Estimation of Urban Tree Crown Volume based on Object-oriented approach and LIDAR Data

Changok, Lim

February, 2007

Estimation of Urban Tree Crown Volume based on Object-oriented approach and LIDAR Data

by

Changok, Lim

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: (fill in the name of the specialisation)

Thesis Assessment Board

Chairman: Prof.Dr. A. Stein
External examiner: Dr.Ir. B.G.H. Gorte
Supervisor: Ms Dr. W. Bijker
Second supervisor: Ir. S.J. Oude Elberink

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

Abstract

Recognizing the importance of the urban trees, the interest in preserving and maintaining trees in urban environment is increasing. However quantifying the value of urban trees for estimating and monitoring is difficult and has little studied so far. With respect to quantification of tree in urban communities, tree crown volume is a practical and meaningful factor.

The core of this research is to estimate urban tree crown volume with reliable processing and analysis techniques. To do this, we used only a rasterised LIDAR image which has great potentials for tree crown volume estimation in terms of improving the accuracy and reducing the complexity of analysis without combination to optical imagery. The study on the estimation of tree crown volume is consist of five main parts: data preparation (rasterized LIDAR image), pre-processing (nDSM and tree mask), tree detection and crown delineation, statistical analysis for crown volume estimation model and validation.

We interpolated LIDAR point data having approximately 1 point / m^2 density into the raster image of 25 cm grid size under consideration of representing tree crowns. Normalized DSM (nDSM) and tree mask was created based on this rasterized LIDAR image (DSM). For generation of nDSM, we produced DTM from DSM by applying object-oriented image classification techniques in order to separate terrain and off-terrain. After successful separation, we re-interpolated using the LIDAR points of only terrain. A tree mask was created by discrimination of trees and buildings based on surface growing algorithm with some manual intervention.

We employed the multi-resolution segmentation algorithm implemented in eCognition software for individual tree detection and crown delineation. By iteration of local maxima and local minima, tree top area detection, instead of finding tree apex point, showed promising possibilities for particularly flat top trees. However, tree detection partially failed when trees have more than two main branches looking like an individual tree crown in the field and in the image. Difficulties caused by dense groups of trees having a similar height are investigated in the crown delineation process. Tree parameters were extracted from image objects in such a way that the maximum pixel value of object corresponds to total height. Crown height was estimated from the regression of total height, and crown diameter was measured by maximum and minimum of X- and Y- coordinates. Crown shape was determined by the ratio of total height to crown diameter based on a certain range of total height which classified geometric crown shape in an objective way.

Finally, we developed tree crown volume estimation model based on a solid geometric volume formula using the predicted crown height and crown diameter by regression. All results were validated by r^2, RMSE and test of Significance such as T-test & Wilcoxon signed rank test. Consequently, object-oriented approach based on rasterized LIDAR image for individual tree crown delineation and tree parameter extraction gave us promising possibilities and reliable tree crown volume estimation model.

Key words: Tree Crown Volume, LIDAR, nDSM, Tree mask, Crown shape,
Multi-resolution segmentation, Correlation and regression, Solid geometric volume formula.

Acknowledgements

I would like to express my sincere appreciate to both my supervisors. My first supervisor, Ms Dr. W. Bijker, gave me unlimited and continuous supports and critical comments. Moreover I do deeply appreciate for her inspiring motivation of this research. Also my second supervisor, Ir. S.J Oude Elberink, offered me constructive advices and encouragements.

I'd learned a lot and expanded my knowledge thank to all lecturers and professors in GFM department. The knowledge what I learned and the way to think scientifically from them will be an invaluable foundation to encourage me for the future work.

Although we have different colors and use different languages, we do understand each others since we are friends. When I stayed here they are my sisters and my brothers. Dear my friends, I will never forget the valuable memories that we shared in ITC.

- Sofanit Girma Araya (Ethiopia)
- Yifru Mamushet Zewuge (Ethiopia)
- Hailemariam Gedlu Kitaw (Ethiopia)

- Mohamed Thamby Mohamed Rafeek (Sri Lanka)
- Gamage Sanka Nirodha Perera (Sri Lanka)

- Benjamin Kishosha Mayala (Tanzania)
- Adam Patrick Nyaruhuma (Tanzania)

- Coco Musaningabe Rulinda (Rwanda)

- Ioannis Delikostidis (Greece)

- Peter Madzudzo (Zimbabwe)

- Karim Faisal (Pakistan)

- Matthew Kwabena Boafo (Ghana)

- Wang Langyue (China)

Especially, I do appreciate all Koreans who spent many moments here, particularly Jung Whan, Yun and Young Hyun, Kim for their kindness and taking care of me. We were a fantastic team for cooking and travelling.

Last but not least, I would like to emphasis that I couldn't finish my thesis successfully without my wife who devoted herself to support me. She delivered my first daughter however I couldn't be with her. Moreover she helped my field survey in spite of being pregnant with my second baby when she visited me for a while. How can I recompense to her? To my love, Sujin, this work is dedicated to you.

Table of contents

List of figures

List of tables

1. Introduction

1.1. Background

Trees are a valuable natural resource that is necessary and essential to our life as we know it. It is hard to imagine that without trees humans would exist. Especially, trees in and near urban communities are very important for people's living because they do not only provide visual joy for people, but also influence directly or indirectly urban environment through their physical characteristics. For example, trees influence urban environmental conditions and energy fluxes [1] and abundance of trees in urban areas may also influence air quality and human health [2].

Trees are major capital assets in various parts of cities such as sidewalks, residential area, public buildings and recreational facilities area. Parts of a community's infrastructure are publicly owned trees. As urban communities grow larger and faster than ever before, urban planners and municipalities become increasingly aware of the interest of trees. Now tree resource management in these areas becomes crucial for achieving sustainable development and maintaining and enhancing the quality of life and the environment. Therefore, it is necessary to know how trees contribute to the urban environment for better comprehension of the importance of tree in urban communities in advance.

Firstly, trees provide ecological and economical services such as; 1) reduced air pollution, 2) storm-water control, 3) carbon storage, 4) improved water quality, and 5) reduced energy consumption [3]. Trees reduce air pollution by trapping particulate matter in their leafy canopies and by absorbing noxious pollution into their leaves. These actions reduce human health problems related to air pollution. Tree canopies particularly intercept large amounts of rain, reducing the amount of runoff that is discharged into streams and rivers and extending the time that a watershed has to absorb rainfall. This reduces flooding and erosion. As trees grow they accumulate biomass that absorbs carbon and nutrients, locking them into a biological cycle that keeps them out of the atmosphere and hydrosphere. The storage of carbon reduces the greenhouse effect that is linked to problems of global climate change. With respect to economical services, trees ameliorate climate by transpiring water from their leaves, which has a cooling effect on the atmosphere. In addition, groups of trees intercept sunlight in summer and use it for photosynthesis. They shade roads, buildings, and other structures as well. Consequently, they help reduce energy consumption.

Secondly, trees provide social benefits. Benefits to society are harder to quantify, but that does not mean they are less important than the ecological and economical services that trees provide. Social benefits include increased job satisfaction, faster recovery time for hospital patients, and improved child development. For example, hospital patients who have a view of trees out of their window recovered more quickly than patients who did not. Properly placed and maintained trees have even been shown to reduce crime and enhance cognitive development in children [4]. In addition, many outdoor recreation activities, such as picnicking, hiking, or even just sitting on a back porch are more

enjoyable around trees. Trees provide homes and are an important component of habitat for many wildlife species.

To date, quantity of trees had been measured like percentage of green surface including trees and ground vegetation within blocks or cells or green area per inhabitant [5]. Quantity of trees based on percentage of green surface provides a means to identify and localize areas of re-planning and restructuring parts of a city. On the other hand, green spaces may also be used for characterizing city image, life quality and attractiveness of city quarters or planning units.

In summary, trees in and near urban areas play a significant role, since urban trees provide many environmental benefits to citizens and to our community directly. Aside from the obvious aesthetic benefits, they protect and enhance city dwellers' health and property. They can even help contribute to a community's economy and way of life. Moreover, urban trees provide great services for the well-being of the inhabitants because of their visual effect meaning that greenness of a city is generally appreciated and their contribution to local biodiversity (e.g. as nesting places for birds, squirrels or insects). Trees in cities are therefore on the job 24 hours every day working for all of us to improve our environment and quality of life.

1.2. Problem description and motivation

As mentioned above, trees are not only beautiful in themselves but add beauty to their surroundings. Beyond aesthetics and emotional well-being, trees perform important functions. For this reason, the importance of preserving and maintaining trees in urban environments is widely acknowledged. For instance, the importance of large trees is reflected in Dutch law for protecting them. Anyone wanting to cut such trees, even on their own land, needs permission from the municipality. Therefore, from an urban greenness point of view, trees in cities are considered a crucial factor for realizing environmental quality goals and fostering sustainable local development.

Problematic issue on urban tree management is a difficulty of sustainable tree management. This is because that the municipality has only figures of yearly number of large trees removed from the permissions granted, on both private and municipal land, and number of trees planted on municipal land. Trees are also planted on private land as well, but no figures exist on their number, nor on their size. It means very often the numbers of 'trees cut' and 'trees planted' are incomparable. Moreover, increase due to natural tree growth and decrease due to disease or death of tree makes it even more complicated.

For those problems encountered, a new quantitative method is required. Quantifying the urban trees for estimating and monitoring is however difficult and has little studied so far. To estimate urban trees at the individual tree level, a more specific and practical indicator should be extracted and used. For instance, simple density of greenness in urban area such as biomass per unit area is not sufficient. Tree species, number, location of trees removed and planted annually, height, stem diameter and tree crown diameter/volume need to be estimated and stored. It means that the quantification of urban tree at the individual tree level is required. Most of all, tree crown volume is essential since it is directly

related to ability of improving environmental benefits and the visual (aesthetic) effect and health of the tree. The size of a tree crown is strongly correlated with the photosynthesis of the tree. The crown displays the leaves to allow capture of radiant energy for photosynthesis. The biomass of the crown and the quantity and quality of the branch material is also of direct interest to ecological studies and research into the effects of trees on pollution. Furthermore, the tree crown also has great visual impact which can make a city more unique and valuable. To summarize, most tree values are completely related to tree crown. With respect to quantification tree value in urban communities, tree crown volume must be the most practical and meaningful factor.

1.3. Research objectives

The central objective of this research is to estimate urban tree crown volume with reliable processing and analysis techniques. This operational method should meet the following requirements:

- With the method, it should be a possible to compare the greenness of previous year with present
- Feasible technical route
- The accuracy could meet user's requirement.

The specific objectives to address the main objective are;

- To explore the strength and limitation of LIDAR for estimation individual urban tree crown volume estimation.
- To improve individual tree crown delineation this is based on an object-oriented approach using LIDAR.
- To develop the best model to determine urban tree crown volume with validation between field and LIDAR derived measurement.

1.4. Research questions

The research objective leads to the following questions;

1. What's the potential and the challenge of LIDAR for estimating tree crown volume?
   (1) What kinds of tree parameters should be determined and extracted?
   (2) What are strengths and limitations of LIDAR to extract these parameters?

2. Is tree crown delineation based on object-oriented approach using only rasterized LIDAR image available?
   (1) How to interpolate LIDAR point clouds for individual tree crown delineation?
   (2) What kinds of data processing are required for crown volume estimation?
   (3) How to extract tree height information and crown diameter?

3. How to estimate individual tree crown volume using field and LIDAR derived measurement?
   (1) Is there correlation between variables of tree parameters which are from field and LIDAR-derived measurement?
   (2) Which regression model is the best to estimate crown volume?
   (3) How to validate the results?
   (4) What kinds of error are expected?

1.5. Assumption and limitation

To avoid ambiguity of expansion in further step, it is essential to limit the boundary of research clearly. The main assumptions are following:

- Large trees (total height: higher than 8m) are the target of analysis.
- The bottom shape of the crown projection of all trees in study area is defined as circle.
- The analysis in this research does not take into account differences in tree species. This research only focuses on an algorithm for individual tree crown volume estimation, regardless of the species of the trees in the study area.

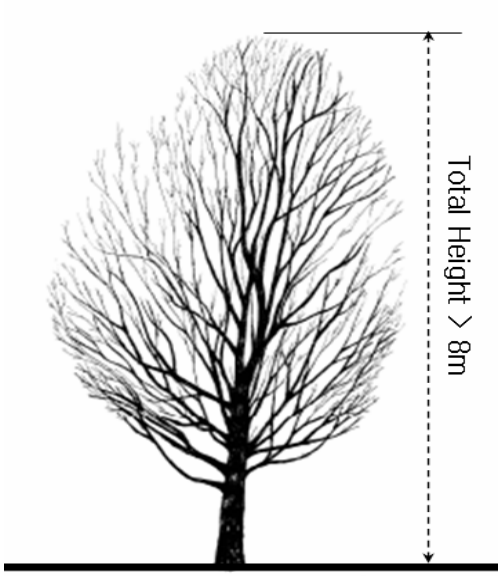


Figure 1.1. Assumption for target and shape of tree

1.6. Study area and Data

's-Hertogenbosch (literally "The Duke's Forest"), colloquially known as Den Bosch, is a municipality in the Netherlands, the capital of the province of North Brabant. It is located in the south-east of the Netherlands, about 80 km south of Amsterdam. 's-Hertogenbosch is characterized by the various landscape types. The municipality of 's-Hertogenbosch sets great green area.
(Source: http://en.wikipedia.org/wiki/%27s-Hertogenbosch)

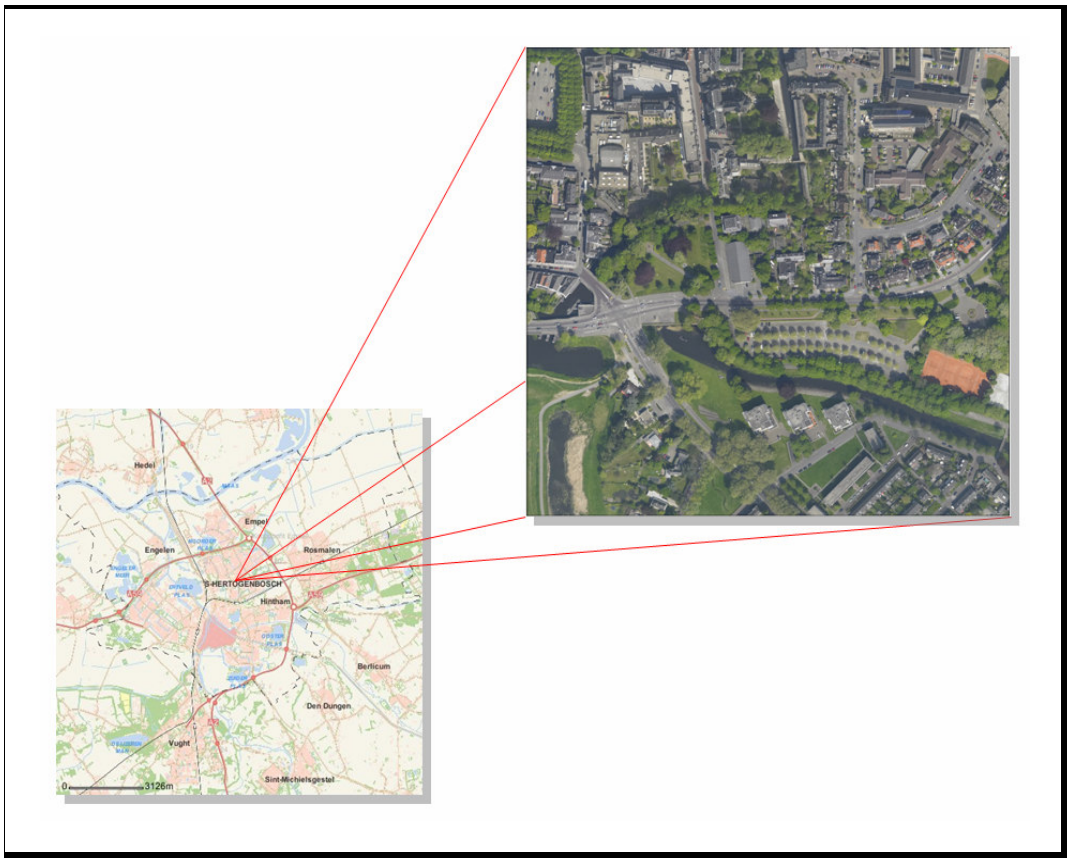


Figure 1.2. Study area: 's-Hertogenbosch, BINNENSTAD OOST, 500 m x 500 m

Table 1.1. City of ’s-Hertogenbosch

's-Hertogenbosch	
Country	Netherlands
Province	North Brabant
Coordinates	51° 42’ N 5° 19’ E
Area	91.26 Km^2
Land	84.63 Km^2
Water	6.63 Km^2
Population (2005)	134,009
Density	1,583 / Km^2

LIDAR to be used in this research was scanned by ATMLM 2050 which has scanning angle of -20 ~ 0 ~ + 20 and height accuracy of 5cm on 8th May 2006. The density of points is approximately 0.9 point / 1m^2. The basic statistics of LIDAR points are shown in table 1.2.

Table 1.2. Basic Statistics of LIDAR Points cloud

Raw LIDAR Points cloud	
No of values	231570
Minimum	1.73
Maximum	40.06
Range	38.33

1.7. Chapter scheme

Chapter 1 provides a general introduction to the thesis. It emphasizes the importance of tree in urban communities and motivates the necessity of tree crown volume estimation as more practical indicator for monitoring and managing urban trees. Also the research objectives and questions are listed.

Chapter 2 explores the challenges and possibilities of LIDAR for individual tree crown delineation through related work and literature review. And the central algorithms for tree detection and individual tree crown delineation have been summarized.

Chapter 3 explores data processing procedures and methods for crown volume estimation that is central process of this research. General approach of this study is introduced. Crown volume estimation is consisting of five main parts: data preparation (rasterized LIDAR image), pre-processing (nDSM and tree mask), tree detection and crown delineation, tree parameter extraction and statistical analysis for crown volume estimation model and validation.

Chapter 4 contains the result of analysis done.

Chapter 5 summarizes the all processes with discussion of the results.

Chapter 6 gives an overview of the main conclusion and recommendations for future work.

2. Related work and Literature review

2.1. Related work : Conceptual approach

Since tree analysis techniques have been conducted in various fields, it can provide promising alternatives and keys to resolve the tree crown volume estimation. In broad point of view on tree analysis, two categories, TROF (Tree resource outside forest) and Biomass estimation are closely related to this research. Of cause, the central techniques of tree analysis must be developed within forest management. The forest management is however too extensive to treat in this paper. Therefore, we benchmark conceptually in this research in terms of characteristics and rules of tree in urban communities.

2.1.1. Quantitative assessment for TROF using remote sensed imagery

By FAO [6], TROF is defined as trees and tree environments on land not defined as forest or other wooded lands. Generally the products, both tangible and intangible, derived from these tree resources help contribute to sustainable agriculture, food security and rural household economies. In certain circumstances they even supply many products and services similar to forests but in different extent; for example timber, fuel wood or recreation [7].

Most of previous studies were for food security, energy sources and useful products such as timber. However, to explain urban trees from a TROF point of view, urban trees give better quality of life environment more than others. Few studies on assessment of trees in urban areas using satellite image data or geographic information systems have been carried out with taking TROF into account. Moreover, they measured simply biomass per unit area in large scale but rarely at individual tree level. A repetitive assessment of urban tree is essential for its proper monitoring and management. With the advancements in remote sensing, it has now become possible to use these techniques to assess individual urban trees.

2.1.2. Biomass estimation for Kyoto protocol

To mitigate the effects of greenhouse gas emissions in the global temperature rise, more countries now seek to comply with agreements under United Nations Framework Convention on Climate Change (UNFCC), clean development mechanism of Kyoto protocol which requires the estimation of carbon sequestration or release in a specific period of time [8]. Above ground productivity derived from biomass assessment can be useful to quantify above ground carbon sequestration by the forest at any particular time interval [9]. Many remote sensing-based methods to estimate biomass have developed not only for Kyoto protocol but also forest type or biodiversity. The methods use satellites imagery with a variety of resolutions however the majority have done by per-pixel analysis [10], [11].

2.2. Literature review: Technical approach

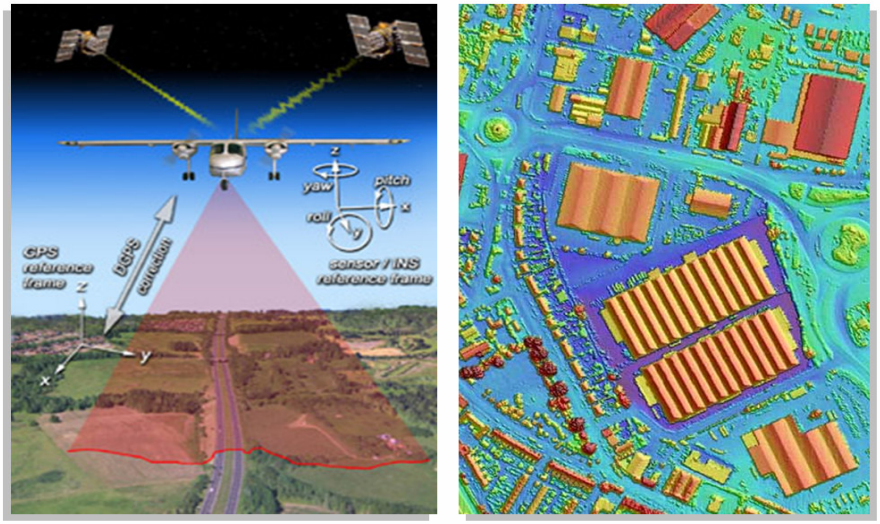
This section deals with methodological issues on tree crown delineation which is a basis of this research. Also, it explores the strengths and limitations of LIDAR to be used for tree crown volume estimation.

2.2.1. Introduction of LIDAR

It is commonly accepted that remote sensing techniques provide potential alternatives to conventional field survey but require the development of methods to easily and accurately extract the required information. Current field survey methods are labor intensive and costly, resulting in low sample coverage and frequency. Therefore remote sensing has potential to provide, at lower cost, forest information with greater coverage than is attainable using field sampling [12].

Amongst remarkable remote sensing techniques, in this section, we discuss the challenges and possibilities of LIDAR for tree crown analysis. Airborne laser scanning (LASER stands for 'Light amplification by stimulated emission of radiation') is a new technology in which several sensors are integrated to obtain 3D coordinates of points in the earth. Such systems are know as LADAR (Laser detecting and ranging) or LIDAR (Light detection and Ranging system) [13].

LIDAR makes use of precise GPS instruments to determine the position of the sensor, inertial navigation system (INS) to determine the attitude of the sensor, and narrow laser beams to determine the range between the sensor and the target points [14] (figure 2.1). Airborne laser scanning is now gaining importance in research and commercial use particularly in the field of mapping and monitoring environmental and ecological changes, urban planning, especially development of 3D city models and disaster management. Facilities for acquiring data both at day and at night including high planimetric resolution (up to 1m) with high accuracy in position enable users to use them in a wide range of applications [15].



(a) (b)

Figure 2.1. Principle of LIDAR: (a) Sensor operation (b) Example of laser scanning, Source: http://www.precisionterrain.com/

In recent years, the use of LIDAR technology to measure tree parameters has been rapidly increasing. LIDAR data combine both surface elevations (z) and accurate plane coordinates (x and y), and processing algorithms can identify single trees or groups of trees to extract various parameters on their three dimensional representation [16]. A review of the rapidly growing literature on LIDAR applications related to tree analysis has been summarized in [17].

2.2.2. The capabilities and limitations of LIDAR for crown volume estimation

This paragraph focuses on the capabilities and limitations of LIDAR to extract tree parameters for crown volume estimation. Tree height and crown diameter is essential for crown volume calculation. In terms of an automatic procedure, single use of optical imagery is still cumbersome and provides relatively low accuracy, although stereo images enable us to extract tree height information. Meanwhile combination of LIDAR and optical imagery complicates processing. Therefore single use of LIDAR could be the optimal alternative in terms of improving accuracy and reducing the complexity of analysis.

Many researches utilize the spectral characteristics of optical images to delineate the individual tree crown. However, optical images are easily influenced by terrain relief and weather condition [18]. For this reason, LIDAR-derived tree delineation processing is less complex than optical imagery since height information derived from LIDAR is an unambiguous physical quantity, while optical imagery, which is easily influenced by illumination conditions and occlusion, requires complex procedures such as atmospheric and geometric correction and similar spectral reflectance between vegetation types increases the complexity of processing as well. Nowadays, LIDAR technology provides horizontal and vertical information at high spatial resolution and high vertical accuracies. Once LIDAR data have been interpolated into regular grided raster and intensity image with height information, tree parameters such as total height and crown diameter can be directly retrieved from it. Moreover, laser scanning systems are active sensors. Therefore they can work day and night. Additionally, shadows do not affect laser data. Short data acquisition and processing times, relatively high accuracy and point density have caused LIDAR to be preferred over traditional aerial photogrammetric techniques [19]. LIDAR is particularly valuable for estimating crown volume since it dose not saturate at high biomass levels in contrast to optical imagery (e.g. NDVI) and LIDAR can easily provide tree crown information.

The accuracy of tree delineation depends on the structure of the tree (number of tree layers, tree species, density), and density of laser points and applied processing techniques [20]. Another of the difficulties in LIDAR data at the tree level is determining the correct geo-location of individual trees. Also the top of trees can be missed by the LIDAR if density of laser points is insufficient. In order to obtain high point density of LIDAR, relative cost is increasing. Besides, laser scanning is usually scheduled once in a few years or longer, which causes difficulties for upgrading.

Taking all considerations into account, LIDAR has ample possibilities for tree crown volume estimation in terms of improving the accuracy and reducing the complexity of analysis without combination to optimal imagery. However, to overcome its limitation, developing of a new approach remains.

2.2.3. Individual Tree Detection and Crown Delineation Algorithms

Indeed, parameters such as the number of tree crowns, the distribution of the crown diameters or the stem density are currently assessed by human interpretation [21]. Algorithms for the automatic extraction of these parameters would also greatly aid tree analysts in their work.

Many isolation techniques of individual tree crowns using optical imagery, LIDAR and their combination have been developed. More specifically, the techniques can be categorized into tree detection and tree delineation. Common algorithms to be used so far are summarized in table 2.1. and 2.2. Most research utilizes various combinations of these. This summary is not complete since algorithms for tree detection and crown delineation had been ongoing research field.

According to the summary of tree detection and delineation algorithms, it has been shown that they are performed based on two distinct spectral characteristic of tree crowns in both optical imagery and LIDAR. The basic tree crown properties are the following:

- Tree apex has local maximum image brightness value, for instance tree apex is shown as high reflectance spot by sun illumination in optical imagery and it has also a relatively bright value in rasterized LIDAR image since tree apex is higher than neighbours. Thus individual trees can be detected by finding local maxima. For this reason, normally, a tree crown appears as a bell shape in 3D surface.

- Delineation of crown boundary utilizes local minima brightness values.

