

Estimation of Carbon Storage in Urban Trees Using Multispectral Airborne Laser Scanning Data

by

Xinqu Chen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Geography

Waterloo, Ontario, Canada, 2016

©Xinqu Chen 2016

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

With the continued growth of global population, urbanization becomes an inevitable trend. As substantial urban expansion undergoes, ecosystem and global land cover have been altered consequently. Urban development becomes the biggest contributor to global carbon emissions while the process of urbanization results in urban heat islands, climate change, and losses of carbon sinks. Urban vegetation has drawn direct attention of city planners and policy makers by considering the importance of vegetation in urban climate modification and energy conservation in different ways. For instance, tree shading and wind shielding effects can attenuate the direct solar heat and air infiltration into individual houses. In city wide, vegetation contributes the largest proportion of carbon storage which reduces climate warming and urban heat island effects by sequestering CO₂ and storing carbon in biomass.

The carbon content stored in individual trees can be estimated by dendrometric parameters such as the diameter at breast height (DBH) using allometry-based models. With the development of airborne laser scanning (ALS) technology, ALS data and very high resolution multispectral imagery have proven to be promising tools for deriving dendrometric parameters in forest. With the emerging multispectral ALS technology, it became possible to obtain both the range and spectral information from a single source meanwhile the intensity of multispectral ALS showed its power in vegetation mapping.

This study aims to develop a workflow that can quantify the carbon storage in urban trees using multispectral ALS data. The workflow consists of four steps: multispectral ALS data processing, vegetation isolation, dendrometric parameters estimation, and carbon storage modeling. First, the raw multispectral ALS data is intensity-rectified and filtered to generate a normalized Digital Surface Model (nDSM) and multispectral ALS intensity information at wavelengths: 532 nm (Green), 1064 nm (Near-infrared, NIR), and 1550 nm (Shortwave Infrared, SWIR), respectively. Vegetation covers are isolated by the support vector machine (SVM) classifier using multispectral ALS intensity information and nDSM in which total six classes including two vegetation classes (grass and tree) are classified. Individual tree crown is delineated by local maxima filtering and marker-controlled watershed segmentation. Tree

height and crown width are derived from the crown segments and compared with field measurements. An ALS-DBH (diameter at breast height) multiple linear regression model is developed to predict field-measured DBH using ALS-derived tree height and crown width and assessed by cross validation. Then the carbon storage in individual trees is calculated by allometric equations using ALS-estimated DBH and height. A total of 40 trees are sampled in the field that four attributes: height, crown width, DBH, and biomass are recorded for each single tree.

The results show that the land cover classification with multispectral ALS intensity images and nDSM achieves above 90% overall accuracy. The result of local maxima filtering is improved by using both multispectral ALS intensity and nDSM as input data. The ALS-derived tree height has a root mean square error (RMSE) of 1.21 m (relative RMSE = 6.8%) and the ALS-derived crown width has a RMSE of 1.47 m (relative RMSE = 16.4%). The prediction performance of the ALS-DBH model achieves R^2 over 0.80 with a RMSE of 4.6 cm. The predicted carbon storage using ALS-modeled DBH corresponded to a RMSE of 142 kg (28.6%) and a bias of 14.4 kg. Results suggest that ALS-based dendrometric parameter estimation and allometric models can yield consistent performance and accurate estimation. Citywide carbon storage estimation is derived in this study by extrapolating the values within the study area to the entire city based on the specific proportion of each land cover type in the entire city. The proposed workflow also reveals the potential of multispectral ALS data in estimating carbon storage at individual-tree level and mapping vegetation in the urban environment.

Acknowledgements

First and foremost, my deepest gratitude is to my supervisor, Professor Dr. Jonathan Li. I am very grateful for his guidance, support, and patience throughout my graduate studies. I am very fortunate to get involved in and work on several real-world applications with him. And I am deeply grateful for having the opportunity to explore different aspects of geomatics technology, in particular, and the state-of-the-art multispectral LiDAR. I am always inspired by his enthusiasm, dedication, and the ideas he brings to me.

I would like to thank my thesis committee members, Dr. Yuhong He, Associate Professor at the Department of Geography and Planning, University of Toronto Mississauga, Dr. Derek Robinson, Assistant Professor at the Department of Geography and Environmental Management, University of Waterloo, and Dr. Michael Chapman, Professor at the Department of Civil Engineering, Ryerson University for serving as my thesis examining committee even at a very busy time or commuting a long way from Toronto to Waterloo to attend my thesis defense in the morning. I would also like to thank my thesis committee members for devoting time in their busy schedules to read and review the thesis and offer me constructive detailed comments and suggestions.

I would like to thank Teledyne Optech for providing me the datasets acquired by their newly released multispectral ALS system TITAN. The datasets allowed me to carry out a novel research in a timely fashion.

This thesis would not have been possible without the support from many individuals. I would like to thank my colleagues Lanying Wang, Shiqian Wang, Haocheng Zhang, Weikai Tan, Menglan Zhou, Yifei Chen, Yuenan Li, Han Jiang, He Zhao, Lingfei Ma, Zilong Zhong, Renfang Liao for accompanying me during the graduate studies. I would especially like to thank Dr. Su-Yin Tan, Dr. Richard Kelly, and Dr. Xiaoliang Zou for providing suggestions on my study, work, and the thesis.

Finally, I would like to thank my parents for their unconditional support and encouragement throughout my six years of study in Canada. I am deeply grateful for having such wonderful parents who care and support me every step of my life. Also, a special thank you goes to my boyfriend, Xianli Li, for being a sane, smart, and supportive company throughout this journey.

Table of Contents

AUTHOR'S DECLARATION.....	ii
Abstract.....	iii
Acknowledgements.....	v
Table of Contents.....	vi
List of Figures.....	viii
List of Tables.....	ix
List of Abbreviations.....	x
Chapter 1 Introduction.....	1
1.1 Motivation.....	1
1.2 Objectives of the Study.....	4
1.3 Organization of the Thesis.....	4
Chapter 2 Related Work.....	5
2.1 Multispectral Airborne Laser Scanning.....	5
2.1.1 Characteristics of A Multispectral ALS System.....	5
2.1.2 Principle of Multispectral ALS System.....	7
2.1.3 Components of A Multispectral ALS System.....	9
2.1.4 Direct Geo-referencing.....	11
2.1.5 ALS-derived DTM and DSM.....	11
2.1.6 ALS-derived Intensity Imagery.....	14
2.2 ALS in Tree-level Inventory.....	15
2.2.1 Canopy Height Model Generation.....	15
2.2.2 Approaches in Vegetation Isolation.....	16
2.2.3 Approaches in Tree Crown Delineation.....	18
2.2.4 Biomass / Carbon Storage Estimation.....	21
2.3 Chapter Summary.....	25
Chapter 3 Study Area and Data Sources.....	26
3.1 Study Area.....	26
3.2 Multispectral ALS Datasets.....	26
3.3 Field Measurements.....	28
3.4 Chapter Summary.....	30
Chapter 4 Estimation of Carbon Storage in Urban Trees.....	31

4.1 ALS Data Processing	31
4.1.1 ALS Data De-noising and Intensity Normalization	31
4.1.2 ALS-derived Intensity Imagery	33
4.1.3 ALS-derived Digital Elevation Model	33
4.1.4 nDSM Preprocessing	34
4.2 Vegetation Isolation.....	35
4.3 Dendrometric Parameter Estimation.....	39
4.4 Carbon Storage Modeling.....	42
4.5 Chapter Summary	45
Chapter 5 Results and Discussion.....	46
5.1 Accuracy Assessment for Ground-point Filtering.....	46
5.2 Analysis of Multispectral ALS Data for Land Cover Classification	46
5.2.1 Spectral Analysis of the Land Cover Classes.....	46
5.2.2 Spectral Analysis of the Tree Class	48
5.2.3 Accuracy Assessment of the Classification Results by SVM Classifier	50
5.3 Performance of the Watershed Segmentation	52
5.3.1 Results for the Local Maxima Selection	53
5.3.2 Accuracy Assessment of Segmentation	55
5.4 Validations for the ALS-derived Dendrometric Parameters	57
5.5 Validations for the ALS-predicted DBH and Tree Carbon	58
5.6 Analysis of the ALS-estimated Results	61
5.7 Analysis of Carbon Storage	62
5.8 Chapter Summary	64
Chapter 6 Conclusions and Recommendations.....	65
6.1 Conclusions	65
6.2 Recommendations	67
References	68

List of Figures

Figure 2.1 Optech Titan system (reprinted from Teledyne Optech Titan, 2015).....	6
Figure 2.2 Titan spectral bands with general spectral behaviors	7
Figure 2.3 Principle of ALS system.....	10
Figure 2.4 Progressive TIN densification.....	13
Figure 2.5 ALS-derived products: (a) DTM; (b) DSM; and (c) CHM.	17
Figure 2.6 SVM in a two-class linear case	19
Figure 3.1: Study area.....	27
Figure 3.2: Location of 40 sampled trees in the field	29
Figure 4.1 Workflow of the proposed methodology.....	32
Figure 4.2 Pits in the raw nDSM.....	34
Figure 4.2 Examples of ALS-intensity-derived indices.....	36
Figure 4.3 Multispectral ALS intensity on waterbodies (green band).....	37
Figure 4.4 Examples of a 3×3 local maxima filtering on CHM	40
Figure 4.5 Measurement of tree height and crown diameter	41
Figure 5.1 Illustration of the ground filtering result and two validation sites	47
Figure 5.2 Histograms for five classes	49
Figure 5.3 Histograms for the tree class	51
Figure 5.4 Trees with dark-colour leaves in Google Earth and the pNDVI	52
Figure 5.5 Classification map of the study area	53
Figure 5.6 Results obtained by local maxima filtering and watershed segmentation	56
Figure 5.7 Residual plots for model-predicted DBH.....	60
Figure 5.8 Scatterplot of the ALS-modeled DBH vs. field-measured DBH	60
Figure 5.9 Scatterplot of the ALS-predicted carbon vs. field-estimated carbon	60
Figure 5.10 Carbon storage map for the Town of Whitchurch-Stouffville, Ontario	63

List of Tables

Table 2.1 Titan configurations	7
Table 2.2 Summary of the methods listed in Section 2.2.4	23
Table 3.1 Summary of the multispectral Titan ALS dataset	29
Table 3.2 Statistics of the field measurements for 40 sampled trees	29
Table 4.1 Description of the land cover types and corresponding training samples	38
Table 4.2 Combinations of input data for SVM classification	39
Table 4.3 Parameters for all-species equations listed in Lambert et al. (2005).....	44
Table 5.1 Accuracy assessment for ground-point filtering.....	47
Table 5.2 Confusion matrix for the SVM classification using six input data.....	54
Table 5.3 Accuracy report for the classification results.....	54
Table 5.4 Generation of tree-isolated nDSM.....	55
Table 5.5 Evaluations of the segmentation result	57
Table 5.6 Validation statistics for the ALS-derived dendrometric parameters	58
Table 5.7 Results of model fitting and model validation.....	59
Table 5.8 DBH model developed by all-field measurements and validation statistics ...	59
Table 5.9 Accuracy of ALS-predicted vs. field-measured results	61

List of Abbreviations

AGC	Automatic gain control
ALS	Airborne laser scanning
CD	Crown diameter
DBH	Diameter at breast height
DTM	Digital terrain model
DSM	Digital surface model
GNSS	Global navigation satellite system
GPS	Global positioning system
FOV	Field of view
H	Height
IDW	Inverse distance weighted
IMU	Inertial measurement unit
kt C	1000 tonne carbon
LiDAR	Light detection and ranging
nDSM	Normalized digital surface model
NIR	Near infrared
pNDWI	Pseudo normalized difference water index
pNDVI	Pseudo normalized difference vegetation index
RBF	Radial basis function
RMSE	Root mean square error
SWIR	Shortwave infrared
SVM	Support vector machine
TIN	Triangulated irregular network
TLS	Terrestrial laser scanning
TOF	Time of flight

Chapter 1

Introduction

1.1 Motivation

With the continuing growth of global population, urbanization becomes an inevitable trend. Urban expansion accelerates as the advancements of socioeconomic activities and population growth. Today, over 50% of the global population resides in cities where by 2050 urban areas will hold up to another 2.5 billion people which in total is equivalent to 66% of the global population (United Nations, 2014). As substantial urban expansion undergoes, ecosystem and global land cover have been altered consequently (Carlson & Arthur, 2000). Land cover transition takes place everywhere from natural environment to diverse urban land types such as impervious surface, build-up areas, recreation areas, and urban green spaces. Meanwhile, substantial urban growth brings pressing problems to the world today such as the urban heat island effect. Intensive urban developments and economic activities increase energy consumption and result in greenhouse gases emissions (Corfee-Morlot et al., 2009; Scatterwaite, 2009; Dodman et al., 2012). Urban development becomes the biggest contributor to global carbon emissions while the energy sector becomes the largest source of greenhouse gases which contributes about 85% of Canada's total greenhouse gas emissions (Environment Canada, 2015). Besides, as urban land covers expand, direct losses of vegetation cover occur. Without vegetation covers act as the largest carbon sink, deforestation becomes the second largest contributor of greenhouse gases (Johnson, 2009).

Increased awareness of high energy use and urban heat islands has led urban municipalities to put conservation efforts in implementing strategies that can protect urban climate and energy use at the city scale. Urban vegetation has drawn direct attention of city planners and policy makers by considering ecological sustainability and energy conservation. Trees play an important role in urban climate modification and energy conservation in different ways. For instance, tree shading and wind shielding effects can attenuate the direct solar heat and air infiltration into individual houses (Sawka et al., 2013). In the context of Canada's climate, there are high demands in both heating and cooling energy for indoor

climate controls (Akbari et al., 2012). By planting vegetation proximate the houses and increasing the albedo of urban surfaces, annual cooling energy use can be reduced 10-19% (Xu et al., 2012; Sawka et al., 2013). Strategic planting can make full use of tree shading and aggregate canopy covers to receive energy conservation benefits. Trees also facilitate evapotranspiration and airflow dynamics (Sawka et al., 2013). Urban thermal comfort can be significantly improved by the effect of evapotranspiration. In city wide, urban trees reduce air pollution through direct dry deposition and also influence the cooling of the ambient temperature which slows the smog formation. Vegetation contributes the largest proportion of carbon storage which in return reduces the rate of climate warming and urban heat islands (Davies et al., 2011). Urban trees both sequester CO₂ and store excess carbon in biomass (71% of total urban carbon storage) which significantly influences environmental quality and human health (Donovan & Butry, 2010; Roy & Pickering, 2012). Preserving carbon storage and also improving green space infrastructure in urban area has significant environmental and substantial benefits (van den Berg et al., 2015). Consequently, estimation and monitoring of urban carbon stocks and green space become important indeed. First quantification of carbon storage in urban trees was conducted in 1990s (Nowak et al., 1993). In Canada, Environment Canada conducts national scale estimations of urban forest carbon sequestration rate and the greenhouse gas emissions and removals at an annual base (Pasher et al., 2014). However, the estimations of tree carbon stocks require excessive field data which could be time-consuming and labor-intensive. Moreover, the estimation of dry biomass often requires destructive measurements. There is a clear need for techniques to efficiently and cost-effectively assess urban trees in both individual and regional scales. Especially, the ability of assessing carbon stocks at individual-tree level is very important since it takes vegetation structures and fine-scale heterogeneous composition into account (Davies et al., 2011). In literature, the carbon content stored in individual trees can be assessed through aboveground dry-weight biomass calculation using allometric equations. Dendrometric parameters such as individual tree height or crown diameters are generally used in the allometric equation to derive diameter at breast height (DBH). Dry biomass then is calculated using allometric

model with DBH as the input and further transformed to carbon storage by a conversion rate around 0.5 (Lieth, 1963; Whittaker & Likens, 1973).

With the development of light detection and ranging (LiDAR) or laser scanning technologies, airborne laser scanning (ALS) and very high resolution multispectral imagery have proven to be promising tools for the derivation of dendrometric parameters in forest. Typical ALS technology provides detailed topographic information through the dense and accurate georeferenced point clouds and ALS-derived digital elevation models (DEM). Meanwhile ALS intensity records the backscattered energy from laser pulses. However, since the response in single wavelength (typical ALS system emits laser pulses at wavelength in 1064 nm) may not be sufficient to identify different targets, the reflectance properties of ALS data are less frequently used in applications. Data fusion technique especially has overcome the limitation of single source data by taking advantages of both spectral information from imagery and range information from ALS point clouds. In terms of vegetation analysis, previous studies have successfully extracted detailed vegetation covers from very high resolution multispectral satellite imagery such as QuickBird and applied it to the ALS-derived canopy height model (CHM) for dendrometric measurements and biomass estimation in both forestry and urban studies (Huang et al., 2013; Schreyer et al., 2014; Pasher et al., 2014; Sousa et al., 2015). However, data discrepancies have always existed due to data inconsistency in terms of cloud occurrence, shadow effects, and inconsistent spatial and temporal resolution. With the emerging multispectral ALS technology, it becomes possible to obtain both the range and multiple reflectance data from a single data source. The first commercial multispectral ALS system, Titan, released by Teledyne Optech Inc. has integrated three laser channels at wavelengths of 532 nm, 1064 nm, and 1550 nm, respectively. This emerging ALS system showed great potential in land cover mapping without the aid of passive multispectral images (Bakula, 2015; Wichmann et al., 2015). This active laser system can largely avoid those factors commonly associated with passive optical sensors such as weather conditions and shadow effects. However, the potential of ALS intensity currently remains undervalued and the applications in use of this newly released system are at an early stage of development. Given the current development of tree inventory

with typical ALS data, this study aims to explore the feasibility of multispectral ALS range and intensity data in carbon storage estimation.

