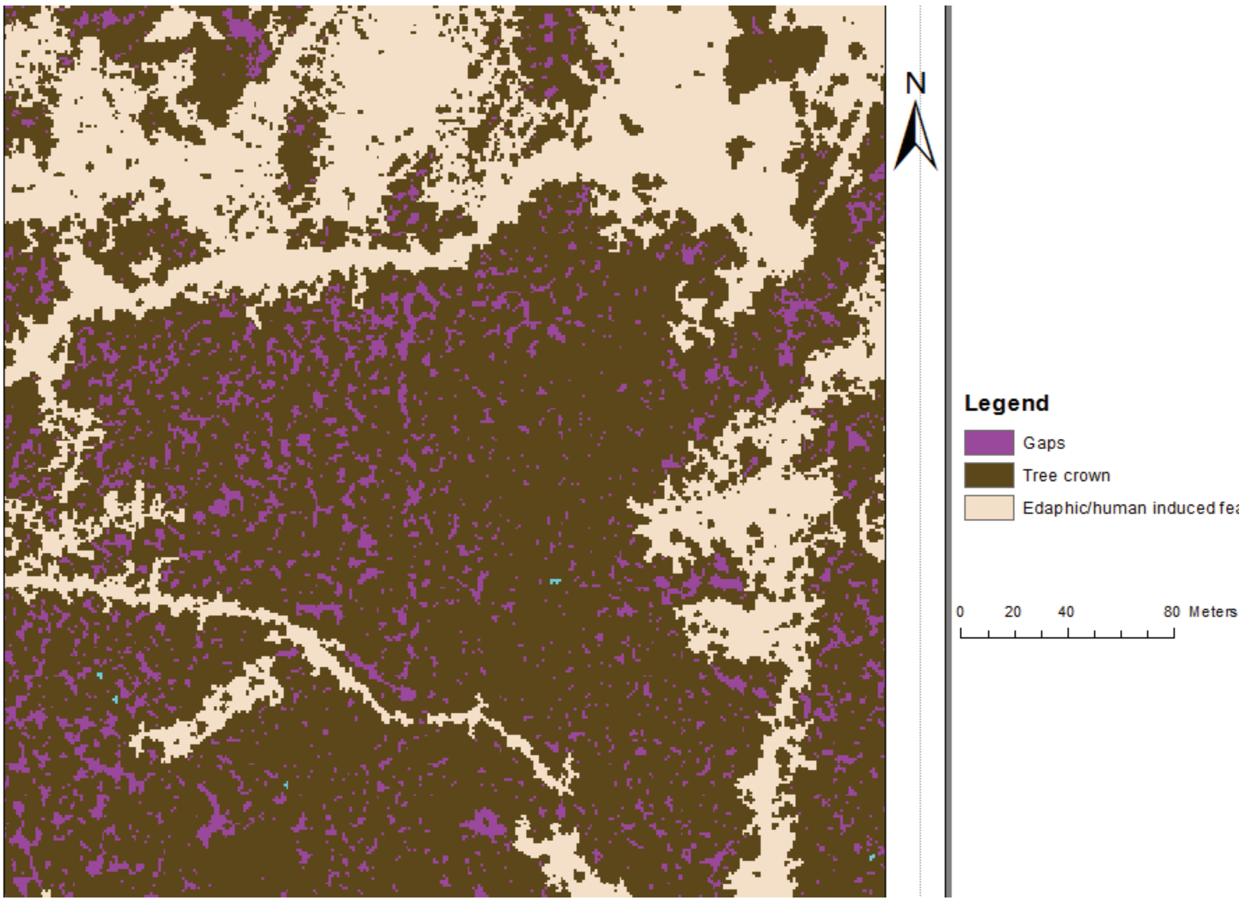



*Fig.2.6. a raster representation of the ECognition output considering gaps only*

### 2.3.5. Mapping Canopy Gaps:

The output from ECognition was:
I) a raster file showing the gaps and tree tops.
II) A shapefile containing polygons of trees and gaps.
III) A point file showing the maximum height value within each polygon.

This shapefile proved to be most useful in the later analysis of gaps. The polygons in that file which were assigned the class 'gaps' were exported as a separate layer to facilitate gap analysis. The human related and edaphic features (such as open areas due to historical farming activities, trails and paths, mountains with no vegetation) had already been classified. Researchers who have previously done gap analysis set various minimum gap area thresholds. Vepakomma (Vepakomma, et al., 2011) had set a $5m^2$ threshold to define a gap. A $5m^2$ minimum size is ample when it comes to analyzing gaps for solely ecological purposes. However, since this research was also to study how gaps are created due to landslide process, a much lower minimum gap size threshold had to be set. One more issue that had to be dealt with was that of *intra-canopy gaps*. These are small areas where the laser scanner penetrated through the canopy, thus showing a very narrow line of no vegetation. Thus, a minimum gap size of $1m^2$ was set for the analysis. Following this, all polygons with an area lesser than $1m^2$ and an area to perimeter ratio of lesser than 0.3 (Since these 'gaps' with a low area-to-perimeter ratio are merely the natural distance between trees) were selected using the select by attributes query of ArcGis. Also, gaps with area more than a $1000m^2$ and area to perimeter ratio more than 4 were selected using the same query. All of the selected polygons were eliminated from the shapefile. The resultant file was a 'gap map' of the study area.

For the beginning of analysis, the shapefile with gaps was clipped with the shapefiles of the landslide and non-landslide areas prepared by Khamarrul.A.Razak by visual interpretation, (2011) to separate the gaps. The gap areas, perimeters, and area-to-perimeter ratios were laid in separate columns in an excel sheet.

### 2.3.6. Extraction of plots and transects:

For various purposes like gap idenfication and delineation, and extracting values of inclination, orientation and Diameter at Breast Height (DBH), it was necessary to extract plots and transects sampled in field. There were two ways of extraction possible:

1) *Clipping from a predefined shape file*:
   This method requires a shapefile(.shp) of the area to be extracted, which is in turn used as a command line in lasclip utility to extract the same very area from the LiDAR point cloud of the whole area. The coordinates of trees and gaps sampled in field were overlayed upon the Canopy Height Model, and the 'create feature class' of the Data Management tools in ArcGis 10 was used to create a shapefile in circular form, encompassing all the trees in a plot. In case of transects, the shapefile was created rectangular, with a length of 25 metres (as the length of a transect sampled in field). Further, in the LASClip utility, this shapefile was used as an input, and the output was a LAS file representing the 3D point cloud of the same area. (For command line, see appendix )

2) *Extracting by coordinates* :
   In this method of extracting, there is no need for defined boundaries or a shapefile. The desired area can be extracted by defining the maximum and minimum X and Y coordinates in the Las2Las utility. Thus, the extent of the desired area (from the coordinates sampled in field) was entered as an input in LAS2LAS and the output was a LAS file of the desired area, plot or transect.

Both methods were utilised in the study for extracting plots and transects, however, both have some flaws:

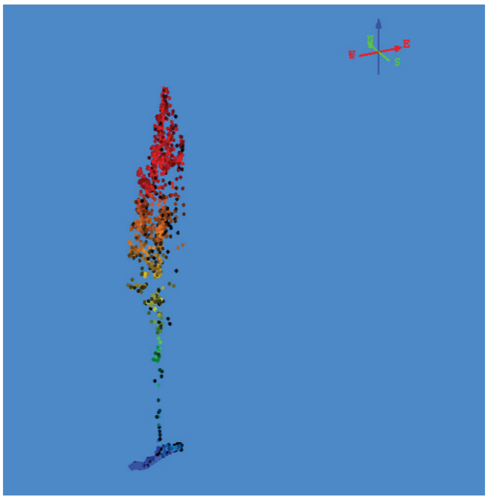
1) Extracting from a shapefile: Since the shapefile is made by the researcher by a manual method, there is a risk of exclusion of a few pixels. Even though a few pixels in a shapefile are not very significant, the same very area might be representing hundreds of points in the LAS file (as the point density in the data is as high as 180 points per m2).
2) Extracting by coordinates: In this method, the maximum and minimum coordinates are, in the end, the coordinates taken at the base of the tree. So, when the plot/transect is extracted and viewed in an advance viewer like Quick Terrain Modeller, it was realised that a large number of points of the canopies of some trees (which were at the boundaries of the plot) were excluded, as the coordinates of the base of the tree fall well inside the canopy.

   Both the methods were employed to extract plots and transects and they were viewed in the LASViewer utility. LASViewer is a simple OpenGL-based viewer for LIDAR in LAS/LAZ/ASCII format that can compute and display a TIN. It also has other options in viewing, such as one to specify the number of points to display and the number of steps to display in. Further, they were also viewed in Quick Terrain Modeller, and it was ensured that no points are excluded.

### 2.3.7. Extraction of single trees:

As explained previously, plots and transects sampled in field were extracted from the LiDAR point cloud. To validate the inclination and orientation of trees sampled in field, it was necessary to extract individual trees sampled in field. An automatic method (like clipping the tree using coordinates from a 2D representation like a CHM) was not advisable as there were tilted trees in the area. Thus, if the canopy polygon was clipped from the point cloud, it could have led to exclusion of points, thus altering the final output. Thus, it was to be done manually. For this job, the Quick Terrain Modeler's trial version was used. This software can open and display LiDAR point clouds with high efficiency. The main advantage of this software is that manual editing and analysis can be done in 3D, unlike LAStools where only viewing in 3D is available. A tool in QT Modeler, called 'Z' or polygon tool, can select a 3D point cloud. The selected plot was visualized in 3D, and the shapefile of the tree coordinates were imported into QTModeler. Each tree was then visualized from every possible angle, it was ensured that all the points belonging to a tree were being extracted, and the select 3D polygon tool was used to select the tree. The crop tool was then used to crop the individual tree and export it as a LAS file. It was also ensured that a considerable amount of ground points were also included in the tree, as they were needed for further processing. However, problems were faced while extracting tilted trees, especially because it was essential to get points on the stem too. Intermingled trees were separated using field and ecological knowledge and referring to field photos.

The extracted trees were to be used as input for SkelTree software. SkelTree skeletonizes the point cloud, and provides information about inclination at 3 different heights (0m, 0.5m, and 1.5m). A skeleton is a line describing the tree shape. Ideally, it is centred within the object and is connected whenever the object is connected. In this study, the 'SkelTree skeleton' was used to extract inclination and orientation. This particular skeleton was previously used to extract the diameter at tree breast height from high density airborne LiDAR data (A. Bucksch, 2011; 2010). So, a total of 211 trees were extracted as point clouds and used as input for SkelTree to extract inclination and orientation values, which would in turn be used for validation of the same values measured in field. An example of an extracted tree is shown in Fig (2.7). However, due to time constraints, this operation could not be completed. These extracted trees though would be the 'raw material' for further research.