Table 2.1. Summary of Individual Tree Detection Algorithms finding local maxima as tree apex

Categories	Description of algorithm	Reference
Enhancement and threshold	After Smoothing and high-pass filtering on global image, then defining ranges to extract tree location based on the resulting pixel brightness values	[22], [23]
Local maxima - filtering	Finding maximum pixel values as tree apex within specified size of moving widow	[24], [25]
Multi - Analysis	Examining the occurrence of edges over several different image to define a region where the brightest pixel value is recognized as tree apex	[16], [26]
Template matching	Matching pre-defined tree crown model and image based on geometrical and radiometric correlations	[27]

Table 2.2. Summary of Individual Tree Delineation Algorithms

Categories	Description of Algorithms	Reference
Valley following	outlining a net-work of minimum image values or growing local maxima out to regions of minima representing shade to determine crown boundaries	[28], [29]
Region growing	Identification of groups of similar neighbouring pixels or growing region until pre-defined parameter is addressed	[30], [31]
Watershed - segmentation	Inverting the images, trees appears as catchment basins	[18],[20] [32]
Edge detection	Finding convex edges and combining them to locate tree crowns	[26], [33]
Marked point - processing	After defining a correct geometrical object to design a tree crown, then fitting the object to the image based on both priori knowledge on the tree and taking into account of interactions existing between points in the process	[21]
Multi-resolution segmentation	Grouping pixels or small objects into a large object such as a tree crown based on spectral similarity, contrast with neighbouring objects and shape characteristics of the resulting object.	[34], [35]

3. Methods

3.1. Overview

Before main data analysis, we reviewed previous works focused on tree analysis to see the importance of urban trees, need for crown volume estimation and the challenges and possibilities of single use of LIDAR which provide the motivation for this research. In order to estimate tree crown volume at individual tree level, we decided to make use of LIDAR providing vertical and horizontal properties of the tree crown. As reviewed in chapter 2, single use of LIDAR has great potentials in terms of improving accuracy and reducing complexity of analysis. As a result, the main objective was addressed using field measurement and data processing of LIDAR data. Field measurement plays significant role in this research since it is used as input data for statistical analysis and for validation of LIDAR-derived results as well. Last but not least, processing using LIDAR data could be central of this research. In specific, it is comprised of five main categorized procedures as below (figure 3.1.).

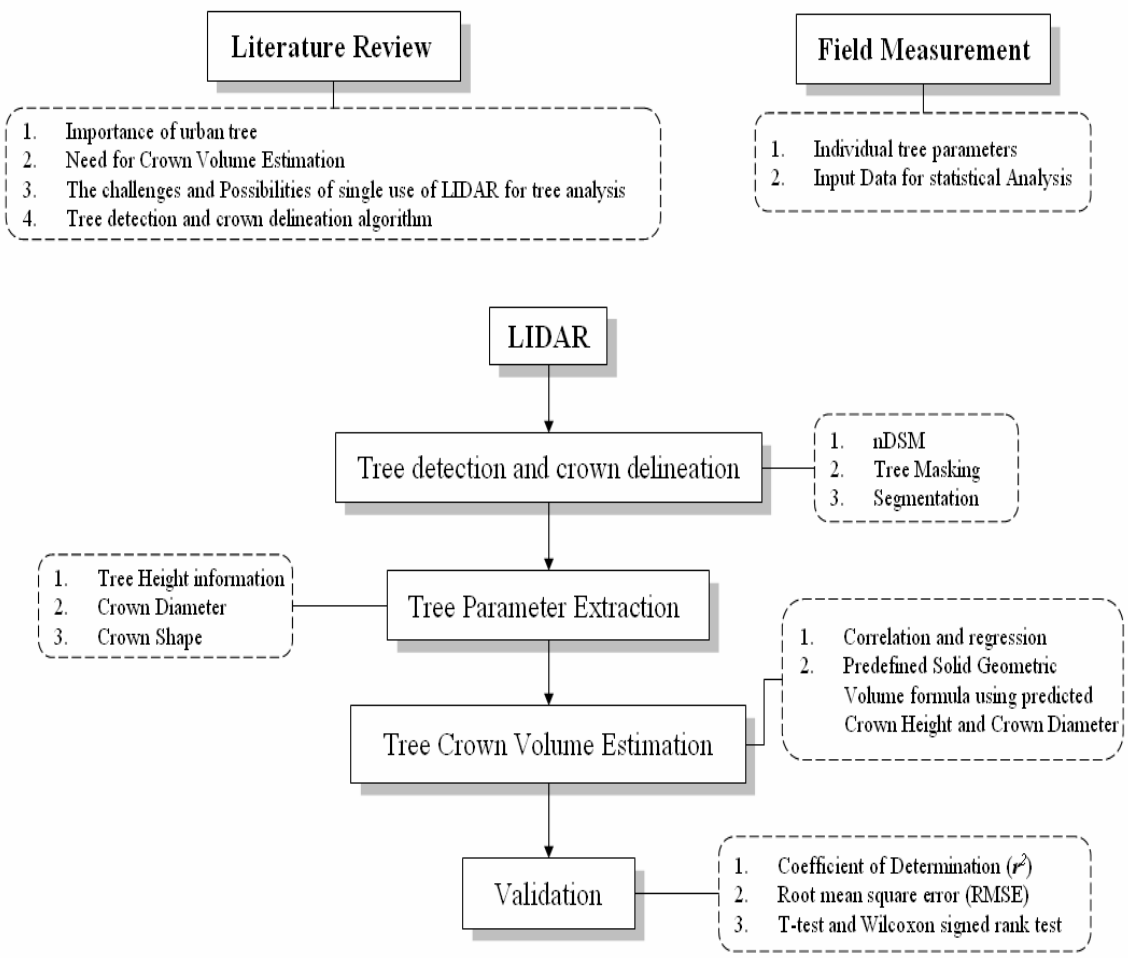


Figure 3.1. General Approach

3.2. Field measurements

Tree height is the critical value that is used to calculate crown volume directly. Tree height is an ambiguous term unless it is clearly defined [36]. Total height is the distance along the axis of the tree stem between the ground and the tip of tree. And Crown height is the distance between the height of first branch and the tip of tree. Eventually, two height measurements in meter were made for each tree by means of HAGA altimeter. The HAGA altimeter consists of a gravity-controlled, damped, pivoted pointer, and hexagonal scales are 15, 20, 25, 30, percentage and topographic scale. Sights are taken through a gun-type peep sight; squeezing a trigger locks the indicator needle, and the observed reading is taken on the scale [36]. First, the total tree height was measured, and, second the height to the first branch. The height to crown variable was defined as the height to first branch. Therefore, crown height is made by subtracting height to first branch from total height in the end. Therefore, we can read height directly from HAGA without further calculation.

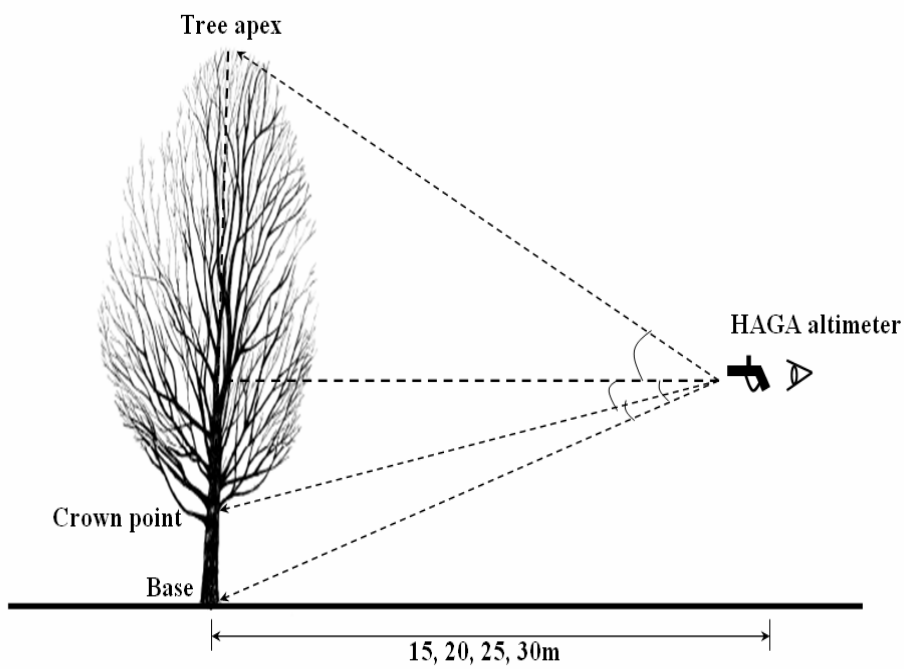


Figure 3.2. Tree Height measurement by HAGA altimeter

With tree height, crown diameter is also directly related to crown volume calculation. Crown diameter was measured by projecting the perimeter of the crown vertically to the ground and taking diameter measurements on this projection. We measured carefully taking into consideration that determination of the outline of the tree crown is difficult because of its irregularity. Crown diameter measurement using identical method in field and from LIDAR is desirable in order to extract accurate and compatible diameter. To do this, maximum and minimum crown diameter are based on projected crown on the ground as measured in field (figure 3.3). In the same way, crown diameter from LIDAR was measured using maximum and minimum of X, Y coordinates to estimate max and min of diameter.

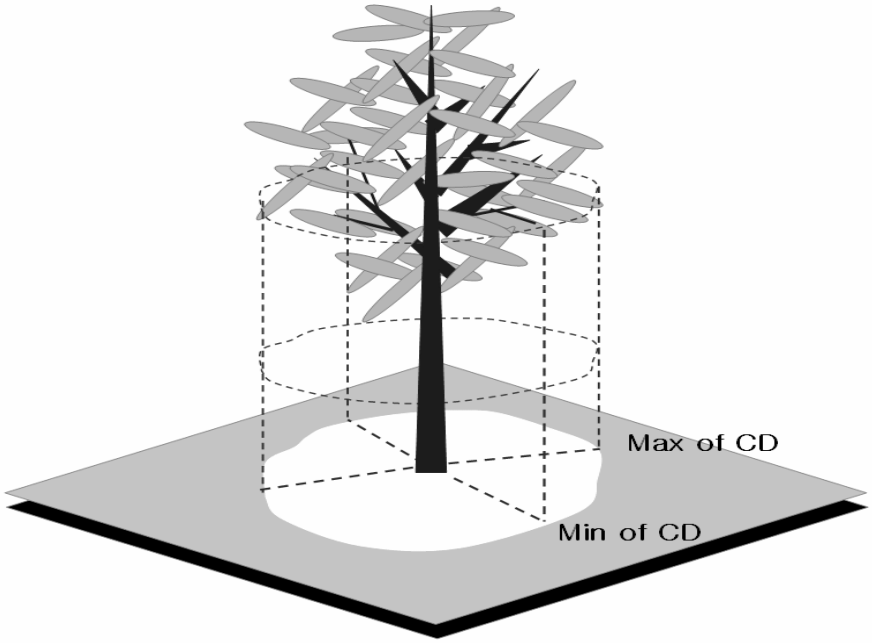


Figure 3.3. Crown Diameter (CD) measurement

3.3. Tree detection and Individual Crown Delineation

Among all processing, the algorithm for delineation of individual tree crowns is the first and most important step in crown volume estimation. An important advantage of tree delineation is the potential to model tree structural variables such as tree height, stem diameter, and biomass [12]. Especially it is essential to extract two variables, such as tree height and crown diameter, to be used directly for crown calculation. In this research, the procedure of individual tree crown delineation consists of four main steps (figure 3.4). Generation of normalized Digital Surface Model (abbreviated to nDSM) and tree mask is the fundamental pre-processing to prepare for segmentation of individual tree crowns. According to these sequences of process, tree parameters can be extracted based on individual segments and crown volume will be calculated in the end.

Intermediate results such as extraction of tree height and crown diameter which we want to extract in final stage, and final result such as crown volume depends on each of the above processes such as nDSM generation, tree mask, tree detection and crown delineation. For instance, undesirable interpolation of LIDAR points cloud will provide under- or over-estimation of tree height and tree crown could be represented in different ways based on resolution as well. This implies that at lower image resolutions crown boundaries becomes less distinct, making them harder to identify. With respect to tree mask, loss of tree periphery during tree masking can cause under-estimation of crown diameter.

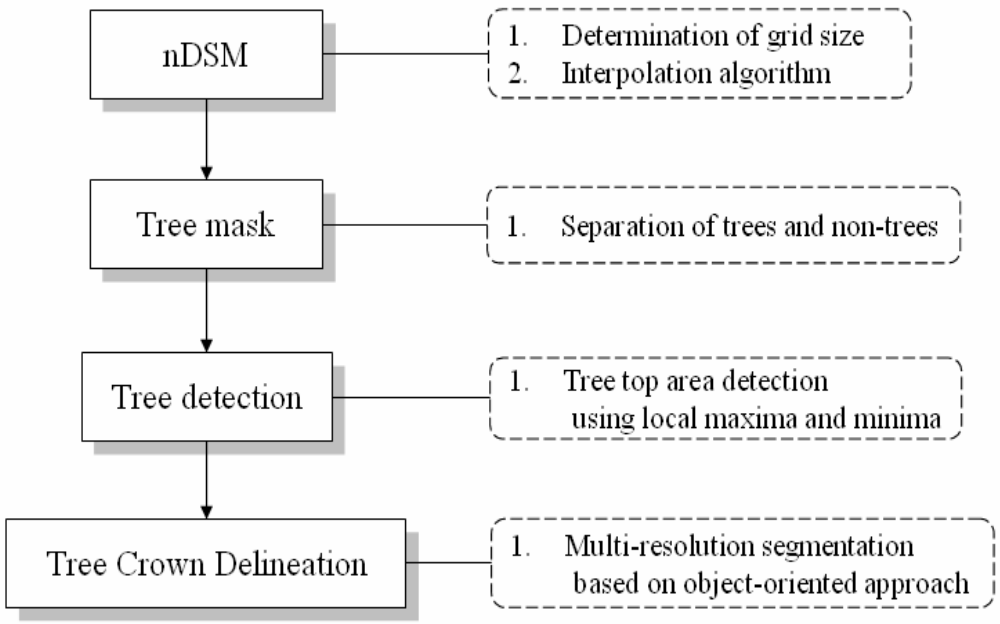


Figure 3.4. Work flow for Individual Tree Crown Delineation

As already mentioned in chapter 2, in the review of the algorithms that had been commonly used (chapter 2.2.3, table 2.1 and 2.2), most researchers used various combination of tree detection and crown delineation algorithms. Eventually, multi-resolution segmentation based on object-oriented approach is applied for individual tree detection and crown delineation in this study. The interpolation of LIDAR points is implemented in Surfer 8 and the full procedure of crown delineation is implemented in eConition professional 5. The specific processing for individual tree crown delineation is described with more detailed sub-processes in sequence.

3.3.1. Grid size and interpolation algorithm for crown delineation

Before interpolation of LIDAR point cloud into regular grid locate height values into individual cells, the grid interval has to be selected in an optimal way under consideration of reducing the information loss and keeping the redundancy to a minimum based on the density of raw LIDAR point cloud. This must be the common consideration for interpolation of raw LIDAR point cloud in any LIDAR application. In addition, the main target object to be studied should be represented well (in this case, it is a tree crown). For these reasons, we need to determine the optimal grid size for interpolation, in order to retrieve ideal height information and crown diameter from rasterized LIDAR data.

In order to resample the raw LIDAR point cloud data into new locations without information loss and redundancy, this can be done if each cell contains one-and-only one raw point. This is because that each cell can have only one value. In other words, if there is more than one point in a cell, they will be averaged into only one new value. Meanwhile, if one selects very small grid size, the number of pixels that contain no LIDAR points will become large. Thus the redundancy and requirement of storage will be increased [14]. According to one research [36], it suggested to use formula for determination of

grid size based on the density of raw LIDAR point cloud in such a way that if one has N LIDAR points distributed as grid in a unit horizontal area, the linear spacing between each pair of points in a row or column is equal to $1/\sqrt{N}$ in order to faithfully represent the laser points in a raster array. In case that the degree of irregularity increases, the average or minimum density of the LIDAR point cloud data can be used to estimate the optimal grid size using the above formula.

With respect to representing tree crowns, another research [12] proposed a more specific way using the ratio of the crown diameter to grid size. Figure 3.5 illustrates how tree crown is represented at different ratios of crown diameter to grid size. It is shown that 3:1 cannot provide distinct crown boundary and 19:1 may contain too much detail within-crown. Finally, it turned out that this ratio can be a good guide to decide the optimal grid size for individual tree crown detection and delineation. The choice of grid size still depends on tree detection / delineation algorithms to be applied and different characteristics of tree structure.

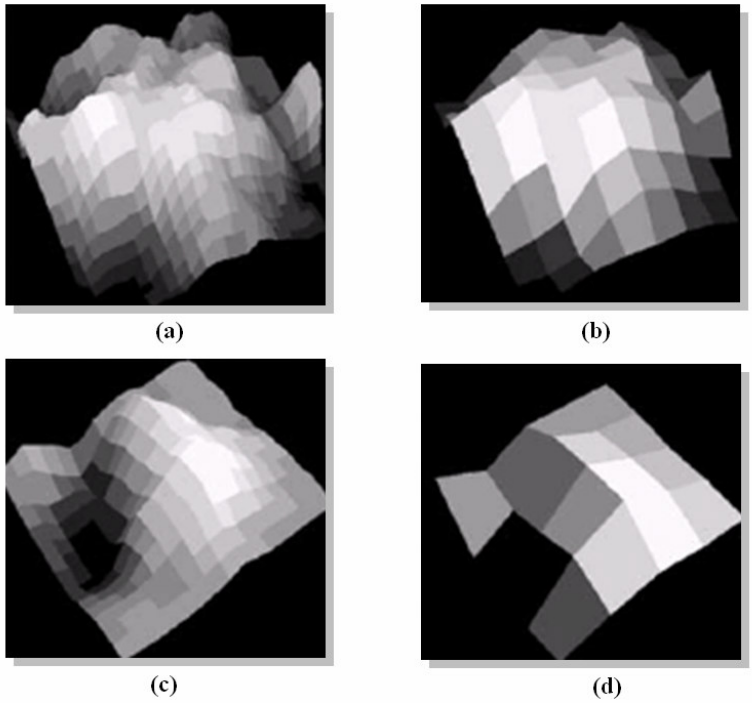


Figure 3.5. Tree Crown representation at different grid size (source form [12])
(a) Crown Diameter: 95.5cm, Grid size: 5cm, **ratio 19:1**
(b) Crown Diameter: 95.5cm, Grid size: 15cm, **ratio 6:1**
(c) Crown Diameter: 43.0cm, Grid size: 5cm**, ratio: 8:1**
(d) Crown Diameter: 43.0cm, Grid size: 15cm, **ratio 3:1**

Determination of grid size is still problematic in case of low density of raw LIDAR point cloud. The density of LIDAR data used in this study is approximately 1 point / m^2. Grid size should be decided as 1m based on this density. But crown boundaries at such a low rasterized LIDAR image resolution will become less distinct, therefore it will be difficult to identify them individually. On the other hand, if rasterized LIDAR image resolution is increased, the numbers of cells that do not contain a LIDAR point also increase. These cells will have interpolated values based on neighbours having LIDAR

points. It means we cannot expect more accurate variation within a crown. For instance, the representation of tree crowns using 10 points / m^2 will delineate height changes within a crown with more detail than representation of tree crown using 1 point / m^2.

There are many ways for interpolation, some of which work properly in some applications and fail in others. There is no single robust way for interpolation. The reason behind is that the interpolation is a prediction of what is not known [14]. Most of the interpolation techniques make the prediction using the direct/close neighbours and fit them into a model. It has to be noted that, the larger the distance between the points, the less contribution in the elevation function will be. Eventually, triangulated irregular network (TIN) which uses the optimal Delaunay triangulation is applied for interpolation in this study since considering that the TIN approach provides rather better results particularly for tree application [37]. Because the original data are used to define the triangles, the data are honored very closely [38].

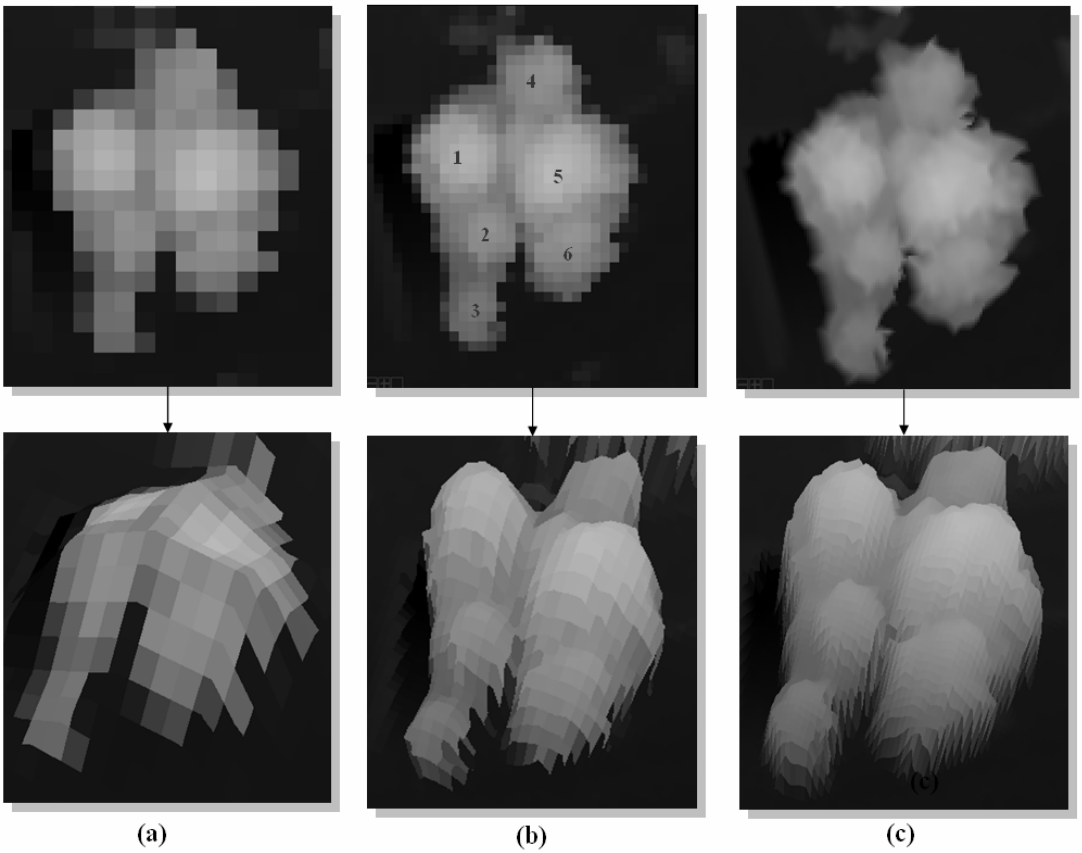


Figure 3.6. Interpolation (TIN) at different grid size: (a) 150cm, (b) 50cm, (c) 25cm
Crown Diameter in figure: 1) 13.95m, 2) 10.6m, 3) 6.65m, 4) 9.9m, 5) 13.8m, 6) 8.8m

Taking all considerations above to determine the grid size and interpolation algorithm in this study, we decide to use TIN and 25cm as grid size in order to avoid interpolation errors and to minimize crown underestimation that has been typically investigated in previous works. Even if 25cm grid size provides higher than 19:1, ratio of crown diameter to grid size, it can not provide more accurate

variation within crown because of the low density of raw LIDAR point cloud. Figure 3.6 shows the results of interpolation using TIN and raw LIDAR point clouds to be used for this study. It says that the smaller grid size is, the more apparent crown boundary is. However it has to be noted that, very small grid size leads to huge storage requirement and very long computing time.

3.3.2. DTM, DSM and nDSM generation

The term 'digital elevation model' generally refers to a digital representation of topographic surface where height values of the terrain are given [15]. According to USGS (U.S Geological Survey) definition [46] "A digital elevation model (DEM) is a digital cartographic / geographic data set of elevation in xyz co-ordinates. The terrain elevations for ground position sampled are regularly spaced horizontal intervals. Therefore, LIDAR data can be used directly as a DEM because it has xyz coordinates for each pixel. As it contains the heights of the upper surface of the terrain and the objects there on, it is referred to digital surface model (DSM) – more specifically it is called a canopy height model (CHM) in forest applications.

Since raw LIDAR data give heights (z) above the sea level, we need further processing to calculate the tree height above the ground. To deduce the actual height of the objects, height of the terrain should be subtracted from the DSM. It is necessary to create a digital terrain model (DTM), where the terrain is represented without the overlying objects. The subtraction of DTM from DSM results in the absolute height values of the objects and the model representing such heights is called a normalized digital elevation model (nDSM) (figure 3.7).

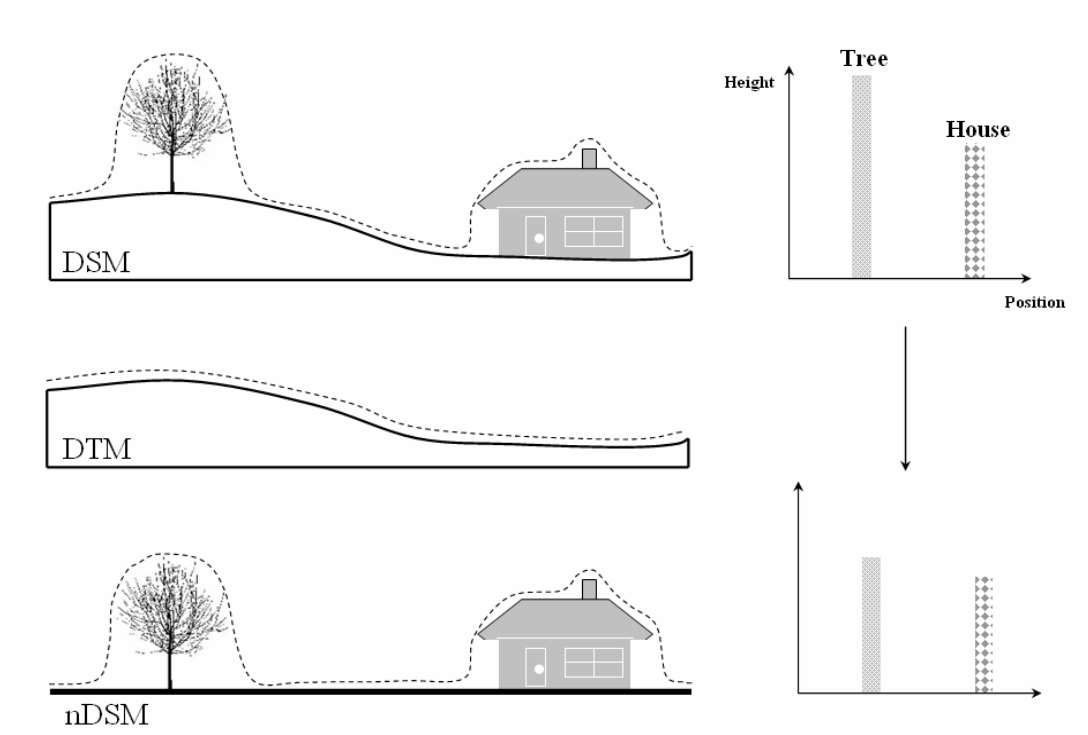


Figure 3.7. Normalized DSM

The usefulness of nDSM for tree height information retrieval depends on the accuracy of the DTM that is derived from the DSM by means of filtering techniques. The term 'filtering' can be defined as performing a removal of off-terrain objects. The basic problem is the separation of terrain points from off-terrain points which are both recorded by the laser sensor [39].

A DTM represents, in contrast to a DSM, the bare terrain surface without any objects such as trees and man-made objects. However the terrain data itself is often not available, it is too expensive or the available accuracy is not sufficient. Therefore the DTM is generated using a DSM, which is often acquired by laser scanning. Normally, DTM generation from DSM is carried out with non-gridded and gridded LIDAR data using segmentation, classification algorithm, various filtering techniques and their combinations. In this study, filtering for DTM generation from DSM is carried out with gridded LIDAR data, 2D rasterized DSM. Figure 3.8 illustrates the general approach. By subtracting DTM from DSM, we generate the nDSM in the end.