1.2 Objectives of the Study

The primary objective of this study was to develop a workflow that can quantify the carbon storage in urban trees using multispectral ALS data. To achieve this, this study aimed to first classify vegetation covers based on multispectral ALS range and intensity data by applying Support Vector Machine (SVM) classifier. Secondly, this study aimed to derive dendrometric parameters such as tree height and crown diameter in order to establish an allometric relationship between ALS-derived measurements (tree height and crown width) and the field-measured parameter (DBH) through regression modeling. Lastly, this study quantifies the carbon storage in urban trees for the Town of Whitchurch-Stouffville, Ontario.

1.3 Organization of the Thesis

This thesis consists of six chapters:

Chapter 2 gives a literature review on the development of multispectral ALS technology with an emphasis on the generation of ALS-derived DEM and related considerations. It also reviews current studies in land cover mapping and biomass estimation in use of both ALS and imagery.

Chapter 3 provides a description of the study area, a summary of the multispectral ALS dataset, and the statistics of the field measurements.

Chapter 4 describes the proposed workflow which consists of four steps: ALS data processing, vegetation isolation, dendrometric parameters extraction, and carbon storage estimation.

Chapter 5 presents results of the study including accuracy assessment of the land cover classification, validation of the tree-crown segmentation, validation of the ALS-DBH linear regression modeling, and the final calculation of the urban tree carbon storage. It also discusses the key findings of this thesis.

Chapter 6 concludes the deliverables of the study and offers recommendations for future works.

Chapter 2

Related Work

In this chapter, a fundamental working principle of the typical ALS system and an introduction to the Teledyne Optech Titan multispectral ALS system are given. The applications of land cover mapping and forest inventory in use of ALS data are reviewed with an emphasis of the approaches in ALS data processing, individual tree segmentation, dendrometric parameter estimation, and allometry-based biomass modeling.

2.1 Multispectral Airborne Laser Scanning

Airborne laser scanning (ALS) is an active-sensor mapping system that provides 3D information of the surface through emitting laser pulses towards the surface and recording the travel time of the return signals. Meanwhile, the tilting direction of the laser beam and the location of the flight are recorded by the system. The ALS system is essentially featured by the penetration capability of laser pulses into the terrain surface. As an active remote sensor, the ALS system can be operated at day and night which successfully avoids the influences by shadow and solar illuminations. Today, ALS technology becomes well established and has opened a new era in the field of forestry, bathymetry, urban management, and so on. This section first introduces the multispectral ALS technology, giving focuses especially on the characteristics of a multispectral ALS system, the principle and system components of the ALS systems, and the direct geo-referencing of the ALS data. The accuracy and generation of ALS-derived data are discussed with specific considerations related to this study.

2.1.1 Characteristics of A Multispectral ALS System

In the development of ALS technology, a new resolution occurs when the operation of lasers with multiple spectral bands becomes available. A typical ALS system usually operates a near-infrared (NIR) laser scanner in topographic mapping where complete absorption occurs in water areas. With the addition of a green-band laser (green lights can penetrate into shallow water), it enabled applications in bathymetry. The multispectral ALS data utilized in this study was generated from the first commercial multispectral active imaging sensor:

Optech Titan (see Figure 2.1). Detailed specifications are listed in Table 2.1. Optech Titan operates three laser channels at wavelength: 1550 nm (shortwave infrared, SWIR), 1064 nm (near infrared, NIR), and 532 nm (green), respectively. These three channels produce independent scan lines by sending pulses with separate forward angles (NIR channel has 0° forward angle; SWIR channel has 3.5° forward angle; green channel has 7° forward angle). It also enabled three independent intensity profiles generated at these wavelengths. Hence, besides the 3D information, the spectral characteristics of objects at different spectral bands became available. The spectral behaviors of basic land cover types at Titan operated spectral wavelengths are illustrated in Figure 2.2. NIR and SWIR lights tend to be absorbed completely at water surface while the green light can penetrate into the water body at certain depths (up to 15 m). Vegetation tends to reflect most at the NIR band but lowest at the green band. Since Optech Titan is released in 2014, its potentials in topographic mapping and land cover classification have been actively evaluated (Bakula, 2015; Wichmann et al., 2015).

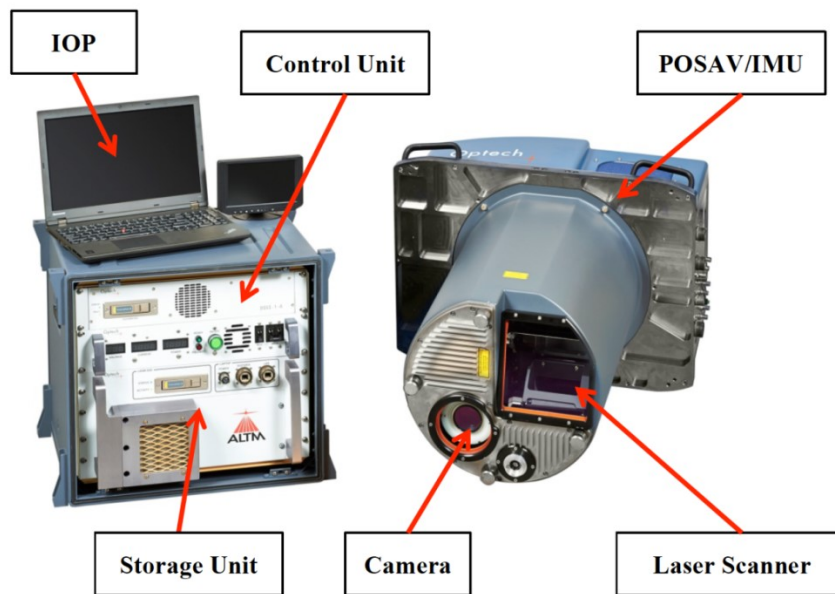


Figure 2.1 Optech Titan system (reprinted from Teledyne Optech Titan, 2015)

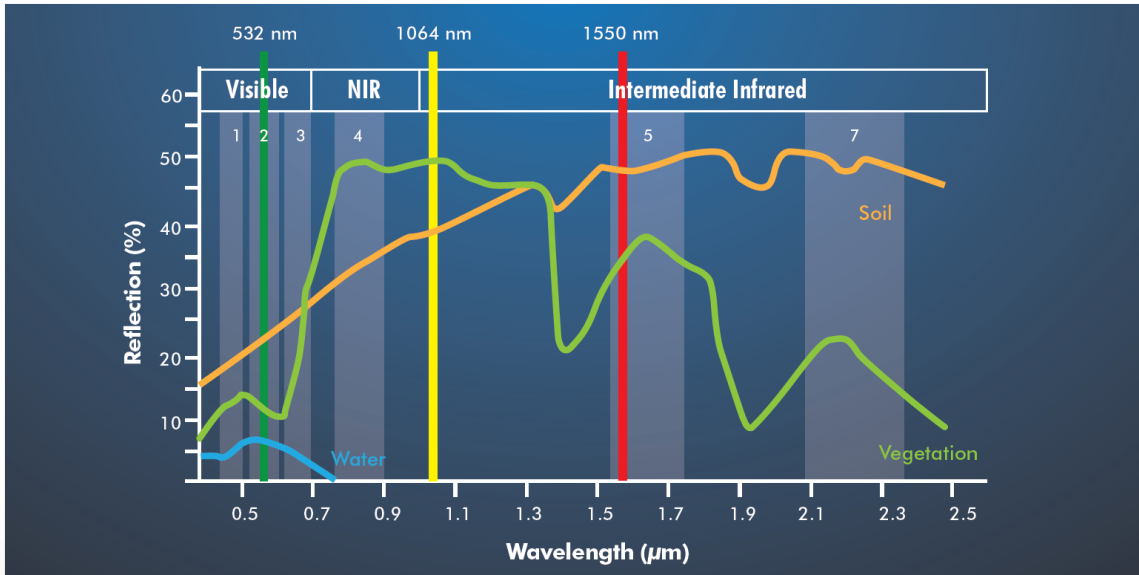


Figure 2.2 Titan spectral bands with general spectral behaviors for vegetation, soil, and water (reprinted from Teledyne Optech, 2015)

Table 2.1 Titan configurations

Laser channels	1550 nm, 1064 nm, 532 nm
Scan angle	0 - 60°
Pulse repetition frequency (PRF)	300 kHz per channel, 90kHz total
Accuracy	Horizontal: 1/7, 500 × altitude, 1 σ Vertical: < 5 - 10 cm, 1 σ
Range precision	< 8 mm; 1 σ

2.1.2 Principle of Multispectral ALS System

A multispectral ALS system follows the standard ALS principles that measures the ranges to an object mainly based on the light transit time estimation method, also known as time-of-flight (TOF) method (Beraldin et al., 2010). Regardless of the scanning platform, the fundamental principle of range measurement by an ALS system can be described by

$$\rho = \frac{c \tau}{n 2} \quad (2.1)$$

where c is the speed of light in vacuum, n is the correction factor that applied if light travels in a medium; ρ is the range, and τ is the time that light travels from the sensor to a target and back to the sensor (also known as TOF). Hence, with $n = 1$ and the flying height is 1000 m, the TOF for a single pulse is 6.7 μs . For precise mapping, the position and orientation of the aircraft have to be recorded at certain level of accuracy. The laser light source is more directional and coherent than sunlight which allows the laser pulses to keep high spatial coherence when it hits and transmits back. The spatial coherence is dependent with the wavelength of the laser source that shorter the wavelength can create longer the coherence length. The multispectral ALS systems now integrate eye-safe lasers with wavelengths in the range of 0.4 to 1.6 μm . The strength of a reflected laser pulse is recorded as intensity values which measures the reflectivity of surface targets at the laser-specific wavelength. The intensity data provides complementary information to the 3D measurement which helps differential objects with similar elevation but different reflectivity. Meanwhile, the scanner emits pulses with certain scanning angles to optimize both the across- and along- track range (see Figure 2.3). A relationship between swath width sw of a scanner and the scanning angle can be established as

$$sw = 2h \tan \frac{\theta}{2} \quad (2.2)$$

where θ is the scan angle and h is the vertical height between the ground and the aircraft. Similarly, the laser footprint which is the size of laser pulse projected on the ground is described by

$$D = 2h \tan \frac{\gamma}{2} \quad (2.3)$$

Where γ is the laser beam divergence (typically ranges from 0.1 to 1 mrad) and h is the vertical height between the ground and the aircraft. Multiple signal returns can be generated by the TOF range system when the rises of signals occur. Therefore, not only the first hit on a tree branch is recorded, the following hits (up to four echoes) penetrated into the canopy will also be recorded.

2.1.3 Components of A Multispectral ALS System

The basic components of a multispectral ALS system include:

- 1) a scanner assembly,
- 2) an airborne Global Navigation Satellite System (GNSS) antenna,
- 3) an inertial measurement unit (IMU),
- 4) a control and data recording unit,
- 5) an operator laptop, and
- 6) a flight management system (see Figure 2.3).

The scanner assembly consists of the lasers and the scanning mechanism. As the aircraft flies, the scanner assembly sends laser pulses continuously towards the terrain. For the Titan system, a point density of 25 points/m² can be reached from a survey with a flying height of 1000 m, PRF of 900 kHz, field of view (FOV) of 30°, and a cruising speed of 60 m/s. The scanning assembly is mounted in the fuselage pointing exposed to the ground. The laser scanner and digital cameras form the imaging components of the ALS system. Range accuracy is mainly affected by the incidence angle of the scan head to the surface of interest (Beraldin et al., 2010). Here, scanners can also be grouped based on two types of range measurement techniques. The TOF scanners calculate the range based on the travel difference of laser pulses. Majority of scanners follow this concept to determine the range. The phase-shift based scanners calculate the phase difference of emitted modulated sinusoidal laser pulse using a phase shift principle. This technique can achieve better ranging accuracy but shorter measurement range. The laser scanners generally collect the amplitude of return points forming the ALS intensity data. The digital cameras can provide additional visual information that often images are geo-referencing with MLS points to depict points with detailed colours.

The GNSS antenna is mounted at the top of the aircraft creating undisturbed connection with GNSS satellites. GNSS has included a series of global navigation satellite system such as U.S. GPS (Global Positioning System) and Russian GLONASS. The GNSS receivers can record: time, position and velocity of the aircraft. By acquiring satellite signals, the GNSS can provide exactly the position and orientation of the ALS system. The antenna typically

receives GPS signals at the frequency of 2 Hz which is equivalent to locate positions with 35 m spacing under an aircraft speed of 70 m/s. To reconstitute the flight trajectory with complete positioning records, IMU system is introduced. IMU records: roll, pitch and yaw angles of the sensor platform. These values can derive the aircraft acceleration rate and rotation rate which could further record the positioning information of the system. The GNSS/IMU

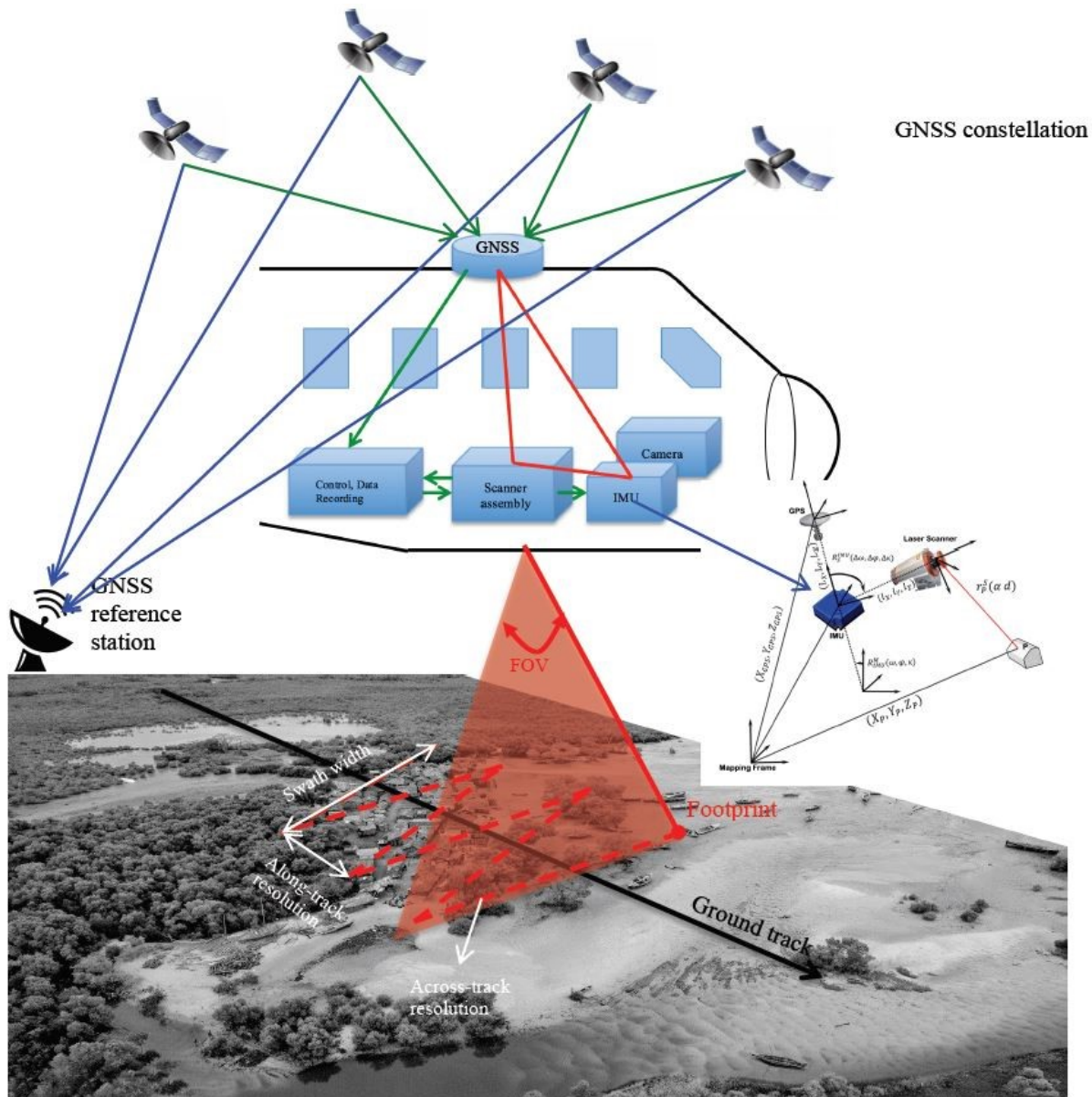


Figure 2.3 Principle of ALS

components control geo-referencing to the point clouds. Once satellite signals are blocked, the IMU will aid in providing location information. Since the GNSS signals have a frequency rate of 1-10 Hz while the IMU typically records orientation and position at rate of 100-2000 Hz, as flight speed increases, the IMU can record more precise positioning information with much smaller timescale. The trajectory is monitored mathematically using algorithmic filtering and smoothing techniques such as Kalman filter (Skaloud & Schwartz, 2000). The control and data recording unit controls the whole system meanwhile it stores ranging and positioning data. The operator laptop displays the system performance. And the flight management system provides real-time survey conditions to the pilot.

2.1.4 Direct Geo-referencing

Before delivering ALS data to end users, the coordinates of laser points have to be assigned by the control unit of the navigational system. This process refers to direct geo-referencing.