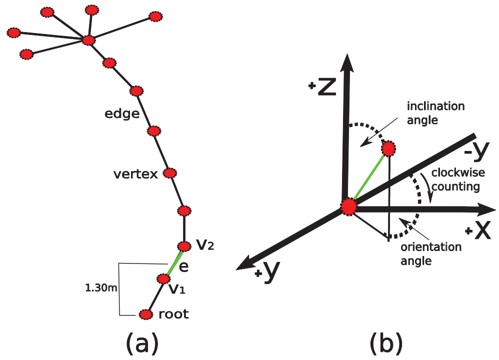



*Fig.2.7. A Tree extracted using QTModeler*
A total of 1458 points in this tree.

*Fig.2.8. a schematic skeleton graph of a delineated tree.(K.A. Razak, et al., 2011)*

### 2.3.8. Extraction of tree heights:

The shapefile of trees delineated in ECognition was opened in ArcGis. Since the 'detect local maxima' operation was already completed in ECognition, each polygon had a point showing the maximum height location. This was to be considered the final tree height extracted from the CHM. However, two facts were taken into consideration before establishing them as tree heights:

1) The delineation done in ECognition was not 100% accurate. Thus, for some polygons as it was observed, more than one tree was included. Thus, while noting down the height, the underlying Canopy Height Model was also referred to, as it was the best representation of the tree boundary. This was done especially for trees sampled in field, because it was necessary to get the height of sampled trees specifically.
2) In landslide affected areas, the tilt of the trees was taken into consideration. For example, in case of a rotational landslide affected area, the tilt of the trees was towards the south.

The heights of trees sampled in field were individually noted down and added to the attribute table of the plot shapefile, which were to be used for tree height validation.
The rest of the tree heights were clipped with the landslide and non-landslide shapefiles to separate them, and were put down in an excel sheet.

### 2.3.9. Validation of extracted forest attributes:

To fulfill one of the research specific objectives, it was essential to assess how accurately LiDAR gives us the information about forest structural attributes. The following attributes had to be and were validated using field data and LiDAR-extracted values:

1) Tree heights: As mentioned before, the tree height values were laid in an excel sheet with the height values of corresponding trees collected in field. A regression analysis was done to get the $R^2$ value of validation.
2) Tree inclination and orientation: These values were collected in field at 4 different heights of the same tree, namely 0, 0.5, 1.5, 2 meters. The trees extracted from LiDAR data were introduced as input in SkelTree, which gave the inclination and orientation values.

3) Canopy Gaps: There were two major factors to be 'validated' in this section, namely gap detection and gap area validation.

i) Gap detection: This section was one of the most important, because detecting gaps is naturally the basic expectation from high point density LiDAR data. The coordinates of gaps sampled in field were added as a shapefile above the CHM in ArcGis, and some specific conditions were set to classify the gap as correctly detected or incorrectly detected, are already explained in section (*gap detection and delineation*). Gap detection accuracy was tested for 3 different files: CHM with 50cm resolution, CHM with 15cm resolution and the 'gap map' produced as output of ECognition.

ii) Gap area validation: Most researchers validate gap area using the area derived from an automatic delineation process. However, there was one little flaw in the automatic delineation that was a major hindrance in validating gap area: interconnected gaps. This problem is illustrated in figure 2.9 .

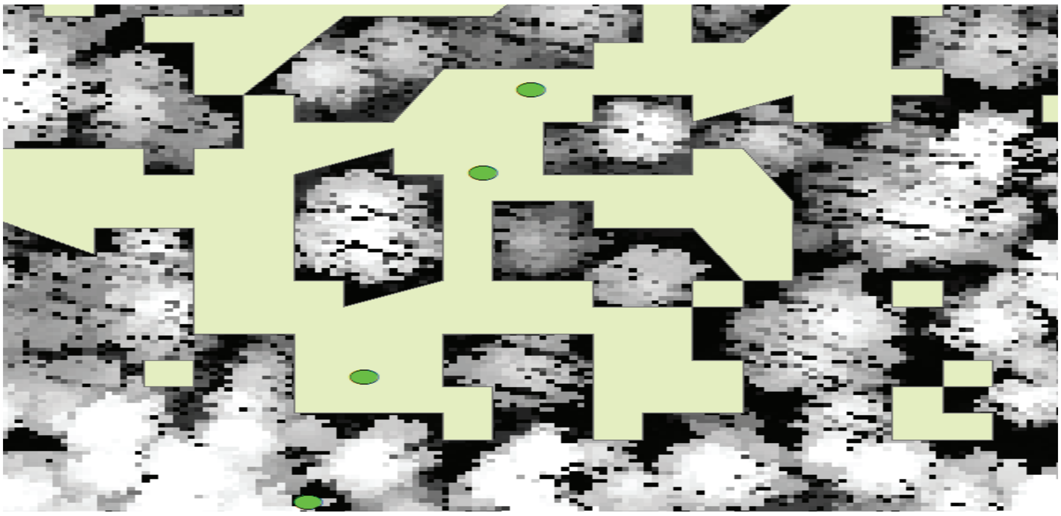


Fig. 2.9 . *Problem of interconnected gaps with automatic delineation. The green points represent coordinates of gaps sampled in field. They have been detected correctly, however, due to a narrow passage of pixels with low vegetation; they are connected, which makes it almost impossible to calculate the area of one single gap.*

Neither the ECognition closing operation nor the shrinking operation in ArcGis could solve this problem, so the manually delineated gaps were used for validation of area in this study.

### 2.3.10. Other parameters for analysis:

During the course of the research, it was realized that some new parameters could also be tested to see whether they can be established as indicators of forest disturbance. One parameter was 'tree height dissimilarity'. This parameter was picked with reference to a study by Razak (K. A. Razak, et al., 2011), who analyzed the same parameter in the same study area. Another parameter chosen was tree density, which is basically the number of trees per unit area (or a 100 m$^2$ as chosen in this study). The analyzing process for both of these parameters is described below:

1) Tree height dissimilarity: The motive behind analyzing this parameter was to see whether the disturbance processes are having an effect on the vertical heterogeneity of the forest. For this purpose, the tree height tables derived from ECognition were used.
2) Tree Density: To analyze this parameter, it was necessary to take samples from the area, as overall tree density in the whole area would not be the right parameter to compare between landslide and non-landslide. The shapefile containing tree polygons was added in ArcGis, and the centroid of each polygon was extracted using feature to point tool of ArcGis. Furthermore, the area was clipped with the landslide and non-landslide area to separate it for comparison later on. A fishnet

of 20x20 meters was overlaid on these files. To generate, random samples, the "Randbetween" function in Excel was used. This operation gave a list of 20 random numbers within the extent of the number of polygons within the fishnet. This table was joined to the attribute table of the fishnet, using the 'FID' column of the fishnet to base the joint upon. The selected polygons were exported as a separate shapefile, and the 'join' function was used again, however, with 'join data from another layer based on spatial location' option. The selected polygons were joined with the table of tree centroid points, and the number of trees that fell within each polygon was noted down. Thus, the tree density of every polygon would be :

Number of trees/400 *100

This gives the tree density per $100m^2$. This was of course, calculated separately for stable and unstable areas, taking 20 samples each in both areas. These density values would be later compared to find out whether or not there was a significant difference between stable and unstable areas.

### 2.3.11. Statistical analyses:

**1) Statistics for validation:**

To validate the values of attributes extracted from LiDAR data, field data was used. Primarily, regression analysis was used to find out how accurately the values have been extracted.

The main purpose was to find out the $R^2$ value, and the standard error in extraction.

This was done for tree heights and gap areas.

2) **Statistics for comparison**:

This study was basically centered on comparing the forest structural attributes in landslide and non-landslide zones. A set of tables with various values like tree heights, canopy gap areas, gap area to perimeter ratios, tree densities were at hand. For comparison, it was necessary to choose the most appropriate statistical test. Some of the properties of these sets of data were:

i) The two 'populations' did not follow any specific parameterized distributions.

ii) The two samples were not equal in size (for example, the number of tree height values in the unstable area was 37414 and that for the stable area was 26690).

iii) The two samples were independent of each other.

iv)

For these very reasons, the Students' T-test was avoided, because it runs on the basic assumption that the samples are normally distributed. A Mann-Whitney U-test was used to compare different structural attributes in the two zones. The Mann-Whitney U test is a non-parametric statistical hypothesis test to assess whether one of the two samples in a set of independent observations tends to have a larger value than the other. One of the biggest advantages in this running this test was that it assumes that under the null hypothesis the distributions of both groups are equal, so that the probability of an observation from one population (*X*) exceeding an observation from the second population (*Y*) equals the probability of an observation from *Y* exceeding an observation from *X*, that is, there is a symmetry between populations with respect to probability of random drawing of a larger observation. Moreover, this test was also used for a similar kind of comparison (tree structural attributes in stable and unstable areas) was used in a previous study, using dendrogeomorphology (Van Den Eeckhaut, Muys, Van Loy, Poesen, & Beeckman, 2009).

For comparing tree density , t-test was run on the 40 samples, 20 in each zone.

**3) Statistics to test forest heterogeneity:**

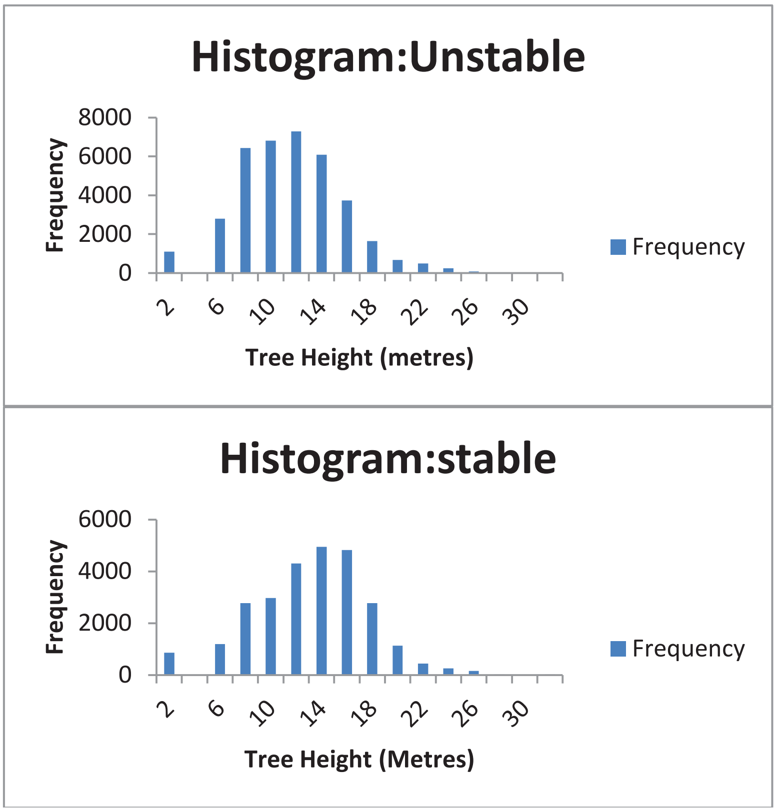
There were in all 37413 tree height values in the unstable area, and 26690 tree height values in the stable area. Forest heterogeneity induced by changes in tree height and density in a forest has a direct effect on gap detection(Zhang, 2008). A Shannon's index was generated for both areas using PAST statistics. Furthermore, a Mann-Whiteney U test was also performed to compare the height distributions in stable and unstable areas.