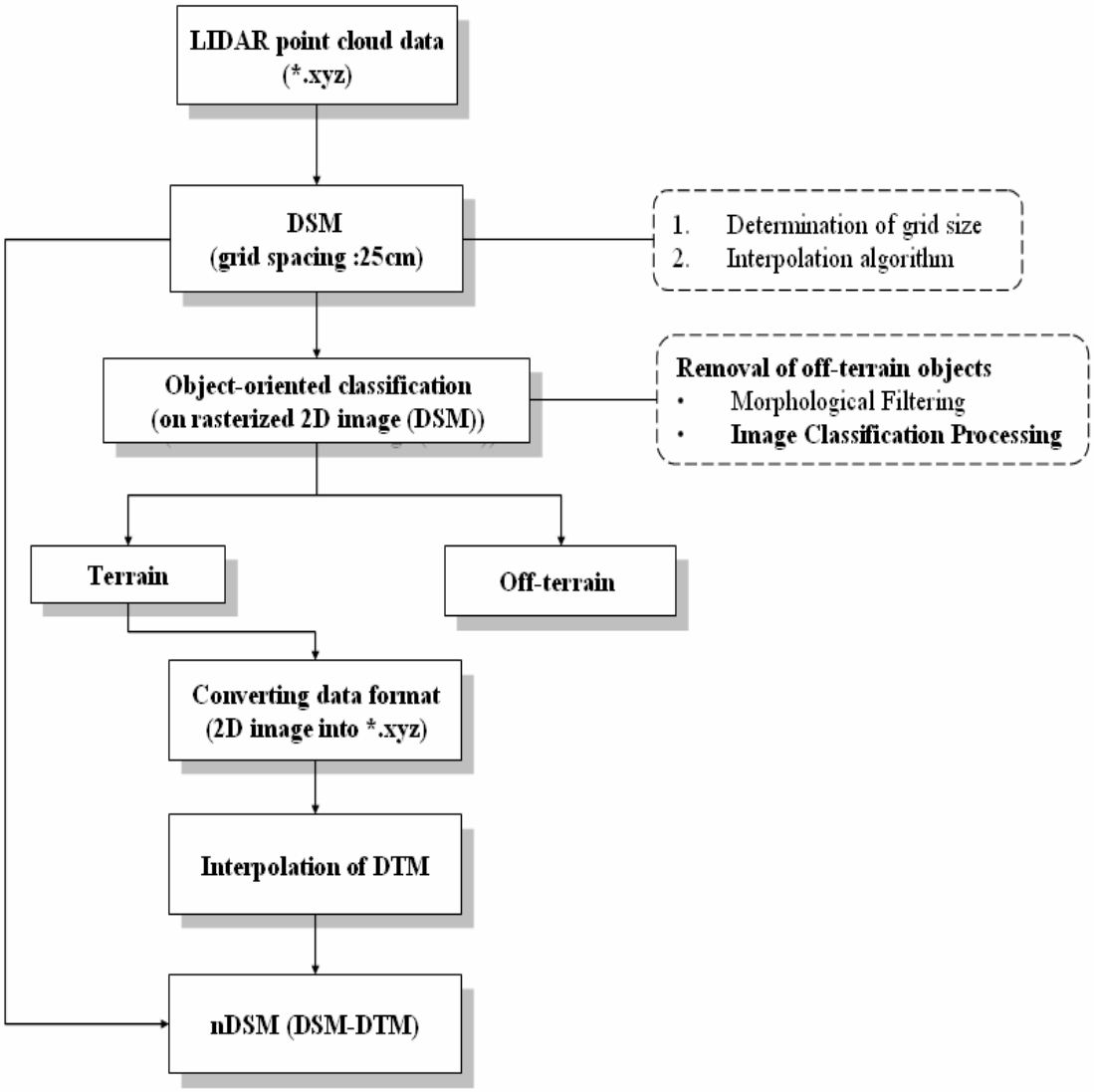


Figure 3.8. General approach for DSM, DTM and nDSM generation

3.3.3. Discrimination of trees and buildings : Tree mask

An important first step in any tree delineation and identification algorithm is the generation of a tree mask. This often defines the outer boundaries of crowns at the interface with non-tree areas. In nDSM which was created in the previous step, trees and man-made objects have to be separated from each other for further processing. A tree mask can be created and applied before and during delineation processing. However it is better to remove non-tree area before, because it reduces complexity and computing time.

Generally, various vegetation indexes which are calculated using spectral information from optical image have been used for tree mask. Also the incorporation of spectral and spatial variation such as image texture is increasingly used. The role of texture information is to separate the building and vegetation when these objects have similar spectral response [37]. Meanwhile, in contrast to using spectral and spatial information in grid data, segmentation of non-grid data, LIDAR points cloud, has been developed to separate objects by means of grouping the neighbouring points having common characteristics [15].

The nDSM to be used in this study has only height information. Tree mask generation is a challenge in itself using only height information. Moreover, the problem becomes more complex when the height of buildings and nearby trees is almost the same. There are still possibilities to discriminate trees and buildings based on grid data such as rasterized nDSM image and non-grid data such as raw LIDAR points cloud. In order to see these possibilities to separate trees and buildings in this study area, we used two methods: Anisotropic texture height measurements [40] using co-occurrence matrix and surface growing using slope based filtering [41], respectively.

Texture is an important characteristic for the segmentation of an image. In a laser scanner image texture is given by local variation of height and height derivatives. The use of texture measures is able to discriminate between buildings and trees. According to anisotropic texture measurement [40], in general, buildings will show a regular, smooth pattern with small variations in height, while trees show an irregular pattern with larger height variations. The knowledge that the orientation and shape of buildings result in orientated height derivatives allows for the use of the contrast texture measure to separate buildings and trees. Trees are supposed to have contrast in horizontal, diagonal, as well in vertical directions. Buildings will show contrast at the edges in only one direction. Therefore, one can use an anisotropic operation to discriminate between orientated and non-orientated features. The minimum of the contrast measures for all directions will have a high value at trees and will show low values at buildings, except for corner pixels.

Meanwhile, a slope based filtering algorithm is developed for the production of Digital Elevation Models by means of separating terrain and off-terrain. The idea is based on the observation that it is unlikely that a large difference in height between two nearby points is caused by a steep slope in the terrain [41]. Similarly, the height change of buildings (especially roof) is unlike to that of a tree crown. The roof of building has regular patterns in terms of height change and angles, but tree crown doesn't. In short, both methods can discriminate building and tree crown based on their different characteristics in height change.

3.3.4. Individual tree detection and crown delineation

In order to represent distinct tree crown boundary, small grid size (25cm), so-called "high resolution", rasterized DSM, DTM, nDSM are used. In other words, it is possible to detect and delineate individual tree crowns more easily in high resolution than in low resolution. However, sometimes due to high resolution, the use of typical tree detection and delineation that normally treats height values at pixel level becomes difficult since the grid size become smaller and the complexity is increasing. A new method that is more flexible and controlled should be developed to make use of rasterized nDSM in high resolution.

A new algorithm called "the fractal net evolution approach", which is a method to describe complex semantics within a largely self-constructing and dynamic network, has apparently achieved this requirement by defining homogeneity in combination with local and global optimization techniques [43]. In this algorithm, a scale parameter is used to control the average image object size and homogeneity criteria are defined by the spectral and spatial information. Moreover, the 'Multi resolution Segmentation' allows to extract image objects in variable resolutions. In order words, it allows homogenous image object extraction in any desired resolution.

Detection and delineation of individual tree crowns is implemented in eCognition software. In the eCognition software, objects are extracted from the image in a number of hierarchical segmentation levels. Each subsequent level yields image objects of a larger average size by combining objects from a level below [47]. This entails the representation of image information on different scales simultaneously. Objects (or pixels at the first level) are grouped into a larger object based on spectral similarity, contrast with neighbouring objects, and shape characteristics of the resulting object. These three characteristics are grouped into a single image parameter called heterogeneity [43]. These can provide a promising solution for individual tree crown delineation using single LIDAR data.

3.4. Tree parameter extraction

3.4.1. Tree height information and crown diameter

Individual tree height information and crown diameter can be extracted during crown delineation process which is applied here in a stepwise manner. In eCognition software, these parameters can be extracted from an image object directly. For instance, the maximum pixel value of object corresponds to total height of tree which is defined as height between ground and tip of tree. As far as crown diameter is concerned, in order to minimize error from diverse measurement methods, we decided to use the same way to extract crown diameter from field and LIDAR derived measurement. To do this, crown diameter is measured as an average of maximum and minimum diameter in the field measurement. From image object, this can be done by means of extract maximum and minimum of X- and Y- coordinates.

3.4.2. Determination of crown shape

Crown height and diameter are variables to be used directly in crown volume calculation formula. Meanwhile, crown shape is a crucial factor strongly related to selection of its formula. Most crown volume estimation models have consolidated all their variation in tree crowns by using calculations for solid geometric object [46]. One single geometric form can not be applied to describe the crown shape of all different type of trees because of variety and irregularity of crown shape. However, if a side view shape and a bottom shape are estimated, the resultant geometric solid can be employed to calculate crown volume.

Appropriate selection of solid geometric object for crown shape has still been a subjective issue. With respect to determination of tree crown shape, the critical issue is how to make it objective. As mentioned above, it seems that it is impossible to directly calculate crown volume without solid geometric formula simplifying the irregular tree crown shape. To do this, it is necessary to pre-define solid geometric shapes which can represent various tree crown shapes. For this, idealized solid geometric shape models [46] are used in this study. We assume that side of crown shape is various but bottom shape is restricted into circle. Next, the side shape of trees in the study area should be classified into pre-defined solid geometric shapes model (figure 3.9). Since visual interpretation by different observers cannot provide a constant result, we should find out more objective indicator to determine the crown shape. Tree species or numerical calculation result could be considered as an objective indicator. For tree species, there is an assumption behind it: One species has one typical shape. However, it is difficult to classify tree species by use of single LIDAR data. Moreover, even same tree species can have various crown shapes based on age, variety and circumstance. Therefore, the use of a numerical calculation result which can be retrieved directly from LIDAR data would be more reliable since this is extracted from current tree situation.

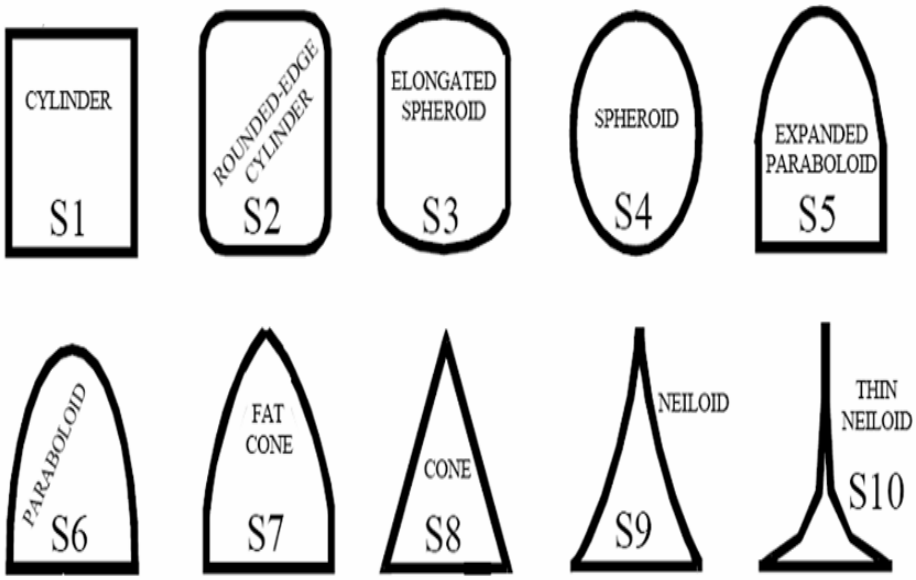


Figure 3.9. Example two-dimensional side view of idealized crown shapes with crown shape factor number and generic name. (source from [46], then modified the sequence of figures)

3.5. Tree crown volume estimation

In the end, tree crown volume estimation model is developed based on solid geometric volume formula using multiplier based on different crown shapes. As shown in figure 3.10, the predicted crown height and crown diameter by regression are used. Therefore, the selection of tree crown volume calculation formula and statistical analysis such as correlation and regression are essential. The main objective of correlation and regression analysis in this study is to see the degree of association between variables and to predict crown diameter and height variables to be used for solid geometric volume formula.

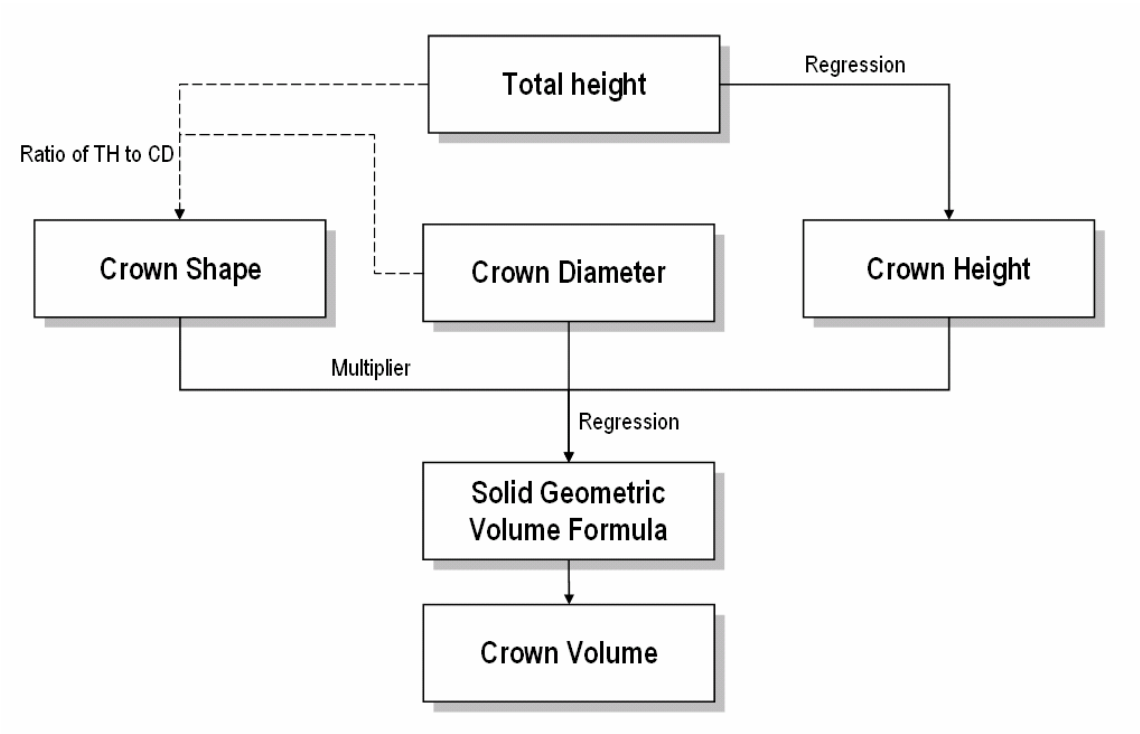


Figure 3.10. Work flow for tree crown volume estimation

3.5.1. Tree crown volume calculation for different crown shape

As mentioned above, estimation of tree crown volume in this study is based on crown height and diameter. These variables are used directly into solid geometric object volume calculation formula. One researcher [46] had designed tree crown volume estimation for different crown shape models using pre-calculated crown shape value, which is a formula multiplier where a right cylinder is 1.0 and the rest of the shape formulas are some fraction of a right cylinder's volume, and two variables. The advantage of using formula multiplier based on ratio of fraction from a right cylinder's volume is that it allows a more simple calculation of complicated shapes which requires tedious and cumbersome work. This model is summarized into table 3.1. All shape are found along a calculation gradient from S1 (multiplier 0.7854) to S10 (multiplier 0.0982).

In the end, tree crown volume is estimated by means of this formula. Variables of crown diameter (CD) and crown height (CH) within this formula will be extracted from LIDAR data. In practice, these variables will be predicted based on LIDAR derived measurement. To do this, we assume that values from field measurement are actual variables to be used for comparison and validation to the predicted variables from LIDAR derived measurement.

Table 3.1. Tree crown volume estimation for different crown shape models (source from [46])

Shape number	Shape Value	Shape Formula	Shape Name
S1	8/8 (1.0)	$(CD)^2 * (CH) * (0.7854)$	Cylinder
S2	7/8 (0.875)	$(CD)^2 * (CH) * (0.6872)$	Rounded-edge cylinder
S3	3/4 (0.75)	$(CD)^2 * (CH) * (0.5891)$	Elongated spheroid
S4	2/3 (0.667)	$(CD)^2 * (CH) * (0.5236)$	Spheroid
S5	5/8 (0.625)	$(CD)^2 * (CH) * (0.4909)$	Expanded parabolic
S6	1/2 (0.5)	$(CD)^2 * (CH) * (0.3927)$	Parabolic
S7	3/8 (0.375)	$(CD)^2 * (CH) * (0.2945)$	Fat cone
S8	1/3 (0.333)	$(CD)^2 * (CH) * (0.2619)$	Cone
S9	1/4 (0.25)	$(CD)^2 * (CH) * (0.1964)$	Neiloid
S10	1/8 (0.125)	$(CD)^2 * (CH) * (0.0982)$	Thin Neiloid

CD: Crown Diameter, CH: Crown Height

3.5.2. Correlation and regression analysis

Correlation and regression analysis are useful in evaluating the association between two or more variables and expressing the nature of relationship. Even if the two kinds of analyses have much in common, there is an obvious difference in their purposes [49].

Correlation determines the degree of association between two variables *X, Y* or more. It expresses relationship between two or more variables to see how closely they are associated. The measure of the degree of association is called the correlation coefficient (r) which has no unit of measurement but just a number [49]. Values of correlation coefficient (r) between 0 to ±1 indicate something between none and perfect correlation, with increasing intensity of association as the correlation coefficient approaches 1 For instance, if positive (+) correlation exists between X and Y, higher values of X tend to result in higher value of Y or low values in X tend to result in lower value of Y. In contrast to positive correlation, if negative correlation exists, higher values of X tend to result in lower values of Y or lower values in X tend to result in higher values of Y. However, correction coefficient (r) measures only linear relationship, it doesn't describe curved relationships between variables. Moreover, in general it is not easy to guess the value of correlation coefficient (r) from the appearance of a scatter-plot because of plotting scales. And correlation is not resistant, so outliers can greatly change the value of correlation coefficient(r) [50].

On the other hand, regression analysis quantifies the relationship between a dependent variable and one or more independent variables [49]. In order words, regression predicts future observations based on the relationship between the variables. If we want to estimate the value of Y from a given value of

X, we would use a regression curve of Y on X, which amounts to interchanging the variables in the scatter diagram so the Y is the dependent variable and X is the independent variable [50]. The quantitative relationship is expressed by an equation and its graphic representation. The simplest form is a linear relationship, $Y = a + bX$. In this regression equation, "a" is the intercept, the value of Y when $X = 0$ and "b" is the slope called the regression coefficient which is the amount by which Y changes when X increase by one unit. The procedure of fitting a regression line is called the method of least squares which gives the best possible fit for the type of the relation chosen [48]. However, there are many methods developed to fit regression line based on the relationship between variables (i.e. linear, curvilinear, simple, multiple regression, etc).

3.6. Validation

In order to evaluate the accuracy of the results, coefficient of determination (r^2), root mean square error (RMSE) and statistical test (T-test and Wilcoxon signed rank test) is conducted. They are commonly used to validate the results but they have specific objectives (table 3.2.). General purpose of validation in this study is to evaluate the accuracy of the tree parameter extraction based on rasterized LIDAR image and object-oriented approach, the reliability of crown volume estimation model.

Although coefficient of determination (r^2) and RMSE have been frequently used for validation, we think that they are not enough to evaluate the accuracy of LIDAR derived measurement and the reliability of crown volume estimation in this study. Generally RMSE is a good estimator when we compare the accuracy or performance of different measurement. For instance, if total height is measured by different equipments A and B, simply RMSE could be an adequate indicator to make the conclusion in such a way, "A measured more accurate than B." However, it doesn't make sense if we compare the accuracy of total height and crown diameter each other by RMSE, since they are completely different subjects for the measurement. Moreover, there is no standard allowable error index to determine the accuracy of total height and crown diameter individually. While coefficient of determination (r^2) cannot estimate directly the accuracy of field and LIDAR derived measurement. Therefore, we additionally used the significance test for evaluation itself.

Table 3.2. Overview of validation methods to be used

Methods	Description	Objective
Coefficient of Determination (r^2)	Squared value of the correlation coefficient	To determine the degree of association between variables.
Root Mean Square Error (RMSE)	Square root of the MSE	To determine the amount of error from LIDAR derived measurement.
T-test	Parametric analysis by difference of mean	To determine if the difference between field measurement and LIDAR derived measurements is statistically significant.
Wilcoxon Signed Ranks Test	Non-parametric analysis by difference of media	To determine if the difference of between field measurement and LIDAR derived measurements is statistically significant.

Correlation and regression are closely connected. The squared of the correlation (r^2), is the fraction of variance of one variable that is explained by least-squares regression on other variable. However, it should be interpreted carefully since the higher coefficient of determination (r^2), does not always show linearity due to outliers. Moreover a correlation can be taken as evidence for a cause and effect relationship in which a change in the value of an independent variable will result in an expected average change in the dependent variable, but cannot indicate precisely what the causal relationship might be [49].

$$r^2 = \frac{\text{Variance of predicted values}}{\text{Variance of observed (actual) values}}$$

RMSE is a frequently-used measure of the difference between values predicted by a model or an estimator and the values actually observed from the thing being modelled or estimated. The error is the amount by which the estimator differs from the quantity to be estimated [53]. The difference occurs because the estimator does not account for information that could produce a more accurate estimate.

$$\boldsymbol{RMSE} = \left(\sum_{i=1}^{n} (e_{1i} - e_{2i})^2 / (n - 1) \right)^{0.5}$$

Where e_{1i} is the result obtained from LIDAR derived measurement for i-th tree, e_{2i} is the corresponding field measurement value and n is the number of trees.

Finally, in order to evaluate crown volume estimation model, statistical significance tests are used. T-test [54] is frequently used for evaluating the probability of occurrence of the difference between two means from data that follow a normal distribution. The Wilcoxon signed rank test [57] is an alternative to the T-test. The difference between them is that Wilcoxon signed rank test evaluates the difference between two medians from data follow a non-normal distribution. For the parametric test, variables of crown volume are transformed so that their distribution can be more nearly normal because variable of crown volume doesn't follow the normal distribution. Logarithm transformation was carried out.

The hypothesis which is used in all tests of significance is the null hypothesis. The null hypothesis is an assumption that there should be no difference between an observation and an expected result [55]. The alternative hypothesis is usually that some suspected cause, which may have stimulated the investigation, is responsible for the difference observation. Consequently, this significance test is to provide objective means for deciding whether or not apparent difference in measurements are likely to have occurred under some hypothesis [49]. Based on this, we can judge the difference between field

and LIDAR derived measurement in terms of statistically not significant, significant and highly significant. After the test, we can compare the p-value with a fixed value that we regard as decisive. The decisive value of p is called the significant level. It is denoted by α, the Greek letter alpha [55]. If p-value is less than or equal to α, we can make a conclusion that alternative hypothesis is true. It should be remarked that the term 'significant' in the statistical sense doesn't mean 'important.' It has to be treated within the sense of probability, for example if we $\alpha=0.05$, we are requiring that the data give evidence against null hypothesis (H_o, there is no difference) so strong that it would happen no more than 5% of the time (1 time in 20) that the null hypothesis is rejected while it is true [55].

4. Analysis and results

4.1. Field measurement and visual investigation

We sampled tree parameters in field so that we can validate the results derived from LIDAR and make use of it as input data to the statistical model as well.

To represent various trees in the study area, we measured trees from single to group and from small to large on the different surface type. Five parameters are measured such as DBH (Diameter at Breast Height), total height, crown height, crown diameter and side shape of tree (figure 4.1). Indeed, among those parameters measured in field, the associations between them and actual values are used to calculate crown volume directly. In order to minimize field measurement error, we measured each parameter three times with same way and collected their average. DBH (Diameter at Breast Height) is not directly related to crown volume calculation. Although it doesn't provide an actual value to calculate crown volume, it is measured to see the association to other parameters. It could be indirect indicator to understand trees in study area.

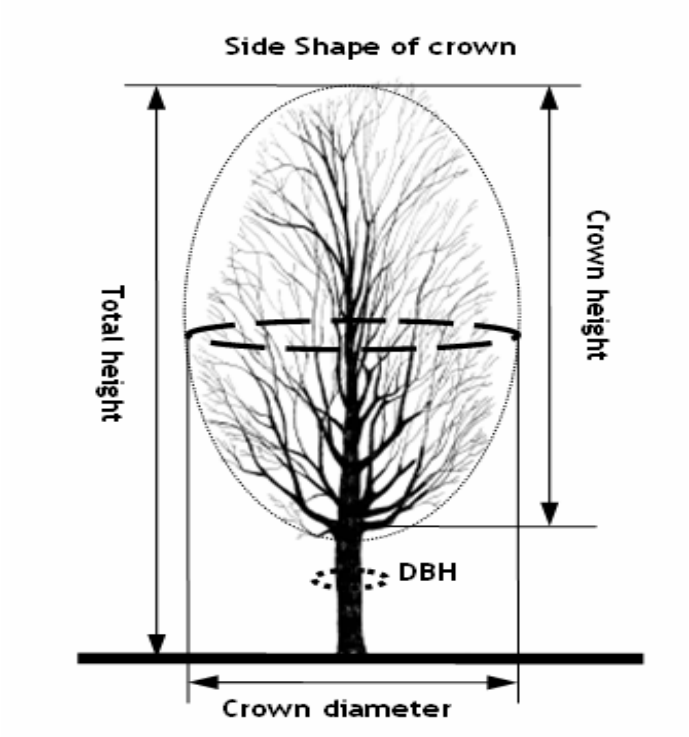


Figure 4.1. Tree parameters to be measured in field

This study area of 25 ha 's-Hertogenbosch, BINNENSTAD OOST contains various trees from small to large and from single to group. Within 25ha, most available situations in which trees are planted in urban area are shown. For instance, we investigate trees that are planted on different characteristic surfaces such as low vegetation (grass), roadway, pavement, parking lots and along a water body.

When it comes to trees planted densely, they are grouped with similar height. For specific reasons, sometimes they are planted very regularly. The dense tree area is however not wide. Almost all trees have rather a flat than a conical apex and are full of leaves. Small trees with low height are planted inside or around the residential area. While large and tall trees can be found around roadway, public buildings, large parking lots and in a small park. Especially, there are a few trees with trimmed branches for beauty and safety around buildings and along the roadway. Furthermore, it is also investigated that two large trees are so much merged that they look as one tree, maybe because of beauty. As conclusion, in comparison to plantation forest, urban trees are planted sparsely and have more heterogeneous characteristics than forest tree in terms of tree species, properties and distributions of small and large trees.

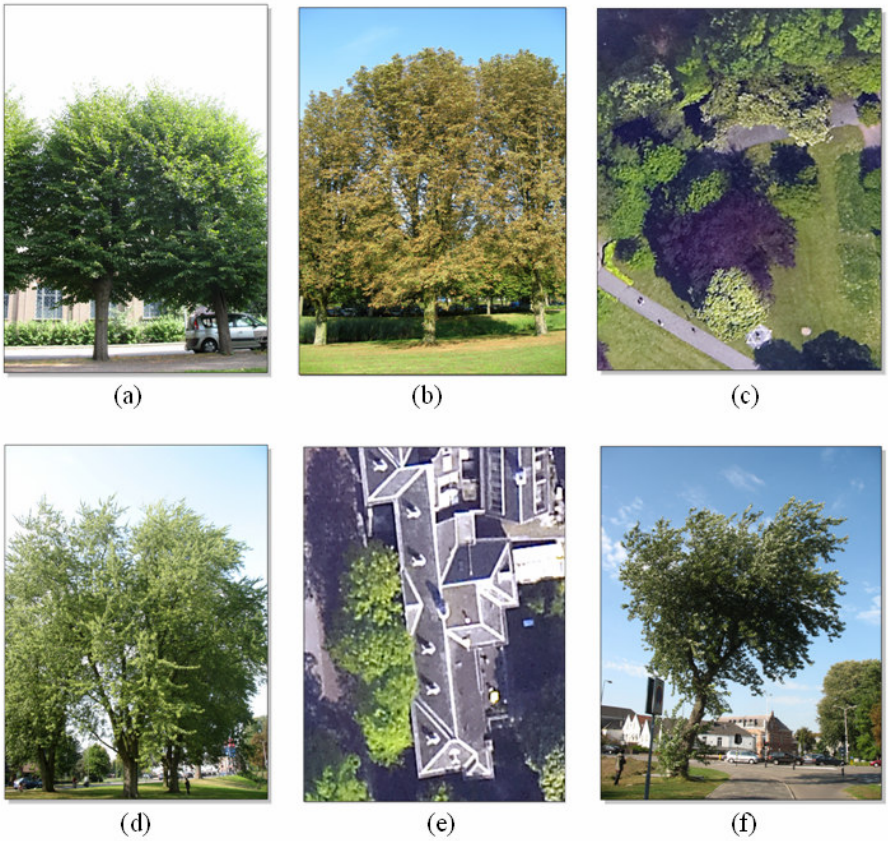


Figure 4.2. Visual investigation of study area
(a),(b) and (c) merged trees, (d) various height distribution in crown surface
(e) trees around building, (f) tree around road

From the visual interpretation of trees in this study area, we can see the challenges that are encountered. Firstly, flat tree apex and merged tree with each other cause the difficulty to detect or to locate individual trees (figure 4.2 (a), (b), and (c)). In addition some single trees have various height distributions on crown surface (figure 4.2 (d)). It gives a high chance to detect several different tree apexes on single tree. Secondly, trees around and close to buildings have similar height with buildings and sometimes cover the buildings (figure 4.2 (e)). It makes discrimination of tree and building more complicated. Finally, the shape of tree is distorted because of beauty or safety (figure 4.2 (f)). It makes the determination of typical tree shape ambiguous.