The coordinate of a target $p(t)$ is calculated by

$$p(t) = p_N(t) + R_N(t) \cdot R_S(t) \cdot r_L(t) \quad (2.4)$$

where $p_N(t)$ refer to the position of the laser scanner at time t in a Cartesian geographic coordinate system, $R_N(t)$ refers to the rotation matrix that describes the orientation of the laser channel in 3D space, $R_S(t)$ refers to the rotation matrix that describes the orientation of the scanning mechanism and the scanning process, $r_L(t)$ refers to the distance measured by the laser scanner at time t (Hebel & Stilla, 2012).

The accuracy of laser points is influenced by several error sources in the ALS system. for example, the positioning errors for a 500-m-height survey can come from the GNSS related errors (3-30 cm), the IMU related errors (10-20 cm), lever arms (2-3 cm), and footprint size (10-20 cm) (Hebel & Stilla, 2012).

2.1.5 ALS-derived DTM and DSM

Besides the point cloud itself, various digital products can be generated from the ALS data; such as digital terrain model (DTM), digital surface model (DSM), and ALS intensity imagery. The DTM refers a digital representation of the bare earth that has ground surface without objects attached (Sithole & Vosselman, 2005). Before, the photogrammetry was

mainly used to derive topographic information for large spatial coverage. In recent years, the generation of DTM from ALS data becomes an available approach. Various factors can affect the quality of the ALS-derived DTM such as ALS point density, flight height, scan angle, surface complexity and so on. Specifically, the amount of returned ground signals decreases as the scan angle increases (Hyypä et al., 2008). In terms of surface complexity, tree canopy and understory vegetation can significantly influence the accuracy of ALS-derived DTM since fewer points can reach the ground. There are several steps and considerations towards the DTM generation by ALS systems. First, filtering is needed to classify points into terrain and off-terrain. Subsequently, interpolation is required to convert terrain points into digital format. And importantly, these steps are required to be automatic in order to process huge amounts of data.

A series of terrain-point filtering methods have been proposed and can be grouped into different streams. Slope-based filtering methods utilized the concept of mathematical morphology to calculate the distance and height difference between two points within a predefined structure element; and by moving the structure element onto each point, the point with height exceeded the predefined threshold was labeled as off-terrain points (Vosselman, 2000; Sithole, 2001). Progressive densification based filtering methods (such as the progressive triangulated irregular network (TIN) densification method) allow the terrain to grow from a coarse approximation and then iteratively refine the criteria. (Axelsson, 2000; Sohn & Dowman, 2002). A TIN was formed by triangulating points iteratively until no points can be added into the TIN if certain constrain of description length, spline function or height difference is met. For example, as shown in Figure 2.5 three angles α_1 α_2 α_3 and the height difference between a point P and the triangle facet are calculated if the point P met all the predefined thresholds, P is added to the terrain TIN.

Surface based filtering methods reconstruct the DTM by defining both the functional model and the weight model. Elmqvist et al. (2001) developed the active shape models that reconstruct the surface by an energy function that is controlled by both inner and external forces. Segment-based filtering methods treat a group of points with similar characteristics as one segment. Segmentation is performed first according to predefined criteria to generate

segments. Then segments were classified by different approaches into terrain and off-terrain segments (Sithole & Vosselman, 2005; Tovari & Pfeifer, 2005). In general, the TIN-based and surface-based filtering methods worked better than others in the DTM extraction (Sithole & Vosselman, 2005).

Once terrain points are filtered out of the point cloud, interpolation methods are applied to create a digital representation of the terrain surface. Four commonly applied interpolation methods in the DTM generation were selected to give an introduction here: point-to-raster interpolation, nearest neighbor interpolation, inverse distance weighted (IDW) interpolation, and linear interpolation. The point-to-raster interpolation method simply assigns cell values based on the statistical properties of the points within each cell. For example, the raster cell value may be defined as the minimum, maximum, or average of the total points within it. This method is computationally efficient. The nearest neighbour interpolation assigns void pixels with the values of its nearest neighbor. This interpolation method is computationally efficient but may generate considerable errors when the data contains gaps. The IDW is the most commonly applied interpolation method that calculates the cell values according to user-defined weighting rules and searching distance. Hence, the result depends on user-tunable parameters and point density variation. The linear interpolation method first computes a TIN surface of the point cloud and finds the triangle that the point located in. The value of the point is linearly interpolated from its barycentric coordinates using the nodes of

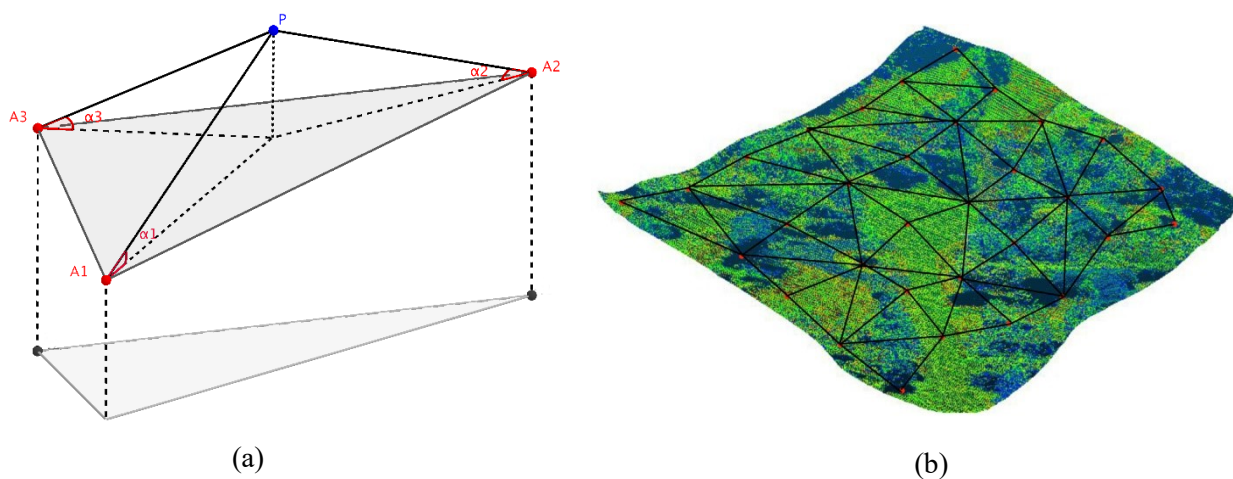


Figure 2.4 Progressive TIN densification

the triangle.

The generation of DSM technically follows the process as for DTM except the selection of interpolation methods. Since DSM encloses the elevations of terrain surface with objects attached, interpolation should be chosen application by application to fit different needs of surface texture. However, due to the vertical variability existed in vegetation structures the interpolation process (such as using the point-TIN-raster method) generally produces underestimation of the surface elevation in vegetated area. Assigning the highest point to the cell and choosing the cell size close to the point spacing can help eliminate this problem (Heidemann et al., 2012).

2.1.6 ALS-derived Intensity Imagery

As discussed previously, the multispectral ALS system records the amplitude of the return pulse and stores it as intensity. The intensity of ALS represents the reflectivity of the target and provides ancillary information to the point cloud. However, unlike reflectivity collected by passive sensors (e.g. QuickBird, camera), the reflectivity data by the ALS system may not look natural and easy to interpret. It is because as an active sensor its light energy may vary 15-20% pulse by pulse (Heidemann et al., 2012). Also, the laser pulses are sent with oscillating incident angles (e.g. 0° to $\pm 30^\circ$) within a single scan swath. For different scan lines, variation in flight heights can also contribute to variation in the returned energy. Comparing to the illumination by Sun, passive sensors can record reflected energy with constant angle of incidence and constant distance relative to the sensor. Hence, extensive variation of recorded intensity can exist within a scan swath or overlap swaths making the intensity calibration difficult. Meanwhile, the ALS system incorporates an automatic gain control (AGC) to adjust the sensitivity of the detector for receiving the return energy (Heidemann et al., 2012). The recorded intensity is modified and compressed into 16-bits accordingly.

The generation of the ALS intensity imagery could follow the same technique for DTM and DSM. ALS-derived intensity imagery refers to the conversion of 3D point into 2D imagery which assigns the ALS intensity instead of elevation as the cell value. Since the laser

band width is extremely narrow (± 20 nm), the typical ALS intensity imagery is similar as one NIR band from hyperspectral imagery (Heidemann et al., 2012). In terms of the multispectral ALS sensors such as Titan, three intensity images can be generated by the NIR, SWIR, and green band. The single NIR ALS intensity showed less importance in previous studies, however, the multiple intensity profiles from the multispectral ALS have showed great values in land cover mapping (Wichmann et al., 2015; Bakula, 2015).

2.2 ALS in Tree-level Inventory

Laser technology has first become popular in forestry sector towards large-scale and cost effective forest inventory since 1980s (Aldred & Bonnor, 1985). Forest inventory mainly covers the information of tree, stand, and large plot profiles. Given the scope of the proposed study, focuses are placed on previous studies conducted at the tree level. Before deriving any estimation from the ALS point cloud, DTMs/DSMs are often generated by classifying points into surface and terrain. Besides forestry studies, the step of vegetation isolation is required in urban vegetation analysis. The tree height and crown diameter are the two most commonly-derived dendrometric parameters from ALS data. From dendrometric parameters, diameter at breast height (DBH), tree volume, and biomass can be derived by using allometric equations. Allometry is the study that finds the relationship between tree species and dendrometric parameters, usually using tree height and stem or base diameter (Heidemann et al., 2012). To derive vegetation measurements such as tree carbon storage, four steps: CHM generation, vegetation isolation, dendrometric parameter calculation, and biomass modeling are needed. Here, this review discussed considerations and different approaches in each of the step.

2.2.1 Canopy Height Model Generation

The normalized digital surface model (nDSM) or the canopy height model (CHM) in forestry is obtained by subtracting the DSM using the corresponding DTM (see Figure 2.6). CHM generally forms in a raster or TIN height format (Hyypä et al., 2008). The difference between the elevation of laser pulse and estimated terrain can indicate the elevation of the above-ground objects, such as trees and manmade objects. The CHM-derived tree height was

reported to have an underestimation in general (Leeuwen & Nieuwenhuis, 2010; Kankare et al., 2013; Estornell et al., 2015). The underestimation of tree height varies from 0.63 m for coniferous trees and 0.43 m for deciduous trees with a sample size of 967 trees (Heurich et al., 2008). The underestimation of tree height is caused by several factors such as point density, flying height, the penetration of the laser pulse, and so on. As the flying height elevated from 400 m to 1500 m, the accuracy of tree height lowered from 0.76 m to 1.16 m (Yu et al., 2004). The accurate ALS-derived raster data such as CHM forms the base of the estimation of dendrometric parameters and the performance of relevant empirical models. Different methods such as distribution-analysis based or individual-tree based techniques can be applied to derive tree information from the laser point cloud. However, all of the methods rely heavily on the CHM data that stores canopy heights in pixels. Hence, the generation of nDSM or CHM acts as the dominant element that provides valuable information towards the derivation of dendrometric parameters.

2.2.2 Approaches in Vegetation Isolation

For urban scenes, the process of vegetation isolation is required to differentiate vegetated area from other land cover types. Due to the lack of spectral information of typical ALS data, data integration that combines multispectral images with ALS data is often applied in order to extract vegetated areas from urban environments. Recent studies in the urban tree carbon storage estimation apply thresholding to the normalized difference vegetation index (NDVI) data which are derived from high-resolution multispectral images (e.g., QuickBird) to identify vegetation areas (Chen et al., 2009; Huang et al., 2013; Raciti et al., 2014; Schreyer et al., 2014; Zhang et al., 2015). The NDVI can be calculated by

$$NDVI = \frac{\rho_{\lambda_{NIR}} - \rho_{\lambda_{Red}}}{\rho_{\lambda_{NIR}} + \rho_{\lambda_{Red}}} \quad (2.5)$$

where ρ refers to the spectral reflectance at wavelength λ .

Raciti et al. (2014) identified vegetated areas from QuickBird imagery by assuming the NDVI value exceeds 0.1. Simple thresholding on NDVI keeps computation of the vegetation isolation process simple and productive. However, fine detailed urban vegetation mapping

generally requires both height and spectral information with the use of advanced classification methods (Yan et al., 2015). The nDSM has shown its value in improving land cover classification accuracy (Huang et al., 2008; Chen et al., 2009; Hartfield et al., 2011). Hartfield et al. (2011) demonstrated the integration of nDSM can increase the overall classification accuracy by 5%. However, the addition of single ALS intensity has been reported to contain certain noises in vegetation areas (as discussed in Section 2.2) which may affect the overall land cover classification performance (Yan et al., 2012).

Various classification methods have demonstrated their functionalities in classifying urban vegetation. These methods can be grouped into two streams: pixel-based classification and object-based classification methods. Lodha et al. (2006) achieved above 90% accuracy by applying the SVM classifier to generate four land cover classes (building, tree, roads, and grass) using ALS-derived height data. The SVM classifier was initially developed to solve two class problems (Kavzoglu & Colkesen, 2009). In a two-class case, SVM searches for an optimal hyper plane that can separate two classes with the maximum margin width (see

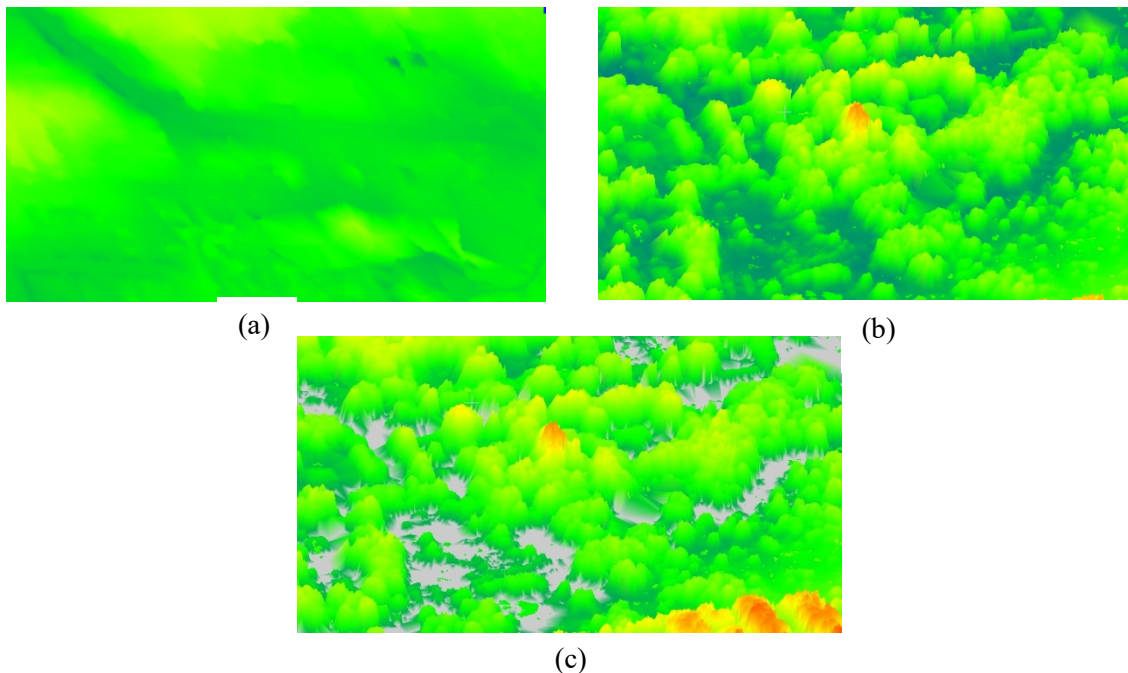


Figure 2.5 ALS-derived products: (a) DTM; (b) DSM; and (c) CHM representing the normalized above-terrain heights.

Figure 2.7). The optimal hyperplane is defined as $w \cdot x + b = 0$ where w is the orientation of the hyper plane, b is the bias of the distance of the hyper plane to the origin, and x refers to the point lying on the hyper plane. Support vectors are the training points lying on two hyper planes which are parallel to the optimal hyper plane. The two hyper planes are defined as $w \cdot x + b = \pm 1$ which can separate points into -1 and 1 classes. Therefore the margin width between the two classes equals to $2/w$ and the optimal hyperplane can be derived. To achieve nonlinear multi-classes separation with remote sensing data, kernel-based mapping functions are introduced and largely dominates the performance of SVM. Kernel types in SVM include namely, linear, polynomial, radial basis function (RBF) and sigmoid kernels.

Guo et al. (2011) examined the feature importance for complex urban environment by applying the Random Forest (RF) classifier on ALS and multispectral aerial imagery. The RF classifier is chosen due to its capacity of handling large dataset, computational efficiency, and its comparable-to-SVM classification accuracy. With the increase of spatial resolutions, object-based classification methods become popular than pixel-based method to classify objects consisted of several pixels. Moreover, object-based methods take the spatial information into account which can largely overcome the salt-and-pepper noises inevitably associated with traditional pixel-based classification methods (Blaschke et al., 2010). Tong et al. (2014) tested the performance of both object-based classification and SVM classification using the data fusion of aerial photo and ALS data. The object-based method consisted of a segmentation process and a decision tree classification. The result outperformed the SVM classification. Zhou et al. (2013) proposed an object-based approach that used ALS-derived height and intensity solely to classify four urban land covers (e.g., building, pavement, tree, and shrub) with an above 90% overall accuracy. However, the performance of object-based classification methods was influenced by the segmentation results or the user-defined rules especially in highly heterogeneous landscapes (Blaschke et al., 2010).