# 3. RESULTS:

## 3.1. From field data:

A total of 13 plots, containing 277 trees and 10 transects with 43 gaps in all were collected from field. Some primary descriptive statistical analysis was performed on the field to get an idea about the distribution of the structural attributes. The distributions of values were studied and Mann Whiteney U tests were performed to compare heights, DBH and gap areas in stable and unstable zones. Following are the results of the descriptive and comparitive statistics:

### 3.1.1. Height (meters):




**Hypothesis Test Summary**

| | Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of Height is the same across categories of Type. | Independent-Samples Mann-Whitney U Test | .002 | Reject the null hypothesis. |

Asymptotic significances are displayed. The significance level is .05.

*Fig 3.1. Results of descriptive and comparative statistics of tree heights in stable and unstable zones, as collected in field.*

### 3.1.2. Diameter at breast height (cm):

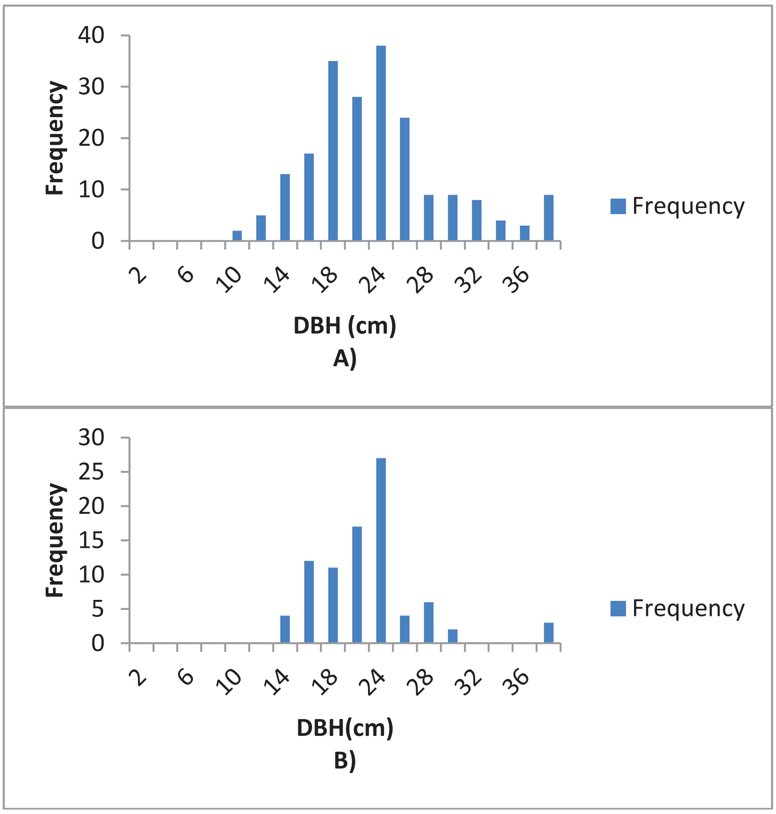



*Fig 3.2. Histograms of DBH of trees in the study area.*

*A) DBH distribution in stable area* *B) DBH distribution in unstable areas.*

Results of the Mann Whitney U test for comparing DBH of trees in stable and unstable zones are as follows:

**Hypothesis Test Summary**

| | Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of DBH is the same across categories of Type. | Independent-Samples Mann-Whitney U Test | .359 | Retain the null hypothesis. |

Asymptotic significances are displayed. The significance level is .05.

### 3.1.3. Inclination (degrees):

| *inclination at 0* | |
|---|---|
| | |
| Mean | 60,81 |
| Standard Error | 2,21 |
| Median | 60 |
| Mode | 60 |
| Standard Deviation | 19,89 |
| Sample Variance | 395,7 |
| Kurtosis | 0,674 |
| Skewness | -0,7 |
| Range | 90 |
| Minimum | 0 |
| Maximum | 90 |
| Sum | 4926 |
| Count | 81 |

| *inclination at 0.5* | |
|---|---|
| | |
| Mean | 68,99 |
| Standard Error | 1,54 |
| Median | 70 |
| Mode | 75 |
| Standard Deviation | 14,12 |
| Sample Variance | 199,3 |
| Kurtosis | -0,22 |
| Skewness | -0,64 |
| Range | 60 |
| Minimum | 30 |
| Maximum | 90 |
| Sum | 5795 |
| Count | 84 |

| *Inclination at 1.3* | |
|---|---|
| | |
| Mean | 72,5 |
| Standard Error | 1,305 |
| Median | 75 |
| Mode | 75 |
| Standard Deviation | 11,96 |
| Sample Variance | 143 |
| Kurtosis | -0 |
| Skewness | -0,67 |
| Range | 50 |
| Minimum | 40 |
| Maximum | 90 |
| Sum | 6090 |
| Count | 84 |

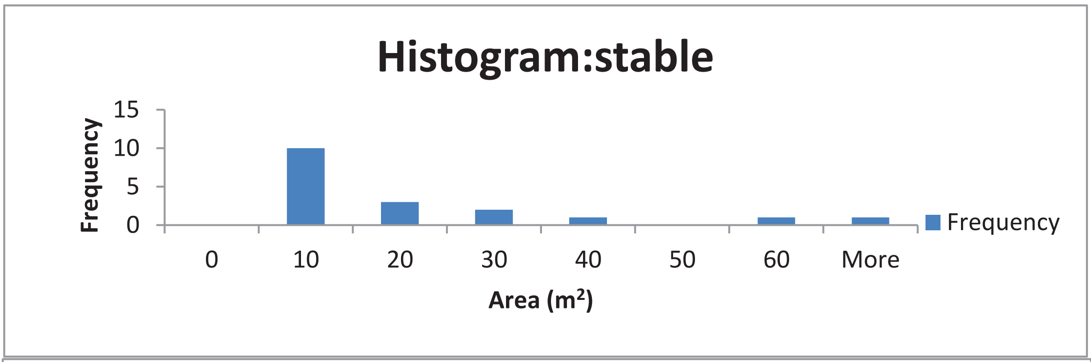
| Inclination at 2 | |
|---|---|
| | |
| Mean | 76,2 |
| Standard Error | 1,22 |
| Median | 80 |
| Mode | 80 |
| Standard Deviation | 11,2 |
| Sample Variance | 125 |
| Kurtosis | 0,24 |
| Skewness | -0,95 |
| Range | 45 |
| Minimum | 45 |
| Maximum | 90 |
| Sum | 6399 |
| Count | 84 |

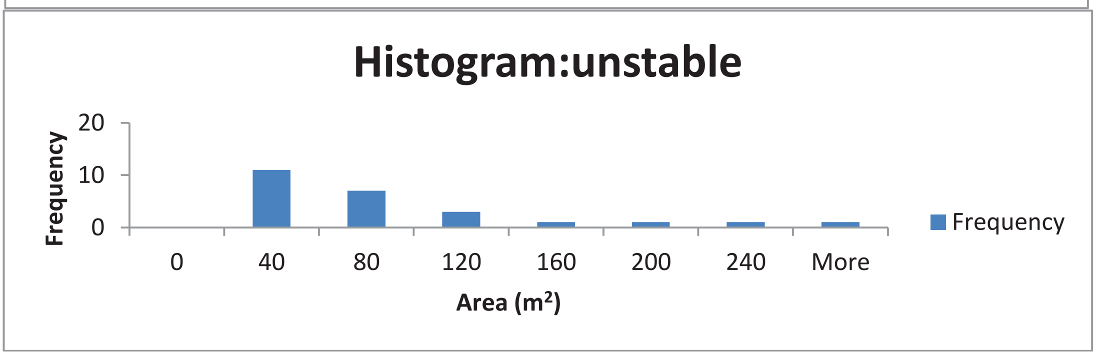
*Fig 3.3. Descriptive statistics of tree inclination values at 4 different tree heights: o metres, 0.5 metres, 1.3 metres and 2 metres*

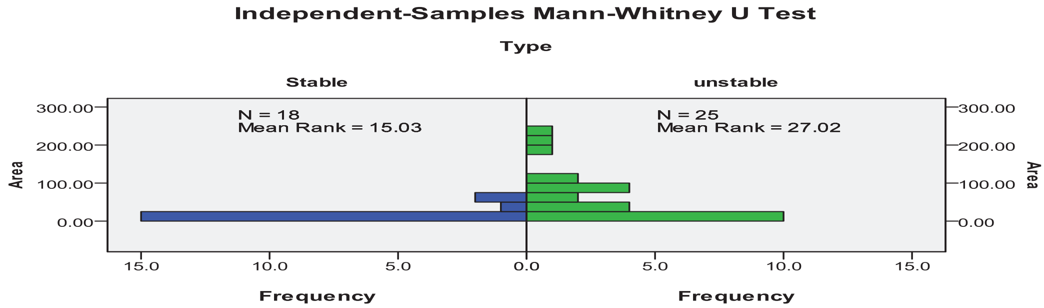
*the count is different at 0 meter height because in field, it was sometimes impossible to reach the base of the tree because of being surrounded by thorny bushes.

### 3.1.4. Canopy gaps:

Descriptive statistics were performed on the gap area, separately for transects in stable and unstable areas. The gap area was calculated using the formula
Area = π (L*W/4), where L is the longer axis and W is the shorter axis.(J.R Runkle, 1992)
The results of the Mann Whitney test are as follows:






| | |
|---|---|
| Total N | 43 |
| Mann-Whitney U | 350.500 |
| Wilcoxon W | 675.500 |
| Test Statistic | 350.500 |
| Standard Error | 40.616 |
| Standardized Test Statistic | 3.090 |
| Asymptotic Sig. (2-sided test) | .002 |

**Hypothesis Test Summary**

| | Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of Area is the same across categories of Type. | Independent-Samples Mann-Whitney U Test | .002 | Reject the null hypothesis. |

Asymptotic significances are displayed. The significance level is .05.