4.2. Generation of normalized DSM

Morphological filtering has been used frequently and in many applications. In order to remove off-terrain objects, it repeats erosion and dilation operation by turns, so-called "opening" until all objects on the bare terrain are removed. The size of the structural element used for the opening operation is a critical parameter, however, there is no single optimal value. Therefore, the use of multiple openings with different sizes of structural elements was suggested [39]. However, it is still problematic to select appropriate size and shape of structure element. If these parameters are not proper, one cannot expect accurate height information retrieval. Figure 4.3 shows the limit of morphological filtering for DTM generation. To see the limitation of morphological filtering, square and disk is used for structure element shape and the size of structure element is increased. It could not remove the large building and could not provide the detail change of terrain height. It has to be noted that figure 4.3 does not mean that morphological filtering always provides poor results. In this case, dominant part is off-terrain objects but terrain, and off-terrain objects have very complex shapes and various sizes. Because of these, it is difficult to remove every object properly, using morphological filtering.

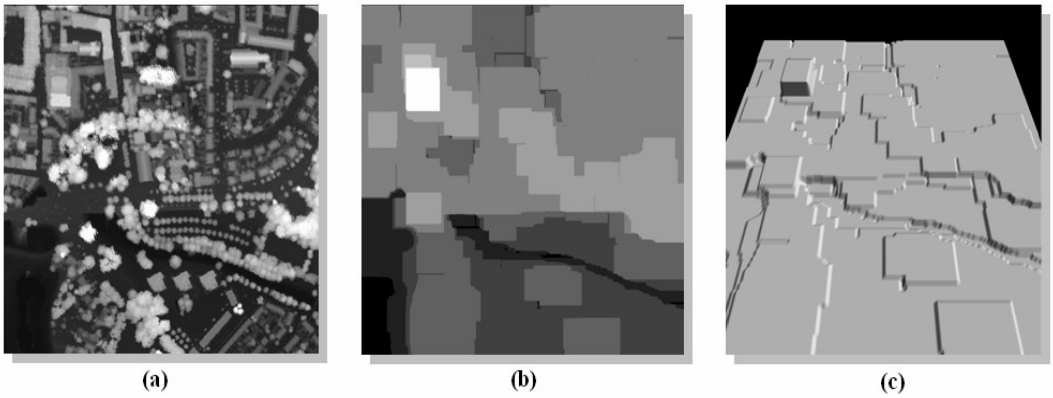


Figure 4.3. DTM generation from DSM using opening morphological filtering
(a) DSM, (b) DTM (2D), (c) DTM (3D)

In order to overcome this limitation, we use an object-oriented image classification technique. Figure 4.4 illustrates the work flow of DTM and nDSM generation. First, DSM is segmented as small as possible (a) then these segments are merged within 0.6m height difference. It creates large plane area that normally represents terrain ((b), red color is one of terrain segments). Finally, we classify DSM into two categories such as terrain and off-terrain using simply mean value and area of segments ((c), blue: off-terrain and green: terrain). Resulting of only terrain class is converted data format into *.xyz for re-interpolating of empty spaces (black in figure 4.4. (d)) which are indicating off-terrain objects. Finally DTM is interpolated (e). The advantage of image classification for DTM generation is that it can provide more simple and controlled operation. Additionally, it can retain much more details.

The final result is illustrated in figure 4.5 and is summarized in table 4.1. It should be noted that the minimum of DSM and DTM is different. All objects segmented into small objects in the DSM rasterized image in prior were merged if their height difference is within 0.6 m in order to detect terrain surface such as road, open area, water body, ground vegetation below 0.6m. Eventually bare terrains that we assumed became larger segments than off-terrain objects such as buildings and trees that could not be merged since they had relatively various height changes and the area of merged

segments such as plane roof, flat tree top was much smaller even if some off-terrain objects were merged. Using this criterion, we classified DSM rasterized image into terrain and off-terrain. But, this algorithm missed some small segments that should be merged into terrain segments during merging processing. By visual investigation, we could check that terrain segments which was not merged into large terrain segments was the part of the small water canal. Because their positions were much lower than road and the area is small, they could not be merged.

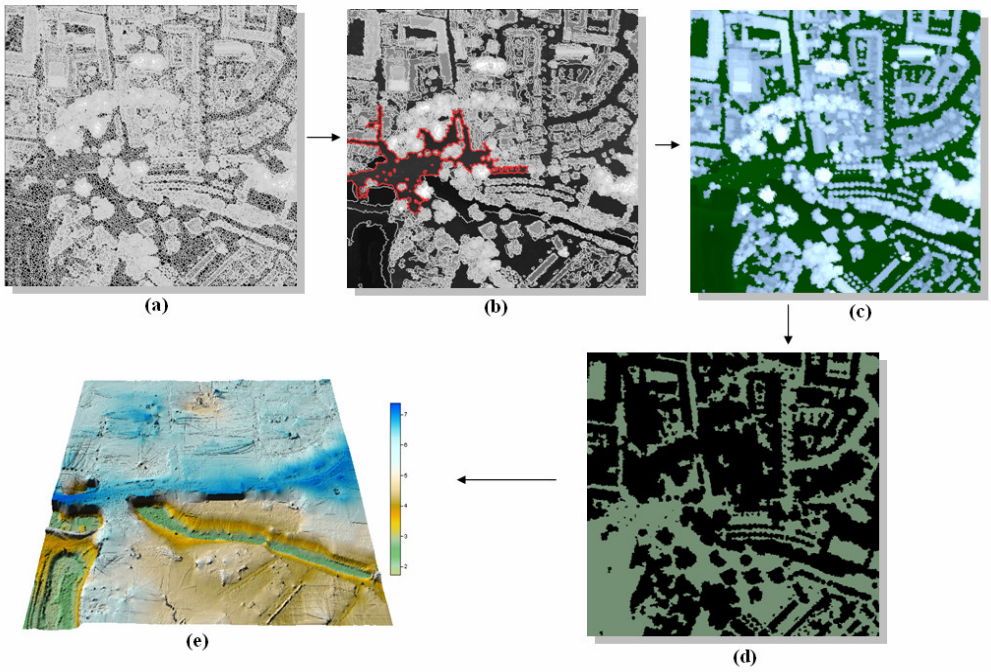


Figure 4.4. DTM generation using object-oriented image classification

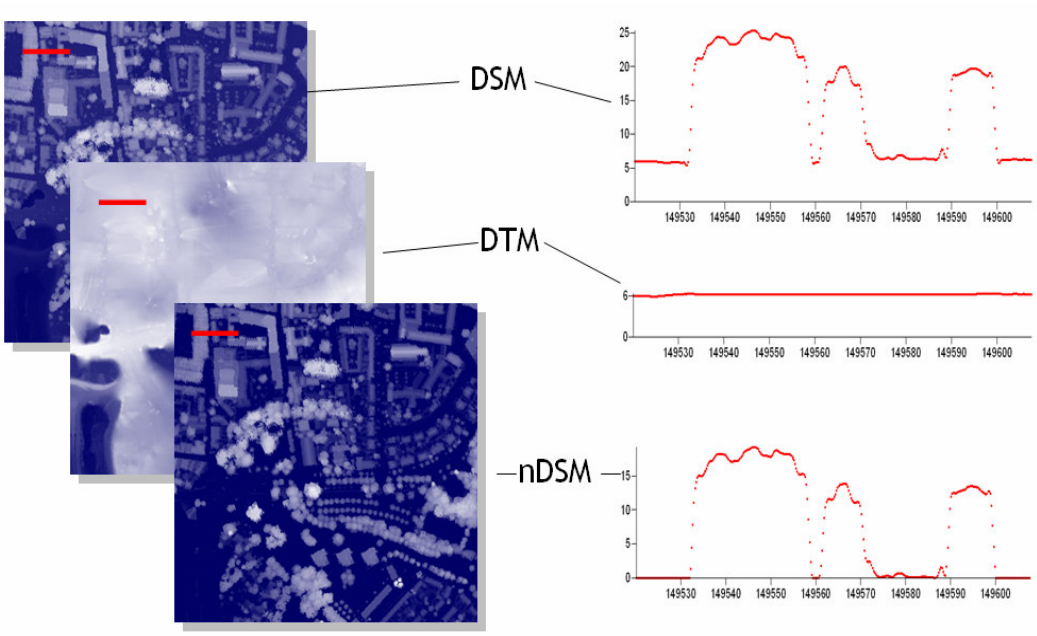


Figure 4.5. Result of DSM, DTM and nDSM generation

Table 4.1. Descriptive statistic of DSM, DTM and nDSM

	DSM	DTM	nDSM
No of values	4000000	4000000	4000000
Minimum	**1.15**	**1.73**	**0**
Maximum	**40.28**	**7.54**	**36.06**
Range	39.13	5.82	36.06
Mean	11.69	5.50	6.33

4.3. Generation of Tree mask

As described in the method section, both texture based and slope based algorithms were tested for separation of trees and buildings. Discrimination of building and tree using texture information on nDSM rasterized image didn't work in this study area since tree and building have a similar texture (compare figure 4.6 (a) and (b)). It is shown in figure 4.6 (c) and (d) that tree cannot have contrast in all directions, therefore it is difficult to separate trees and buildings. This is because trees are full of leaves and have a relatively flat top looking like the roof of building. In particular, the density of raw LIDAR point cloud is not sufficient to represent height changes in detail even if the grid size is small (high resolution). Finally, a surface growing algorithm is applied to discriminate building and tree.

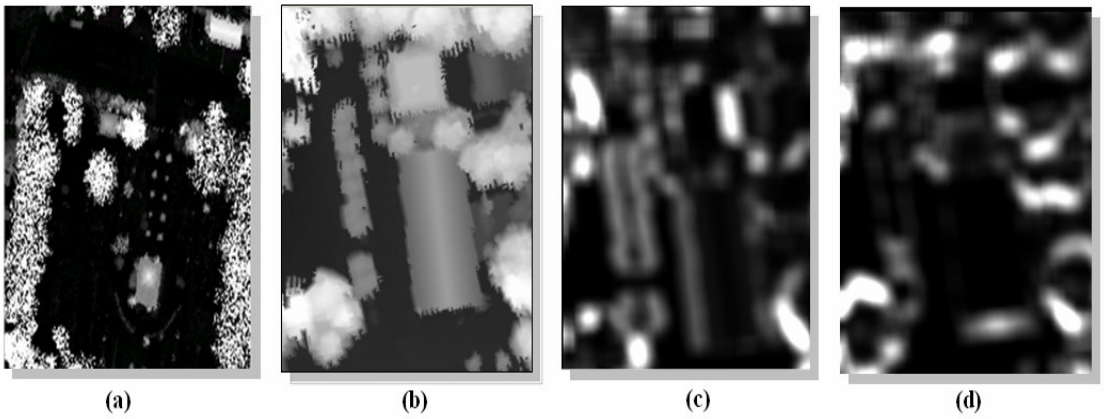


Figure 4.6. Anisotropic texture measurement to discriminate between orientated (buildings) and non-orientated (trees) features in study area.
(a) Example of different textures between building and tree (source from [12])
(b) nDSM of study area, (c) Contrast in vertical of nDSM, (d) Contrast in horizontal of nDSM

In contrast to texture information in the rasterized nDSM image, the raw LIDAR point cloud retains original information of individual xyz points. It allows measurement of distance, height difference and angle between neighbouring points. Figure 4.7 illustrates work flow to generate tree mask. First, the nDSM is converted into *.laser format in order to use this in PCM software which is designed for LIDAR point segmentation applying a slope based filtering algorithm. In order to reduce complexity and computing time, terrain points that have approximately zero values are removed by means of a threshold before importing into PCM. Next, PCM detects plane area by defined parameter such as number of seed points, maximum slope angle, maximum distance to surface and etc.

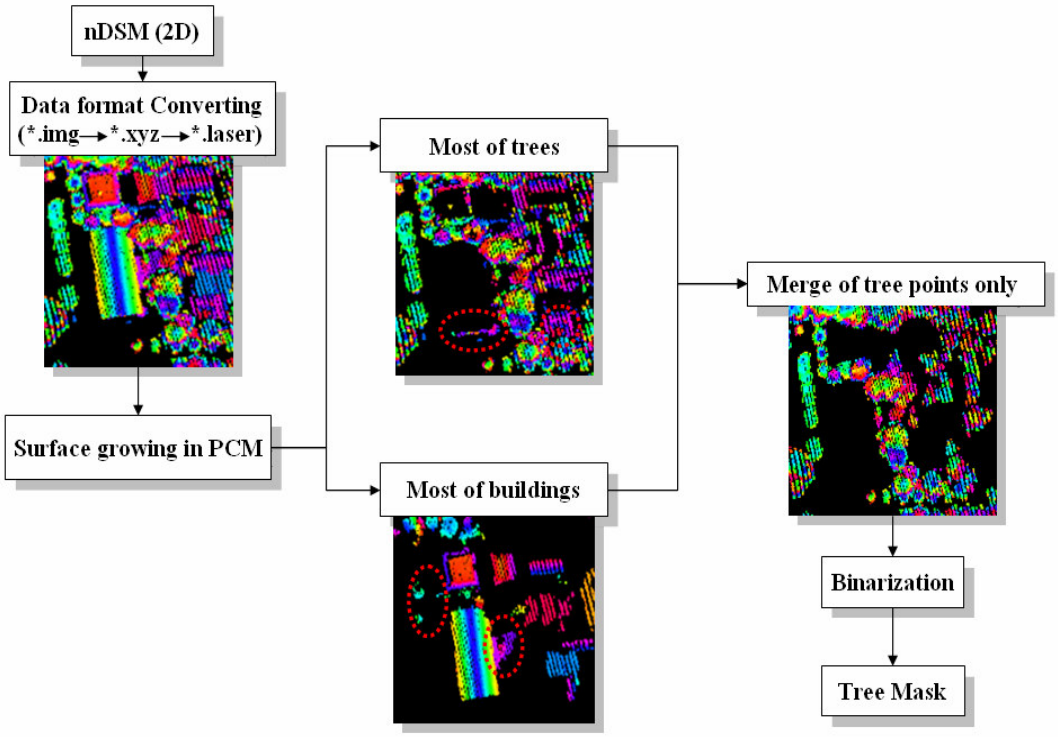


Figure 4.7. Work flow for tree mask generation

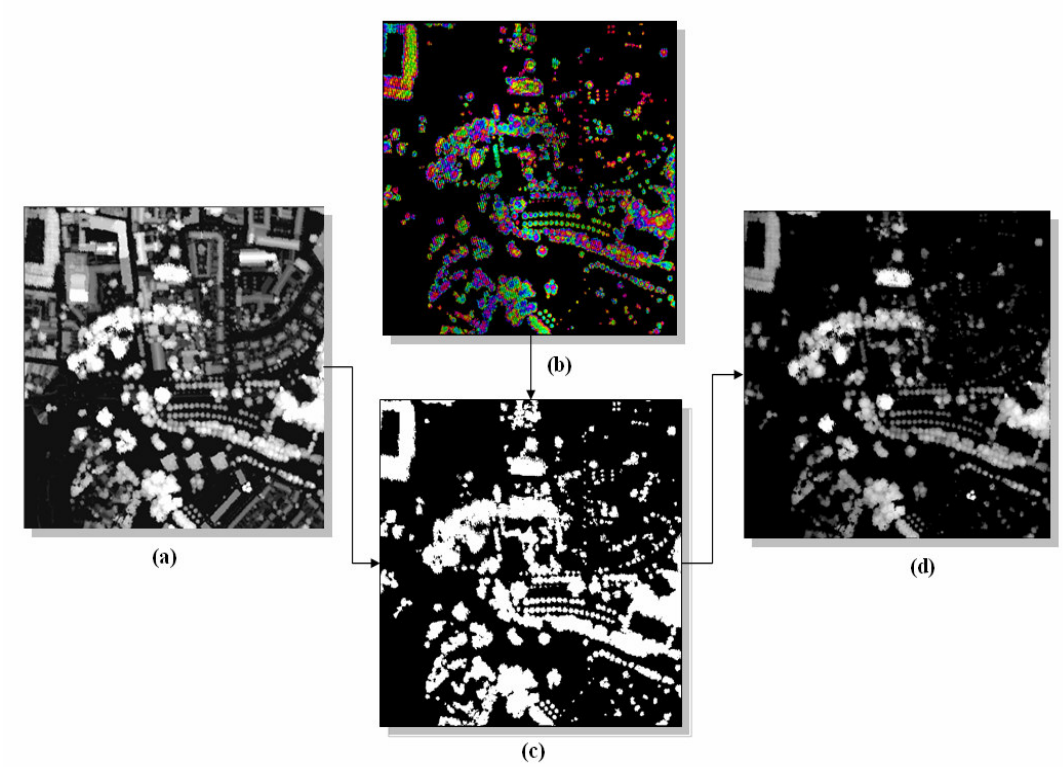


Figure 4.8. Results of tree mask

(a) Rasterized nDSM image (trees and buildings), (b) Collection of only tree LIDAR points in PCM
(c) Binarization (tree: 1, non-tree: 0) (d) nDSM (only trees)

As a result, most of buildings are segmented as plane area, while tree crowns cannot have a plane on it. However, even if most of the buildings have a plane, at the edges of objects, the measured points are noisy. One reason for this is the diameter of the laser beam. A part of the laser beam is reflected at the edge and another part of the laser beam is reflected at another object behind it [42]. Also, some flat top areas of trees are segmented as planes. In conclusion, most of the buildings and most of the trees are separated. However, there are still a few buildings in the tree category and a few trees in the buildings category. Therefore, removing points of buildings from the tree category and collecting points of trees from the building category is done in a manual way. When manual removing was carried out, aerial image of study area was used as a reference in order to minimize the error. Finally only tree points are used to make the binary tree mask. Additionally, figure 4.8 shows final result of tree mask.

4.4. Tree detection and crown delineation

4.4.1. Selection of segmentation parameters

In eCogntion software, the scale parameter is important for segmentation because it determines the maximal allowed heterogeneity of the objects and thus indirectly influences the tree crown size to be extracted. Generally, the larger the scale parameter the larger the objects become. The final decision of scale parameters is often made by an interpreter based on visual inspection of the image. But it is very time-consuming to conduct segmentation with all the possible scale parameters. Thus it is better to investigate maximum area of tree crown based on scale. Figure 4.9 shows the result of crown diameter to be extracted based on different scale parameters. This can be used to determine the optimal scale without cumbersome work.

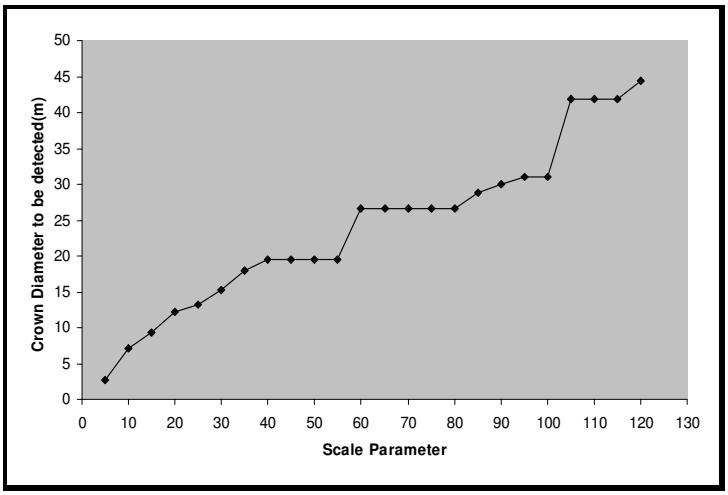


Figure 4.9. Optimal scale parameter selection based on crown diameter

Other significant parameters are color / shape and smoothness/compactness. It is important to note that the term 'color' refers to the digital value of the parameter of color in eCognition, which indicates spectral information in the optical imagery. In a similar way, the color can be interpreted as height information in rasterized LIDAR image since in the rasterized LIDAR image z-value is represented as in the grey tone intensity. With these parameters the influence of color versus shape homogeneity on the object generation can be adjusted. The higher the shape criterion, the less spectral homogeneity influences the object generation. When the shape criterion is larger than 0 the user can determine whether the objects shall become more compact or smoother [44]. For homogeneity, one chooses the relative weight to apply spectral versus shape criteria to reduce heterogeneity. A higher smoothness emphasis would be used to have objects observed to greater variability between features. The compactness weight made it possible to separate objects that had quite different shapes but not necessarily a great deal of color contrast [45]. The structure of segmentation parameter is summarized in figure 4.10.

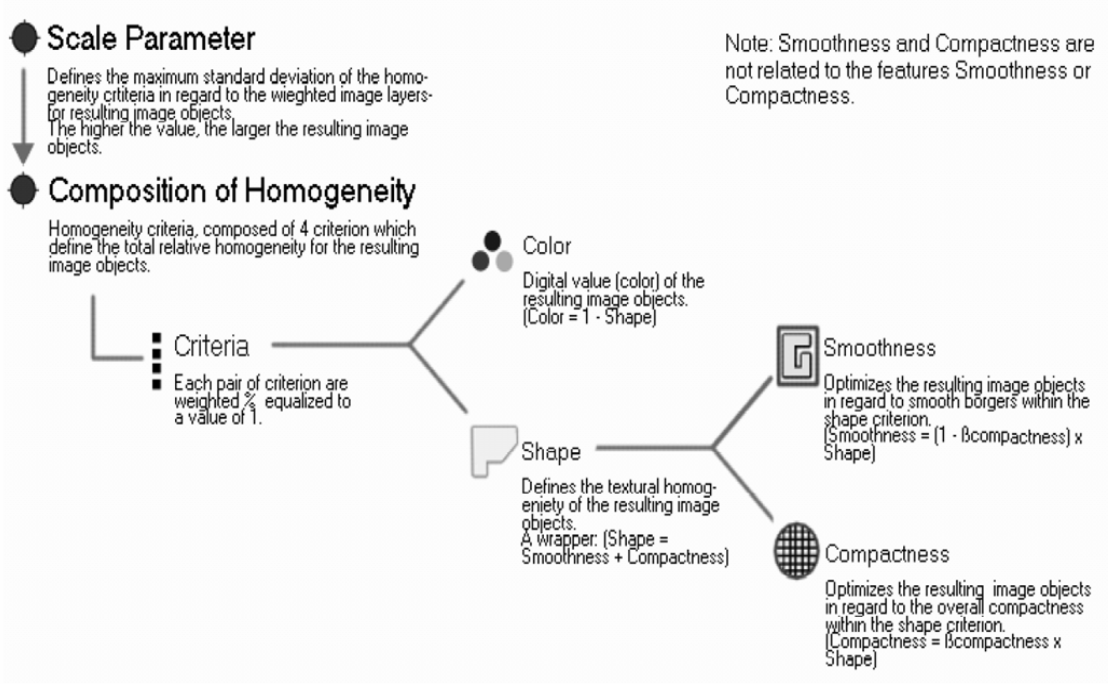


Figure 4.10. The structure of segmentation parameter in eCognition (source from [47])

With consideration to general characteristic of each parameter, Figure 4.11 shows that color, shape, smoothness and compactness criterion are applied in a mixed form to define homogeneity for the tree crown in the study area. From the figure 4.11 can see that the tree crown is delineated like a contour line when the parameter of color and smoothness increase. Contour lines around and inside of tree crowns is not helpful to delineate tree crown correctly.

In summary, individual tree crowns can be delineated as objects through multi-resolution segmentation. These objects are polygons of roughly equal size exhibiting interior homogeneity. In the segmentation process, one specifies the scale parameter based on different crown size by determination of the maximum allowed heterogeneity for resulting objects. In other words, by modifying the value of the scale parameter, we can vary the size of the individual tree crown. Color in rasterized nDSM image represents height variation. Therefore it is better to reduce this parameter for large tree crown which has various height distributions. For the same reason, the larger compactness in the shape parameter makes it possible to delineate large tree crown well.

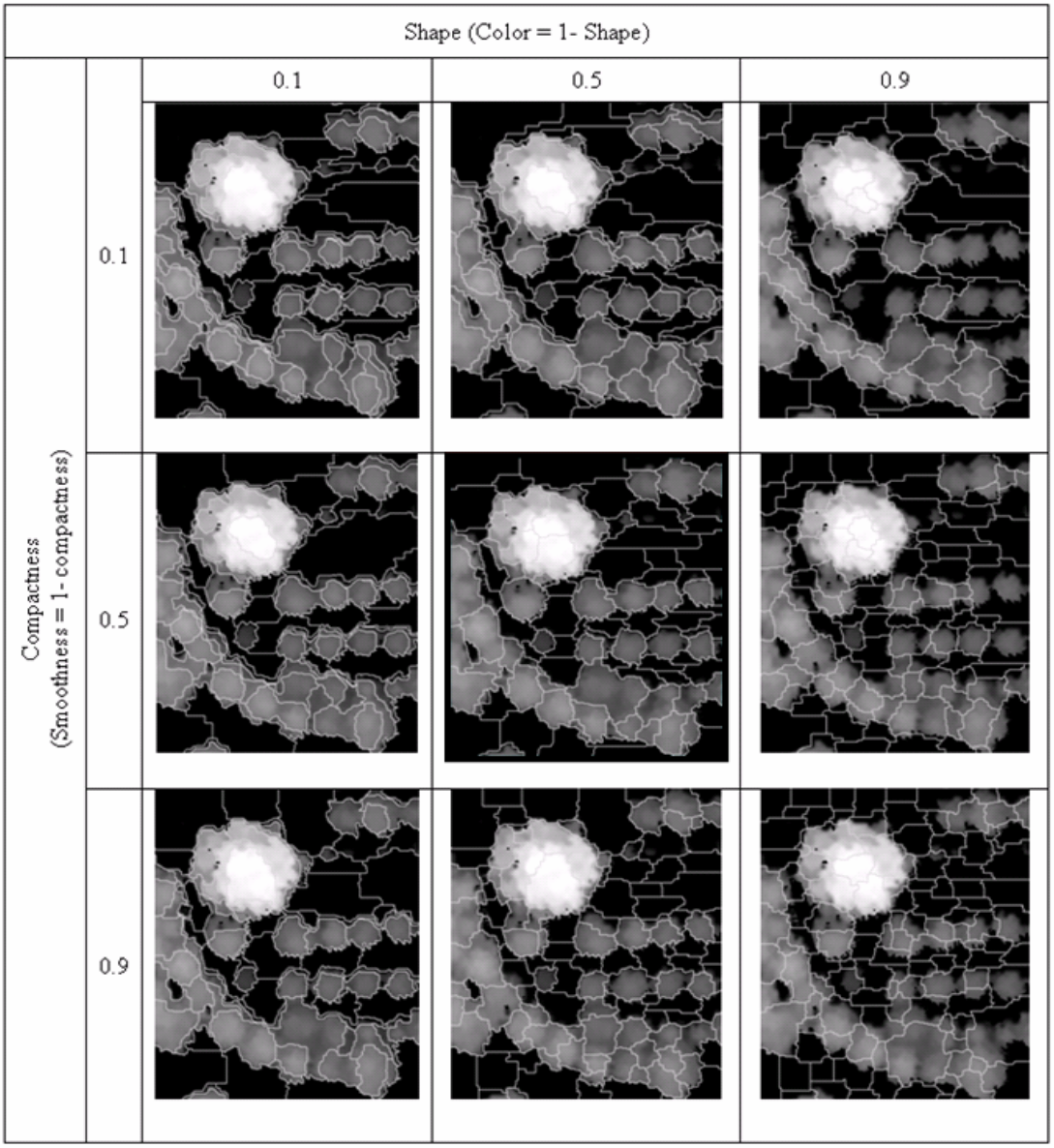


Figure 4.11. Combination of homogeneity criteria (color/shape and smooth/compactness) with fixed scale parameter (25)

4.4.2. Individual tree detection

As emphasized before, for calculation of tree crown volume, individual tree detection and crown delineation processing are essential and have a strong influence on the final result. For tree detection, local maxima are commonly used and many segmentation algorithms are employed for crown delineation. In this study, the typical procedure, tree detection then crown delineation, is followed. The way based on object-oriented approach of tree detection and crown delineation is unique and focuses on improving accuracy and feasibility of operation.