2.2.3 Approaches in Tree Crown Delineation

Once the vegetation can be classified and isolated from the study area, the vegetation-isolated CHM can provide a representation of the surface of canopy only. In order to obtain information at the individual tree level, tree segmentation methods are implemented to

identify individual tree crown. Numerous tree detection and delineation algorithms have been proposed to first detect treetops and then delineate tree crowns from optical images in the past two decades. These algorithms have also been applied and shown good performance on the ALS-derived CHMs. Local maxima filtering (Culvenor, 2002; Wulder et al., 2000; Popescu et al., 2003), template matching (Pollock, 1996), valley-following algorithm (Gougeon, 1995), region growing segmentation (Solberg et al., 2006), and marker-controlled watershed segmentation (Meyer & Beucher, 1990) especially have drawn great attentions.

Local maxima filtering has been widely adapted by many crown delineation methods as the first step to locate the position of treetops (Chen et al., 2006; Zhao et al., 2009; Dalponte et al., 2011; Xu et al., 2014; Zhang et al., 2015). Treetop positions assumed to have higher values in the CHM were identified by finding the local maxima within a moving window with fixed size. The local maxima filtering was largely influenced by the selection of window size and the degree of smoothing of the dataset (Koch et al., 2014). Commission errors might be presented by detecting non-treetop points as treetops (Chen et al., 2006). Valley-following tree crown delineation algorithm finds local minima as the valley instead of finding the local

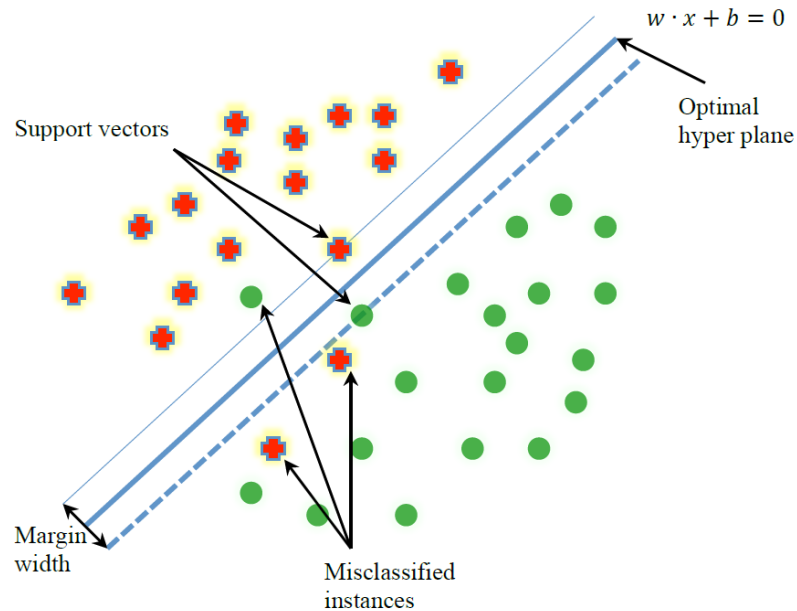


Figure 2.6 SVM in a two-class linear case

maxima. The valley pixels can grow and separate the mountain-like tree crowns. The region-growing segmentation first placed seeds at locations generated by the tree detection results. The seeds acted as the starting position where a region can grow as long as it met the criteria to stop. Solberg et al. (2006) implemented the region-growing based algorithm using the ALS-derived canopy surface model to derive individual tree heights with around 66% detection rate. Dalponte et al. (2011) performed seed region growing method on a convoluted CHM to delineate tree crowns where the seeds were selected by thresholding. Raciti et al. (2014) used object-based segmentation algorithms to hierarchically segment nDSM height metrics into various sizes of vegetation objects where a region growing method was used to select the seed within each object. The watershed segmentation is one of the traditional region-based segmentation approaches introduced by Beucher and Lantuejoul (1979). The fundamental idea of this watershed segmentation is to consider the greyscale image as a surface consisting of catchment basins and watershed lines. Water tends to flow from the regional minima across the entire image. Once water from different regional minima is about to merge, dams are built along the boundary of different catchment basins which forms the watershed lines. The presence of many regional minima in the digital images gradients may usually lead to overestimating the number of catchment basins. This may due to the noise or local irregularities in the image (Ke & Quackenbush, 2011).

To reduce the oversegmentation, Meyer & Beucher (1990) introduced the marker-controlled watershed segmentation. This marker-controlled watershed process could be explained under the water flow scenario as well. Water is raised from the marker-present regions with different colors. Water tends to flow over a catchment basin if there is no color associated with it. Catchment basins with distinct markers are separated with a watershed line (Najman & Schmitt, 1996). Chen et al. (2006) isolated individual treetops from ALS-derived canopy maxima model using the marker-controlled watershed segmentation with a 64% absolute accuracy. Xu et al. (2014) implemented a watershed segmentation method to delineate individual tree segment on the CHM to derive tree height and crown diameter. The crown diameter was measured by the mean of the maximum crown diameter and the one took at the perpendicular direction. Casas et al. (2016) identified tree crowns in a burned

forest using watershed segmentation algorithm combined with local maxima filtering where the local maxima points were defined by searching in eight neighborhoods.

Kaarinen et al. (2012) tested nine tree detection methods and concluded that the detection result was mainly dependent on the applied algorithm. The authors also concluded that for the ALS point density closed to 8 pts/m² local maxima filtering combining with either region growing or watershed segmentation could yield better tree detection results. Koch et al., (2014) pointed out that the detection of local maxima was the vital step for most of the algorithms. Vauhkonen et al. (2012) concluded that instead of the applied algorithms, forest types, tree density, and the degree of canopy overlapping became important factors for the result of tree detection.

2.2.4 Biomass / Carbon Storage Estimation

Biomass is conventionally measured in the field by destructive cutting the tree off, dry weighing its components, and interpolating its total dry mass. The carbon stored in wood and bark is about 50% of the biomass (Lieth, 1963; Whittaker & Likens, 1973). To avoid destructive sampling, allometry was introduced to establish the relationship between dendrometric parameters and biomass. By measuring dendrometric parameters such as tree height, DBH, and crown base height in the field, foresters can build a predictive model that takes accurately-sampled trees as input to predict the stem volume or biomass at plot or stand level. However, since in-situ measurements are labor-intensive and time consuming for large-scale vegetation inventory, ALS remote sensing became a well-established technology to derive dendrometric parameters in forestry. Allometry is also very important in biomass estimation with the use of ALS technology that it established empirical relationships between both direct and indirect measurements in the ALS data and the biomass. Many studies have confirmed the agreement between ALS-derived dendrometric parameters with field samples (Popescu, 2007; Zhao et al., 2009). The feasibility of using the ALS-derived tree structure parameters to estimate biomass or carbon storage has sequentially been proved (Hauglin et al., 2014; Schreyer et al., 2014; Raciti et al., 2014). Biomass estimation with ALS can be done by both area based methods and individual tree crown based methods (Vauhkonen et al.,

2014). Given the scope of this study was to derive carbon storage at single tree level, focuses are given to previous studies using individual tree crown based methods (see Table 2.2).

Popescu (2007) developed linear regression model to derive DBH and biomass for individual pine trees from ALS-derived crown diameter and height. 43 trees were sampled for DBH and allometry-derived biomass. Field-measured DBH and biomass then were applied to construct regression equations with ALS-derived variables. Strong correlation ($R^2 = 0.87$) was found between the ALS-derived height, crown width and the field-measured DBH. The regression model showed R^2 of 0.87 and a RMSE of 4.9 cm for DBH estimation; also R^2 of 0.88 and a RMSE of 162.72 kg for biomass estimation. Zhao et al. (2009) developed scale-invariant biomass models using the canopy height distribution (CHD) as the independent variable to map biomass at a range of scales. The CHD is defined as the probability distribution of tree heights over a unit area. The reason to use the CHD instead of conventional allometric variables like DBH was that the CHD can be directly measured by the ALS. Zhao et al. (2009) established linear relationships for both pines and deciduous trees between field-measured DBH from 117 sampled trees and ALS-derived height, crown width, and crown base height. The classification of tree species was performed on a QuickBird image with an overall accuracy of 89.7%. The tree height and crown width were identified from the CHM that assuming the crown width was a linear function of height. The crown base height defined as the height of the crown was estimated by a voxel-based height bin approach. The ALS-derived DBH was calculated by the regression model and further the distribution of the ALS-derived DBH was used to model the distribution of CHD at the plot level. For the individual-tree level biomass estimation, they achieved R^2 of 0.8 and 0.88 with RMSE of 237 kg and 138 kg for pines and deciduous species, respectively. Hauglin et al. (2014) estimated the crown biomass of Norway spruce using a similar modeling approach presented by Zhao et al. (2009) and Popescu (2007). Hauglin et al. (2014) constructed two ALS crown biomass model that one was using ALS-derived crown volume as the independent variable to predict the terrestrial laser scanning (TLS) derived crown biomass; another one was using the ALS-derived tree height and crown diameter to predict the TLS-measured DBH and predicting crown biomass using an existing allometric model.

Meanwhile, a total of 85 field samples were split into 68 trees for model fitting and 17 trees for model validation. The R^2 of the ALS-DBH model was 0.84 with a RMSE% of 35.1. Kankare et al. (2013) evaluated the prediction of different combinations of the ALS-derived variables to model four biomass components (stem wood, living branch, canopy, and total biomass). They found that the total biomass was best estimated by tree species, DBH, height, crown area, and the percentage of signal returns at 40% of total tree height. The DBH was calculated from an existing allometric model using tree height. The overall accuracy of the ALS-derived height reported with a RMSE of 2.4 m and RMSE% of 12.6. The estimates of DBH was reported a RMSE of 5.053 cm with RMSE% of 24.95. The estimates of total biomass of pine trees had an R^2 of 0.74 with a RMSE of 41.43 kg and spruce species had an R^2 of 0.82 with a RMSE of 86.58 kg. Casas et al. (2016) proposed a Gaussian process regression model to estimate snag DBH using tree height and crown area which achieved a higher R^2 as 0.81 with a RMSE of 10.6 cm compared with a commonly used allometric approach (R^2 of 0.73 and a RMSE of 16.5 cm). Moreover, Casas et al. (2016) explored the use of ALS intensity and height information to classify snag and live trees. Estornell et al. (2015) estimated the pruning biomass of olive trees by creating separate regression models for both field measurements and ALS data. The regression model established for field measurements used sampled tree heights and crown diameters as independent variables and the sampled pruning biomass as the predicted variable. The ALS-based regression model used the ALS-derived crown area and maximum intensity as independent variables with an achieved R^2 of 0.89 and a RMSE of 2.78 kg.

Studies of mapping carbon storage in urban trees with ALS technology have been actively conducted in recent years (Raciti et al., 2014; Schreyer et al., 2014). Raciti et al. (2014) estimated carbon storage in urban trees for the City of Boston by developing an ALS-height-only regression model to estimate carbon storage across the city. A total 404 accurately segmented tree crowns from the nDSM were split into 284 samples for model fitting and 120 samples for validation. The reason to use a simple linear regression of tree biomass and height was to avoid the influence of crown segmentation results. An R^2 of 0.79

Table 2.2 Summary of the methods listed in Section 2.2.4

Authors	Model Type	ALS-derived Variables	Predicted Variable
Popescu (2007)	linear regression	<ul style="list-style-type: none"> • Height, crown width; • DBH 	<ul style="list-style-type: none"> • Field-measured DBH • Allometry-based biomass
Model Validation: DBH: $R^2 = 0.87$, RMSE = 4.9 cm Biomass: $R^2 = 0.88$, RMSE = 162.72 kg			
Zhao et al. (2009)	linear regression	Height, crown width, crown base height	Field-measured DBH
Model Validation: DBH: Pines: $R^2 = 0.87$, RMSE = 5.62 cm Deciduous: $R^2 = 0.89$, RMSE = 4.80 cm Biomass: Pines: $R^2 = 0.8$, RMSE = 237 kg Deciduous: $R^2 = 0.88$, RMSE = 138 kg			
Kankare et al.(2013)	linear regression	Tree height, canopy size, and canopy density	Biomass by destructive sampling
Model Validation: Height: Bias% = -2.44%, RMSE%= 12.60% DBH: Bias% = -18.51%, RMSE%= 24.85% Biomass: Pines: $R^2 = 0.74$, RMSE% = 26.3% Spruce: $R^2 = 0.82$, RMSE% = 36.79%			
Hauglin et al. (2014)	linear regression	Height, crown diameter	Allometry-based crown biomass
Model Validation: Biomass: $R^2 = 0.84$, RMSE% = 35.1%			
Raciti et al. (2014)	linear regression	Height	Field-estimated biomass
Model Validation: Biomass: $R^2 = 0.79$			
Estornell et al. (2015)	linear regression	Crown area, maximum intensity	Pruning biomass by destructive sampling
Model Validation: Pruning Biomass: $R^2 = 0.89$, RMSE = 2.78 kg			
Casas et al. (2016)	Gaussian process regression	Height, crown area	Field-measured DBH
Model Validation: DBH: $R^2 = 0.81$, RMSE =10.6 cm			

was found between field-estimated biomass and model predicted biomass. This study further linked carbon storage in urban trees to the demographic characteristics that no strong correlations were found in Boston neighborhoods. Schreyer et al. (2014) estimated the carbon storage in urban trees and its distribution was extrapolated to the entire city of Berlin in terms of land use types. This study did not propose a regional-specific allometric model for the study area but applied the ALS-DBH model developed by Zhao et al. (2009) and a carbon allometric model using DBH as the only independent variable. 87% of the modeled DBH showed an underestimation which was further calibrated by a weighted arithmetic average DBH. The carbon storage in urban trees was calculated as half of the model-based biomass regardless of the genus. Meanwhile the crown base height was assumed to be half of the ALS-derived height and the crown width was calculated in 16 directions with a series of criteria instead of the conventional estimation of crown width in 4 directions as proposed by Popsecu et al. (2003) or as used by Xu et al. (2014).

2.3 Chapter Summary

This chapter first introduced the characteristics of the Optec Titan multispectral ALS system and the fundamental backgrounds of the ALS technology in terms of the working principle, system components, direct-georeferencing, and special considerations related to ALS-derived products. This chapter also reviewed the applications of ALS in both forest inventory and vegetation mapping with an emphasis on discussing the methods of vegetation isolation, individual tree delineation, and allometric biomass modeling. A summary of the methods of biomass modeling mentioned in Section 2.2.4 can be found in Table 2.2. Noted by carefully reviewing the literature, it was worth trying to examine the feasibility of the multispectral ALS intensity data in land cover classification. Meanwhile, both the multispectral ALS intensity and the ranging information could show their powers in the carbon/biomass estimation at the individual tree level.

Chapter 3

Study Area and Data Sources

This chapter consists of three sections. Section 3.1 provides a description and a map of the study area. Section 3.2 summarizes the multispectral ALS data used in this study. The field measurements are described in Section 3.3 the sampling process and the statistics of the field data.

3.1 Study Area

The study area of this study is located in the Town of Whitchurch-Stouffville, Ontario (43°58'00"N, 79°15'00"W), which is a municipality in the Greater Toronto Area. Whitchurch-Stouffville is 24 km north of Toronto with a population of 37,628 in 2011 (Statistics Canada, 2015). The study area has a total area of approximately 1 km². The study area is characterized by a typical residential landscape (see Figure 3.1). It contains two water bodies (Musselman Lake and Winsor Lake) and three land cover types: residential area, open area (grassland and woody area) and park and recreation (DMTI Spatial, 2015). The residential area consists of single detached dwellings with mature street and backyard trees planted at least ten years ago (Google Earth, 2004). There are different types of trees that live in the study area including deciduous trees such as maple, ash, oak, elm, black cherry, basswood and conifers (Brook McIlroy, 2002). The study area has a hilly topography where near half of the residential areas located on the uphill side and the rest located on the downhill side near the Musselman Lake.