*Fig 3.4 Histograms of gap areas in stable and unstable and results of the Mann Whitney U test.*

As an additional analysis, the percentage of transect length in gaps was also calculated, which goes by the formula:
Percentage of total land area in gaps =(transect distance in gaps /total transect distance)*100 (J.R Runkle, 1992)
The results are as follows:

| S No | Type of landslide | Stability type | Gap percentage (%) |
|---|---|---|---|
| 1 | None | Stable | 52 |
| 2 | Rotational | Unstable | 90 |
| 3 | None | Stable | 29 |
| 4 | Translational | Unstable | 64 |
| 5 | None | Stable | 30 |
| 6 | Rotational | Unstable | 58 |
| 7 | Translational | Unstable | 55 |
| 8 | None | Stable | 34 |
| 9 | Earth flow | Unstable | 92 |
| 10 | Rockfall | Unstable | 60 |

*Table 3.1. Gap percentage of transects sampled in field*

### 3.2. Canopy height model

The first result coming from gridding in LAStools was a canopy height model. The grid size was tested at different values (0.1, 0.15, 0.5, 0.6 and 1 meters) and two sizes were selected as most appropriate for different purposes. CHM with a grid size of 50cm was the most appropriate for delineation in ECognition (since the software was incapable of handling a lower grid size) and one with a grid size of 15cm was most appropriate for canopy gap detection. Fig  shows a zoomed in version of both the canopy height models:

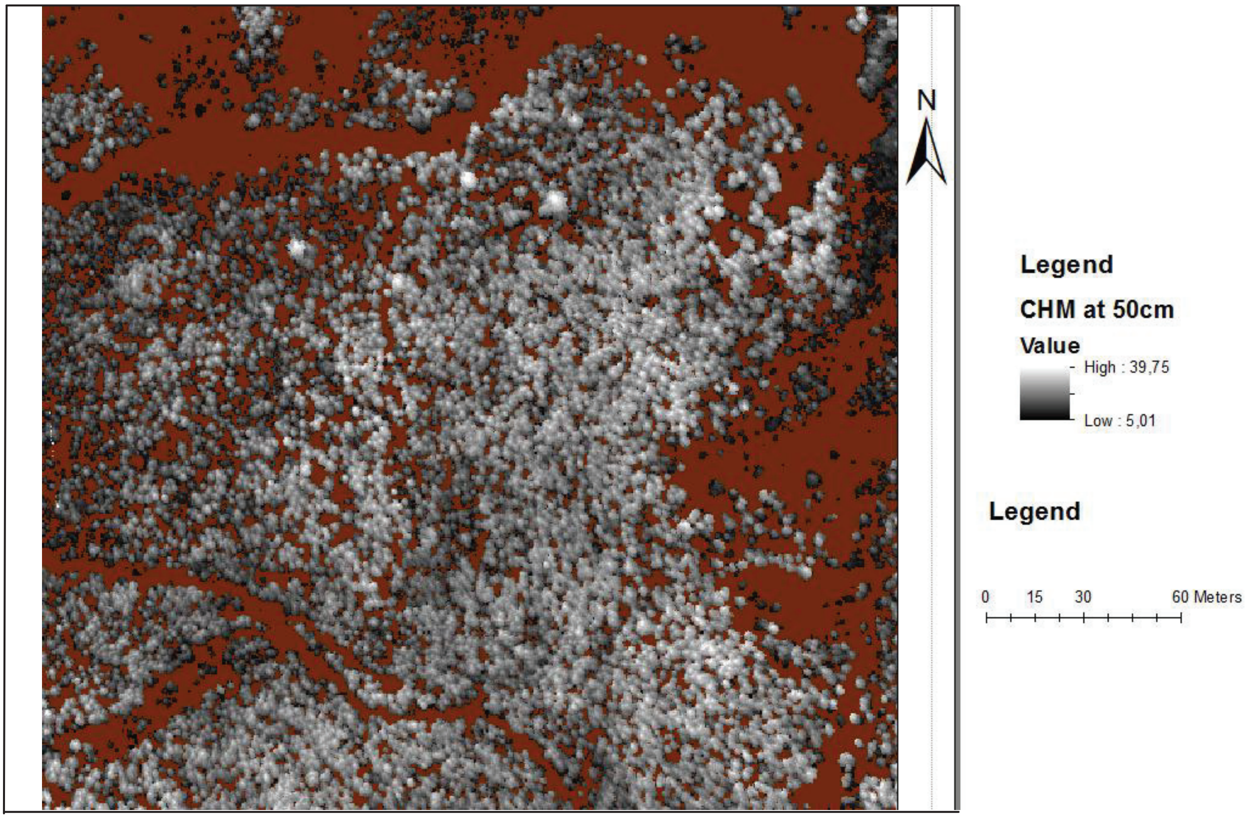



*Fig.3.5. Canopy Height Model at 50cm. All the pixels with a value below 5 have been assigned as 'nodata'.*

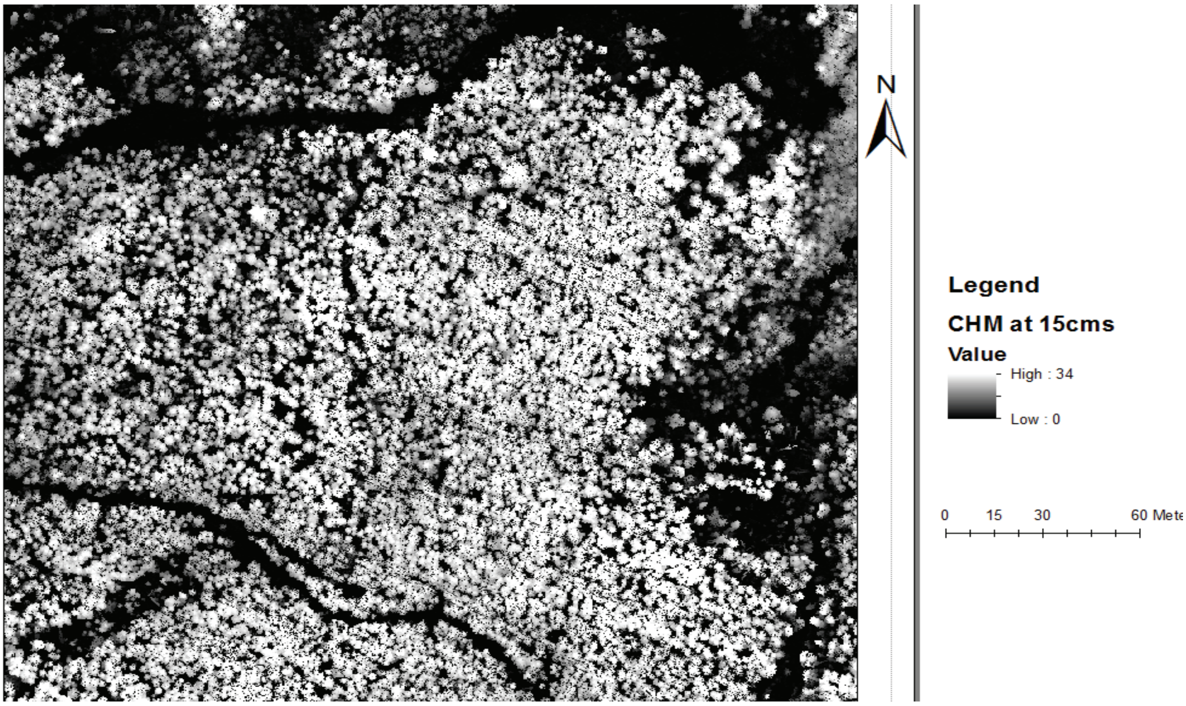



*Fig.3.6. Canopy height model at 15 cm grid size.*

### 3.3. Plot extraction:

All of the 13 plots and 10 transects were extracted as .las files. Fig  shows an example of a plot extracted using LASClip operation:

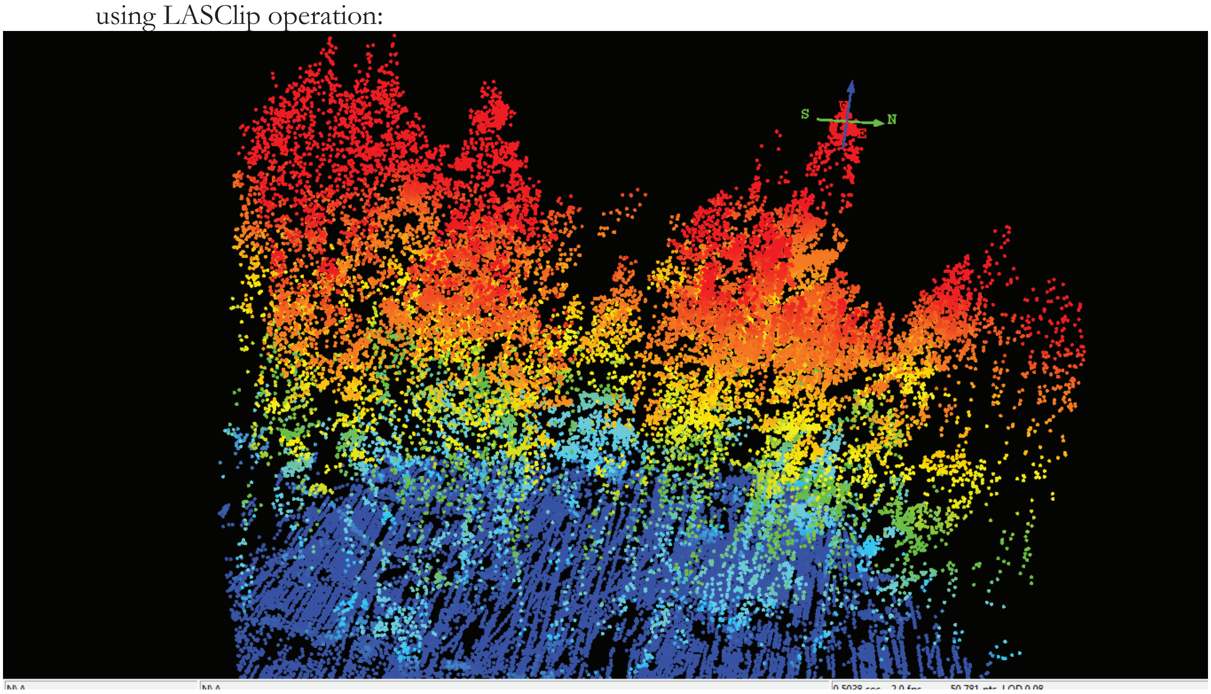


*Figure 3.7: A field plot extracted from the LiDAR point cloud*

## 3.4. Gap delineation:

The output of delineation process was a shapefile containing polygons with their respective classes. Out of these, the polygons belonging to the class 'gap' were selected and exported as a shapefile. After eliminating the polygons with areas falling below and above the lower and upper threshold respectively, a 'gap map' of the whole area was produced, a zoomed-in version of which is shown in fig . Also, the manual delineation of gaps sampled in field in shown in fig