By the tree mask, men made objects are removed from the rasterized LIDAR image in order to reduce complexity and computing time. Thus the rasterized LIDAR image to be classified has two categories such as tree and ground. Then classification is conducted to separate tree and non-tree (ground) area (figure 4.12), large and small trees as well (figure 4.12), initially. To make sure how many segments should be done, tree top area detection is carried out. Based on this evidence and the already existing tree database from the municipality or visual investigation which provides a general view, segmentation for individual tree crown delineation is performed. Next, tree parameters such as crown diameter, height information from well-delineated tree crowns as objects are extracted. Objects extracted are excluded in next step. Finally, identical procedure is repeated until all tree crowns are delineated.

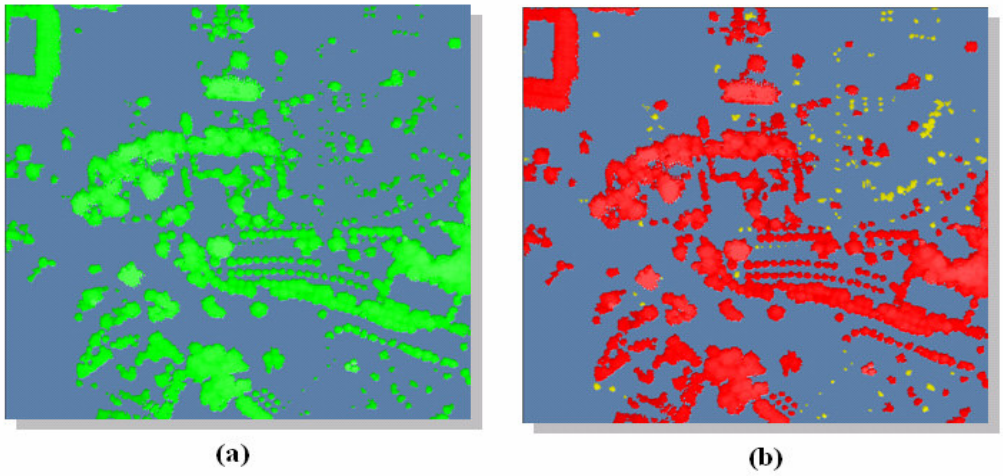


Figure 4.12. Initial classification
(a) Tree (green) and Non-tree (blue), (b) Large tree (red) and Small tree (yellow)

Traditionally, to locate and detect individual trees, the use of finding local maxima is common. Besides locating and detecting trees, local maxima points are used for segmentation as 'seed points' to delineate tree crown. However, it is not feasible to apply this method to flat top trees since tree tops cannot be identified easily. It means many local maxima points indicating tree apex are identified within one tree crown. This phenomenon occurs more frequently in single large trees and flat top trees. For instance, with fixed sized ordinary local maximum filters, increasing the sample window size increased omission error, whereas decreasing the window size increased commission error [12].

Instead of finding local maxima point as tree apex, tree top area as an object is detected because most of the trees in the study area have a top which is more flat than conical which is properly detected as local maxima point by moving window. Unlike traditional algorithms, result of tree detection will not be used directly for tree crown delineation. For instance many researches have delineated individual tree crown based on local maxima point as seed point such as region growing segmentation. However, in this study, the result of tree top area detection, similar to local maxima, provides simply knowledge on how many crowns should be segmented. In other words, it allows to draw virtual guide lines for segmentation by the operator.

To do this, after classification of tree and non-tree area, the tree area is segmented again. Image domain in eCognition allows re-segmentation with different scale on certain class. Tree top area instead of local maxima can be detected by means of iteration of local maxima and local minima by turns. Eventually, the performance of searching the local maxima and minima by terns grows area of higher and lower within tree crowns. Finding local maxima and minima is repeated until the tree crown is filled by them. Experimental work, the sequence of performance and the ratio of iteration is important. Iteration ratio of finding local minima and local maxima which provides the optimal result in case is 2:1or 3:1 when finding local minima is carried out first. The result of these iterations is approximate individual tree crowns (figure 4.13). Although this performance cannot delineate individual tree crown properly, it is good enough to identify the number of crowns to be segmented. Even if small local maxima points are shown around the detected tree top area (figure 4.13, (b) and (c)), unlike searching local maxima points, they don't make identifying the tree top areas confused. Based on this information, individual tree crown will be segmented by means of adjusting parameters of scale, color/shape and smoothness/compactness.

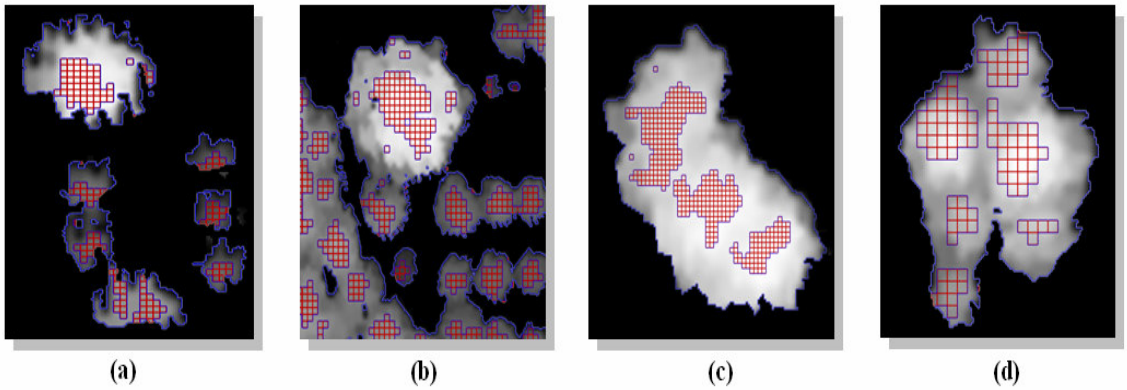


Figure 4.13. Tree top area detection ((a), (b): single trees and (c), (d): group trees)
Red color indicates local maxima as tree top area. For better visualization, local minima area is merged.

Although finding tree top area for locating and detecting individual tree is more promising than using typical local maxima, it is problematic in case of trees that have many sub-crowns (because of the absence of a central stem in the crown) and more than two stems. (figure 4.14) This is very difficult to identify individual tree crowns correctly by an automatic algorithm. Tree detection accuracy has been well researched and is commonly performed at an individual tree level using reference data consisting

of tree locations visually interpreted from the imagery [16]. Even if visual interpretation is available in this case, it still gives wrong information. From figure 4.15, it illustrates one example that shows the limitation of tree top area detection algorithm applying iteration of local maxima and local minima. As mentioned above, over detection of tree top area is more often investigated on trees that have more than two sub-crowns. (a) and (b) in figure 4.15 shows how tree top areas are detected in wrong way. One way to solve this problem is the robust additional knowledge on trees in study area. Once correct number of tree is known, we can segment individual tree crown properly (figure 4.15, (c)) by adjusting segmentation parameters. This additional information can be obtained from tree database already existing in the municipality or from field survey. However, due to the difficulty and associated costs of identifying tree location in the field, we have to rely on the visual observation that could be a subjective investigation by the operator.

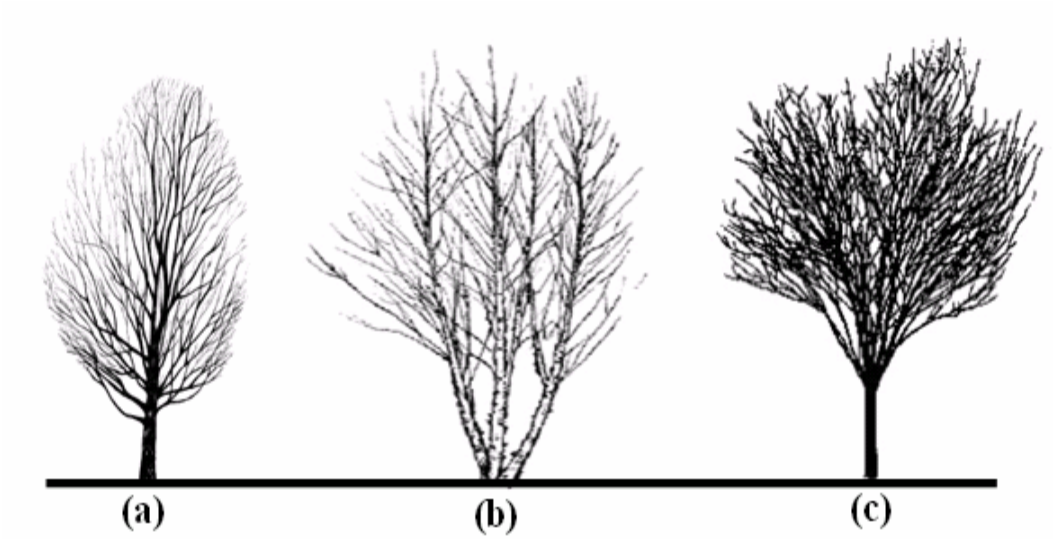


Figure 4.14. Comparison of tree crowns
(a) one crown with one central stem
(b) sub-crowns from different stems in one tree
(c) sub-crowns in one tree (the absence of central stem within crown)

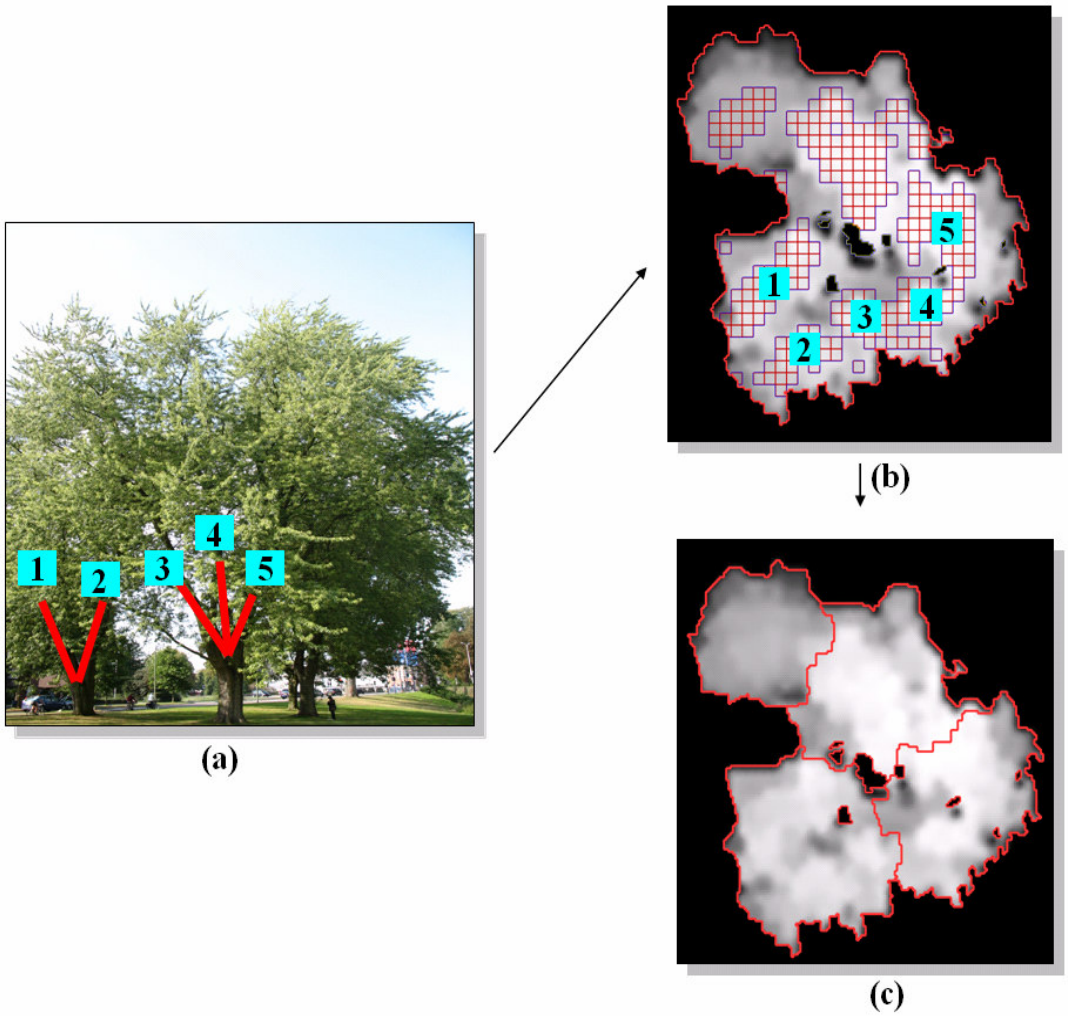


Figure 4.15. Limitation of tree detection
(a) Tees having more than two sub-crowns
(b) Result of tree top area detection in wrong way
(c) Actual tree crowns

4.4.3. Tree crown delineation

As detail, individual tree crown delineation has been carried out based on single and group trees. For single trees, in first phase, to exclude small trees that are not in the scope of this study, nDSM that masked non trees is classified into large and small trees. We assumed trees are higher than 8m can be defined as large tree. Most of single trees are well delineated during initial classification. These single trees are excluded in further step after extracting crown height information and diameter to be used for crown volume estimation (figure 4.16). In case of figure 4.16, it is necessary to segment completely isolated single in further step. However, if single trees are connected slightly, sometimes they are merged with each other in initial classification. Thus it is necessary to re-segment by optimal segmentation parameters based on characteristics of trees to be segmented.

Meanwhile, two categories of group of trees are present in this study area: high and low density. Group of trees are represented as connected objects in the image. To segment individual trees, it is better to apply different segmentation parameters according to different type of group. In the experiment, it is impossible to find and apply the single segmentation parameter globally on the image. Therefore all trees are carefully segmented into individual objects according to the different type of group with different parameters in a stepwise manner (figure 4.17). Based on the result of tree top area detection (figure 4.17 (b)), after first segmentation is performed, two tree crowns within the group are well-delineated. Next, excluding these objects, re-segmentation is carried out based on only the remaining group trees with different scale parameter. These procedures are repeated (figure 4.17 (c)). Multi-resolution segmentation at different scale makes it possible to extract individual tree crowns with high accuracy. In addition, it allows operators to control the segmentation process.

It turned out that the individual tree crown delineation developed based on object-oriented approach doesn't work in case of a group of trees with high density that have almost the same height (figure 4.18). It is even difficult to identify individual tree crowns visually. Those groups of trees look like a big single tree and have an almost plane tree top. It makes it more difficult to keep crown shape or height distribution. It is impossible to delineate individual tree crowns directly. Therefore, in order to extract crown diameter, a regression model based on tree height from detected tree top area is used.

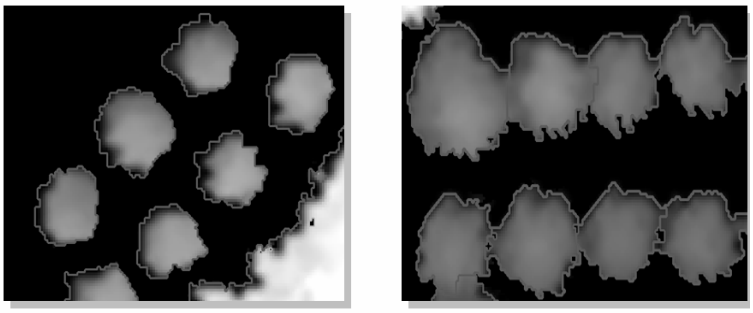


Figure 4.16. Single tree delineation after initial classification

Tree delineation accuracy has not commonly been evaluated because of the difficulty of precisely measuring tree crowns in the field [12]. Field measurements are subject to errors relating to how well field personnel can project the crown boundary and identification of a suitable boundary point to measure for tightly overlapping or irregular crowns. Thus tree crown delineation accuracy in this study will be evaluated indirectly in the end by statistical analysis.

In summary, the new concept of using an image domain in eCognition allows to focus on the pre-classified tree crown. It leads to a performance with possibility of different segmentation algorithms to extract various tree crowns size in one level. In order to detect tree top area, pre-classified tree objects are broken down to pixel-sized objects so-called "chessboard segmentation" in eCognition. Based on the result of this, individual tree crowns can be delineated by use of iteration performance, "segmenting and excluding". However the iteration performance of finding local maxima and minima to detect tree top area for counting the number of tree crowns to be segmented partially failed when trees have more than two main branches looking like individual tree crown in visual and in the image. And problems in dense group of trees having a similar height are investigated for crown delineation process.

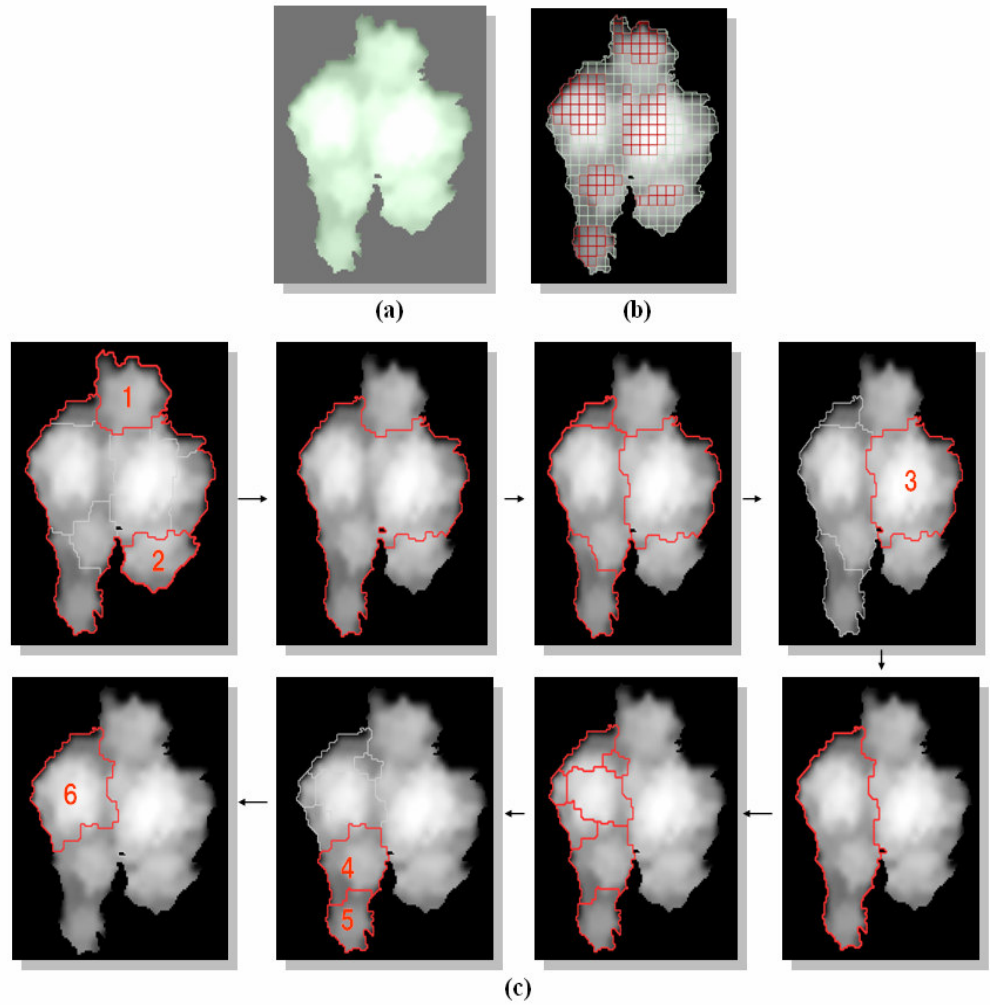


Figure 4.17. Group tree delineation in stepwise manner
(a) Classification of tree and non-tree, (b) Tree top area detection, (c) Segmentation with different scale

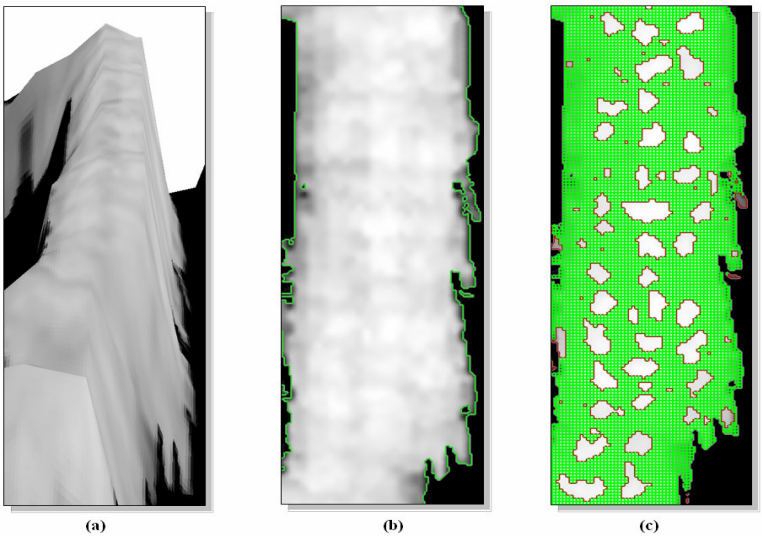


Figure 4.18. Difficulty of tree crown delineation in high dense group trees.
(a) 3D view of high dense group tree, (b) 2D view of high dense group tree, (c) Tree top area detection

4.5. Tree parameter extraction

4.5.1. Tree height information and crown diameter

After tree crown delineation (segmentation), tree height information and crown diameter were extracted from an image object (figure 4.19). The maximum pixel value of object corresponds to total height and As crown diameter is measured as an average of maximum and minimum diameter in the field measurement, the maximum and minimum of X- and Y- coordinates were extracted from image object. All measurements were saved in spread sheet for statistical analysis.

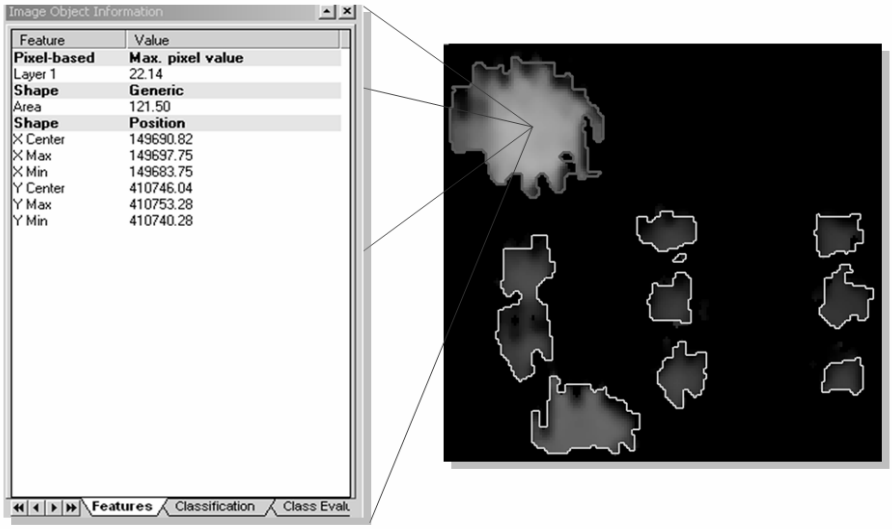


Figure 4.19. Tree Height and Crown Diameter extraction from image object.

- Total height = maximum pixel value
- Crown diameter = average(maxi, mini of X- and Y- coordinates)
- Crown height by regression model of total and crown height

However, we need a crown height but total height for crown volume estimation. Indeed, crown height has so far not been estimated directly from LIDAR data. For calculation of crown volume, two variables, crown height and crown diameter are essential. By extracting minimum pixel value from object, crown height can be substituted. However it is not always guaranteed when crowns of adjacent trees are in dense stands where the crowns interfere and because of various crown shape, for instance, it is very often that laser beam can not hit bottom of spheroid tree crown. Therefore, to induce crown height from LIDAR data, we use regression model based on significant correlation between total height and crown height. This would be more reliable way to extract crown height.

4.5.2. Tree crown shape

In practice, ratio of total height to crown diameter is used as indicator for crown shape determination since total height and crown diameter are unambiguous variables which can be extracted from LIDAR data directly. During field measurement, side of crown shape was also investigated. According to this, the dominant crown shapes (figure 4.20) are spheroid (S4 in figure 3.9) and parabolic (S5, S6 in figure 3.9). For parabolic, it is subdivided into simple (S6, 0 < total height <20) and expanded (S5, 20

<= total height) according to height. It should be noted that figure 3.9 graphically defines idealized tree crown shapes which have the same diameter and height. However, in practice, there would be many different shapes, within the height and diameter. Eventually, these are pre-defined as solid geometric objects in order to represent crown shapes in the study area.

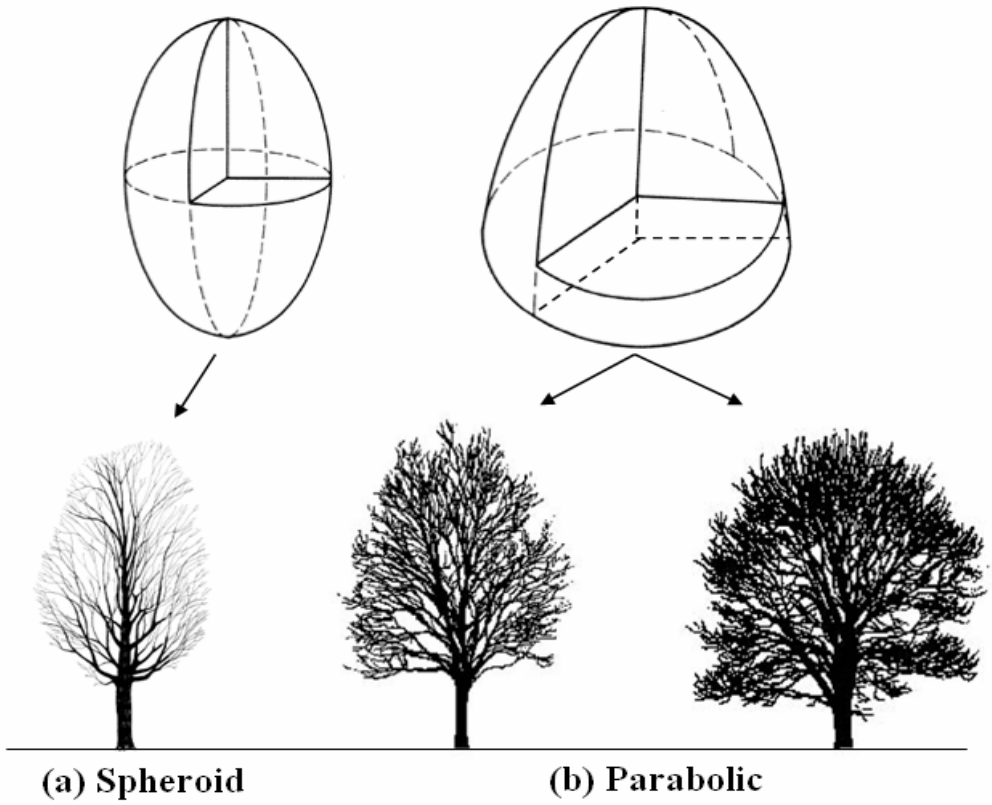


Figure 4.20. Dominant crown shapes investigated in the study area

The ratio of total height to crown diameter is calculated. It is necessary to remark that only values from relatively clear spheroid and parabolic trees are used for ratio calculation to minimize the error by visual investigation of different personnel. Then we investigated the ratio and crown shape in order to see how the ratio can be linked into crown shape. Finally, we found out that each crown shapes had a unique mean value of ratio range based on a certain range of total height (table 4.2.).