3.2 Multispectral ALS Datasets

Two multispectral ALS datasets were acquired by the Teledyne Optech Titan multispectral ALS system on July 2, 2015. The multispectral ALS datasets had two flight lines that covered and intersected at the study area. For each laser channel the georeferenced point clouds were stored and delivered separately in the LAS files. Total six LAS files and two KML files of the trajectory were included. Each LAS file stored seven attributes of the point

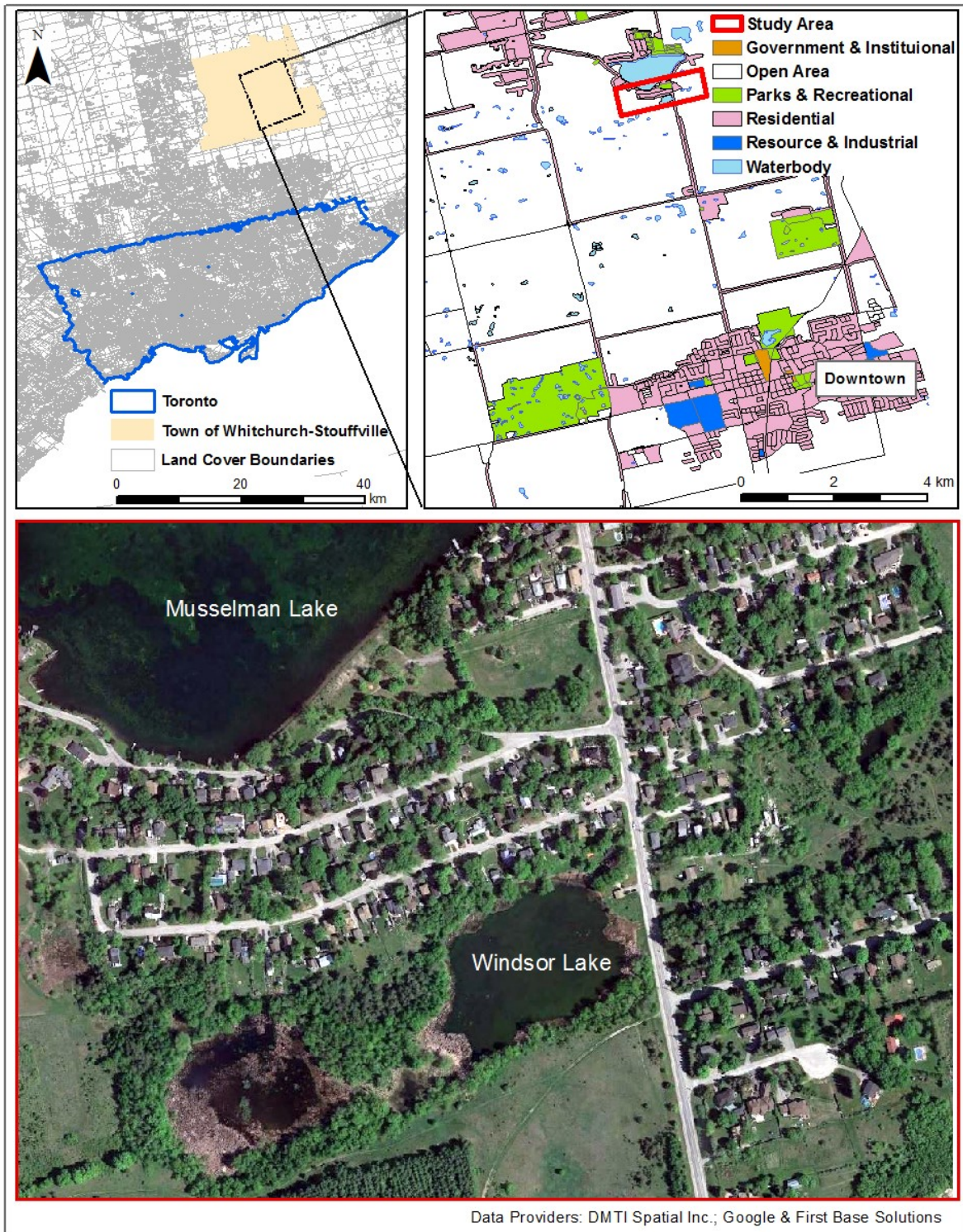


Figure 3.1: Study area

clouds including scan angles, flightline edge, return numbers, number of returns, scan direction, and the intensity collected by each of the three laser channels. No camera images were collected during the survey. The initial survey strips were first cropped to the size of the study area and data pre-processing has been done for the two ALS data separately before merging them together. The survey was planned primarily for system calibration but not for forest inventory or vegetation mapping. The flight altitudes were above 1000 m with the pulse frequency of 100 kHz for each channel, yielding an average point spacing of 0.8 m per point and average point density of 7.7 points / m² (two strips combined). Detailed summary of the acquired Titan ALS datasets can be found in Table 3.1.

3.3 Field Measurements

Field data were collected on February 9, 2016 (see Figure 3.2). A total of 40 trees are sampled in the field that four attributes: height, crown width, DBH, and biomass are recorded for each single tree. Tree heights were measured using a hypsometer in units of meter. DBH was measured with a diameter tape in units of centimeter. Since the field measurement was conducted during the leaf-off season, the third attribute, crown width was measured using the aerial photos from Google Earth. The crown width is defined as the mean of the maximum crown diameter and the diameter measured at the direction perpendicular to the maximum using the Ruler tool in Google Earth in units of meter. The fourth attribute, single tree carbon storage is estimated by plugging the field-measured DBH and tree height in the Canadian national aboveground all-species biomass equations (Lambert et al., 2005). The equations calculate the dry aboveground biomass by relating tree height and DBH to each biomass components such as wood, bark, and foliage, with a uniform relationship: $Biomass = \beta_1 DBH^{\beta_2} H^{\beta_3}$; where β_1 , β_2 and β_3 are the parameters generated for the all-species group with different values according to the tree compartments. The carbon storage in sampled trees was then defined as half of the sum of the dry aboveground biomass in each compartment. The descriptive statistics of the field measurements and the allometry-based biomass are given in Table 3.2

Table 3.1 Summary of the multispectral Titan ALS dataset	
Laser channels	532 nm; 1064 nm; 1550 nm
Flight of View (FOV)	30°
Pulse repetition frequency (PRF)	100 kHz per channel, 300 kHz in total
Flight height	Strip 1: 1030 m; Strip 2: 1043 m
Flight Speed	69-70 m / s
Flight Date	July 2, 2015
Number of ALS strips	2
Number of returns	4
Number of points	Channel (532 nm): 3,665,956 Channel (1064 nm): 3,692,211 Channel (1550 nm): 3,072,895
Average point density	7.7 points / m ² in total

Table 3.2 Statistics of the field measurements for 40 sampled trees				
	Height (m)	Crown Width(m)	DBH (cm)	Carbon (kg)
Maximum	26.90	16.57	98.0	2172.14
Minimum	9.70	4.64	27.0	105.31
Mean	17.48	9.11	48.8	481.85
Std	4.57	3.07	14.6	382.40



Figure 3.2 Location of 40 sampled trees in the field

3.4 Chapter Summary

This chapter provided the description of the study area and data sources. The study area was characterized by single detached houses with mature street trees planted over ten years. The multispectral ALS data were acquired on July 2, 2015 and the field measurements were collected in February 2016. A total of 40 trees were selected and matched with the ALS-derived results. Four attributes of each sampled tree were collected in the field.

Chapter 4

Estimation of Carbon Storage in Urban Trees

The methodology of this study consists of four parts: ALS data processing, vegetation isolation, dendrometric parameter extraction, and carbon storage modeling. Detailed processes in each of the step are illustrated in Figure 4.1.

4.1 ALS Data Processing

4.1.1 ALS Data De-noising and Intensity Normalization

Before grouping the raw ALS points into ground and non-ground points, the outliers collected by each laser channel were removed from the raw point clouds. An ALS intensity filtering was performed first to threshold the intensity values within three standard deviations (Wichmann et al., 2015). This de-noising process removed points from specular reflections that had extreme high intensity in the raw ALS point clouds. Then a statistical outlier removal filter was applied to remove air points that were isolated and far from their neighbors. The filter calculated the mean distance from a point to its six neighbors and removed the point if the mean distance exceed the global distance mean plus one standard deviation. All of the de-noising processes were completed separately for each flight strip.

The filtered raw ALS intensity was then normalized by the range normalization equation (Korpela et al., 2010):

$$I' = I_{raw} \cdot \left(\frac{R}{R_{ref}} \right)^a \quad (4.1)$$

where I' is the range normalized intensity, I_{raw} is the raw ALS intensity, a is the exponent defined based on the target geometry. The raw intensity is normalized according to the ratio of flight-to-target range, R , to the mean reference range, R_{ref} . In literature, the selection of exponent, a , was examined with the conditions of surface objects that $a = 2$ was optimal for large surfaces such as trees with broad leaves and $a = 3$ was optimal for point-like objects such as conifer crowns (Korpela et al., 2010). In this study, a was selected empirically by testing $a \in [2, 4]$ with an interval of 0.5. Very little changes were observed by altering a

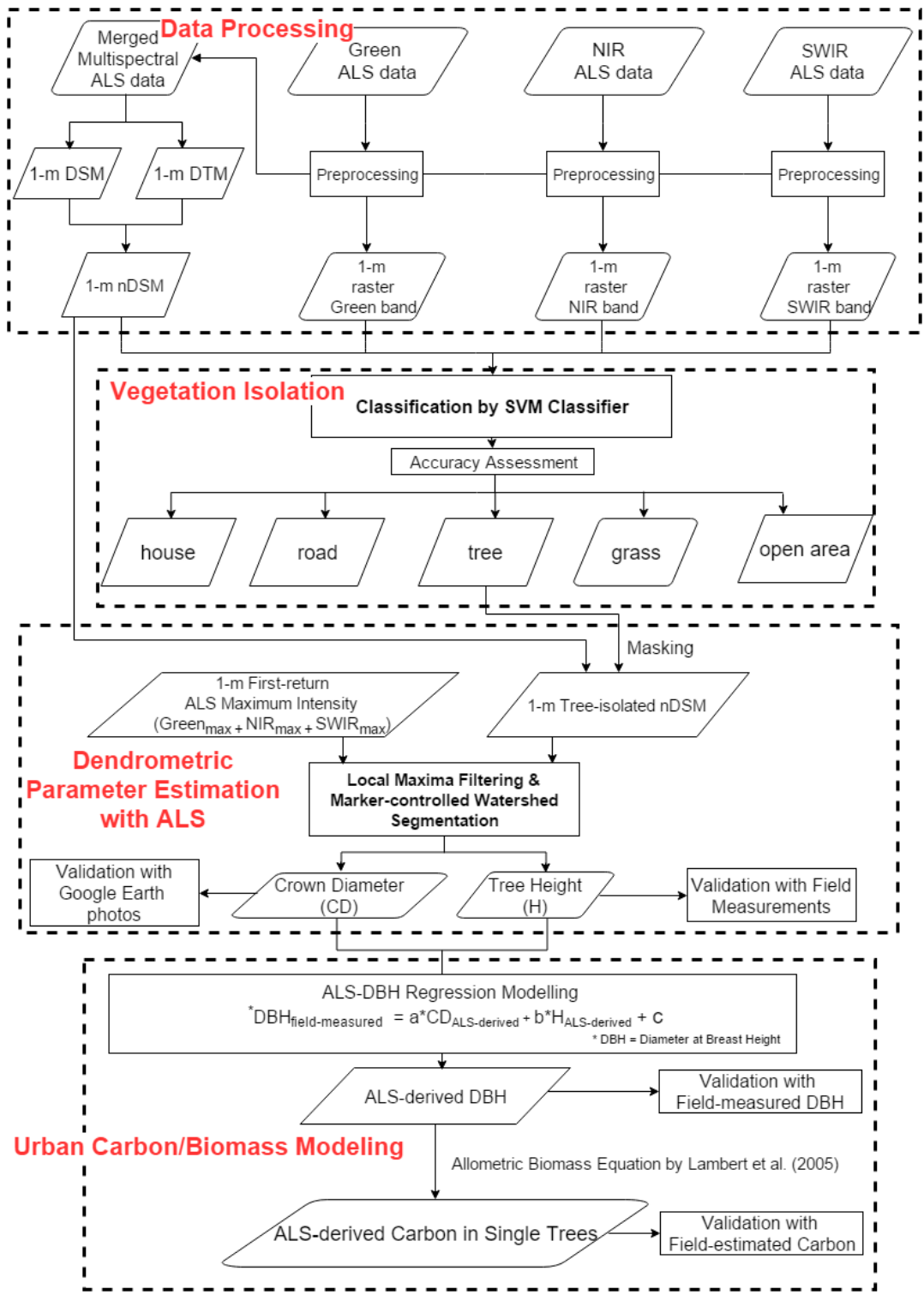


Figure 4.1 Workflow of the proposed methodology

from 2 to 4 in which both the mean value and standard deviation were barely changed while the minimum value varies around 0.05. Hence, $a = 2$ was applied in Equation (4.1) to normalize the ALS intensity.

4.1.2 ALS-derived Intensity Imagery

After removing outliers and rectifying the intensity values, point clouds from two flight strips were merged together which in total three ALS point clouds were acquired by the laser channels. Then each point cloud was rasterized into an intensity image with the ground resolution of 1 m. The pixel size was selected according to the point spacing of the dataset (0.8 m / point in Section 3.2). By selecting the pixel size close to the point spacing, most of the pixels can contain at least one point and the vertical distribution of the points can be largely highlighted. In this way, points were grouped into 1-m grids and the pixel values were assigned by the mean intensity of the points within the grid. For the grids which had no point filled in, the grid values were interpolated linearly by searching the neighbors. Here three mean intensity raster data were generated for Channel 532 nm, 1064 nm, and 1550 nm, respectively.

4.1.3 ALS-derived Digital Elevation Model

Besides the generation of ALS intensity images, DSM and DTM were created from the raw ALS by a ground-point filtering and rasterization process. First, the point clouds generated in Section 4.1.1 were merged together. The last returns in the whole point clouds were inputted in the progressive TIN densification algorithm proposed by Axelsson (2000) to classify ground and non-ground points. This method was selected because it has been widely applied and showed its robustness in both urban and forest conditions (Sithole & Vosselman, 2004; Meng et al., 2010; Tong et al., 2014). The whole ALS dataset was classified into ground and non-ground classes.

To assess the accuracy of the ground-point filtering result, two residential sites with the size of 90 m \times 150 m per site located at the left-part and right-part of the study area respectively were selected and evaluated by manual classified ground and non-ground results. Type I error was calculated by the percentage of misclassified ground points over total

ground points. Type II error was calculated as the percentage of misclassified non-ground points over total non-ground points. Overall accuracy was obtained by calculating the corrected-classified points over total points.

A DTM raster data was generated by rasterizing all the ground points into 1-m grids based on linear interpolation method. The DSM was generated in a similar way by using the non-ground class and the maximum height within the grid was assigned to the pixel values. In this way, the points that represent the treetops could be largely reserved. The linear interpolation method was selected due to its computational efficiency; moreover, considering the hilly topography of the study area and the relatively sparse ground points, linear interpolation was more suitable than other interpolation methods. Lastly, a nDSM raster data was acquired by subtracting the DTM from the DSM.

4.1.4 nDSM Preprocessing

Holes or pixels with negative values were found with the initial ALS-derived nDSM (see Figure 4.2). The holes were commonly presented in ALS-derived nDSM causing from processes in rasterization, the ground-point filtering, and interpolation (Ben-Arie et al., 2009). A hole-filling algorithm was applied to fill a pixel with the mean of its eight neighbors if this pixel has more than 2 m elevation difference with its eight neighbors (Zhang, 2015). The hole-filled nDSM and three multispectral ALS intensity data were applied to the next section as the input data.

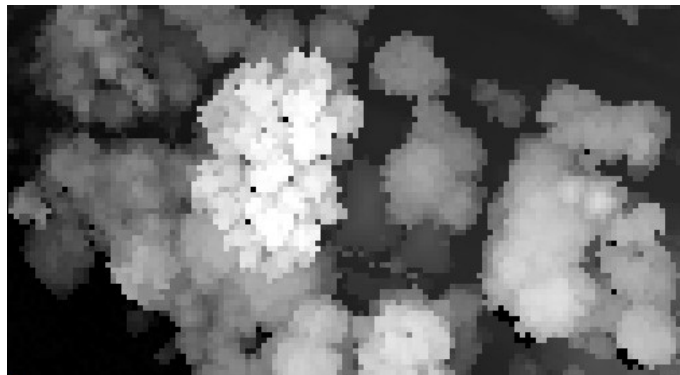


Figure 4.2 Pits in the raw nDSM

4.2 Vegetation Isolation

Besides the multispectral intensity and nDSM data described in Section 4.1, two additional indices are derived as follows

$$pNDWI = \frac{C_{Green} - C_{NIR}}{C_{Green} + C_{NIR}} \quad (4.2)$$

$$pNDVI = \frac{C_{NIR} - C_{SWIR}}{C_{NIR} + C_{SWIR}} \quad (4.3)$$

where C_{Green} , C_{NIR} , C_{SWIR} refer to the laser channels at 532 nm, 1064 nm, and 1550 nm, respectively.

By visually examining the pNDWI and pNDVI indices (see Figure 4.2), both the pNDVI and pNDWI showed good discrimination for artificial objects. These two indices can facilitate the manual selection of training samples and work as the ancillary data in the classification process. The contribution of these two indices to the overall accuracy of classification was analyzed. Total six input data, including 1) Green channel intensity; 2) NIR channel intensity, 3) SWIR channel intensity, 4) nDSM, 5) pNDWI, and 6) pNDVI were generated and inputted into the classification scheme.

Because the study area was characterized by simple residential landscape, it can be grouped into six land cover types including water, house, road, grass, tree, and open area. Description of the land cover types are presented in Table 4.1. However, due to the bathymetry capability of the channel at 532 nm, water points in the study area tended to have irregular intensity that was induced by the interaction of laser points with both water surface and organic matters underneath. It appeared noisy in the intensity data (see Figure 4.3) and might potentially influence the classification result. Hence, water bodies including Musselman Lake, Windsor Lake and one small water region were masked out of the dataset resulting in only five land cover types were trained and classified in this study. Interpreting land cover types from ALS intensity were not as easy as from passive optical images. Training samples are selected based on all the input data and examined by the aerial photos from Google Earth (see Table 4.1). Certain land cover types such as grass and open area



Figure 4.2 Examples of ALS-intensity-derived indices

could be distinguished from only one or two intensity data and can be hardly identified from the rest. For each land cover type, a minimum of 600 pixels were selected as training samples and for vegetation covers who occupied a large portion of the study area more than 1000 pixels were selected. Total 5535 pixels were collected as training samples and the pixels selected for each land cover type were evenly distributed across the study area.

The SVM classifier was selected to perform land cover classification with multispectral ALS-derived data due to the popularity of SVM in typical ALS-related classification studies which makes the classification result of this study comparable to previous studies. In literature, there is no clear clue to use one kernel function rather than others (Mather et al., 2011). So the RBF kernel, $e^{-\gamma\|(x-x_i)\|^2}$ which has been commonly applied in the classification with remote sensing data was selected with a user-defined kernel width, γ as 0.167 (defined as one over the number of input data) and the regularization parameter, C as 100.