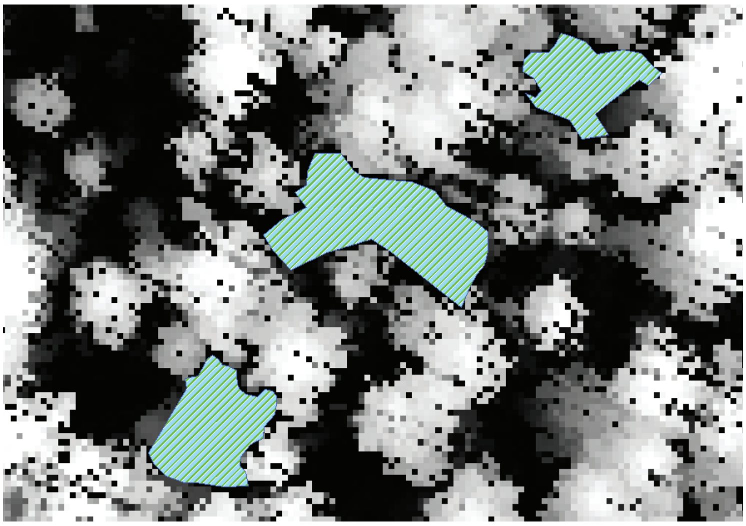


*Fig.3.8. Output of manual gap delineation*

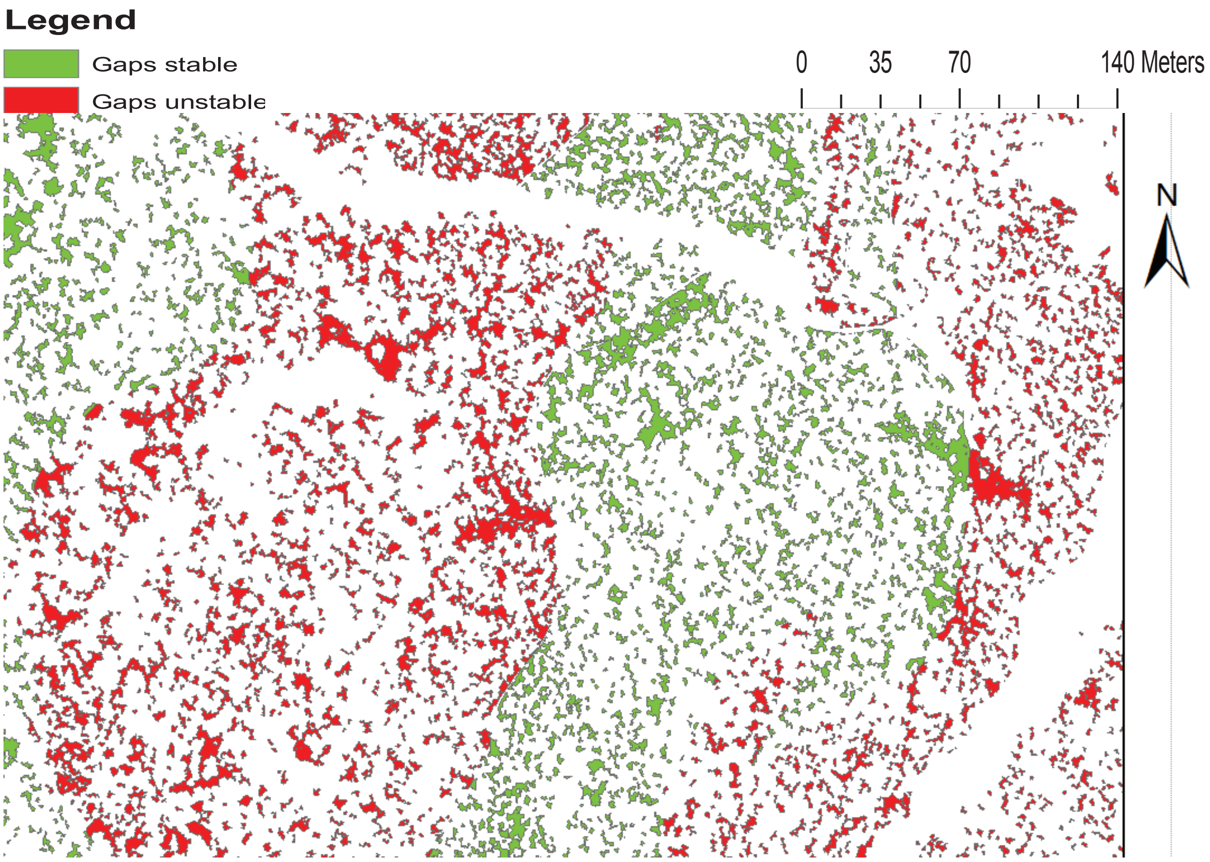



*Fig 3.9. A portion of the 'gap map' of the area showing both landslide and non-landslide zones*

As mentioned in the methodology, the 'shrink' approach for automatic delineation of gaps as used by Blackburn(Corresponding & Blackburn, 2004) was also tried out in this study. This method, did not prove to be successful in this study. The output of the shrink method is shown in a zoomed in version in fig

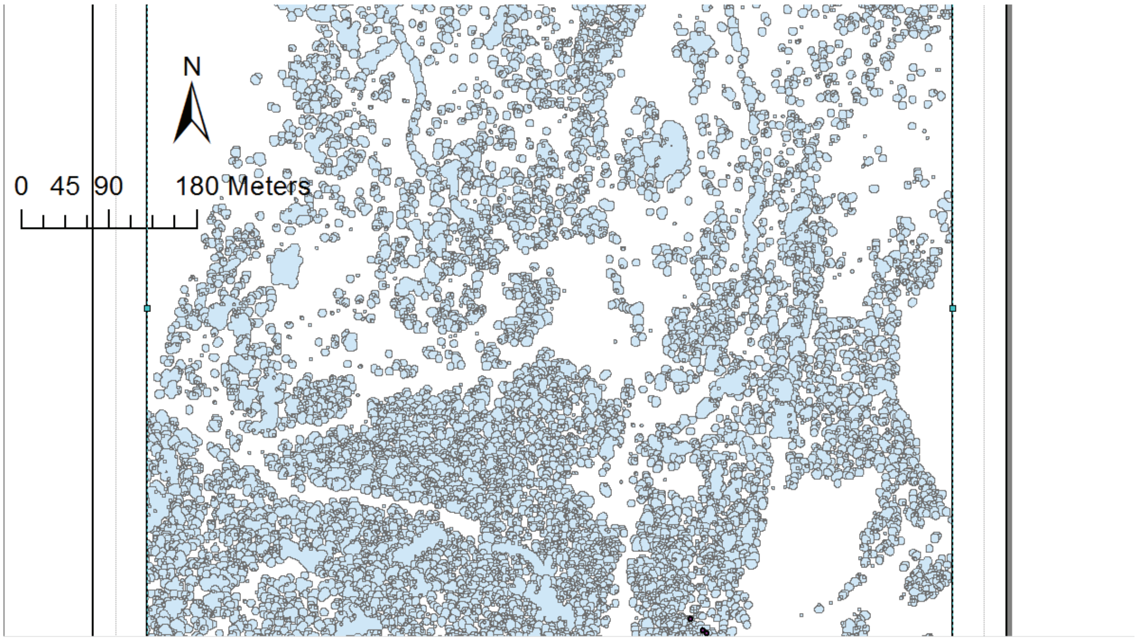



*Fig.3.10. Output of the shrink approach for delineating gaps*

### 3.5. Validation:

i) Gap detection:

A total of 43 gaps were sampled. Following are the number of gaps detected correctly and incorrectly in the two CHMs at two different grid sizes and the gap map:

| File | CHM (50cm) | CHM (15cm) | Gap Map |
|---|---|---|---|
| Correctly detected | 25 | 35 | 31 |
| Incorrectly detected | 18 | 8 | 12 |
| Accuracy % | 58 | 81 | 72 |

*Table 3.2. Gap detection accuracies in 3 different raster files*

ii) Gap area:

Since the automatic delineation method was not valid for estimating, individual gap area, the area of manually delineated gaps was used to run the regression equation. Firstly, a histogram was generated with all of the 43 gaps. After removing the outliers, a regression analysis was done. A summary output of the regression is shown in table :

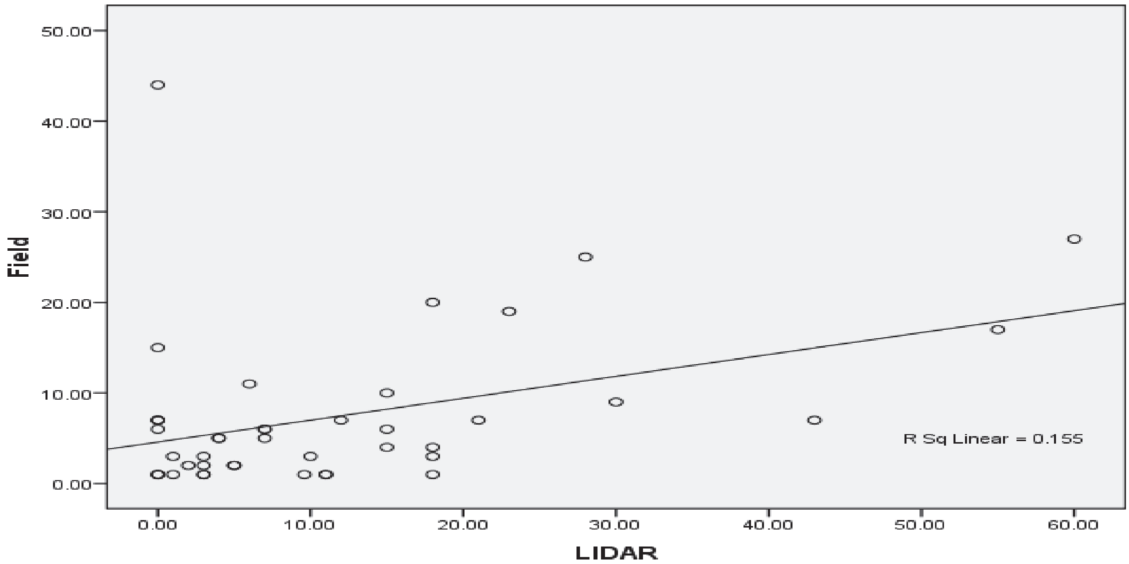



***Fig3.11. Scatter plot of gap areas. On Y axis are the gap areas collected in field and on X axis, gap areas extracted from the CHM. Points lying vertically above the zero value of X axis are the outliers, the gaps not detected correctly, been assigned the value zero.***

*Table3.3. Gap area validation, regression statistics after removing the outliers*

| *Regression Statistics* | |
|---|---|
| Multiple R | 0,70 |
| R Square | 0,50 |
| Adjusted R Square | 0,48 |
| Standard Error | 10,30 |
| Observations | 34 |

iii) Tree heights:

A regression equation was run for height validation using CHM-extracted tree heights as the dependent variable. A summary of the output of the regression is presented in table

| *Regression Statistics* | |
|---|---|
| Multiple R | 0,78 |
| R Square | 0,61 |
| Adjusted R Square | 0,61 |
| Standard Error | 1,69 |
| Observations | 277 |