In table 4.2., it is shown that the spheroid shape has a rather narrow crown because it has a higher mean value than the parabolic at every range of total height. Like minimum distance classification method, one ratio belonging to a certain range of total height will be subtracted from both mean values. Finally, by comparison, the crown shape can be classified into one solid geometric object. For instance, if there is a tree that has a total height of 17m and a ratio of 1.5, it is likely to have parabolic(S6) shape since the difference from parabolic mean is smaller (1.5 - 1.44 = 0.06) than that from spheroid (1.68 - 1.5 = 0.18). The method of using ratio for determination crown shape is meaningful in terms of reducing errors by different personnel investigation and by applying single geometric object to all trees.

In short, the ratio of total height to crown diameter based on a certain range of total height is very helpful to classify the crown shapes more specific into spheroid, parabolic and expanded parabolic. For instance, if we don't consider the pattern of ratio based on the range of total height, spheroid will have '1.85' mean value of ratio and '1.60' is for expanded parabolic. Therefore spheroid, which is below than 20 m, is likely to classify into parabolic.

Table 4.2. Statistics of ratio (= Total Height / Crown Diameter) for dominant crown shape in study area

Range of Total height	**Spheroid**			**Parabolic**			
	Min	Max	**Mean**	Min	Max	**Mean**	Specification
0 < TH* <10	1.38	1.91	**1.62**	1.33	1.37	**1.34**	Parabolic
10 <= TH < 20	1.36	1.93	**1.68**	1.38	1.52	**1.44**	
20 <= TH	1.38	2.32	**1.86**	1.29	1.98	**1.60**	Expanded parabolic

* TH : Total Height

4.6. Tree crown volume estimation model

4.6.1. Descriptive statistics

Basic descriptive statistic is fundamental to overview/describe the characteristic of variables (in this study, tree parameters from field and LIDAR derived measurement), to justify the correlations between them and to develop the predictable crown volume calculation model. Moreover, all results could be evaluated statistically based on these. It is essential to examine data before statistical analysis in order to avoid misleading results. In other words, reliable results can come from making use of adequate methods based on convincing data.

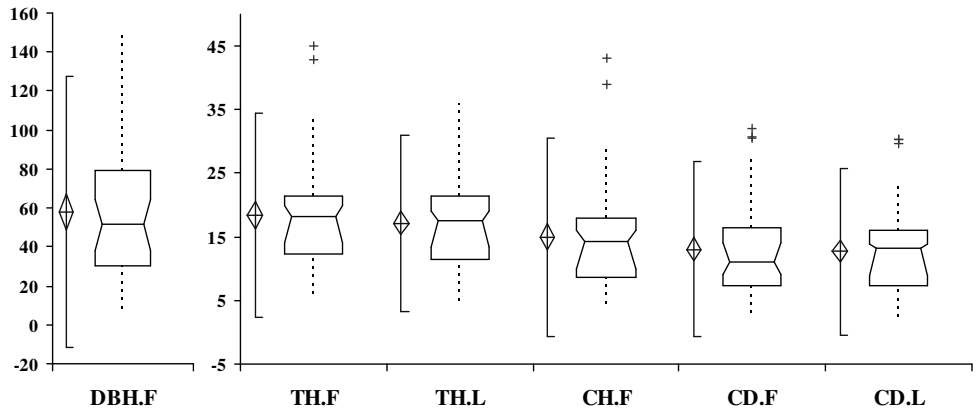


Figure 4.21. Comparative descriptive of variables from field and LIDAR derived measurement

Individual summary statistics is available in appendix A-2.

Abbreviations

- TH (Total height), CH (Crown Height), CD (Crown Diameter), CV (Crown Volume)
- .F (Field measurement), .L (LIDAR derived measurement)

Interpretation of figure (appendix A-1)

- Diamond : Mean and confidence interval of mean
- Notched – lines : Parametric percentile range
- Notched box : Median, lower and upper quartiles and confidence interval of median
- Dotted-lines : The nearest observations within 1.5 IQRs of the lower/upper quartile
- Cross : Possible near outliers, between 1.5 and 3.0 IQRs (inter-quartile ranges) away

In figure 4.21, box-whisker plots graphically show the central tendency and dispersion of the individual observations of tree parameters from both field and LIDAR derived measurement. It is easer to understand and to compare differences between the variables thought the vertical box-whisker plots which are shown side-by-side. Especially, parametric (blue line in figure) and non-parametric (notched box and whiskers in figure) summary statistics are presented for each variable in order to indicate the central tendency and the dispersion of the observations.

From box-plots, red crosses and circles provide very meaningful information. They indicate the near and far outlier, respectively, which is a single observation "far away" from the rest of the data. In order words, they may indicate problems in sampling or data collection or transcription [48]. Because of this, outliers need to be attended carefully. Outliers are investigated from total/crown height (TH/CH) and crown diameter (CD). In case of total height, it can be said that measurement error might occur during total height measurement in the field since the total height measured from LIDAR data doesn't have outliers. The big difference of maximum value (TH.F: 45m and TH.L: 35.85m) between them makes it more acceptable. Especially, the way of height measurement in the field survey tends to overestimate a large (tall) flat tree (figure 4.22). Therefore, it makes sense that outliers of total height could be considered as a measurement error. In contrast to outlier of total height, it is however unreasonable to say that the outlier of crown diameter also may indicate measurement errors. It should be noted that statistics derived from data sets that include outliers will often be misleading [48]. In case of crown diameter, both of field and LIDAR derived measurement have outliers which may be indicative of observations that belong to a different population than the rest of the sample set. And the outliers of crown height are out of question since it is simply derived from total height by subtracting crown live height from ground. It means crown height is likely to have error from total height inherently.

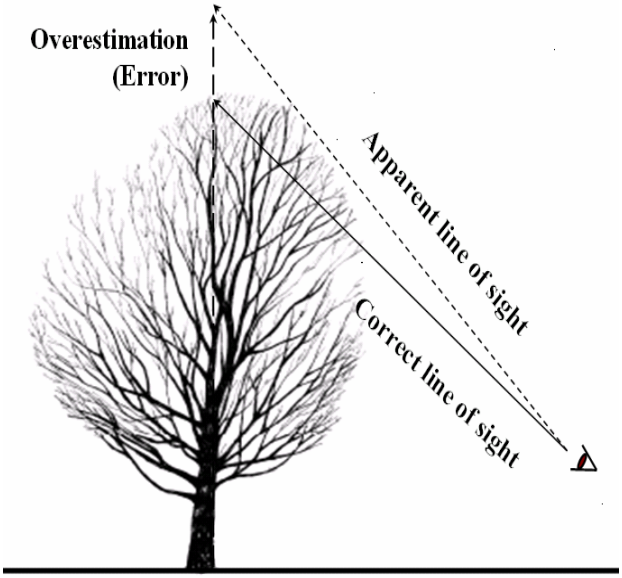


Figure 4.22. Overestimation of Height measurement

In order to determine whether statistical analysis method is optimal, it must be useful to find out if variable to be analyzed is normally distributed [52]. Normality of individual variables is tested on Shapiro-Wilk and Kolmogorov-Smirnov test. Corresponding to the results, if Shapioro-Wilk statistic is too small, distribution is likely to be non-normal and if Kolmogorov-Smirnov statistic is less than critical value of 90, 95, or 99%, it has a normal frequency distribution. In addition, kurtosis and skewness which is a measure of the peak and the symmetry of a distribution, respectively, can help describe the normality.

Finally, it turned out that every variable is normally distributed well and there are more observations in the left-tail (skew > 0) (table 4.3). Generally, variables of LIDAR derived measurement tend to follow normal distribution rather than variables of field measurement. This result could be the outline to determine optimal statistical techniques such as parametric and non-parametric, or to find out alternatives if assumptions of the statistical test used are not satisfied.

Table 4.3. Normality test of Shapiro-Wilk and Kolmogorov-Smirnov

Statistics	DBH.F	TH.F	TH.L	CH.F	CD.F	CD.L
Shapiro-Wilk stat.	0.9284	0.9209	0.9668	0.8971	0.9029	0.9301
Kolmogorov-Smirnov stat.	0.095	0.122	0.076	0.12	0.136	0.14
Critical K-S stat, (alpha=.10)	0.155	0.155	0.155	0.16	0.155	0.16
Critical K-S stat, (alpha=.05)	0.172	0.172	0.172	0.17	0.172	0.17
Critical K-S stat, (alpha=.01)	0.207	0.207	0.207	0.21	0.207	0.21
Skew	0.841	1.127	0.547	1.33	1.074	0.86
Kurtosis	0.125	1.631	-0.01	2.3	0.619	0.48

- Detailed descriptive with histogram of frequency and normal probability plot is available in appendix A-3.
- Kurtosis interpretation: = 0 is normal, > 0 is with fewer in the tails, < 0 is more observations in the tails.
- Skewness interpretation: = 0 is symmetric, > 0 is more observations in the left-tail, < 0 is more observations in the right-tail.

4.6.2. Correlation analysis

In order to see the degree of association between variables, correlation analysis is carried out. To make the best estimation of correlation, Pearson correlation [56] is used to measure how linearly the variables are related.

If the variables are not associated, there is no linear pattern between variables, they are independent of each other [50]. The degree of association is expressed by the correlation coefficient (r). The results are graphed on the lower left and their correlation coefficients (r) listed on the upper right. Each square in the upper right corresponds to its mirror-scatter plots in the lower left (figure 4.23). As mentioned above, it should be kept in mind that the higher correlation coefficient (r) cannot always be interpreted as high linearity (appendix A-4). In terms of the vulnerable relationship between correlation coefficient (r) and linearity, the mirror- scatter plot corresponding to correlation coefficient (r) is very helpful to identify actual linearity since the type of line in scatter-plots is more informative to determine the regression line.

From scatter-plot and correlation coefficient (r), all variables shows they are related very closely (r= 0.83 ~ 0.99) and their scatter plots illustrate their linearity relationship visually. Especially, total

height and crown height in field measurement (r=0.99), total height from field and LIDAR derived measurement (r=0.96) have almost perfect linear relationship and their scatter plots show also convincing visual evidence. Since crown height information cannot be extracted from LIDAR data directly, it is possible to extract crown height using these relationships. In other words, crown height information from LIDAR data will be predicted based on regression model between total and crown height in the field measurement and total height from LIDAR derived measurement.

Consequently, the result of correlation analysis can help to measure not only the degree of association but also determining the type of function to be used for fitting regression line during regression analysis. In order to avoid the misleading of single correlation coefficient (r), visual inspection of scatter-plot is important.

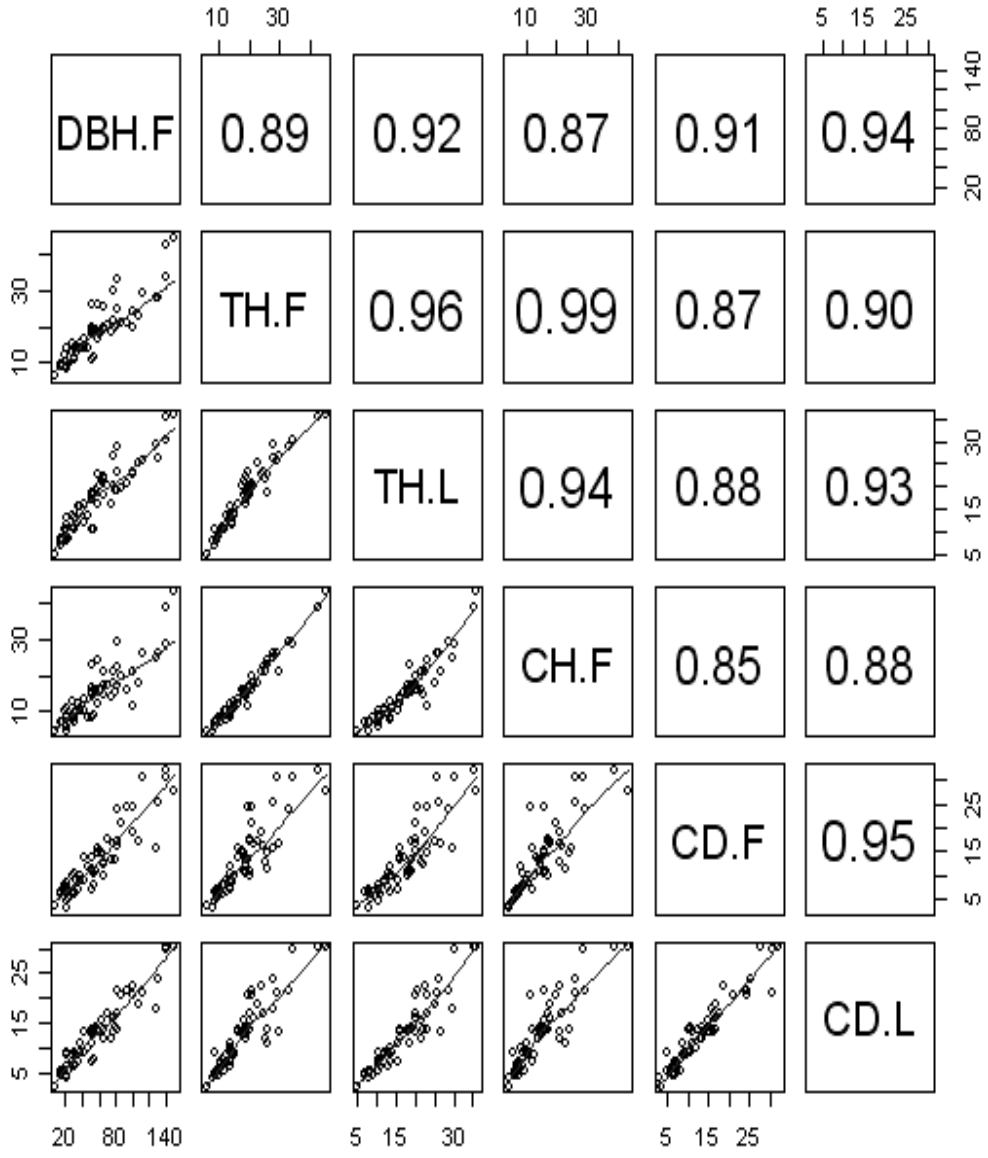


Figure 4.23. Mirror-scatter plot corresponding to correlation coefficient.
Scatter-Plot: Lower Left, Correlation Coefficients (r): Upper Right

4.6.3. Regression analysis

The main objective of regression analysis in this study is to predict crown diameter and height variables to be used for solid geometric volume formula. To do this, regression analysis is carried out between variables from field and LIDAR derived measurement respectively. According to this, crown height information which cannot be extracted directly from rasterized LIDAR image can be obtained. Finally, crown volume, which is calculated based on field measurement, becomes a response variable (also called dependent, explained or predicted variable, named Y). And crown height and crown diameter from LIDAR derived measurement become predictor variables (also called independent, explanatory or control, usually named $X_1, X_2 \ldots$) to predict the Y variable.

Strictly speaking, the values of field measurement are considered as true (actual) values to be compared with predicted values (Y) based on extracted values from rasterized LIDAR image to be used for predicting true values (values of field measurement).

Many functions are available to analyze the relationship between a response variable and 1 or more predictor variables. Although many variations are available, the determination of type of potential regression line, which can make prediction errors as small as possible in sense of least-squares, is the central work. After investigation of scatter-plot and correlation coefficient(r), straight line relationship between field and LIDAR derived measurement is the most convincing.

Simple linear regression is used to validate the results between field and LIDAR derived measurement. Table 4.4 summarizes the result of simple linear regression with regression line plot which illustrates the fitted regression line, with simultaneous confidence and prediction intervals. By examination of observations relative to the regression line, the result of regression can be evaluated visually in various ways. For instance, if the observations deviate widely from fitted regression line, then the line may be a poor line fit since the vertical distance of the observations (points in plot) from the line can be considered as prediction errors. The scatter of observations around regression line should be roughly constant over the range of X. If the observations converge towards the line, then the variance is not constant across the range of X. It could be said that if long runs of observations appear on either side of the fitted of regression line, the relationship may not be linear (straight line). Additionally, if the prediction intervals which is range likely to contain future observation with probability determined by the requested confidence level (95% in this study) is wide, it indicates the regression line does not have good fit, and equation will not accurately predict future values based on X [51].

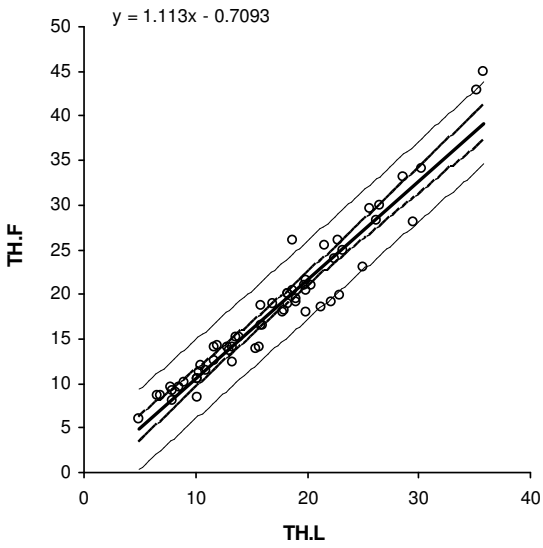
According to the coefficient of determination (r^2=0.88~0.97), the predicted values based on LIDAR derived measurement describe well the value of field measurement which is considered as actual. More specifically, the coefficient of determination (r^2) gives us the variance of the predicted responses as a fraction of the variance of the actual responses [48]. Finally, crown height information is obtained by the regression model between TH.F and CH.F. In the regression equation such as "CH.F=0.9573*TH.F-2.5944", TH.F is substituted by TH.L in order to predict CH.L.

Table 4.4. Statistic for Simple regression between field and LIDAR derived measurement

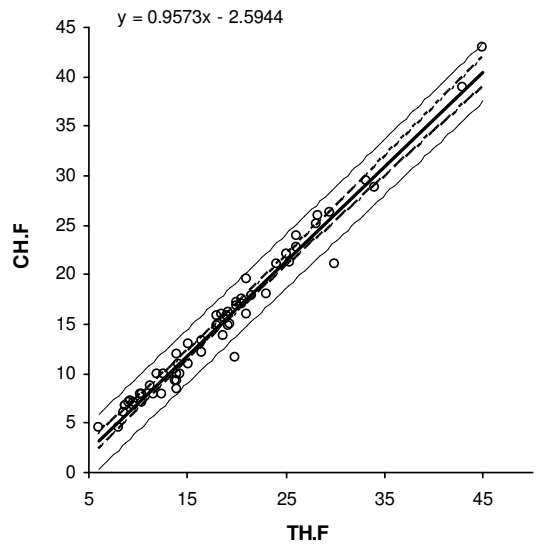
Predicted variable(Y)	Variables(X)	r^2	Regression equation = 0	Graph
TH.F	TH.L	0.93	1.1130 * TH.L - 0.7093 - TH.F	11-1
CH.F	TH.F	0.97	0.9573 * TH.F - 2.5944 - CH.F	11-2
CD.F	CD.L	0.90	0.9998 * CD.L +0.3477 - CD.F	11-3
CH.F	CH.L*	0.88	0.9876 * CH.L + 0.1866- CH.F	11-4

Note :

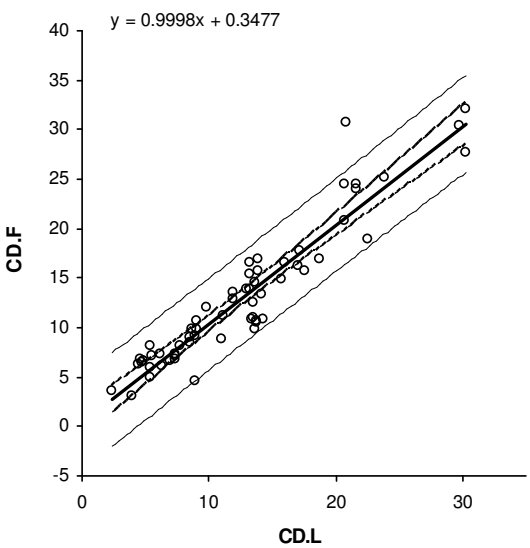
- Regression coefficient table and residual plot are available in appendix A-5.
- CH.L* = 1.0654 * TH.L - 3.2734
- Every regression coefficient is significant (p = <0.000.1)



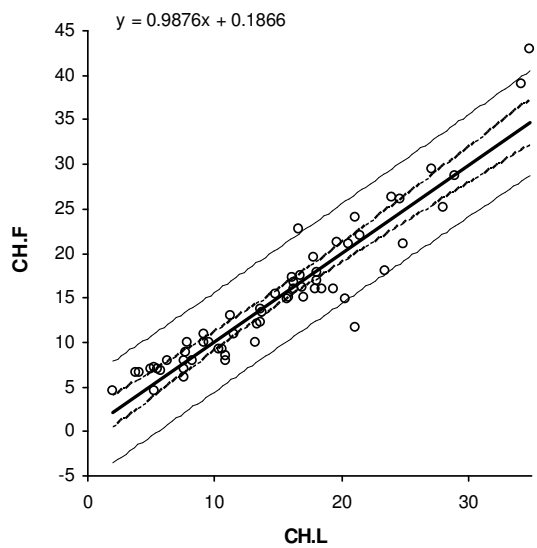
(11-1)



(11-2)



(11-3)



(11-4)

- Center is Regression line, middle is Confidence interval line, exterior is Predicted interval line.
- Confidence interval for a mean response and prediction interval for future observation

4.6.4. Tree crown volume estimation model

In order to develop tree crown volume estimation model at individual tree level, we believe that the best way to calculate crown volume is using a solid geometric volume formula and the same is true when crown volume is estimated actually with values of field measurement. According to regression, crown height and crown diameter to be used as variables for calculating crown volume are available. Based on crown volume estimation model proposed in this study (figure 4.24), crown volume is calculated. Predicted crown height and crown diameter substitute for crown height and crown diameter in predefined solid geometric volume formula. For this reason, TH.L is used in final crown volume equation since CH.L is predicted by TH.L (table 4.4).

Depending on Crown Shape

Crown Volume = (Crown Height) * (Crown Diameter)^2 * (Multiplier)

Predicted CD by regression model

Predicted CH by regression model

Crown Shape	Multiplier	Equation
Spheroid	0.5236	(Multiplier) * (0.9998*$\mathbf{CD.L}$ + 0.3477)^2 * (1.065* $\mathbf{TH.L}$ – 3.273)
Parabolic	0.3927	
Expanded parabolic	0.4909	

Figure 4.24. Crown Volume Estimation Model

4.7. Validation

Table 11 and 12 summarize the result of validation between field and LIDAR derived measurement and regression model respectively. By comparison between table 4.5 and 4.6, crown volume estimation model is validated as promising. We used predicted crown height and crown diameter instead of values extracted from rasterized LIDAR image. RMSE was reduced (total height from 2.61 to 2.15 and crown diameter from 2.30 to 2.37). Furthermore, p-value shows that there is no difference statistically significant with high probabilities when we compare p-value as α=0.05 significant level. According to the result of validation between field and LIDAR derived measurement (table 4.5), it is shown that the difference of total height measurement is highly statistically significant. Probably, measurement error and low density of LIDAR point cloud cause this difference. The rest of all lead to the same practical conclusion: There is no evidence for difference between field and LIDAR derived measurement, actual and predicted values by regression model. In other words, the methods to be used for measurement and the values for crown volume estimation model are reliable and convinced.

In summary, object-oriented approach based on rasterized LIDAR image in order to extraction tree parameters such as tree height information and crown diameter gives promising possibilities. And the proposed crown volume estimation model is also reliable.

Table 4.5. Result of validation between field and LIDAR derived measurement

			Two-sampled T-statistics		Wilcoxon signed rank test	
Variables	r^2	***RMSE***	**T-statistic**	**2-tailed p**	**Wilcoxon's W-statistic**	**2-tailed p**
TH.F-TH.L	0.93	2.61	4.12	0.0001	1450.5	<0.0001
CD.F-CD.L	0.90	2.30	1.17	0.2453	1054	0.2021
CV.F-CV.L	0.92	1224.994	1.21	0.2301	1061	0.2825

- F : Field measurement, L: LIDAR derived measurement
- More details are available in appendix A-6 and A-7.

Table 4.6. Result of validation for regression

			Two-sampled T-statistics		Wilcoxon signed rank test	
Variables	r^2	***RMSE***	**T-statistic**	**2-tailed p**	**Wilcoxon's W-statistic**	**2-tailed p**
TH.F-TH.P	0.93	2.15	0.002	0.9983	983	0.6167
CH.F-CH.P	1	0.09	-0.09	0.9292	793	0.3691
CD.F-CD.P	0.90	2.27	0.00	0.995	935	0.8829
CV.F-CV.P	0.93	1086.7	-0.241	0.8083	970	0.6856

- F : Field measurement, P: Predicted values by regression model
- More details are available in appendix A-6 and A-7.

5. Summary and Discussion

This study showed how to estimate tree crown volume of individual urban trees from LIDAR data using an object-oriented approach, without use of additional optical image. The achieved results demonstrate the potential of object-oriented image processing, applying segmentation and classification of LIDAR data for individual tree detection and crown delineation.

The study consists of five main parts, leading to the extraction of crown height, crown diameter and determination of crown shape: 1) data preparation (rasterized LIDAR image); 2) pre-processing (nDSM and tree mask); 3) tree detection and crown delineation; 4) statistical analysis for crown volume estimation model and 5) validation. The full processing is reviewed in table 5.1.

5.1. Interpolation of LIDAR point cloud

We decided to use LIDAR data to extract tree parameters since LIDAR has ample possibilities for tree crown volume estimation in terms of improving the accuracy and reducing the complexity of analysis without combination of optimal imagery. The size of grid is the most sensitive and important because the accuracy of height information and crown periphery depends on the resolution of the rasterized LIDAR image. First, the LIDAR point cloud was interpolated under consideration of the grid size necessary to represent tree crown. We reviewed the representation of tree crowns at different resolution (grid size). We choose 25cm of grid size although the density of LIDAR point cloud is low (approximately 1 point / m^2). Finally it turned out that high resolution of the rasterized LIDAR image is better to represent tree crowns and improve the accuracy in terms of location and height measurement but increasing computing time and storage requirement should be considered.

5.2. Generation of nDSM and tree mask

In order to extract true crown height information, normalized DSM (nDSM) should be created. nDSM can be obtained simply by subtracting DTM from DSM. However DTM generation from DSM is not simple. We used the object- oriented image processing technique to classify the terrain and off-terrain on rasterized LIDAR image. This method separated them successfully. After removing the off-terrain, the empty space was re-interpolated using only terrain LIDAR point cloud which was converted from rasterized image into *.xyz. The advantage of object-oriented image processing technique for DTM generation is that it can provide simpler and more controlled operation.

Although a strict accuracy assessment for interpolation was not carried out, visual inspection showed that it can retain much more details. Another essential part of data pre-processing is the generation of a tree mask. We discriminated trees and buildings based on surface growing algorithm with some manual work since the low density of LIDAR point did not allow the use of texture based separation, which is commonly used on rasterized LIDAR images. In order to minimize the error by manual work, we compared the result to the aerial photograph. According to our experience, it was better to remove all objects on the nDSM image except for trees in order to reduce complexity and computing time for the tree detection and crown delineation process.