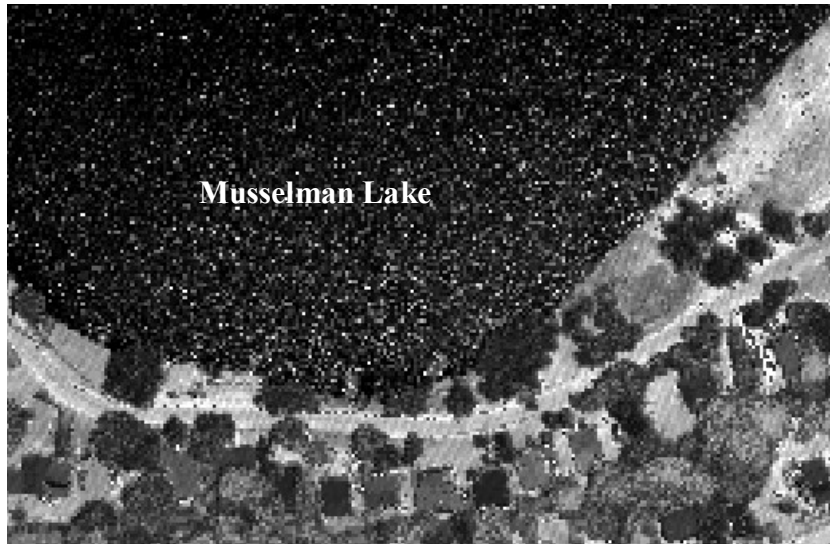


Figure 4.3 Multispectral ALS intensity on waterbodies (green band)

The SVM classification is performed on three combinations of the input data shown in Table 4.2. The selection of input data was mainly designed for showing the benefits of multispectral ALS in land cover classification (Selection 2) compared with typical ALS data (Selection 3). Meanwhile, Selection 1 was designed to examine the contribution of multispectral-intensity-derived indices in overall classification accuracy. To assess the quality of the classification result, total 522 reference pixels were selected by stratified random sampling across the study area and manually labeled into land cover types. For each class, a minimum of 60 reference pixels were labeled and for vegetation covers who occupied a large portion of the study area more than 100 reference pixels were included. Lastly, the error matrix was generated to examine the performance of the classification results according to the commission and omission errors. Kappa coefficient was not used in this study due to its reported problems in accuracy assessment (Pontius & Millones, 2011). Kapa coefficient was reported to be highly correlated with the overall accuracy and have problems with the correction for chance agreement used in Kappa formula (Olofsson et al., 2014).

Table 4.1 Description of the land cover types and corresponding training samples

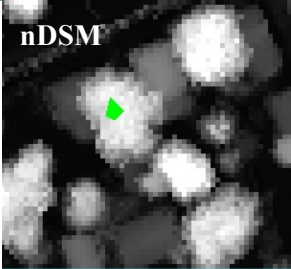
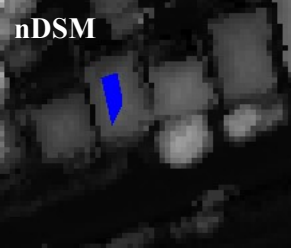
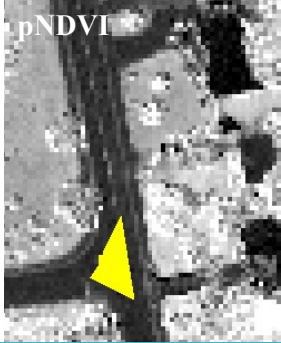
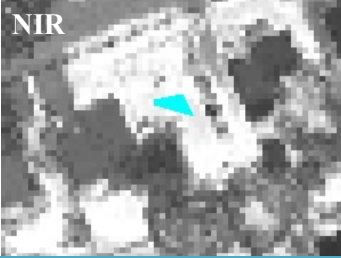
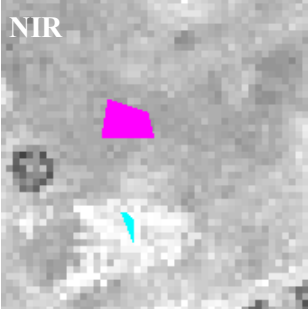
		 <p>nDSM</p>
House	Single family houses that are visible from nDSM and pNDWI	 <p>nDSM</p>
Road	Both main road and paths connected to individual houses were included and examined based on pNDVI	 <p>pNDVI</p>
Grass	The grassland that has been well-maintained by the residents and showed high reflectivity in the NIR intensity data but were invisible from nDSM	 <p>NIR</p>
Open Area	The public green space that were located in the park and recreation zones which showed a mixture of soil and grass by viewing in aerial photos	 <p>NIR</p>

Table 4.2 Combinations of input data for SVM classification	
Selection 1	Green + NIR + SWIR + nDSM + pNDWI + pNDVI
Selection 2	Green + NIR + SWIR + nDSM
Selection 3	NIR + nDSM

The tree class was further extracted and employed iteratively by a 3×3 majority filter for five times. The majority filtering can help reduce the salt-and-pepper effect which is commonly existed in the pixel-based classification. The number of iterations was tested empirically to generate an optimal tree mask that can remove isolated pixels and reserve the shape of the crowns for the dominant trees. By this method, pits within a tree crown can be filled; meanwhile the filtered tree class was used for masking the nDSM to generate a tree-isolated nDSM.

4.3 Dendrometric Parameter Estimation

The tree-isolated nDSM was also referred to as the CHM, a model displaying tree positions by tree crowns in a top-down view and storing the height values in pixels. A 3×3 local maxima filter was first employed on the CHM to detect treetops. A pixel with the highest value than its eight neighbors was defined as the treetop. To eliminate the commission errors associated with the local maxima filtering (see Figure 4.5) and detect the true treetop pixels, the local maxima in the CHM were further filtered by the ALS intensity data. The ALS intensity values were dependent not only on the reflectivity of the object but also on the range between the sensor and the object. A true treetop pixel would have both high intensity and height values. Under this assumption another 1-m raster data was generated here as the sum of the maximum intensity of the first return in each channel. The clusters of local maxima in the CHM which had more than fifteen pixels together were further extracted and only the pixels that also were the local maxima in the maximum intensity layer were retained in the final treetop results. Previous studies relied on changing the window size and shape of

the filter to refine the treetops (Chen et al., 2006; Zhao et al., 2009). These approaches were not suitable here because the CHM resolution (1 m) generated in this study was relatively coarse so that increasing the window size of the local maxima filter would result in excluding small tree crowns.

The Marker-controlled watershed segmentation was applied to segment the CHM into individual tree crowns by defining the pre-detected treetops as the markers. In this way, every pre-detected treetop would have one closed segment. The performance of the segmentation in isolating tree crowns was evaluated by the absolute accuracy calculated as:

$$\text{absolute accuracy}_{\text{tree isolation}} = \frac{n_{1,1}}{n_{\text{total}}} \quad (4.4)$$

where $n_{1,1}$ is the number of detected crown segments which have a one-to-one relationship to the ground truth; n_{total} is the number of tree crowns in the ground truth. Two sites with street trees located along the sides were selected as the ground truth. Each site had an area of 60 m by 300 m. The watershed segmentation result was visually compared with the original tree-isolated nDSM and the aerial photos in Google Earth meanwhile the total number of tree crowns identified in the ground truth was recorded.

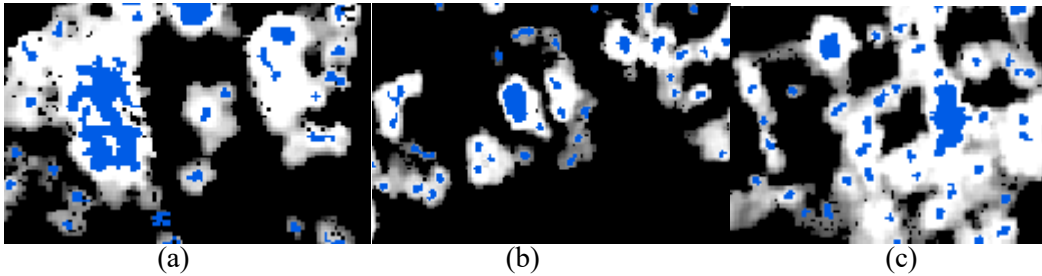


Figure 4.4 Examples of a 3×3 local maxima filtering on CHM

Tree height is defined as the average of the local maxima within each segment and the crown diameter is defined as the average of the maximum crown diameter passing through the center of the local maxima and the one measured at the perpendicular direction (see Figure 4.6).

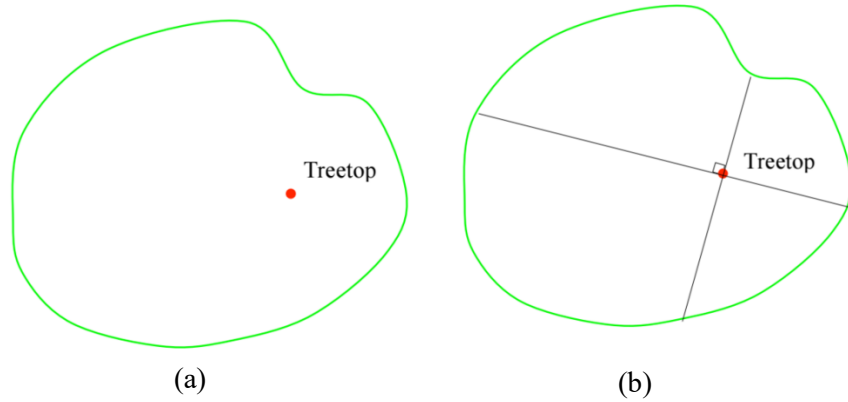


Figure 4.5 Measurement of tree height and crown diameter: (a) location of the pixel with maximum value within a tree crown, (b) illustration of the crown width measurement

To evaluate the accuracy of the ALS-derived dendrometric parameters, the ALS-derived tree height and crown width were compared with the field measurements. The crown segments generated from the ALS data were matched with the 40 field-sampled trees and the RMSE and bias are calculated to compare the ALS-derived dendrometric parameters with field samples:

$$\text{Bias} = \frac{\sum_{i=1}^n X_{ALS,i} - X_{field,i}}{n} \quad (4.5)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (X_{ALS,i} - X_{field,i})^2}{n}} \quad (4.6)$$

$$\text{RMSE}\% = \frac{\text{RMSE}}{\bar{X}_{ALS}} \quad (4.7)$$

where n is the number of field samples which equaled to 40 trees in this study; X refers to the values of dendrometric parameters (height or crown width) measured either in the field or from the ALS data and \bar{X}_{ALS} is the arithmetic mean of the ALS-derived measurements.

Moreover, a linear regression model was fit to the ALS-derived tree height and crown width to determine if there was a strong correlation between these two variables.

4.4 Carbon Storage Modeling

In order to predict the carbon storage in trees, a multiple linear regression model was developed empirically to predict the field-measured DBH by the ALS-derived tree height and crown width as the independent predictors (Popescu et al., 2007). The empirical equations derived from the ALS-DBH linear regression model are:

$$DBH_{Field} = a \cdot CD_{ALS} + b \cdot H_{ALS} + c \quad (4.8.1)$$

$$a = \frac{(\sum H_{ALS}^2)(\sum CD_{ALS} \cdot DBH_{Field}) - (\sum CD_{ALS} \cdot H_{ALS})(\sum CD_{ALS} \cdot DBH_{Field})}{(\sum CD_{ALS}^2)(\sum H_{ALS}^2) - \sum CD_{ALS} \cdot H_{ALS}^2} \quad (4.8.2)$$

$$b = \frac{(\sum CD_{ALS}^2)(\sum H_{ALS} \cdot DBH_{Field}) - (\sum CD_{ALS} \cdot H_{ALS})(\sum H_{ALS} \cdot DBH_{Field})}{(\sum CD_{ALS}^2)(\sum H_{ALS}^2) - \sum CD_{ALS} \cdot H_{ALS}^2} \quad (4.8.3)$$

$$c = \overline{DBH} - a \cdot \overline{CD}_{ALS} - b \cdot \overline{H}_{ALS} \quad (4.8.4)$$

where CD is the crown width in m and H is the canopy height in m; a, b, c refer to the parameters derived from the empirical model fitting.

The 40 field-sampled trees were split into two datasets that 20 trees used for model parameterization and the remaining 20 trees reserved for validation. To eliminate the influence of tree locations in model fitting, each 10 adjacent tree samples were grouped together under one sampling location which results in total four sampling groups. Six combinations of training and validation datasets were chose by selecting two sampling locations out of four for model development and using the remaining two locations for validation. All six models were developed at a 0.05 significance level and were fitted through a cross validation process. The parameters generated for each model were collected. The predictive power of the regression models and the performance were inspected by the coefficient of determination (R^2) and the accuracy of the prediction was examined by the RMSE of the predicted parameters

$$R^2 = \frac{\sum_{i=0}^n (y_{fitted} - \bar{y}_{field})^2}{\sum_{i=0}^n (y_{field} - \bar{y}_{field})^2} \quad (4.9)$$

where y refers to the values of dendrometric parameters (height or crown width) measured either in the field or from the ALS data.

Comparing R^2 and RMSE of the six ALS-DBH regression models, the one with high R^2 in the model fitting and low RMSE in the validation was selected to predict DBH. In addition, a multiple regression model was constructed by using the height and crown width of the 40 field samples to predict field-measured DBH. This field-DBH model was first evaluated by R^2 of modeling fitting and was applied to estimate DBH using ALS-derived parameters. The strength of relationship between the ALS-predicted DBH and field-measured DBH was also examined by linear regression.

After selecting the ALS-DBH regression model, the ALS-estimated DBH and height were plugged into the Canadian national above-ground biomass equations proposed by Lambert et al. (2005) to estimate the carbon storage in trees. In Lambert et al. (2005), the set of equations based on DBH and height for all species (Equations 4.10.1-4.10.6) is selected to calculate the biomass here since no genus or species information is available in this study. The above-ground biomass was estimated as the sum of biomass in tree compartments (foliage, branch, wood, and bark). And the carbon stored in trees was estimated as the half of total biomass. The carbon storage predicted by the ALS-derived parameters was compared with that estimated by the field-measured DBH and height and evaluated by the RMSE and R^2 .

$$y_{foliage} = \beta_{foliage1} D^{\beta_{foliage2}} H^{\beta_{foliage3}} + e_{foliage} \quad (4.10.1)$$

$$y_{branches} = \beta_{branches1} D^{\beta_{branches2}} H^{\beta_{branches3}} + e_{branches} \quad (4.10.2)$$

$$y_{wood} = \beta_{wood1} D^{\beta_{wood2}} H^{\beta_{wood3}} + e_{wood} \quad (4.10.3)$$

$$y_{bark} = \beta_{bark1} D^{\beta_{bark2}} H^{\beta_{bark3}} + e_{bark} \quad (4.10.4)$$

$$y_{total} = y_{foliage} + y_{branches} + y_{wood} + y_{bark} \quad (4.10.5)$$

$$\text{Carbon} = \frac{y_{total}}{2} \quad (4.10.6)$$

where y is the dry biomass compartment in kg, D is the DBH in cm, H is the height in m, and the β is the model parameter with defined values in each compartment (see Table 4.3); and e is the error terms.

Table 4.3 Parameters for all-species equations listed in Lambert et al. (2005)			
Species	Parameter	Value	Standard Deviation
All Species	b_{wood1}	0.0283	0.0004
	b_{wood2}	1.8298	0.0075
	b_{wood3}	0.9546	0.0101
	b_{bark1}	0.012	0.0003
	b_{bark2}	1.6378	0.017
	b_{bark3}	0.7746	0.0233
	$b_{branches1}$	0.0338	0.0008
	$b_{branches2}$	2.6624	0.0182
	$b_{branches3}$	-0.5743	0.0233
	$b_{foliage1}$	0.1699	0.0036
	$b_{foliage2}$	2.3289	0.0184
	$b_{foliage3}$	-1.1316	0.0235

To show whether carbon storage in trees varied with land cover types, the ALS-derived tree-crown segments were first converted into vector data in GIS with the amount of carbon storage stored in the attributes. The carbon storage within each land cover type was calculated by adding up all the carbon storage in trees and dividing by the area of the land cover type. For the Town of Whitchurch-Stouffville, the carbon stocks were extrapolated by multiplying the specific carbon amount per unit area with the total area of each land cover type. And for those land cover types such as the industrial sites and government and institutional area which were excluded in the study area, the carbon stored in government and institutional areas was given the same amount per unit area as the residential but zero for the industrial sites. Then a citywide carbon storage map was created.

4.5 Chapter Summary

This chapter described the methodology of the study in details. It presented the stepwise procedures in data processing, vegetation isolation, dendrometric parameter extraction, and carbon modeling. A series of methods were conducted in the present study including SVM classification, watershed segmentation, and allometry-based multiple linear regression. The rationale behind the method selections was also discussed. Principles and applications in the literature can be found in Chapter 2. And the results in each of the step would be presented in the next chapter.

Chapter 5

Results and Discussion

This chapter presents and discusses the results generated using the proposed methodology. A data analysis of the multispectral ALS data is also included to give a detailed introduction to this new sensor. The performance of the ground filtering and SVM classifier are evaluated. The accuracy of ALS-derived dendrometric parameters are assessed by comparing ALS-derived dendrometric parameters with field measurements. The parameterization process of the allometry-based ALS-DBH models is also included to estimate tree carbon stocks in the study area. Lastly, a map showing the estimated carbon storage in the Town of Whitchurch-Stouffville is presented.