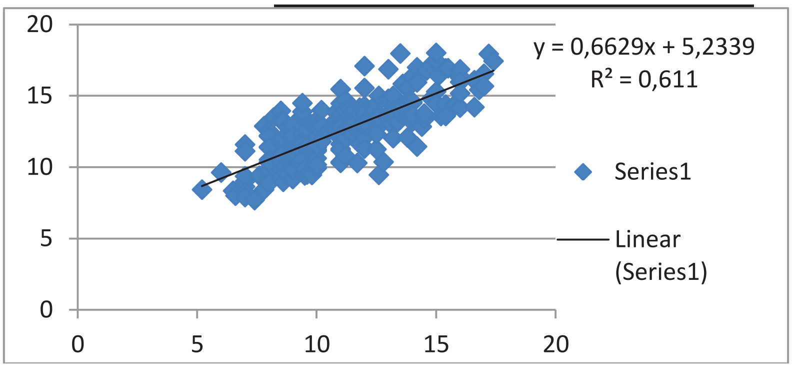



Fig.3.12. *Regression statistics for tree heights, without excluding outliers*

| *Regression Statistics* | |
|---|---|
| Multiple R | 0,851634 |
| R Square | 0,725281 |
| Adjusted R Square | 0,724204 |
| Standard Error | 1,475905 |
| Observations | 257 |

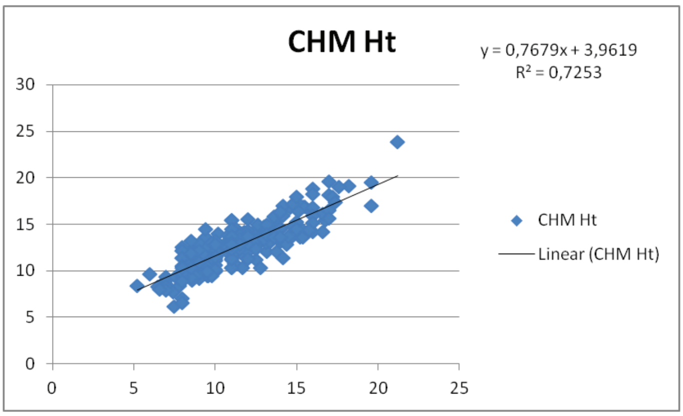



*Fig.3.13. Regression statistics for tree heights, excluding outliers*

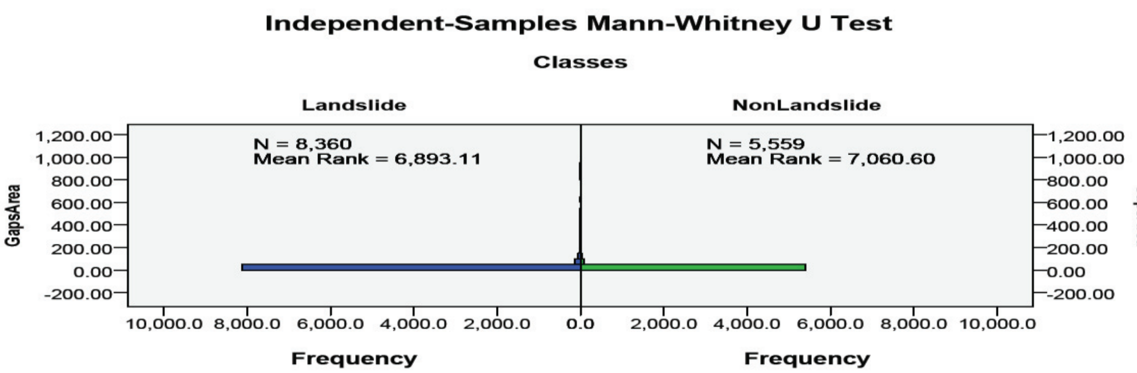
### 3.6. Gap analysis:

The gap analysis was done entirely for the 'gap map' produced, as it covered the whole area and was by far, the most accurate representation of canopy gaps throughout the area. For comparison of gap area in stable and unstable plots, a Mann-Whiteney U test was done, the results of which are as follows:

**Hypothesis Test Summary**

| | Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of GapsArea is the same across categories of Classes. | Independent-Samples Mann-Whitney U Test | .016 | Reject the null hypothesis. |

Asymptotic significances are displayed. The significance level is .05.




| | |
|---|---|
| **Total N** | 13,919 |
| **Mann-Whitney U** | 23,795,835.500 |
| **Wilcoxon W** | 39,249,855.500 |
| **Test Statistic** | 23,795,835.500 |
| **Standard Error** | 232,183.028 |
| **Standardized Test Statistic** | 2.409 |
| **Asymptotic Sig. (2-sided test)** | .016 |

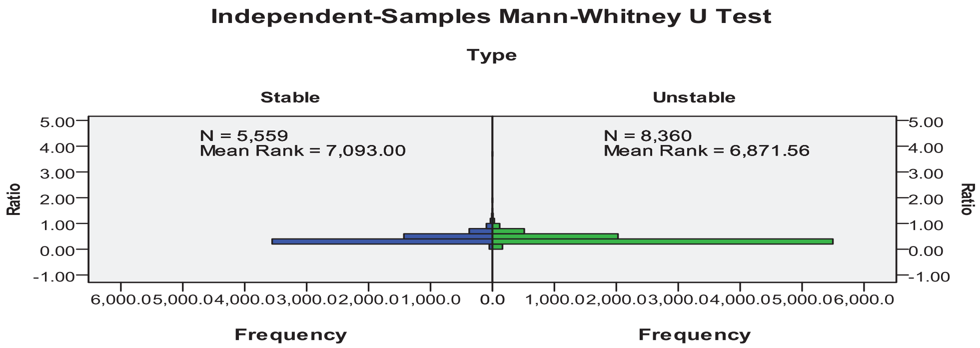
*Fig3.14. Results of comparison between gap area of stable and unstable zones using a Mann Whitney U test*

A similar kind of Mann-Whitney U test was performed for comparing the area-to-perimeter ratio, which is roughly a numeric representation of shape. The results of this test are as follows:

**Hypothesis Test Summary**

| | Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of Ratio is the same across categories of Type. | Independent-Samples Mann-Whitney U Test | .001 | Reject the null hypothesis. |

Asymptotic significances are displayed. The significance level is .05.




| | |
|---|---|
| **Total N** | 13,919 |
| **Mann-Whitney U** | 22,497,271.000 |
| **Wilcoxon W** | 57,446,251.000 |
| **Test Statistic** | 22,497,271.000 |
| **Standard Error** | 232,183.027 |
| **Standardized Test Statistic** | -3.184 |
| **Asymptotic Sig. (2-sided test)** | .001 |

*Fig.3.15. Results of comparison of gap shape (area to perimeter ratio) between stable and unstable zones using a Mann Whitney U test.*

### 3.7. Height analysis:

For analyzing the distribution of heights in stable and unstable areas, primarily a histogram was generated for both zones. The histograms are as follows:

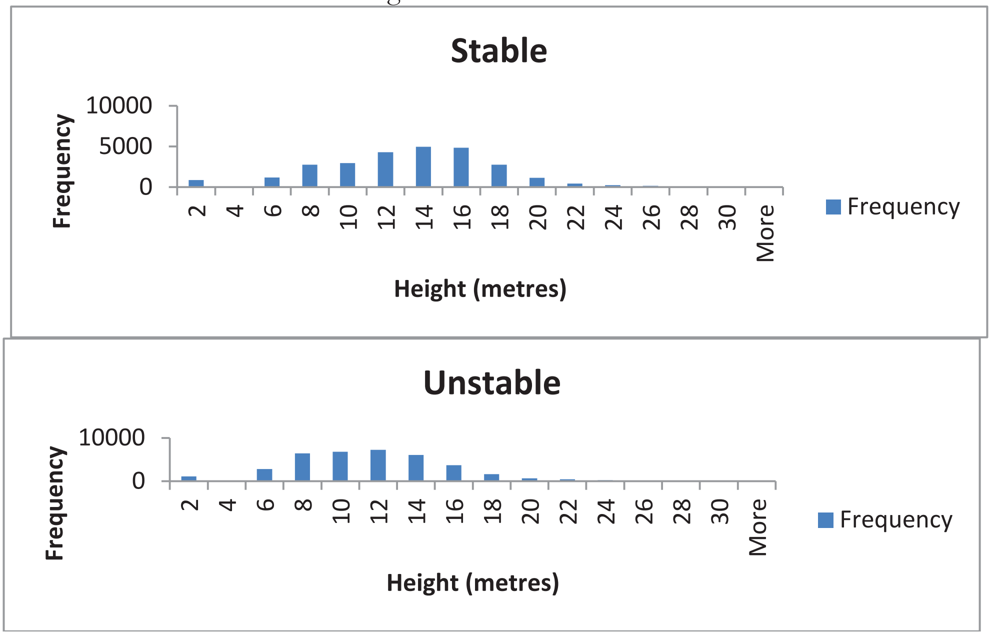



*Fig.3.16. Histogram of tree height distributions in stable and unstable zones*

Furthermore, to analyze the diversity in tree heights, a Shannon's index was generated. The Shannon's index for unstable area was 10.44 and 10.11 for stable area.

To compare the tree height distributions in both areas, a Mann-Whitney U test was performed, the results of which are displayed below:

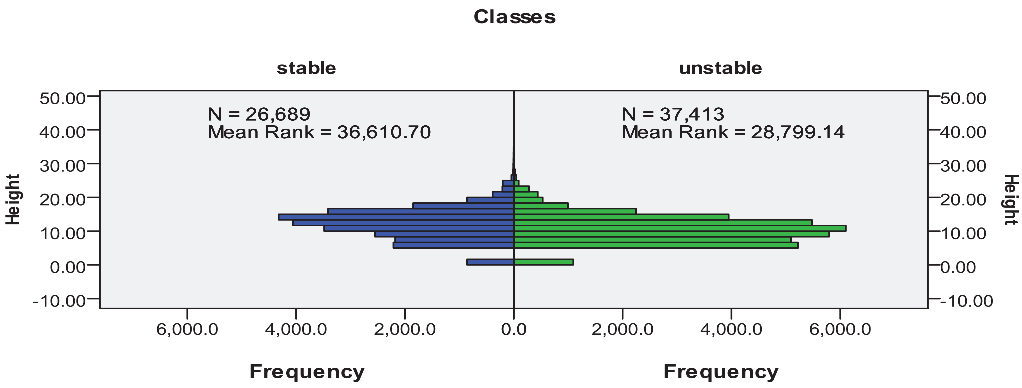
**Hypothesis Test Summary**

| | Null Hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of Height is the same across categories of Classes. | Independent-Samples Mann-Whitney U Test | .000 | Reject the null hypothesis. |
| 2 | The medians of Height are the same across categories of Classes. | Independent-Samples Median Test | .000 | Reject the null hypothesis. |

Asymptotic significances are displayed. The significance level is .05.