If a multi-pulse (first and last) of LIDAR point cloud is available, a nDSM can be generated by subtracting the last pulse from the first pulse and also trees and buildings can be separated. Poor resolution of LIDAR data results in poor height information and crown diameter measurement, because there is less chance to hit the tree apex, but it is likely to hit the shoulder of the tree instead of the top and the distinct crown periphery cannot be represented. Consequently, taking advantage of a high density LIDAR point cloud will improve accuracy for the measurement of tree height and crown diameter of individual trees. Furthermore, in a very dense LIDAR point cloud, texture measurement can be used for an accurate, automatic separation of trees and buildings.

5.3. Individual tree detection and crown delineation

Based on nDSM image which had only trees and ground, multi-resolution segmentation technique was applied for individual tree detection and crown delineation. We detected tree top area instead of the apex point by iteration of local maxima and local minima. It was useful to detect flat top trees which were difficult to be detected by a local maxima point algorithm. Multi-resolution segmentation by adjusting segmentation parameters successfully delineated single trees and trees in low density group but worked less well for the trees in high density group trees having similar height. Probably, this is because of the low density of the LIDAR point cloud. The density was not enough to represent in very dense groups. An object-orientated classification technique is more powerful in urban area than in a forest environment to delineate crowns of the single trees particularly since trees are planted more individually in an urban area.

5.4. Tree parameter extraction

The tree parameters to be used in solid geometric volume formula were extracted from image objects. We assumed segmented image objects could be identified as individual tree crowns. Therefore, the maximum pixel value of an object corresponds to total height. Finally, crown height was estimated from the regression of total height. Total heights extracted from field and LIDAR derived measurement and total height and crown height showed strong correlation to each other. Because of strong correlation between the variables of tree height parameters, Crown height could be estimated from total height, crown diameter was measured by maximum and minimum of X- and Y-coordinates. Like for height measurement, the limitation of tree crown delineation for the high density group trees having a similar height could be resolved by a regression model of total height and crown diameter. When it comes to crown diameter measurement, it should be noted that positional errors of the field measurement of crown projections must be the error source inherent in our data. Next, crown shape was determined by the ratio of total height to crown diameter based on a certain range of total height. This method was very helpful to classify geometric crown shape in an objective way.

5.5. Tree crown volume estimation and validation

Finally, the tree crown volume estimation model was developed based on a solid geometric volume formula using the predicted crown height and crown diameter by regression, the multiplier based on ratio of fraction from a straight cylinder's volume. This model was validated by the coefficient of determination (r^2), root mean square error (RMSE) and test of significance such as T-test and

Wilcoxon signed rank test. It should be noted that the visual inspection of coefficient of determination (r^2) is essential to avoid bias. A RMSE measures the amount of error occurring from the measurement to be able to compare the practical accuracy between different approaches. Finally, a significance test can provide the evidence, whether or not the difference is significant. As far as p-value is concerned, there is no distinct boundary to determine however it is very useful to show results are convincing.

Table 5.1. Summary of processing for crown volume estimation

Task	Description (Algorithms & methods)	Source
DSM	Triangulated irregular network (TIN) under consideration of representing tree crowns	LIDAR point cloud
DTM	Classification of terrain and off-terrain by object-oriented image classification technique and re-interpolation	Rasterized DSM image
nDSM	Subtracting DTM from DSM	Rasterized DSM & DTM
Tree mask	Discrimination of trees and buildings based on surface growing algorithm with some manual work	Converted LIDAR point cloud
Tree detection	Detecting tree top area instead of apex point, by iteration of local maxima and local minima	Rasterized and non-tree masked nDSM image
Crown delineation	Multi-resolution segmentation by adjusting segmentation parameters	
Extraction of Crown height & crown diameter	Extraction from image objects • Total height = maximum pixel value • Crown diameter = average (maxi, mini of X- and Y- coordinates) • Crown height by regression model of total and crown height	Delineated crown segments
Crown shape determination	Like minimum distance classification using the ratio of total height to crown diameter based on a certain range of total height	Total height & Crown diameter
Descriptive statistic & Normality test	Overview & describe the variables to justify the correlations between them, to develop the crown volume prediction model and to determine the optimal statistical analysis methods.	Tree parameters from field and LIDAR derived measurement
Correlation & regression	Determination the degree of association and regression line based on mirror-scatter plot corresponding to correlation coefficient and simple linear regression based on least-squares to expect crown diameter and height variable	
Crown volume estimation	Solid geometric volume formula using predicted crown height and crown diameter by regression and multiplier based on ratio of fraction from a straight cylinder's volume	
Validation	• Coefficient of determination (r^2) • Root mean square error (RMSE) • Test of Significance : T-test & Wilcoxon signed rank test	

6. Conclusion and Recommendation

6.1. Conclusion

This study proved that it is possible to estimate tree crown volume of individual urban trees from LIDAR data using an object-oriented approach, without use of additional optical image. It demonstrated detection of individual trees and delineation of crowns applying segmentation and classification of LIDAR data.

The model that was developed shows very promising results for individual tree crown volume estimation using solid geometric volume formula with linear regression. Also, this research shows it is possible to use only a rasterized LIDAR image in crown volume estimation. Object-oriented approach based on a rasterized LIDAR image for individual tree crown delineation and tree parameter extraction gave us promising results and reliable tree crown volume estimation model (r^2=0.93, RMSE=1086.7, the difference is not significant at α= 0.05 and 0.01). It should be noted that for best crown volume estimation, the same way based on field and LIDAR derived measurement should be used. We employed the solid geometric volume formula with multiplier, which is ratio of fraction from a straight cylinder's volume, based on different crown shapes. In this formula, predicted crown height and crown diameter by regression between field and LIDAR derived measurements were used. Therefore, in order to estimate tree crown volume, the information of crown height, crown diameter and crown shape is particularly important.

6.2. Recommendation

The problems or limitations to be accounted for in this research are the following;

- Determination of grid size for interpolation
- Appropriate density of LIDAR point cloud for tree analysis
- Generation and accuracy assessment of nDSM and Tree mask
- Determination of crown shape
- Outliers in regression model

These should be attended to in the future work, in terms of controlling the error sources in order to improve the accuracy.

In phase of data preparation, we didn't propose the specific interpolation algorithm, grid size and density of LIDAR point cloud for tree crown volume estimation. They are very important since representation of individual tree crown with accurate location and height information depends on them. Therefore, it is recommendable to study these factors more specifically in such a way,

comparing the degree of tree crown representations at different grid size in conjunction with different density of the LIDAR point cloud. Additionally, the computing time and interpolation error should be taken into account.

When it comes to nDSM and tree mask generation, they are directly related to the accuracy of tree height and crown diameter estimation. In order to reduce the error, one should separate accurately terrain and off-terrain for nDSM, trees and buildings for tree mask respectively. In this study, we did the visual investigation, but a more strict and precise accuracy assessment for nDSM and tree mask should be used. Thus, a method of accuracy assessment, specified for nDSM and tree mask, is required. From a broader point of view, increasing the density of the LIDAR point cloud is probably the most efficient way to improve accuracy of individual tree parameter extraction. More research, in this field therefore required to take advantage of increased capacities of the laser scanning systems as technology improves.

We proposed a specific method to determine the crown shape. Although it was helpful to classify crown shape into a geometric object shape, it was limited to only dominant tree shapes. In order to reduce the error by applying a few geometric shapes to all trees, a more objective method which can determine individual tree crown shapes should be developed. Finally, to make the best regression model, the control of outliers is recommended.

Reference

[1] C. Small, 2001, "Estimation of urban vegetation abundance by spectral mixture analysis," International journal of remote sensing, Vol 22, pp. 1305-1334.

[2] D. M. Wagrowski, R A. Hites, 1996, "Polycyclic aromatic hydrocarbon accumulation in urban, suburban and rural vegetation," Environmental science and technology, Vol. 31, pp 279-282.

[3] Arthur Plotnix, 2000, "The Urban Tree Book: An Uncommon Field Guide for City and Town," Three River Press, 1st Edition.

[4] Frances E. Kuo, 2003, "The role of arboriculture in healthy social ecology," Journal of Arboriculture, F.E., pp. 148-155.

[5] E. Schopfer, S.Lang, and T.Blaschke, 2004, "A green index incorporating remote sensing and citizen's perception of green space," www.isprs.org/commission8/workshop_urban/schoepfer.pdf visited on 22nd Oct., 2006.

[6] FAO. 1998, "FRA2000 terms and definitions, FAO", http://www.fao.org/forestry/fo/fra/docs/FRA_Wpleng.pdf, visited on 22nd Oct., 2006.

[7] Kleinn, C., 2000, "On large area inventory and assessment of trees outside forests," Unasylva 200 Vol. 51. pp.3-10.

[8] Susan R. 2003, "Fletched, Global Climate Change: The Kyoto Protocol, International Environmental Policy, Resources, Science and Industry Division", CRS Report for Congress.

[9] Garzuglia, M. and M. Saket, 2003, "Wood Volume and Woody Biomass: Review of FRA 2000 Estimates", Forest Resource Assessment Program, FAO, working paper 68, Rome, Italy.

[10] Darius S. culvenor, 2003, "Extracting individual tree information: A survey of techniques for high spatial resolution imagery," Remote sensing of forest environments: concepts and case studies, pp. 254-277.

[11] Lu, D., Batistella, M. and Moran, E., 2005, "Satellite estimation of aboveground biomass and impacts of forest stand structure", Photogrammetric Engineering and Remote Sensing, pp. 967–974.

[12] D.A. Poliot, D.J. King, F.W Bell and D.G. Pitt, 2002, "Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration", Remote Sensing of Environment 82(2002) pp. 322-334.

[13] Takashi Fuji, Tetsuo Fukuchi, 2005, "Laser Remote Sensing", CRC, pp. 3-15.

[14] Michel Morgan, 2002, "Interpolation of LIDAR data and automatic building extraction", ACSM-ASPRS 2002 Annual Conference Proceedings.

[15] J. Dash, E.Steinle, R.P. Singh, and H.P Bahr, 2004, "Automatic building extraction from laser scanning data: an input tool for disaster management", Advances in Space Research 33 (2004), pp. 317–322.

[16] Sorin C. Popescu, Randolph H. Wynne, and Ross F. Nelson, 2003, "Measuring individual tree crown diameter with LIDAR and assessing its influence on estimating forest volume and biomass", Can. J. Remote Sensing, Vol 29, No 5, pp. 564-577.

[17] Ward W. Carson, Hans-Erik Anderson, Stephen E. Reutebuch and Roubert J. McGaughy, 2004, "LIDAR applications in forest : An overview", ASPRS Annual Conference Proceedings, Denver, Colorado.

[18] Chen L., T. Chiang and T. Teo, 2005, "Fusion of LIDAR Data and High Resolution Images for Forest Canopy Modeling", Asian Conference on Remote Sensing, Hanoi, Vietnam.

[19] Amin P. Charaniya, Roberto Manduchi, and Suresh K. Lodha, 2004, "Supervised parametric classification of aerial LIDAR data," IEEE Workshop on Real-Time 3D Sensors, pp. 25–32.

[20] Juha Hyyppä, Teemu Mielonen1, Hannu Hyyppä, Matti Maltamo, Xiaowei Yu, Eija Honkavaara and Harri Kaartinen, 2005, "Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds", ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, the Netherlands.

[21] Guillaume Perrin, Xavier Descombes and Josiane Zerubia, 2004, "Tree crown extraction using marked point processes", http://www.eurasip.org/content/Eusipco/2004/defevent/papers/cr1086.pdf , visited on 14th Oct. 2006.

[22] Dralle K. and Rudemo M., 1997, "Automatic estimation of individual tree positions from aerial photos", Canadian Journal of Forest Research, pp.1728– 1736.

[23] Walsworth N. and King D. J., 1998, "Comparison of two tree apex delineation techniques", In D. A. Hill, D. G. Leckie (Eds.), Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, pp. 93– 104, Victoria, BC: Canadian Forest Service, Pacific Forestry Centre.

[24] Wulder M., Niemann K. O., and Goodenough D. G., 2000, "Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery", Remote Sensing of Environment, pp. 103– 114.

[25] Niemann K. O., Adams S. and Hay, G., 1998, "Automated tree crown identification using digital orthophoto mosaics", In D. A. Hill, & D.G. Leckie (Eds.), Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, pp. 105-113.

[26] Brandtberg T. and Walter F., 1998, "Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis", Machine Vision and Applications, pp. 64–73.

[27] Pollock R., 1998, "Individual tree recognition based on a synthetic tree crown image model", In D. A. Hill, D. G. Leckie (Eds.), Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, pp. 25 – 34, Victoria, BC: Canadian Forest Service, Pacific Forestry Centre.

[28] Gougeon F. A., 1995, "A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution images", Canadian Journal of Remote Sensing, pp. 274– 288.

[29] Darius S. Culvenor, 2002, "TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery", Computer & Geosciences 28 (2002) pp. 33-44.

[30] Dirk Tiede, Christian Hoffmann, 2006, "Processing oriented object-based algorithms for single tree detection using laser scanning", 3D Remote Sensing in Forest, 15th Feb 2006 Vienna, Austria.

[31] Dirk Tiede, Gregor Hochleitner and Thomas Blaschke, 2005, "A full GIS-based workflow for tree identification and tree crown delineation using laser scanning", IAPRS, Vol. XXXVI, Part 3/W24, Vienna, Austria, 29th – 30th Aug 2005.

[32] Christioher Mei and Sylvie Durrieu, "Tree crown delineation from digital elevation models and high resolution imagery", http://citeseer.ist.psu.edu/759333.html visited on 22nd Oct 2006.

[33] Pinz A., 1999, "Tree isolation and species classification", In Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, 10th–12th Feb 1998, Victoria, B.C. Pacific Forestry Center, Canadian Forest Service, Natural Resources Canada.

[34] Peter Bunting and Richard Lucas, 2005, "The delineation of tree crows in Australian mixed species forests using hyperspectral Compact Airborne Spectrographic Imager (CASI) data", Elsevier Inc, Remote Sensing of Environment, pp. 230-248.

[35] Curtis A. Collins, Robert C. Parker, David L. Evans, 2004, "Using multispectral imagery and multi-return LIDAR to estimate tree and stand attributes in a southern bottomland hardwood forest", ASPRS Annual Conference Proceedings, Denver, Colorado.

[36] Bertram Husch, Thomas W. Beers, John A. and Kershaw Jr., 2003, "Forest mensuration", John Wiley & Sons, 4th Edition, pp 85-113.

[37] Tee-Ann Teo, Liang-Chen Chen, 2004, "Object-based building detection from LIDAR data and high resolution satellite imagery", 25th ACRS 2004 Poster Session 5-5.

[38] Surfer 8 user's guide, 2002, Golden Software Inc, p135.

[39] H. Arefi and M. Hahn, 2005, "A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data", ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, the Netherlands.

[40] Oude Elberink, S. and H.-G. Maas, 2000, "The use of anisotropic height texture measures for the segmentation of lasers canner data", International Archives of Photogrammetry and Remote Sensing, vol. 33, part B3/2, pp. 678-684.

[41] Vosselman G., 2000, "Slope based filtering of laser altimetry data', IAPRS, Vol XXXIII, Part 3B, Amsterdam, the Netherlands, pp 935-942.

[42] T. Aschoff and H. Spiecker, 2004, "Algorithms for the automatic detection of trees in laser scanner data", International Archives of Photogrammetry, Remote Sensing and Spatial Information, Vol XXXVI 8/W2.

[43] Baatz M., Benz U., Dehghani, S. And Heynen, M., 2004, "eCognition user guide 4", Definiens Image GmbH, Munchen, Germany.

[44] Hofmann P., 2001, "Detecting urban features from IKONOS data using an object-oriented approach", RSPS 2001 Proceedings.

[45] Sohel Syed, Paul Dare and Simon Jones, 2005, "Automatic classification of land cover features with high resolution imagery and LIDAR data : An object-oriented approach", Proceedings of SSC2005 Spatial Intelligence, Innovation and Praxia, the national biennial conference of the spatial science institute, September, 2005, Melbourne, Spatial science institute.

[46] USGS, Rocky Mountain Mapping Center, http://rockyweb.cr.usgs.gov/elevation/dpi_dem.html visited on 2nd Oct 2006.

[47] Definiens Professional 5 Reference Book, 2006, pp 9-15.

[48] David S. Moore, and George P. McCabe, 1998, "Introduction to the practice of statistics", W.H Freeman and Company, 3rd Edition, pp. 128-148.

[49] Bertram Husch, Thomas W. Beers and John A. Kershaw Jr., 2003, "Forest mensuration", John Wiley & Sons, pp 48-52.

[50] Murray R. Spiegel and Larry J. Stephens, 1999, "Theory and problems of statistics", McGraw-Hill Companies Inc., 3rd Editions, pp 281-286.

[51] Murray R. Spiegel and Larry J. Stephens, 1999, "Theory and problems of statistics", McGraw-Hill Companies Inc., 3rd Editions, pp 661-694.

[52] Wikimedia Foundation, http://en.wikipedia.org/wiki/Normality_test, visited on 2nd Sep 2006.

[53] Wikimedia Foundation, http://en.wikipedia.org/wiki/RMSE, visited on 22nd Jan 2007.

[54] David S. Moore and George P. McCabe, 1998, "Introduction to the practice of statistics", W.H Freeman and Company, 3rd Edition, pp. 504-519.

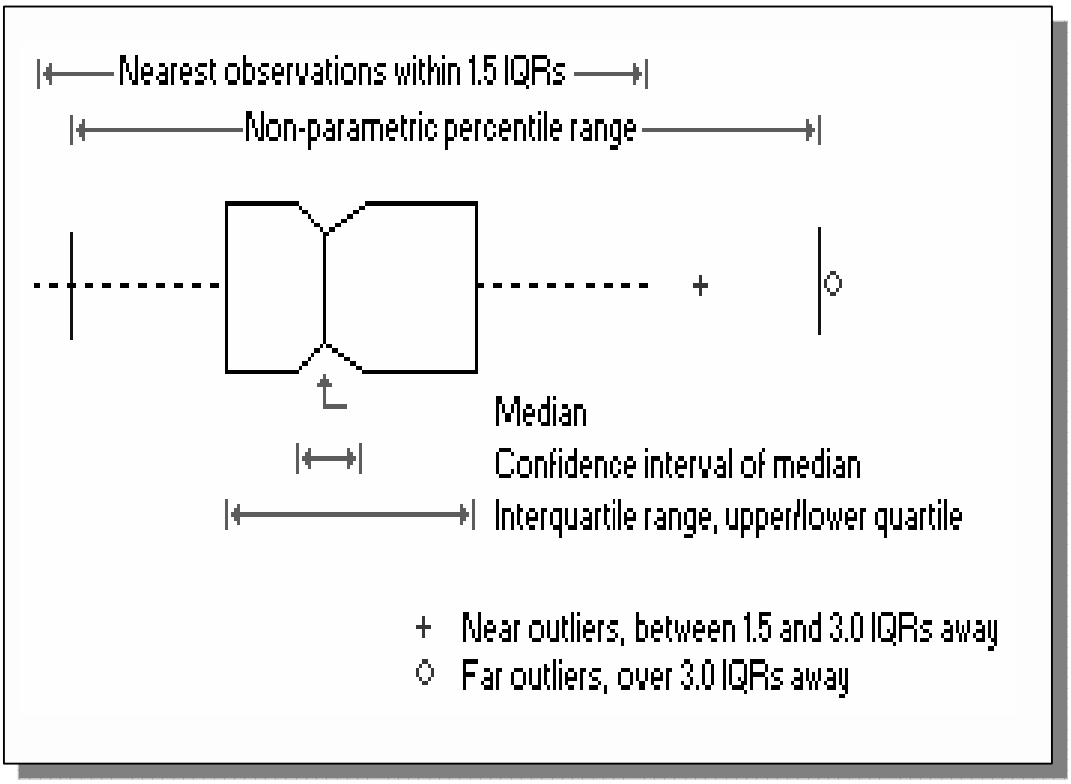
[55] David S. Moore and George P. McCabe, 1998, "Introduction to the practice of statistics", W.H Freeman and Company, 3rd Edition, pp. 453-466.

[56] Wikimedia Foundation, http://en.wikipedia.org/wiki/Correlation, visited on 22nd Jan 2007.

[57] David S. Moore and George P. McCabe, 1998, "Introduction to the practice of statistics", W.H Freeman and Company, 3rd Edition, CD-ROM, chapter 14.2. Wilcoxon signed ranked test.

Appendix

A - 1. Symbol interpretation in Box-whisker plot



A - 2. Statistics of summary

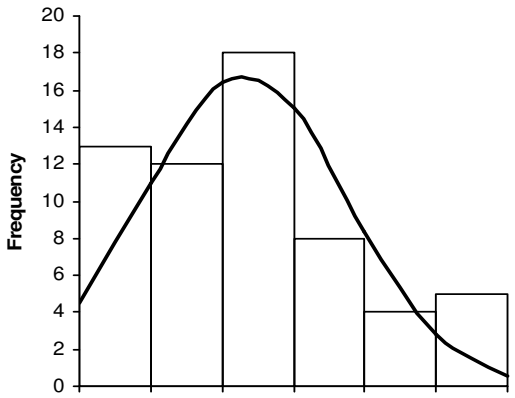
	DBH . F	TH . F	TH . L	CH . F	CH .L*	CD . F	CD . L	CV . F*	CV . L*
Number of values	60	60	60	60	60	60	60	60	60
Sum	3473	1100	1027	897	826.9	782.3	762	138900.9	117205.4
Minimum	8.1	6	4.945	4.5	2.141	3.1	2.38	22.64	6.32
Maximum	149.8	45	35.85	43	31.72	32.1	30.3	19727.3	14248.6
Range	141.7	39	30.91	38.5	29.58	29	27.9	19704.6	14242.3
Inter-quartile range	49.10	9.10	10.03	9.23	9.60	9.11	8.56	2365.3	1859.3
Mean	57.88	18.33	17.11	15	13.78	13.04	12.7	2315.0	1953.4
Median	51.25	18.1	17.42	14.3	14.08	10.93	13.1	931.2	970.1
First quartile	29.7	12.15	11.26	8.65	8.18	7.3	7.38	242.3	216.6
Third quartile	78.4	21.25	21.42	17.9	17.9	16.38	15.9	2583.9	2085.4
Standard error	4.593	1.059	0.918	1.03	0.879	0.909	0.86	513.1	390.8
95% CI of mean	48.69 to 67.07	16.21 to 20.45	15.27 to 18.95	12.90 to 17.02	12.02 to 15.54	11.22 to 14.86	10.97 to 14.41	1288.20 to 3341.83	1171.29 to 2735.55
95% CI of median	38.00 to 64.40	14.00 to 19.80	13.20 to 19.07	10.00 to 16.00	10.16 to 15.66	9.10 to 13.95	8.88 to 13.75	410.35 to 1404.31	323.39 to 1454.60
Variance	1266	67.29	50.56	63.5	46.32	49.6	44.4	15799487.7	9166817.4
Average deviation	28.09	6.184	5.732	6	5.486	5.504	5.17	2439.6	1918.9
Standard deviation	35.58	8.203	7.111	7.97	6.806	7.043	6.66	3974.8	3027.6
Coefficient of variation	0.615	0.447	0.416	0.53	0.494	0.54	0.53	1.71699	1.54993
Skew	0.841	1.127	0.547	1.33	0.547	1.074	0.86	2.898	2.848
Kurtosis	0.125	1.631	-0.01	2.3	-0.01	0.619	0.48	8.558	8.658
Kolmogorov -Smirnov stat	0.095	0.122	0.076	0.12	0.076	0.136	0.14	0.282	0.26
Critical K-S stat, (alpha=.10)	0.155	0.155	0.155	0.16	0.155	0.155	0.16	0.155	0.155
Critical K-S stat, (alpha=.05)	0.172	0.172	0.172	0.17	0.172	0.172	0.17	0.172	0.172
Critical K-S stat, (alpha=.01)	0.207	0.207	0.207	0.21	0.207	0.207	0.21	0.207	0.207

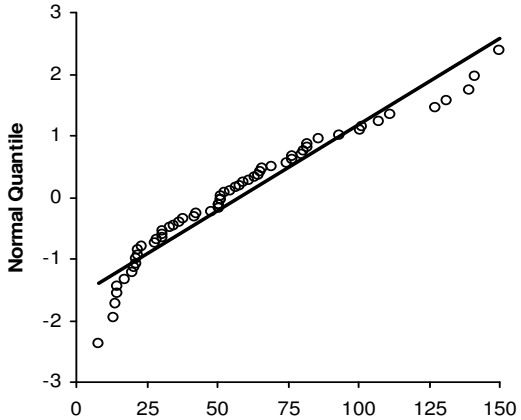
*NOTE:

- CH.L, CV.F and CV.L was added after regression analysis and crown volume calculation

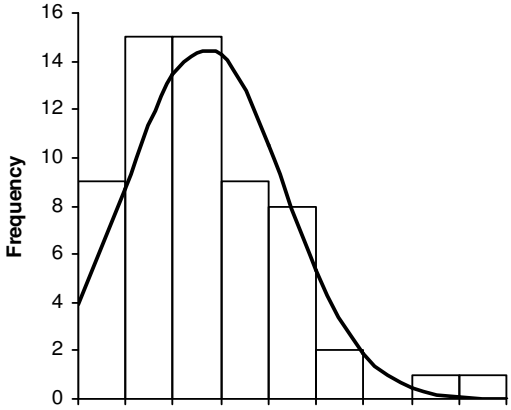
A - 3. Histogram of frequency and normal probability plot

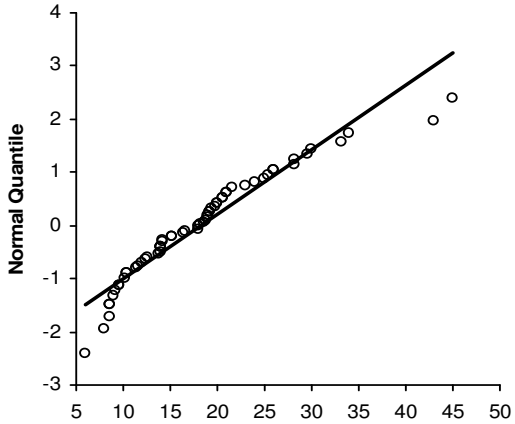
A-3-1. DBH from Field measurement (DBH.F)



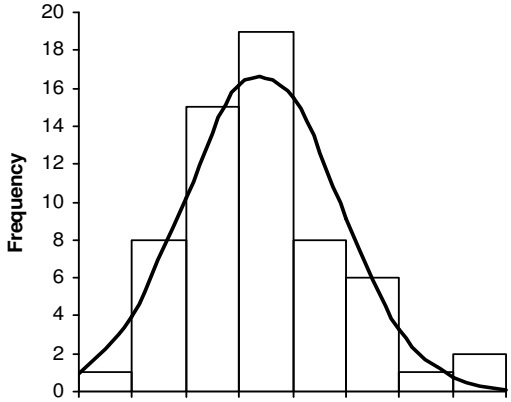


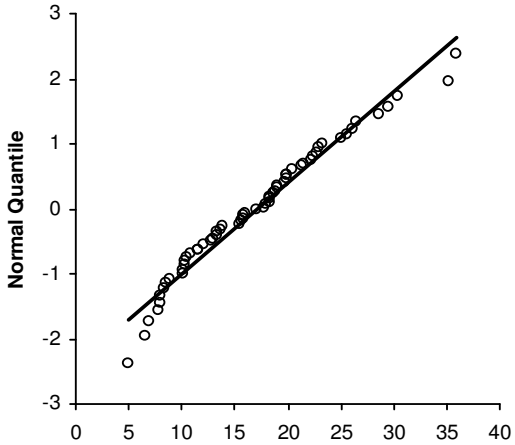
A-3-2. Total Height from Field measurement (TH.F)