5.1 Accuracy Assessment for Ground-point Filtering

Two sites with a size of 90 m × 150 m per site are selected to evaluate the performance of ground-point filtering using the progressive TIN densification algorithm (see Figure 5.1). For the two validation sites, ground and non-ground points were manually classified as the reference data and were compared to the automatic filtering result. There were total 193,316 points in the two sites. And both the number of ground points misclassified in the non-ground class and the number of non-ground points misclassified in the ground class are counted and showed in Table 5.1. Type I error was 2.8% and Type II error was 4.2%. Larger Type II error was due to the filter includes low vegetation points into the terrain. Low vegetation was also causing difficulties to the manual ground point selection. The overall accuracy for ground-point filtering was 96.4% which was calculated as the total correctly-classified points over the total points.

5.2 Analysis of Multispectral ALS Data for Land Cover Classification

5.2.1 Spectral Analysis of the Land Cover Classes

To analyze the potential of the intensity of the Titan multispectral ALS data, the spectral patterns of the five defined classes in each channel are plotted in Figure 5.2. The intensities

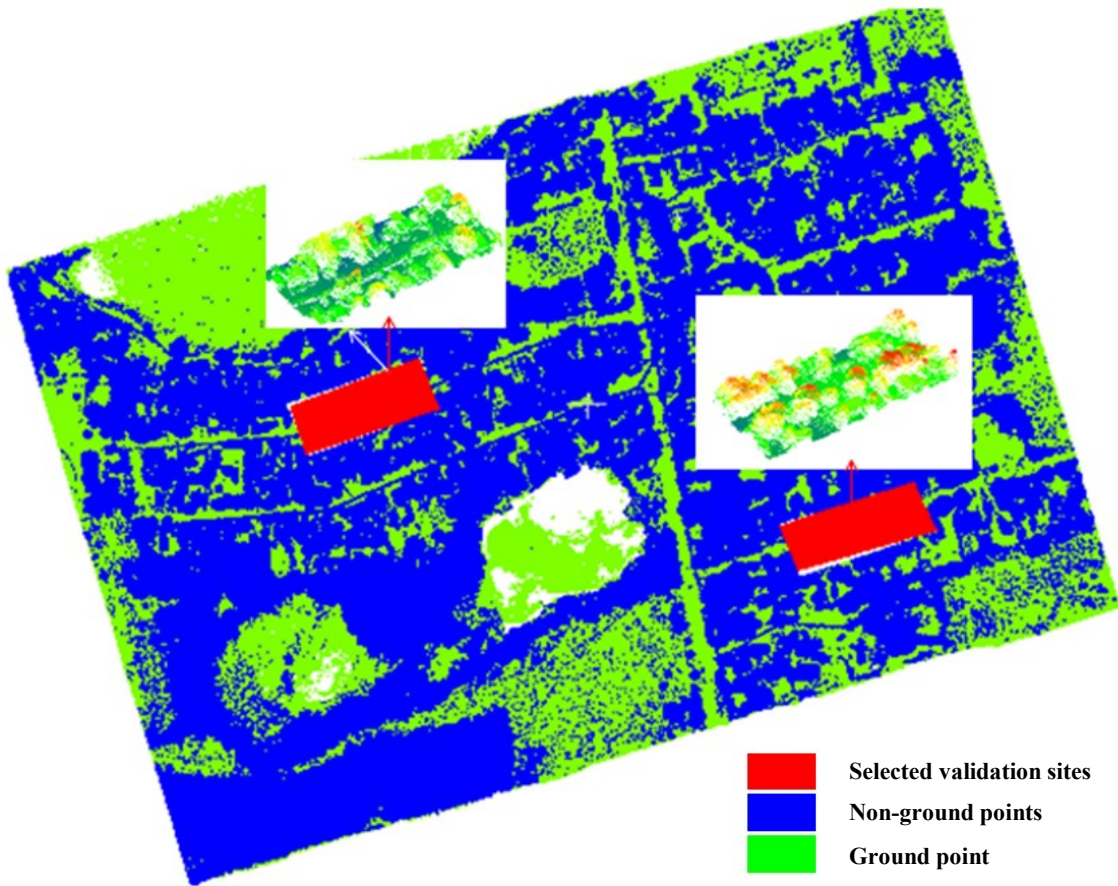


Figure 5.1 Illustration of the ground filtering result and the location of two validation sites

Table 5.1 Accuracy assessment for ground-point filtering

Reference	Filtered		Total	Error (%)
	Ground	Non-ground		
Ground	76,404	2,212	78,616	Type I = 2.8%
Non-ground	4,817	109,883	114,700	Type II = 4.2%

were stored in 8-bits integers giving the range of values from 0 to 255. The histogram for each class is generated according to the training samples selected during the classification process. In addition to the original channels, histograms on the nDSM, pNDVI, and pNDWI are also shown in Figure 5.2.

The tree class contained spectral peaks at low intensity (< 30) for all three bands. This class also had a peak at the highest elevation among the classes (> 10 m). The tree class had positive pNDWI and pNDVI values. For house class, low intensities were observed in both NIR and SWIR bands. The house class also had a peak at medium intensities for the green band, negative pNDVI, and positive pNDWI. For all the low-elevation classes such as road, grass and open area, high intensities were observed in the green band. Grass and open area tended to have a peak at low intensities in the SWIR band while the road class had a low-intensity peak in the NIR band. The road class followed the same trend as the house class in the spectral indices that had negative pNDVI and positive pNDWI. The grass class had all positive values in both spectral indices where no distinct separation was observed. The open area class appeared to have two positive peaks in the pNDWI and one positive and one negative peak in the pNDVI. Except for the house class, all classes had a wider intensity distribution on the green band. Manmade objects such as road and houses showed clear discrimination on pseudo indices. The elevations of the house class varied from 3 to 10 m. The elevations of grass and open area were more fluctuant than the road class. The spectral patterns observed in the grass class were similar to those found in optical imagery that grasses tended to reflect more at the near-infrared light. This was not observed in the tree class which was mainly because tree class contained multiple returns and as the pulses energy decreased as it transmitted farther. The house class was associated with the lowest intensities in the NIR band compared with other classes. And the road class had the highest reflectivity in the SWIR band.

5.2.2 Spectral Analysis of the Tree Class

Because the classification at tree-genus or -species level is important especially for precise biomass estimation, the spectral patterns of the tree class are examined further to find if

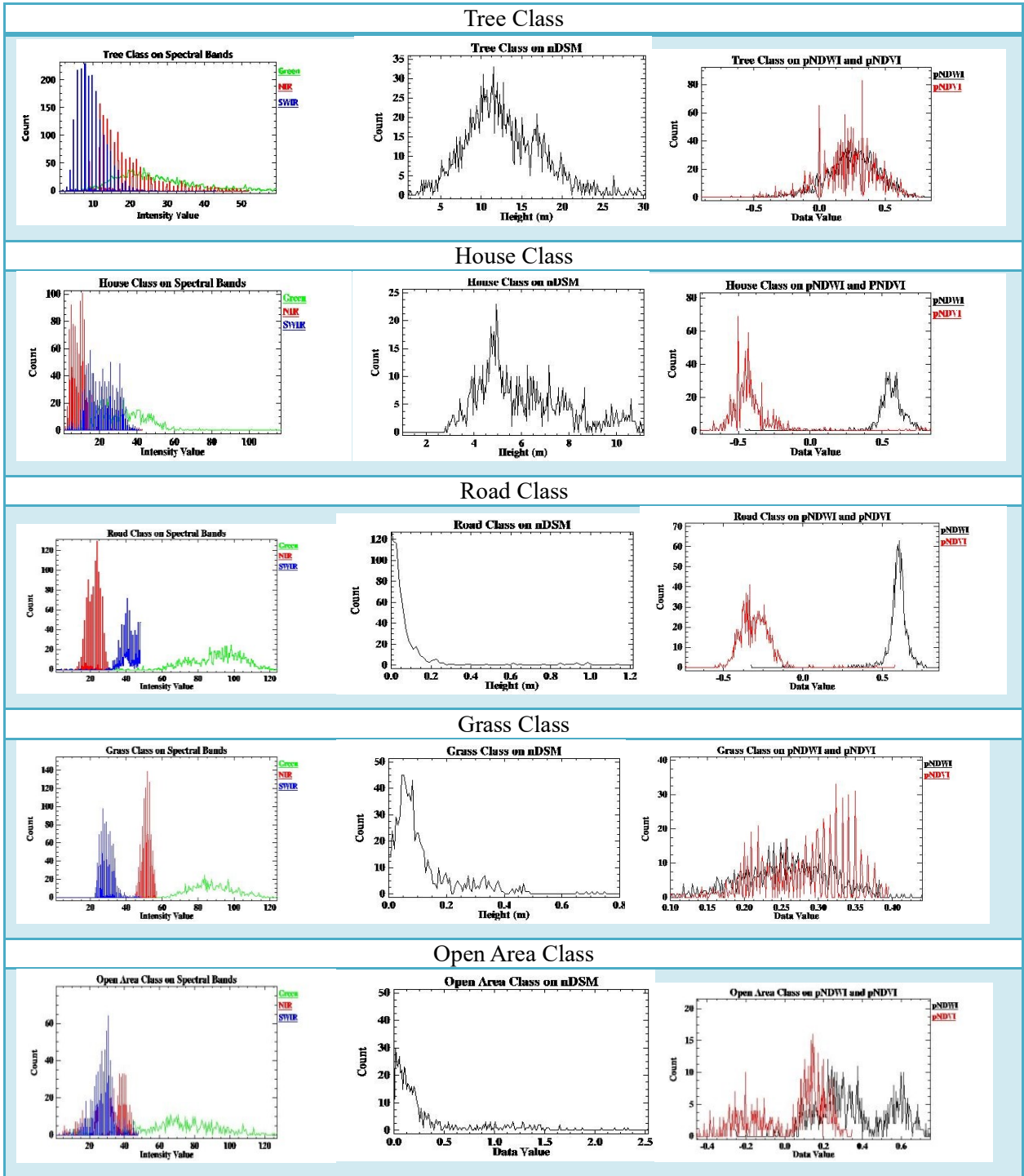


Figure 5.2 Histograms for five classes

current multispectral ALS datasets can distinguish the tree class into conifer and deciduous trees. By confirming the tree types in Google Earth, separate subsets for conifers and deciduous trees are created and plotted in Figure 5.3. There are no obvious distinctions between the tree genus on the histograms and by visual observations. However, through close visual observation to the pNDVI dataset, the trees with dark-colour leaves such as the Crimson King Maple tree could stand out from the tree class (see Figure 5.4). The spectral pattern for trees with dark-colour leaves is also plotted in Figure 5.3. No clear spectral distinctions between trees with different leaf colours are observed from the laser channels (see Figures 5.3 (a), (c), and (e)). But a close-to-1 peak is observed in the pNDVI dataset for trees with dark-colour leaves meanwhile both conifers and deciduous trees have a pNDVI peak close to 0 (see Figure 5.3 (f)). The multispectral ALS datasets generated in the study may not be sensitive enough to provide separate classification for conifers and deciduous trees. However, the analysis presented here shows that the multispectral ALS intensity may be influenced by factors such as the colour/reflectivity of the objects which in turn will be beneficial to the studies of tree mortality, rooftop solar energy, and so on.

5.2.3 Accuracy Assessment of the Classification Results by SVM Classifier

Accuracy assessment was conducted on the classification results in order to examine the classification performance of multispectral ALS data. Table 5.2 lists the confusion matrix of the classification result using all the six input data (Selection 1 in Table 4.2). For each class, User's accuracy and Producer's accuracy were calculated. The tree class achieved both high User's accuracy and Producer's accuracy (>95%). This class is mainly misclassified with the house class which was not surprising since they have similar spectral patterns and elevations as shown in Figure 5.2. A construction site with sandy ground located in the lower right corner of the classification map also caused problems to the classifier (see Figure 5.5). This area is misclassified as a mixture of roads and open area because the spectral behavior is more as the road class. Meanwhile the term "open area" defined in this study is more as public green space which are natural and vegetated areas. The grass class and the open area class were not very discriminable from each other which both were having the lowest User's accuracy and Producer's accuracy. Besides these two classes, all the classes achieved the

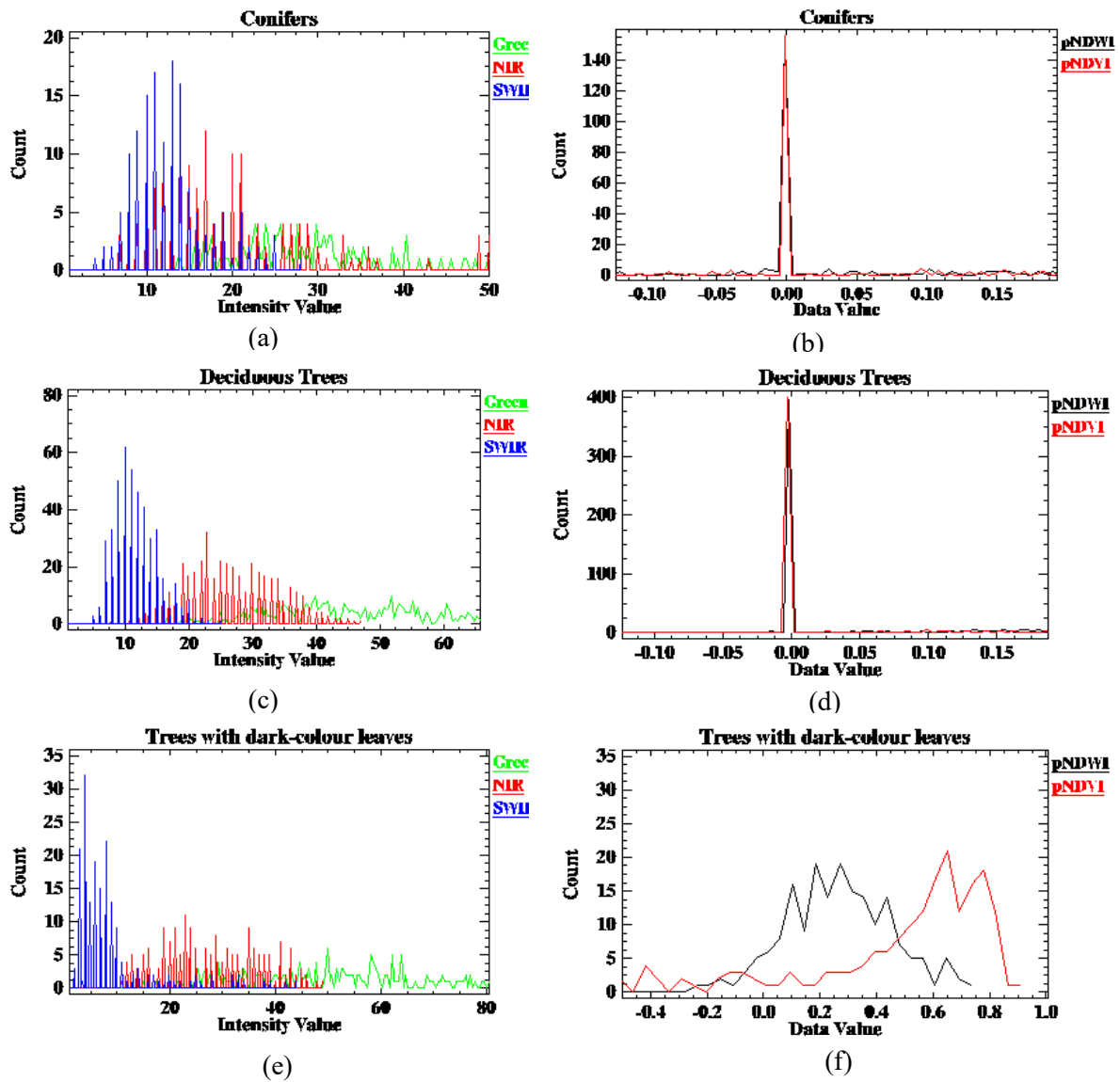


Figure 5.3 Histograms for the tree class



(a)(b)(d), and (e) reprinted from Google Earth (2015)

Figure 5.4 Trees with dark-colour leaves in Google Earth and the pNDVI

high User's and Producer's accuracy (>90%). The overall accuracy of this data input was around 90%. Figure 5.5 shows the final classification map. This combination of input data achieves highest overall accuracy among the three selections (see Table 5.3). The 89% overall accuracy was achieved by using the ALS-derived raster data along which indicated the contributions of the two calculated indices to the overall accuracy was not significant. The comparison between typical ALS data and multispectral ALS data on land cover classification was also conducted that the same classification process was applied to the nDSM and NIR band only. The overall accuracy of classification in use of typical ALS data was around 79%.

5.3 Performance of the Watershed Segmentation

In order to distinguish individual tree crowns in the CHM, a three-step segmentation process was conducted by first isolating the nDSM into CHM according to the tree class generated from the classification result. Then, the second step was to find treetops in the CHM. Lastly

the watershed segmentation was performed to create a crown segment for each treetop. The generation of CHM is illustrated in Table 5.4.