*Fig.3.17 a. Results of comparison of tree heights between stable and unstable zones using a Mann Whitney U test.*

**Independent-Samples Mann-Whitney U Test**




| | |
|---|---|
| **Total N** | 64,102 |
| **Mann-Whitney U** | 377,577,315.500 |
| **Wilcoxon W** | 1,077,462,306.500 |
| **Test Statistic** | 377,577,315.500 |
| **Standard Error** | 2,309,509.199 |
| **Standardized Test Statistic** | -52.687 |
| **Asymptotic Sig. (2-sided test)** | .000 |

*Fig.3.17.b. Results of comparison of tree heights between stable and unstable zones using a Mann Whitney U test.*

## 3.8. Tree density:

*Table 3.4. : Results of the T-test comparing tree densities in stable and unstable zones.*

| | *Variable .* | *Variable 2* |
|---|---|---|
| Mean | 0,0473 | 0,042125 |
| Variance | 0,0004 | 0,000551 |
| Observations | 20 | 20 |
| Pooled Variance | 0,0005 | |
| df | 38 | |
| t Stat | 0,7274 | |
| P(T<=t) one-tail | 0,2357 | |
| t Critical one-tail | 1,686 | |
| P(T<=t) two-tail | 0,4715 | |
| t Critical two-tail | 2,0244 | |

# 4. DISCUSSION:

This study incorporated three different fields: Forest structure analysis, LiDAR data analysis and forest disturbance analysis. Originally, it was an objective to extract structural attributes as indicators of forest disturbance. However, it was realised that it is possible to characterize certain forest attributes or tree structural anomalies, but more field data and survey is required before a certain attribute is called an indicator of disturbance. Field data collection could not be done extensively enough because of the following reasons:

1. Due to unavoidable circumstances, there was no time to make a sampling design or study the area before leaving for fieldwork. This made accessibility a big issue, because many areas were with a rugged terrain damaged by landslides.
2. The 3 different researchers in the study area had different objectives, so sampling had to be done considering everyone's objectives.
3. The Leica GPS system 1200 accompanied by a total station is a fairly heavy piece of equipment. Moreover, readings could be taken only at locations visible from the established GPS station. Thus, due to time restrictions, only one or two GPS station establishments were possible in one day.

These shortcomings made the field data collection a bit lacking. However, this study did prove that gap area can potentially be established as an indicator of forest disturbance. A major point to note here is that these gaps are not treefall gaps or gaps created by direct erosion of the super surface layer. Those are obvious indicators of disturbance. Interesting fact here is that the landslides in Bois Noir are *slow moving,* which means that they do not just uproot trees in their path. They just cause some growth anomalies, thus creating the phenomenon what the Siberians call drunken trees. Thus, these growth anomalies would have otherwise been very hard to detect from satellite imagery or any other low resolution remote sensing data. LiDAR opens up an entirely new window of opportunities for analysing forest structure in detail. Some prime properties of the forest stand in Bois Noir which make this study more interesting are:

a) The species diversity of the forest is fairly low, containing a majority of *Pinus uncinata* and some *Pinus Sylvertris* (Scots Pine) and very few *Larix decidua* (Larch) trees. This implies that the heterogeneity in heights or gap sizes is not due to species difference.
b) The tree density of the forest is high, ranging between 300 to 800 trees per hectare, which falls into the medium to high category as described in a study regarding tree density in regular coniferous forest stands (Hudak et al., 2006). This also implies that this tree density leaves very less room for canopy gaps, which is the prime issue in this study.
c) Owing to the high tree density and concluding from a previous study, it can be said that there is no factor of wind induced disturbance in the study area. A previous study which related climatic conditions to the landslides in the same study area (Bois Noir) concluded that the climatic conditions (such as rainfall, snow and wind) merely trigger landslides in the area, but do not cause any disturbances themselves.

The $R^2$ value of tree height regression was 0.72, which is still fairly low compared to other researchers who have used LiDAR data. Popescu(2002) got an $R^2$ value of 0.9 for estimating tree height, with data of much lower point density. Considering the high point density of the data, the $R^2$ value should have been much higher. The most obvious reason for this could be regarding field data.The accuracy of field data was also a question. Traditionally, field data is considered to be most accurate when using low resolution remotely sensed data, because the standard error in metres does not matter in such a case. However, LiDAR data used in this study had an average point spacing in centimetres, which means that the tree height estimation error could be only measured in centimetres. It was observed in field that 3 different researchers estimating tree height had a semi-consistent difference of 1 meter in measuring height of the same tree, using the Nikon laser rangefinder. This raises an important question that given LiDAR data of high point density, is field data collected by humans accurate enough to validate LiDAR data?

# 5. CONCLUSIONS

## 5.1. From field:

The major conclusion from field data collection was that the sampling design should be decided before starting fieldwork, especially in landslide affected mountainous areas, because of accessibility issues. While sampling in plots, the tree density should also have been noted down.

While sampling in transects, the predefined sampling scheme is even a larger issue. The predefined transect length of 50 metres could not be used in reality because of inaccessible terrain or extreme slope. Furthermore, a camera with fish-eye lenses should be used to take a picture of each gap, because it captures the gap in 180 degrees and can be used to calculate gap length and area.

From the descriptive statistics of field data, the following conclusions could be drawn:

1. The DBH of trees does not differ in stable and unstable areas, as concluded from the Mann Whitney U test with a significance of 0.359.
2. Canopy gap area was found to be much more in unstable areas than stable areas, as concluded from the Mann Whitney U test with a significance of 0.002.
3. Gap length percentage of transects in unstable areas (mean 69.83) was much more than that in stable areas (mean 36.25).

## 5.2. From data:

5.2.1. In relation to research question 1

(Which tree structural parameters differ significantly in affected and unaffected forests?), it can be concluded that canopy gap area, canopy gap shape (area to perimeter ratio) and tree height distributions differ significantly in stable and unstable areas.

5.2.2 In relation to research question 2 (What is the relation between structural anomaly and the type of the landslide affecting the tree?) it can be concluded **only from field observations** that a rotational landslide induces a backward tilt in trees, and the inclination and orientation values change at different tree heights. In contrast to this, a translational landslide produces a forward tilt in trees, but the tree inclination and orientation stays uniform. However, this could not be validated, so it remains as a mere field observation.

5.2.3. In relation to research question 3 (To what extent can we derive tree parameters from the high density LiDAR data?) the following conclusions can be made:

i) Tree height: These can be derived very easily, both from the LiDAR point cloud and from a gridded canopy height model.

ii) Canopy projection area: This can be derived from a gridded Canopy Height Model, by digitizing or automatic delineation in ECognition.

iii) Diameter at Breast Height: This can be derived if an individual tree is extracted and used as input for Skeltree software, provided the tree has enough LiDAR points on its stem. However, this requires a large amount of time and manual work.

iv) Tree inclination and orientation: This can be derived by extraction of tree from LiDAR point cloud and using as input for Skeltree. However, it must be ensured that the tree has enough points on the stem, and some ground points are retained.
However, due to time constraints, tree inclination and orientation and DBH could not be extracted.

5.2.4. In relation to research question 4 (How reliable is the LiDAR data for analysing the tree structure?) it can be said that tree heights can be derived with a high amount of accuracy ($R^2$ value of 0.72). Also, canopy gaps can be detected with a fairly high accuracy (81%) which increases with grid resolution.

Overall, it can be concluded that "canopy gap area differs significantly between stable and unstable forests, and can be considered an indicator of landslide induced forest disturbances".

# 6. RECOMMENDATIONS:

Though this study stopped at proving that certain forest structural attributes differ significantly, it opens up a new topic for further research, perhaps one relating these anomalies to landslides in deeper detail, with explicit knowledge of landslide and the underlying terrain. Some of the tree attributes are discussed in detail below:

1. Tree height: It was concluded that though there is no major difference in tree heights between stable and unstable areas, there is indeed a major difference in the distribution of tree heights in stable and unstable areas. This opens up a new question to investigate: Is this a sign of ecological disturbance or mechanical disturbance?

2. Canopy Gaps: This forest structural attribute has been studied for decades now, James Runkle (J.R. Runkle, 1982) being the 'pioneer' of studying canopy gaps with minute details. LiDAR makes gap analysis more interesting because now every detail of the gap sampled in field can be verified using the point cloud. Some factors associated with gap analysis are discussed below:

i) Gap definition: The initial concept of a forest canopy 'gap' was as an area of disturbance created by a single treefall (K.A. Razak, et al., 2011; J.R Runkle, 1992). However, the definition of gap should differ based on the motive of the study. Kneeshaw (Vepakomma, et al., 2008) set a 5 meter vegetation height threshold for defining gaps. However, in this study, the topic of interest was not canopy gaps due to treefall, but gaps created due to slow moving landslides. There are major agents like hurricanes, earthquakes, volcanic eruptions which cause huge gaps in canopy. However, since slow moving landslides are not the major agents of gap creation, an area threshold had to be set to define gaps. Surrounding trees were also a major issue, because there were ridges and earthflow zones there was a continuous passage of no vegetation. However, these could be detected by calculating the area to perimeter ratio. It could be said that area to perimeter ratio of gaps says more about disturbance patterns than just gap area, because it also speaks about the shape of the gap. Also, owing to the high point density of LiDAR data, the intra canopy gaps were also visible, which had to be eliminated from the analysis. Thus, the definition of gaps (limited to this study) would be "an opening between multiple canopies, with an area of more than $1m^2$, and an area to perimeter ratio of more than 0.4". This also implies that the largest fitting eclipse within the gap occupies a minimum 40 % of the total gap area.

ii)

ii) Sampling gaps and recording gap properties: As mentioned earlier, due to time constraints, a perfect sampling scheme could not be made. It is important to mention here that while sampling for gaps, accessibility and location should be the prime factors to be taken into consideration. While recording gap properties, namely major and minor axis of the gap, it was a dilemma whether to consider the end of surrounding canopies as the boundary of the gap or the base of the surrounding trees. Both methods are valid, however, differ significantly in measurements. Moreover, the gap axes were measured from the ground, so it was a "view from below", whereas while validating the size of gaps, the measurements are from "above". This factor could also be the main reason for a low $R^2$ value for gap area validation. Once again, the accuracy of human measurement vs. LiDAR accuracy comes into question, as the axes measured in field were done by approximately starting from the end of one canopy to the opposite one in the axis.

iii) Gap detection: Going with the parameters mentioned above for defining gaps, it is necessary to analyze the number of gaps that can actually be detected from the LiDAR data. Gridding is essential as the view in two dimensions gives a better idea about the size and shape of the gap. It was also observed that gaps are better detected when the grid resolution is high. However, gaps could be even better detected from the LiDAR point cloud itself. A previous study concluded that gap detection accuracy increases by 16% when done using the LiDAR point cloud instead of gridding and interpolation(Gaulton & Malthus, 2010). However, this method could not be implemented in this study because of failure on the part of Ecognition to handle such a high point density. At a 15cm grid size, 81% of the gaps were detected. But, the expected accuracy was much higher because: The LiDAR data had a point density as high as 180 points per $m^2$, and the coordinates of the gap location had been taken with a Leica GPS system 1200 coupled with a total station, which gives accuracy up to millimeter level. This opens up a new possibility that the gap was created in the period between LiDAR data acquisition and field data collection. Also, at some places, the location collected in field, fell within the canopy of a tree when overlaid on the CHM. A 1 meter threshold was set for such cases. However, this could also mean that the canopy has shifted, indicating tree disturbance. Both these cases point to a forest disturbance which has occurred between 2009 and 2011.

iv) Gap delineation: This proved to be the toughest job of all in gap analysis. Previous researchers came up with different methods to delineate gaps, yet no research can answer the question "when two gaps are interconnected, what could be considered as the boundary between two gaps"? The easiest way out of this problem is to consider gap fraction, so there is no question of determining the boundary between two gaps. However, it still proves to be a hindrance while validating gap area. Manual digitizing is also an option, but it is open to criticism because of human-induced bias.