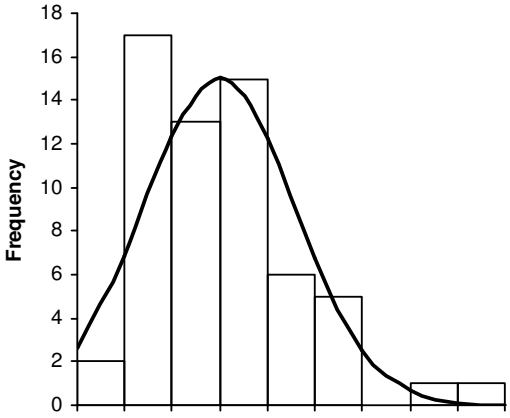
A-3-3. Total Height from LIDAR derived measurement (TH.L)

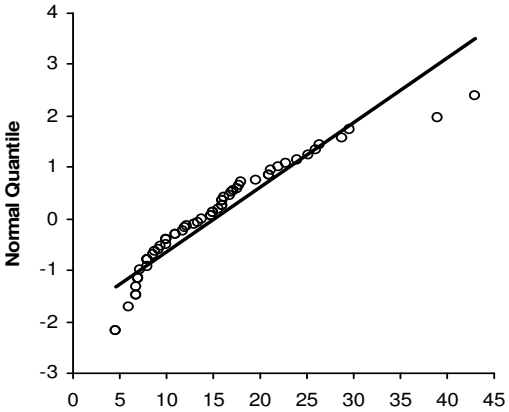




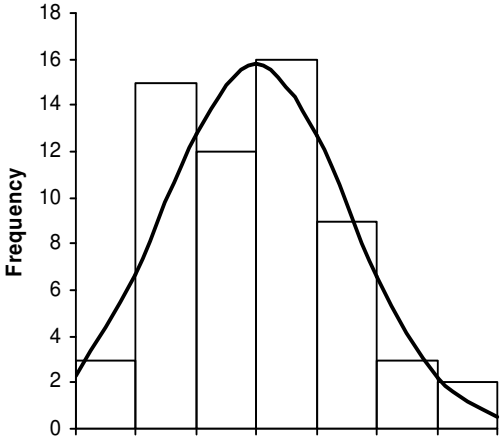
A - 3. Histogram of frequency and normal probability plot (continued)

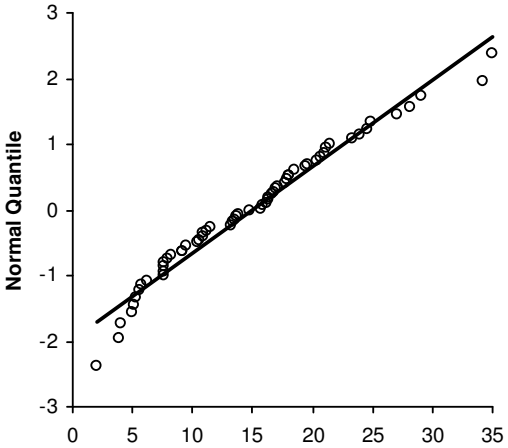
A-3-4. Crown Height from Field measurement (CH.F)



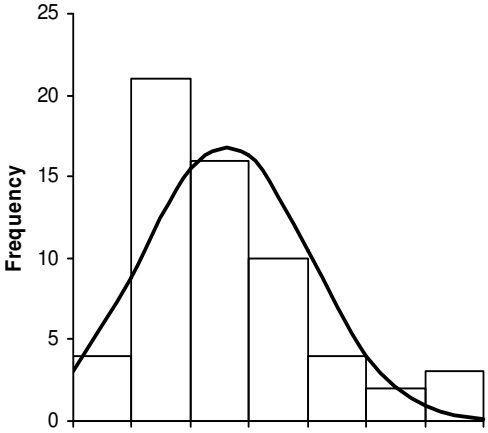


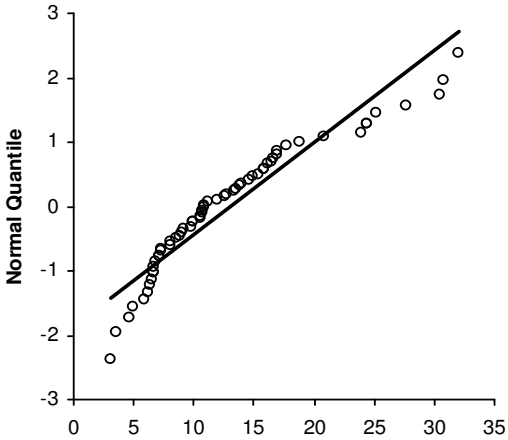
A-3-5. Crown Height from LIDAR derived measurement (CH.L)





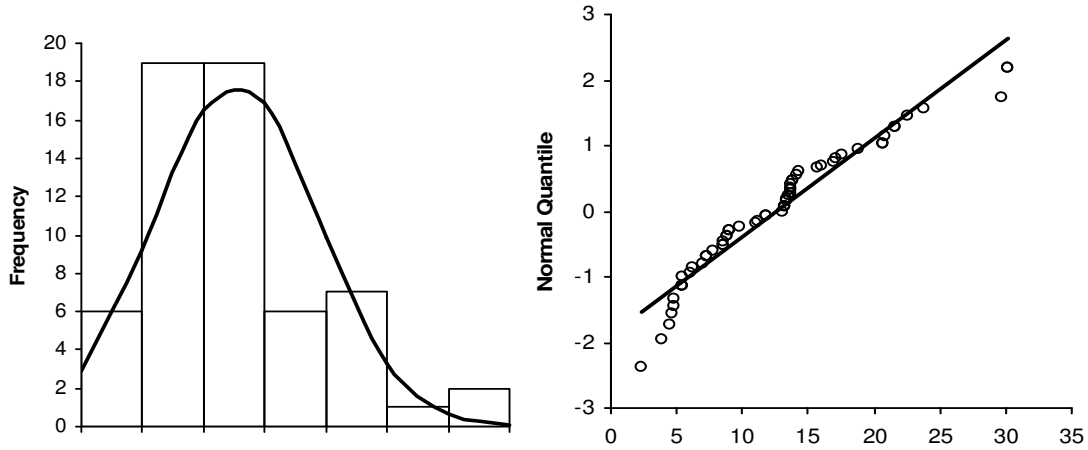
A-3-6. Crown Diameter from Field measurement (CD.L)



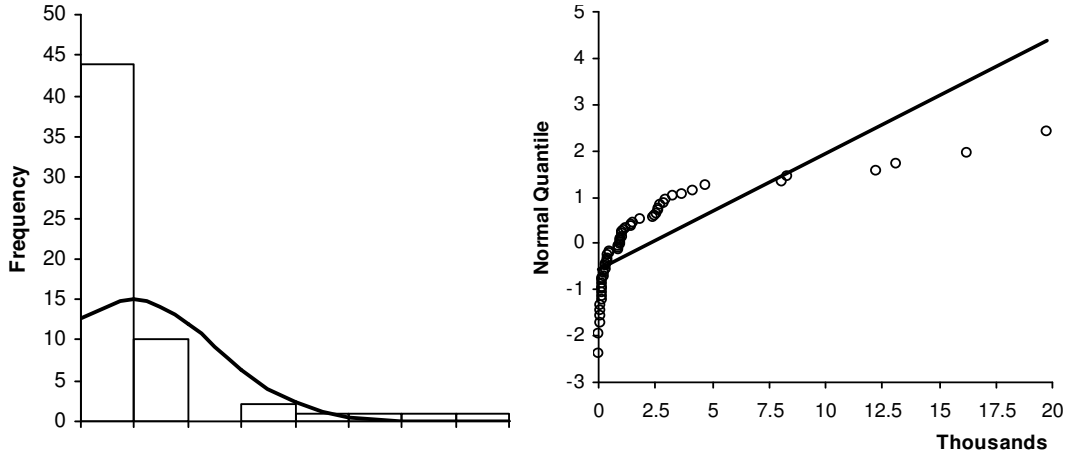


A - 3. Histogram of frequency and normal probability plot (continued)

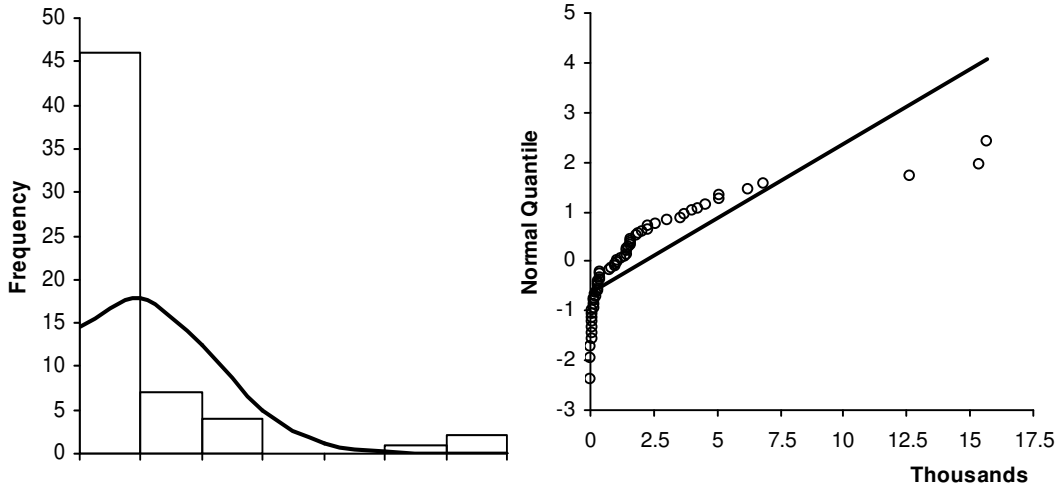
A-3-7. Crown Diameter from LIDAR derived measurement (CD.L)



A-3-8. Crown Volume from Field measurement (CV.F)



A-4-9. Crown Volume from LIDAR derived measurement (CV.L)

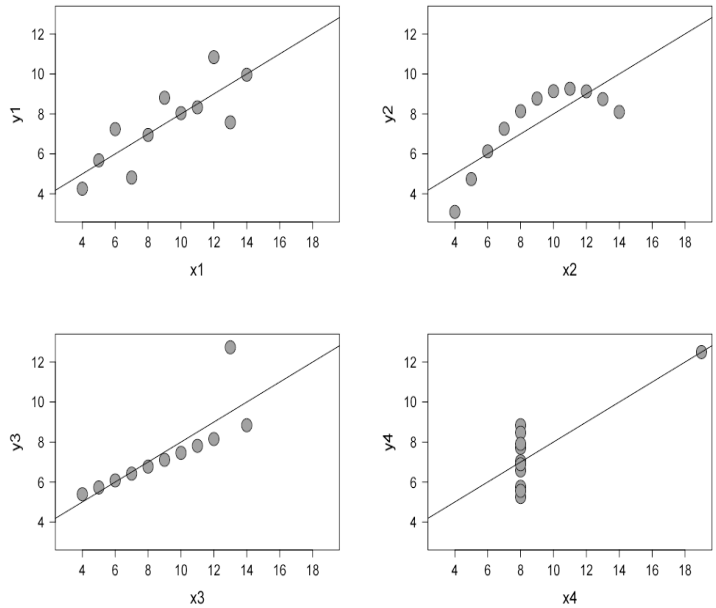


A - 4. Correlation coefficient(r) and linearity

Four sets of data with the same correlation of 0.81, as described by F. Anscombe. While Pearson correlation indicates the strength of a linear relationship between two variables, its value alone may not be sufficient to evaluate this relationship, especially in the case where the assumption of normality is incorrect.

The image on the right shows scatter-pots of four different pairs of variables, first described by Francis Anscombe. The four *y* variables have the same mean (7.5), standard deviation (4.12), correlation (0.81) and regression line ($y = 3 + 0.5x$). However, as can be seen on the plots, the distribution of the variables is very different. The first one (top left) seems to be distributed normally, and corresponds to what one would expect when considering two variables correlated and following the assumption of normality. The second one (top right) is not distributed normally; while an obvious relationship between the two variables can be observed, it is not linear, and the Pearson correlation coefficient is not relevant. In the third case (bottom left), the linear relationship is perfect, except for one outlier which exerts enough influence to lower the correlation coefficient from 1 to 0.81. Finally, the fourth example (bottom right) shows another example when one outlier is enough to produce a high correlation coefficient, even though the relationship between the two variables is not linear.

These examples indicate that the correlation coefficient, as a summary statistic, can not replace the individual examination of the data.



This graphic represents the four datasets defined by Francis Anscombe for which some of the usual statistical properties (mean, variance, correlation and regression line) are the same, even though the datasets are different.

Source: Wikimedia Foundation, Inc, http://en.wikipedia.org/wiki/Correlation visited on 6th Feb 2007

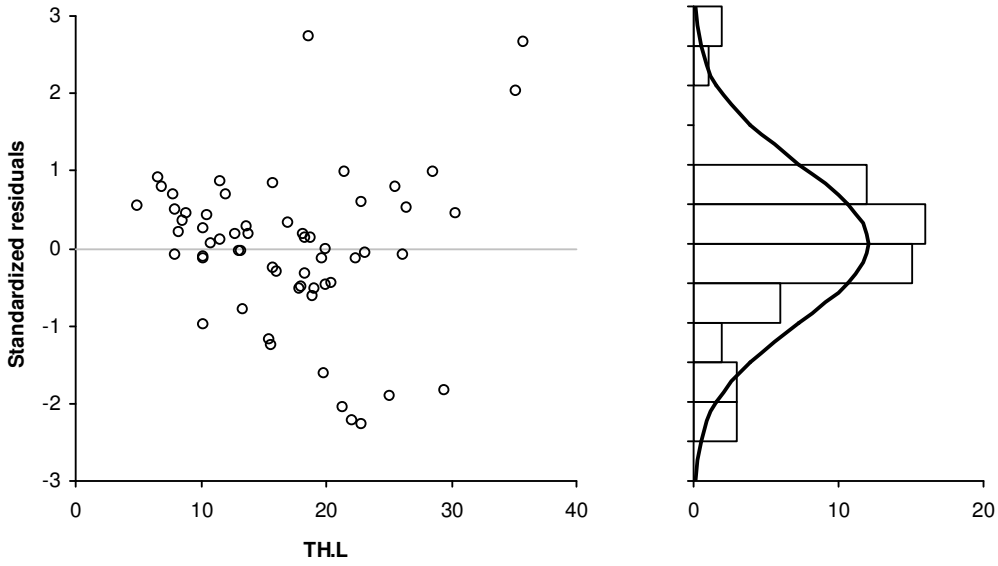
A - 5. Regression coefficient table and residual plot

A-5-1. Regression of Total Height from Field and LIDAR derived measurement

n	60
R^2	0.93
Adjusted R^2	0.93
SE	2.1758

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	-0.7093	0.7372	0.3400	-2.1849 to 0.7664
Slope	1.1130	0.0398	<0.0001	1.0333 to 1.1928

Source of variation	SSq	DF	MSq	F	p
Due to regression	3695.518	1	3695.518	780.62	<0.0001
About regression	274.576	58	4.734		
Total	3970.093	59			



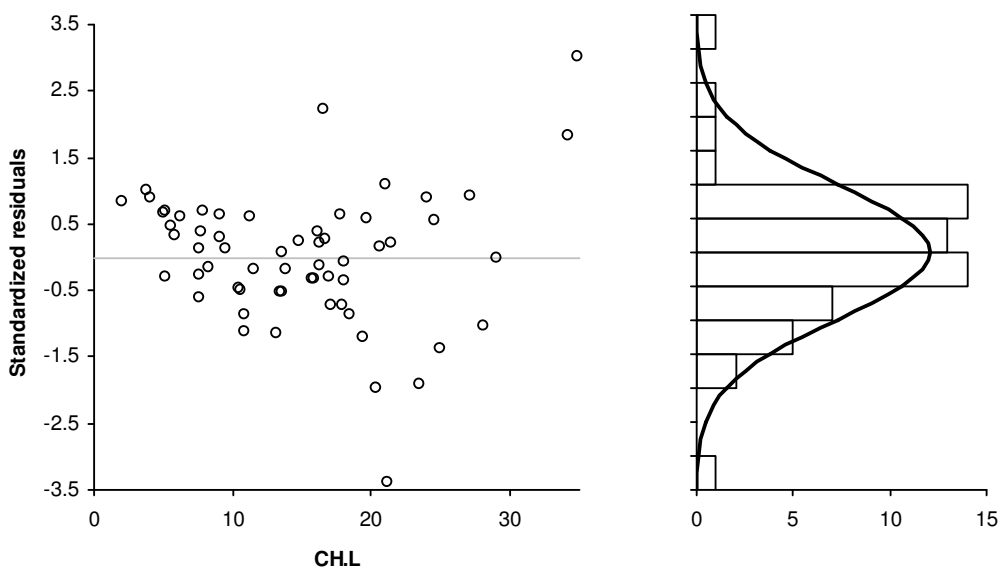
A - 5. Regression coefficient table and residual plot (continued)

A-5-2. Regression of Crown Height from Field and LIDAR derived measurement

n	60
R^2	0.88
Adjusted R^2	0.88
SE	2.7583

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	0.1866	0.7933	0.8148	-1.4013 to 1.7746
Slope	0.9876	0.0474	<0.0001	0.8927 to 1.0825

Source of variation	SSq	DF	MSq	F	p
Due to regression	3302.970	1	3302.970	434.13	<0.0001
About regression	441.277	58	7.608		
Total	3744.247	59			



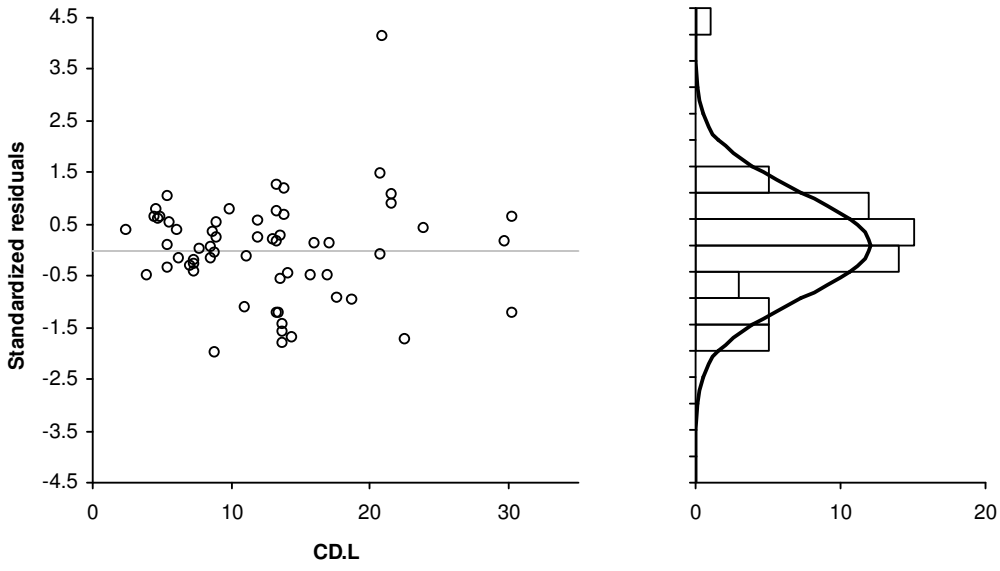
A - 5. Regression coefficient table and residual plot (continued)

A-5-3. Regression of Crown Diameter from Field and LIDAR derived measurement

n	60
R^2	0.90
Adjusted R^2	0.89
SE	2.2992

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	0.3477	0.6427	0.5906	-0.9388 to 1.6341
Slope	0.9998	0.0449	<0.0001	0.9099 to 1.0897

Source of variation	SSq	DF	MSq	F	p
Due to regression	2619.894	1	2619.894	495.61	<0.0001
About regression	306.600	58	5.286		
Total	2926.493	59			



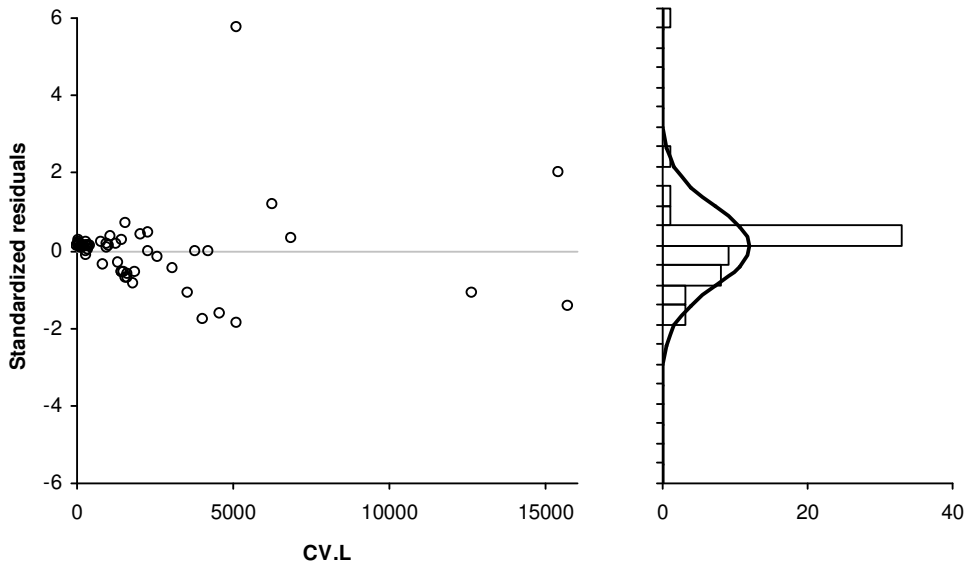
A - 5. Regression coefficient table and residual plot (continued)

A-5-4. Regression of Crown Volume from Field and LIDAR derived measurement

n	60
R^2	0.92
Adjusted R^2	0.92
SE	1119.5937

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	-136.2374	172.2438	0.4322	-481.0208 to 208.5459
Slope	1.1461	0.0438	<0.0001	1.0585 to 1.2338

Source of variation	SSq	DF	MSq	F	p
Due to regression	859148295.179	1	859148295.179	685.40	<0.0001
About regression	72702420.944	58	1253490.016		
Total	931850716.122	59			



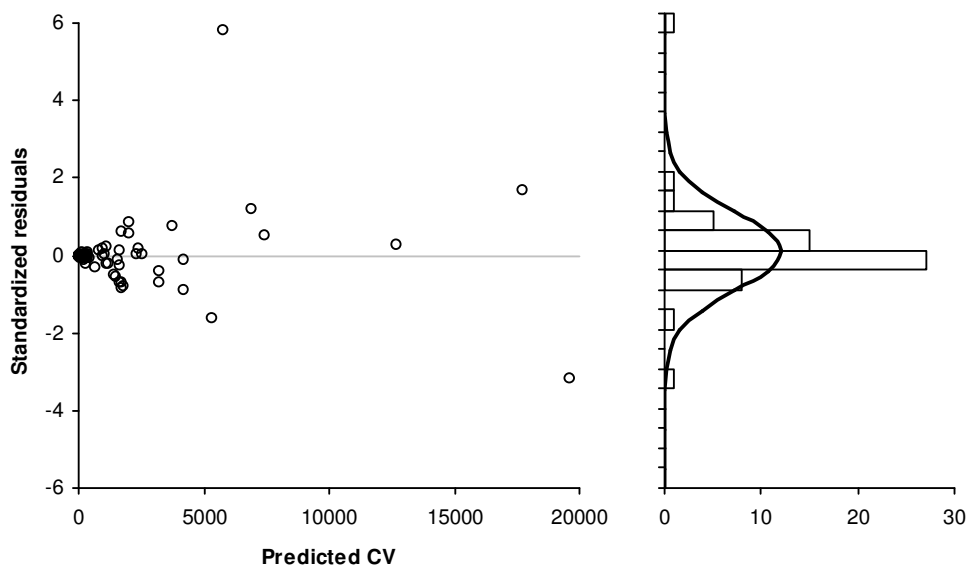
A - 5. Regression coefficient table and residual plot (continued)

A-5-5. Regression of Crown Volume from Field measurement and prediction (model)

n	60
R^2	0.93
Adjusted R^2	0.92
SE	1094.3922

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	50.6584	164.5939	0.7594	-278.8120 to 380.1288
Slope	1.0035	0.0374	<0.0001	0.9287 to 1.0784

Source of variation	SSq	DF	MSq	F	p
Due to regression	862384442.073	1	862384442.073	720.04	<0.0001
About regression	69466274.050	58	1197694.380		
Total	931850716.122	59			



A - 6. Two-paired T-statistics

	n	Mean	SD	SE
TH.F	60	18.333	8.203	1.0590
TH.L	60	17.109	7.111	0.9180
Difference	60	1.225	2.302	0.2972
95% CI	0.630 to	1.819		
t-stat.	4.12			
2-tailed p	0.0001			
CD.F	60	13.038	7.043	0.9092
CD.L	60	12.692	6.665	0.8604
Difference	60	0.345	2.280	0.2943
95% CI	-0.244 to	0.934		
t-stat.	1.17			
2-tailed p	0.2453			
CV.F(transformed)	60	6.660	1.564	0.2019
CV.L(transformed)	60	6.576	1.721	0.222
Difference	60	0.084	0.5239	0.0696
95% CI	-0.055 to	0.224		
t-stat.	1.21			
2-tailed p	0.2301			
TH.F	60	18.333	8.203	1.0590
TH.P	60	18.333	7.914	1.0217
Difference	60	0.001	2.157	0.2785
95% CI	-0.557 to	0.558		
t-stat.	0.00213			
2-tailed p	0.9983			
CH.F	60	14.957	7.966	1.0284
CH.P	60	14.958	7.868	1.1057
Difference	60	-0.001	0.099	0.0128
95% CI	-0.027 to	0.024		
t-stat.	-0.09			
2-tailed p	0.9292			
CD.F	60	13.03750	7.043	0.9092
CD.P	60	13.03730	6.664	0.8603
Difference	60	0.00020	2.280	0.2943
95% CI	-0.589 to	0.589		
t-stat.	0.00068			
2-tailed p	0.99946			
CV.F(transformed)	60	6.660	1.564	0.2019
CV.P(transformed)	60	6.672	1.579	0.2039
Difference	60	-0.012	0.383	0.0494
95% CI	-0.111 to	0.087		
t-stat.	-0.24			
2-tailed p	0.8083			

A - 7. Wilcoxon signed ranked test

	n	Rank Sum	Mean rank
TH.F – TH.L			
Positive	50	1450.5	29.01
Negative	10	379.5	37.95
Zero	0		
Difference between medians	1.093		
95.1 % CI	0.680 to 1.515		
W-stat./ 2-tailed p	1450.5 / < 0.0001		
CD.F – CD.L			
Positive	36	1054.0	29.28
Negative	23	716.0	31.13
Zero	1		
Difference between medians	0.400		
95.1 % CI	-1.192 to 0.900		
W-stat./ 2-tailed p	1054 / 0.2021		
CV.F-CV.L			
Positive	37	1061.0	28.68
Negative	23	769.0	33.43
Zero	0		
Difference between medians	48.933		
95.1 % CI	-86.404 to 122.743		
W-stat./ 2-tailed p.	1061 / 0.2825		
TH.F – TH.P			
Positive	31	983.0	31.71
Negative	29	847.0	29.21
Zero	0		
Difference between medians	0.107		
95.1 % CI	-0.369 to 0.518		
W-stat./ 2-tailed p	989 / 0.6167		
CH.F – CH.P			
Positive	26	793.0	30.50
Negative	34	1037.0	30.50
Zero	0		
Difference between medians	-0.012		
95.1 % CI	-0.035 to 0.014		
W-stat./ 2-tailed p.	793 / 0.3691		
CD.F-CD.P			
Positive	33	935.5	28.33
Negative	27	895.0	33.15
Zero	0		
Difference between medians	0.054		
95.1 % CI	-0.540 to 0.553		
W-stat./ 2-tailed p	935 / 0.8829		
CV.F-CV.P			
Positive	33	970.0	29.39
Negative	27	860.0	31.85
Zero	0		
Difference between medians	10.216		
95.1 % CI	-77.813 to 73.934		
W-stat./ 2-tailed p	970 / 0.6856		