Two new possibilities came up after gap delineation and area validation was done:

a) If the gap area sampled in field was lesser than that derived from the CHM, this could mean that gap closing is in progress. As mentioned earlier, regeneration was out of question because the time period between LiDAR acquisition and field data collection was just 2 years. Thus, the most probable reason for gap closing is that the trees have been subjected to more disturbance of some kind, causing them to incline more or change their orientation at certain heights.

b) If the gap area sampled in field was more than that derived from the CHM, this very clearly points towards disturbance, which caused the gap to widen.
Both these possibilities lead us into the field of gap dynamics. Gap dynamics have been studied by nearly every researcher who studies canopy gaps. Gap dynamics tell us about the ecological conditions in the forest, and also whether the forest is in quasi-equilibrium(http://www.csc.noaa.gov/crs/tcm). Since LiDAR data was available for only the year 2011, studying gap dynamics was impossible in this case. However, since it can be concluded from this study that gap area shows a significant difference in landslide and non-landslide areas, gap dynamics could easily be

associated with the landslide processes that occur in the study area. It opens up an entirely new field of study, especially for the Bois Noir landslide. Previous studies dealing with analysis of trees for reconstruction of landslides mainly dealt with tree ring analysis (Lopez Saez, et al., 2011). Now, with high resolution data like LiDAR available, it is well possible to relate various tree characteristics with landslide processes. In fact, for a landslide researcher or geomorphologist, it is also possible to reconstruct past landslides using gap dynamics.

Thus, it can be confidently stated that " *canopy gap area, gap shape do indicate disturbance caused due to landslides, however, studying these two attributes at multiple points in time (that is gap dynamics) could potentially be a major indicator of landslide processes in the Bois Noir region"*.

It is hoped that if LiDAR data for Bois Noir is collected again, some research will come up with the gap dynamics and their association with landslide processes.

Last but not the least, it is also important to mention that LiDAR is emerging as an excellent source of data, especially to study forest structure. Given efficient software like Quick Terrain Modeler, an endless horizon of possibilities arises from studying the LiDAR data. For decades researchers have struggled with low resolution remotely sensed data for analyzing forest structure, but LiDAR opens a new window of opportunities to study the forest. To describe the efficiency of LiDAR data, it would be appropriate to quote William Blake here, "*if the doors of perception are cleansed, everything would appear to man as it is: infinite"*.

## LIST OF REFERENCES


Baker, W. L. (1990). Species richness of Colorado riparian vegetation. *Journal of Vegetation Science, 1*(1), 119-124.

Bucksch, A. (2011). *Revealing the skeleton from imperfect point clouds*: Verlag Dr. Hut.

Bucksch, A., et al. (2010). SkelTre: Robust skeleton extraction from imperfect point clouds. *Vis. Comput., 26*(10), 1283-1300.

Corresponding, S. K., et al. (2004). Quantifying the spatial properties of forest canopy gaps using LiDAR imagery and GIS. *International Journal of Remote Sensing, 25*(15), 3049-3072.

Flageollet, J. C., et al. (1999). Landslides and climatic conditions in the Barcelonnette and Vars basins (Southern French Alps, France). *Geomorphology, 30*(1-2), 65-78.

Gaulton, R., et al. (2010). LiDAR mapping of canopy gaps in continuous cover forests: A comparison of canopy height model and point cloud based techniques. *International Journal of Remote Sensing, 31*(5), 1193-1211.

Hais, M., et al. (2009). Comparison of two types of forest disturbance using multitemporal Landsat TM/ETM+ imagery and field vegetation data. *Remote Sensing of Environment, 113*(4), 835-845.

http://www.cs.unc.edu/~isenburg/lastools/download/lasheight_README.txt. from http://www.cs.unc.edu/~isenburg/lastools/download/lasheight_README.txt

http://www.csc.noaa.gov/crs/tcm. from http://www.csc.noaa.gov/crs/tcm

Hudak, A. T., et al. (2006). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing, 32*(2), 126-138.

Hudak, A. T., et al. (2009). LiDAR utility for natural resource managers. *Remote Sensing, 1*(4), 934-951.

Hyyppä, H., et al. (1999). Comparing the accuracy of laser scanner with other optical remote sensing data sources for stand attribute retrieval. *The photogrammetric journal of Finland, 16*(2), 5-15.

John F, S., Jr. (1978). Dendrogeomorphological analysis of mass movement on Table Cliffs Plateau, Utah. *Quaternary Research, 9*(2), 168-185.

Lopez Saez, J., et al. (2011). Dendrogeomorphic reconstruction of past landslide reactivation with seasonal precision: the Bois Noir landslide, southeast French Alps. *Landslides*, 1-15.

Mackey, B., et al. (2009). Long-term kinematics and sediment flux of an active earthflow, Eel River, California. *Geology, 37*(9), 803-806.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment, 80*(1), 88-99.

Nelson, R., et al. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment, 24*(2), 247-267.

Persson, A. (2001). Extraction of individual trees using laser radar data. *2001*.

Popescu, S. C., et al. (2002). Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture, 37*(1-3), 71-95.

Popescu, S. C., et al. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing, 29*(5), 564-577.

Razak, K. A., et al. (2011). *Characterizing tree growth anomaly induced by landslides using LiDAR*.

Razak, K. A., et al. (2011). Airborne laser scanning of forested landslides characterization: Terrain model quality and visualization. *Geomorphology, 126*(1-2), 186-200.

River, S. (2005). Physiography, Climate, and Land Use. *Rivers of North America*, 152.

Runkle, J. R. (1982). Patterns of disturbance in some old-growth mesic forests of eastern North America. *Ecology, 63*(5), 1533-1546.

Runkle, J. R. (1992). Guidelines and sample protocol for sampling forest gaps. *U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station*.

Schuster, R. L., et al. (2003). Impact of landslides and innovative landslide-mitigation measures on the natural environment. *Geologic Hazards Team, US Geological Survey, Denver, Colorado*.

Sithole, G., et al. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing, 59*(1-2), 85-101.

St-Onge, B., et al. (2004). Assessing forest gap dynamics and growth using multi-temporal laser-scanner data. *Power, 140*, 200uJ.

Thiery, Y., et al. (2007). Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. *Geomorphology, 92*(1-2), 38-59.

Thomas, N. E., et al. (2011). Validation of North American Forest Disturbance dynamics derived from Landsat time series stacks. *Remote Sensing of Environment, 115*(1), 19-32.

Van Den Eeckhaut, M., et al. (2009). Evidence for repeated re-activation of old landslides under forest. *Earth Surface Processes and Landforms, 34*(3), 352-365.

Vehmas, M., et al. (2011). Using airborne laser scanning data for detecting canopy gaps and their understory type in mature boreal forest. *Annals of Forest Science, 68*(4), 825-835.

Vepakomma, U., et al. (2008). Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote Sensing of Environment, 112*(5), 2326-2340.

Vepakomma, U., et al. (2011). Response of a boreal forest to canopy opening: assessing vertical and lateral tree growth with multi-temporal lidar data. *Ecological Applications, 21*(1), 99-121.

Yu, X., et al. (2004). Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment, 90*(4), 451-462.

Zhang, K. (2008). Identification of gaps in mangrove forests with airborne LIDAR. *Remote Sensing of Environment, 112*(5), 2309-2325.

# APPENDICES

**Appendix 1: Field data collection form:**

i) Plots:

| Tree ID | Tree Position | | | Tree Height (m) | Tree orientation at () Heights (m) | | | | Inclination angle at () Heights | | | | DBH (m) | Tree Species | Crown Diameter (m) | Type of deformation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | x | y | z | | 0 | 0.5 | 1.3 | 2 | 0 | 0.5 | 1.3 | 2 | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |

i) Transects:

Field Observation Form, Barcellonnette

Transect no: 6

Transect length: 50m

Type of landslide: Rotational+Earthflow

| Gap ID | Photo id | Number of surrounding trees | Major Axis Length (m) | Minor Axis Length (m) | x | y |
|---|---|---|---|---|---|---|
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |

**Appendix 2: List of command lines used in LAStools:**

1. For LASGround:

lasground -lof file_list.txt -merged -o lasground1.laz

2. For LASHeight:

lasheight -i *.las -replace_z

3. For Gridding:

lasgrid -i *.las -otif -step 0.1 -elevation -lowest -fill 3

4. For clipping from shapefiles (lasclip) :

lasclip –i *.las –poly transect.shp –o transect.las

5. For clipping by coordinates (las2las):

las2las -i *.las -o out.las -clip 630250 4834500 630500 4834750

6. For viewing a .las file:

lasview –i transect7.las

Note: The specific file names or coordinates mentioned in these command lines are just as an example to facilitate understanding on the part of the reader/user.

## Appendix 3: Rule set used in ECognition:

- segmentation
  - chess board: 2 creating 'level1'
  - shadow masking
    - unclassified with Mean Layer 1 < 5 at level1: gap1
    - gap1 at level1: merge region
    - gap1 with Area >= 2000 Pxl at level1: remove objects into gap1 (merge by shape)
  - crown delineation
    - gap1 with Area >= 1000 Pxl at level1: Gaps_use
    - gap1 with Elliptic Fit > 0.2 at level1: closing: gap1
    - at level1: export object shapes to Tree Heights
    - tree top1 with Max. pixel value Layer 1 > 5 at level1: enclosed by domain: tree top1 +
    - unclassified with Mean Layer 1 >= 3 at level1: Trees1
    - Trees1 at level1: tree top1
    - Trees1 at level1: min Layer 1 : seed1
    - seed1 at level1: merge region
    - seed1 with Rel. border to Trees1 < 1 at level1: Trees1
    - Trees1 at level1: max Layer 1 : tree top1
    - tree top1 at level1: merge region (*)
    - 15x: tree top1 with Existence of seed1 (1) < 1 at level1: <- Trees1 Mean diff. to brighter neighbors Layer 1 <= 10