5.3.1 Results for the Local Maxima Selection

The second step of the methodology was to find the local maxima in the tree-isolated nDSM (the so-called CHM) as the treetop candidates. In this study, a new approach was proposed to reduce the number of pixels generated by the original 3×3 local maxima filter and to find the true treetop pixels. Figure 5.6 illustrates the comparison between the original local

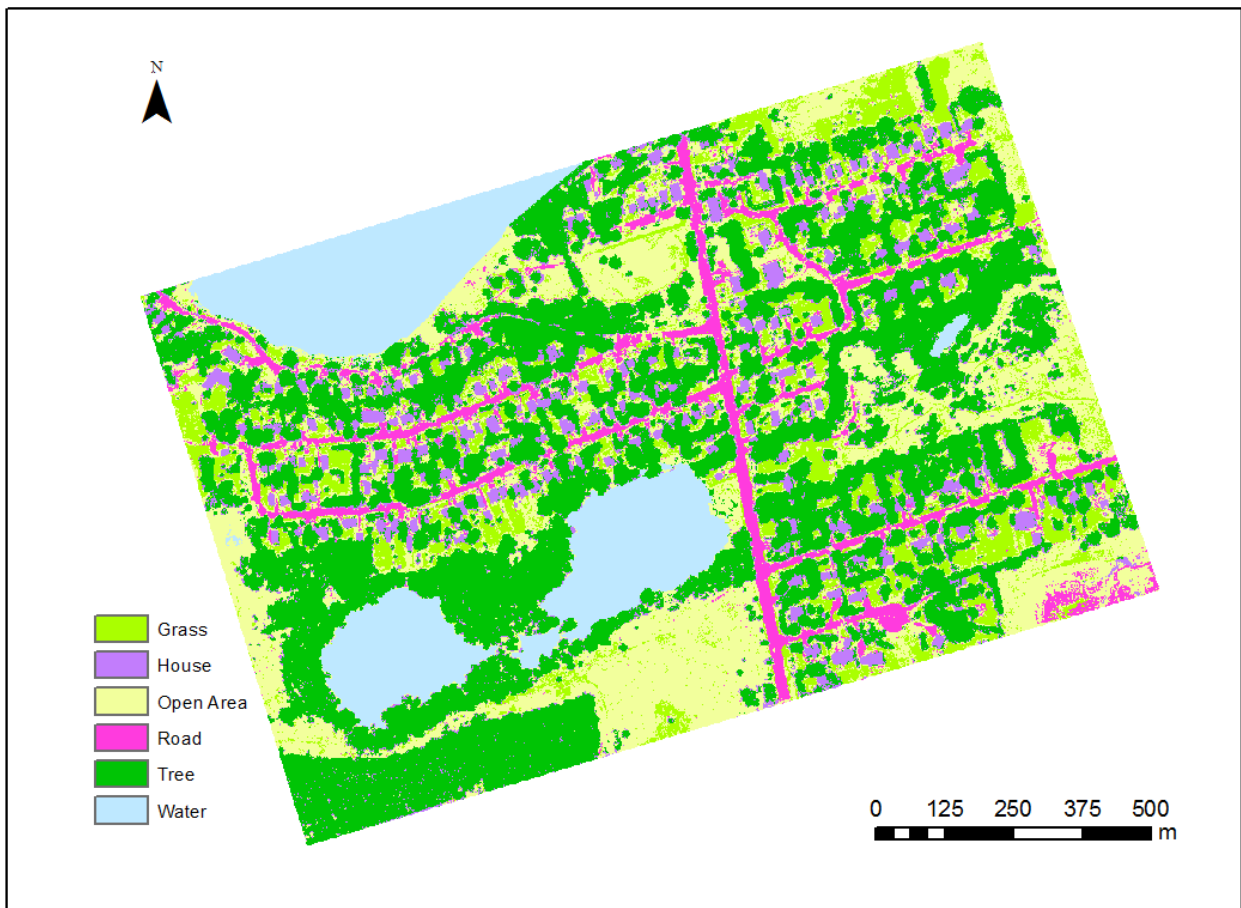

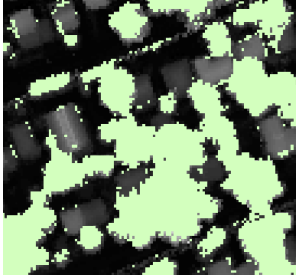
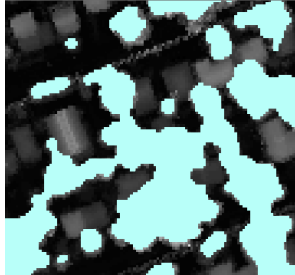
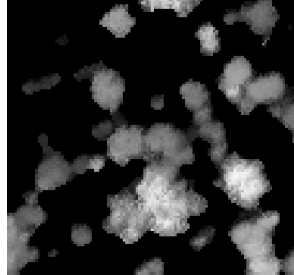


Figure 5.5 Classification map of the study area

Table 5.2 Confusion matrix for the SVM classification using six input data							
	Grass	Tree	House	Road	Open Area		
Classified Pixels						Total	User's Accuracy
Grass	99	0	2	0	11	112	88.39%
Tree	0	139	3	0	2	146	95.21%
House	1	4	62	0	1	66	93.94%
Road	4	0	1	82	4	91	90.11%
Open Area	15	1	0	2	89	107	83.18%
Total	119	144	68	84	107	522	
Producer's Accuracy	83.19%	96.53%	91.18%	97.62%	83.18%		
Overall Accuracy = 90.23%							

Table 5.3 Accuracy report for the classification results							
	Green	NIR	SWIR	nDSM	pNDWI	pNDVI	Overall Accuracy
Selectio n 1	√	√	√	√	√	√	90.23 %
Selectio n 2	√	√	√	√			89.12%
Selectio n 3		√		√			79.04%

Table 5.4 Generation of tree-isolated nDSM			
nDSM	Tree Class on nDSM	Tree Mask by Majority Filtering	Tree-isolated nDSM
			

maxima result performed on the CHM and the refined result which only retained the pixels that were having a local maximum on both the maximum intensity data and the CHM. It was noticeable that the clusters of local maxima in the original result were eliminated by the proposed method. Then the watershed segmentation algorithm is applied to generate crown segments by using the refined local maxima as the markers (see Figure 5.6 (c)).

5.3.2 Accuracy Assessment of Segmentation

To evaluate the segmentation result, two validation sites are selected to count the number of tree crowns that can be visually identified from aerial photos in Google Earth and compared with the automatically generated crowns (see Table 5.5). The evaluation focused on examining the capability of the segmentation algorithm in delineating individual tree crowns. Hence the accuracy was calculated as the percentage of the number of detected crowns that were having a one-to-one relationship to the ground truth over the total number of the crowns. For Site 1, there were total 65 trees that can be identified in the aerial photo while 49 trees that had a single crown delineated by the watershed segmentation algorithm. For Site 2, there were 114 trees in the ground truth and 80 trees were identified as individual crowns in the segmentation result. The accuracies for these two sites were 75.4% and 70.2%, respectively. The achieved accuracy rates were not surprising since previous studies have reported that due to the nature of ALS system the dominant trees located in the upper portion of the canopies

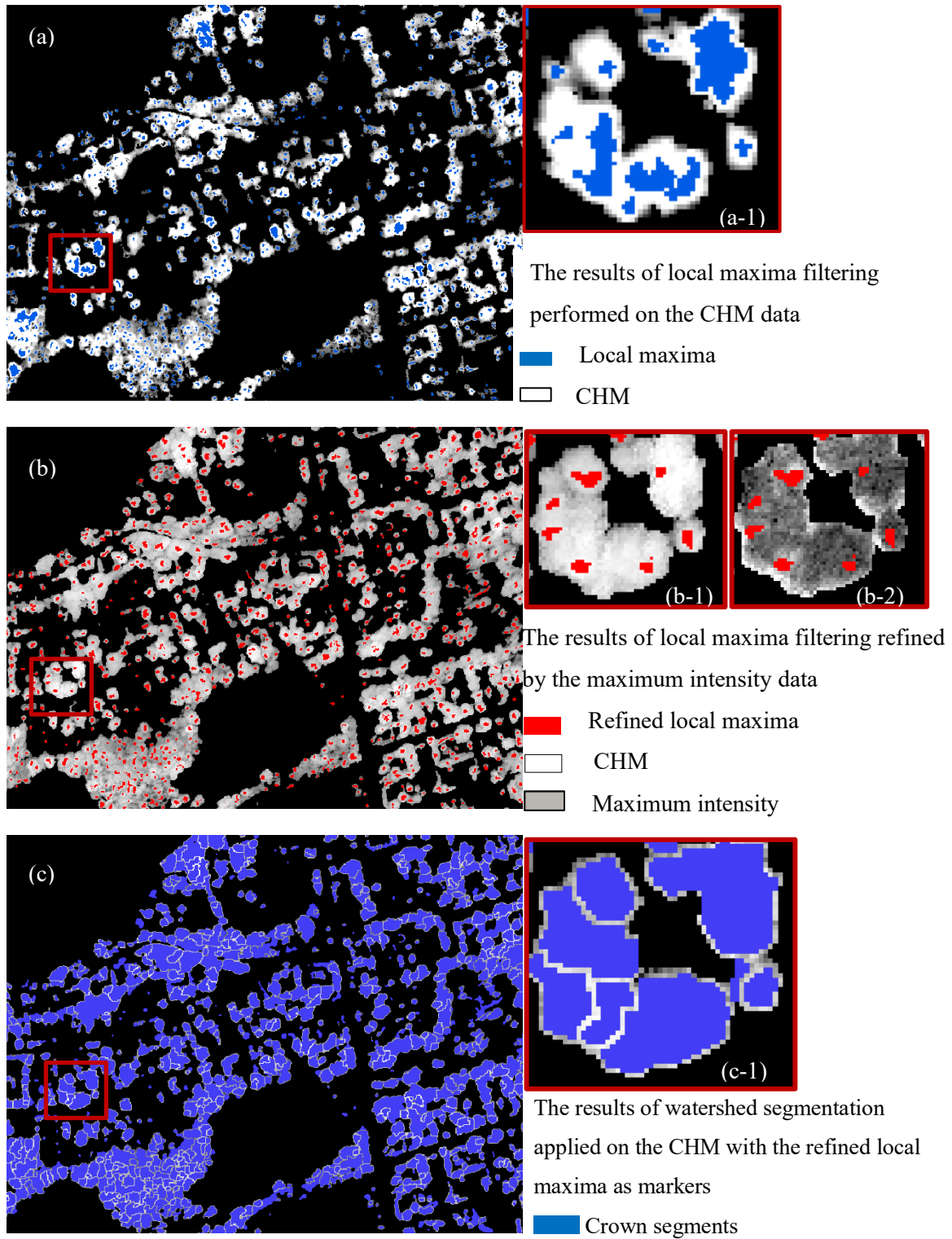






Figure 5.6 Results obtained by local maxima filtering and the marker controlled watershed segmentation

Table 5.5 Evaluations of the segmentation result

	Site1	Site2
Reference		
Results		
N_{1,1}	49	80
Total	65	114
Accuracy	75.4%	70.2%

were better detected than the canopies nearby or underneath (Kaartinen et al., 2012). Due to the relative coarse resolution of the dataset and the nature of the ALS system, the segmentation result mainly included the dominant trees in the study area.

5.4 Validations for the ALS-derived Dendrometric Parameters

The accuracies of ALS-derived tree height and crown width were assessed by the samples measured in the field. For the ALS-derived tree height, a RMSE of 1.21 m (relative RMSE = 6.8%) and a negative bias of 0.2 m (relative bias = -0.1%) is given in Table 5.6. For the ALS-derived crown width, a RMSE of 1.47 m (relative RMSE = 16.4%) and a negative bias of 0.18 m (relative bias = -2%) were observed. The tree height was underestimated because of the undergrowth and the factors related to the flying height (1000 m) and the point density (7 to 8 points / m²). The relative RMSE of crown width was mainly caused by the resolution of the CHM and the results of crown segmentation.

Table 5.6 Validation statistics for the ALS-derived dendrometric parameters				
Parameter	RMSE	RMSE%	Bias	Bias%
Height (m)	1.21	6.8%	-0.20	-0.1%
Crown Width (m)	1.47	16.4%	-0.18	-2%

5.5 Validations for the ALS-predicted DBH and Tree Carbon

The results of model fitting and validation of the six ALS-DBH regression models are listed in Table 5.7. The 40 field-measured samples were previously split into 4 groups and 10 trees in each group. The models were iteratively fitted by 20 trees selected from two groups out of four and were validated by the rest 20 trees. Model 2 was selected as the overall best model to predict DBH in this study because it has a relatively high coefficient of determination ($R^2=86\%$) from model fitting and a relatively low RMSE (5.6 cm) from the validation. The regression equation is as:

$$DBH = -11.2792 + (-0.2958) \times CD + 3.2637 \times H \tag{5.1}$$

To determine the number of decimal places for the coefficients in the regression equation, the residuals between the DBH values predicted by the coefficients with eight decimal places and that predicted by coefficients rounded to two, three, and four decimal places, respectively are compared in Figure 5.7. To keep high prediction accuracy, coefficients in the regression model are rounded to four decimal places. For comparison purposes, Table 5.8 shows the field-based DBH model that is constructed by using all the field-measured height and crown width as explanatory variables and the validation result generated by comparing the field-measured DBH with the modeled DBH using the ALS-derived height and crown width. The R^2 between field-measured DBH and field-measured dendrometric parameters (height and crown width) was 0.79. The RMSE of the validation was 6.4 cm. The accuracy of the field-based model was approximately the same as the average accuracy of the six models. No significant bias was observed. Moreover, the R^2 between field-measured tree height and crown width was calculated as 0.34 indicating insignificant correlation between these two variables. Though DBH cannot be directly measured on the CHM, all the generated ALS-

DBH models showed that the DBH correlated well with ALS measurements. The accuracies of the ALS-modeled DBH and ALS-derived carbon storage are given in Table 5.9. The predicted DBH using ALS-derived parameters corresponded to a RMSE of 6.4 cm (relative RMSE = 13.1%) and a bias of 0.4 cm. The relationship between field-measured DBH and ALS-modeled DBH is plotted in Figure 5.8. ALS-derived carbon storage was obtained by plugging the ALS-modeled DBH into the Canadian national biomass equations (see Chapter 4, Equation 4.10). The results are compared with the reference carbon storage estimated by field-measured DBH and plotted in Figure 5.9. The R^2 of both DBH and carbon storage were both above 0.80. The predicted carbon storage using ALS-modeled DBH corresponded to a RMSE of 142 kg (28.6%) and a bias of 14.4 kg.

Table 5.7 Results of model fitting and model validation				
Model	Model Fit R^2	Model Fit RMSE (cm)	Validation R^2	Validation RMSE (cm)
1	0.83	5.35	0.80	6.82
2	0.86	6.60	0.76	5.60
3	0.86	3.86	0.71	8.25
4	0.75	7.89	0.77	4.82
5	0.78	5.59	0.85	6.55
6	0.81	7.00	0.83	5.20

Table 5.8 DBH model developed by all-field measurements and validation statistics			
# of training samples	Model Equation	Model R^2	Validation RMSE (cm)
40	$DBH = -1.4394 + 0.6193 \times CD + 2.5520 \times H$	0.79	6.41

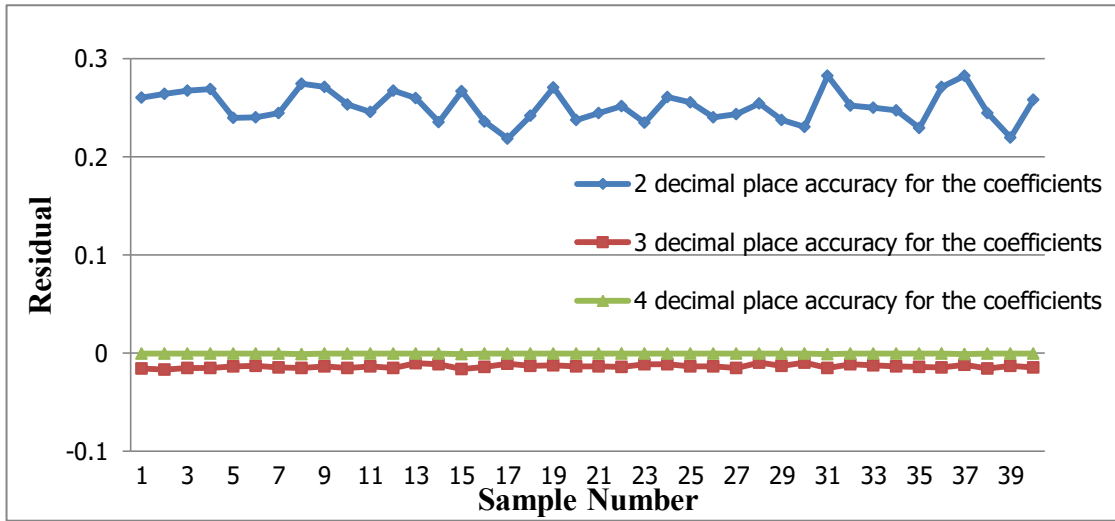


Figure 5.7 Residual plots for model-predicted DBH generated by coefficients with 2, 3, and 4 decimal places

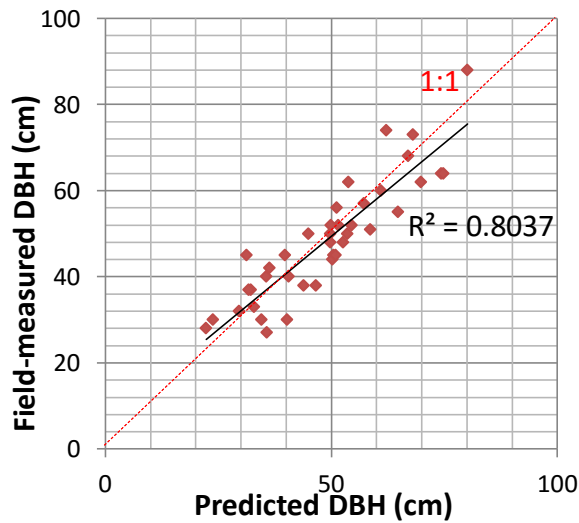


Figure 5.8 Scatterplot of the ALS-modeled DBH vs. field-measured DBH

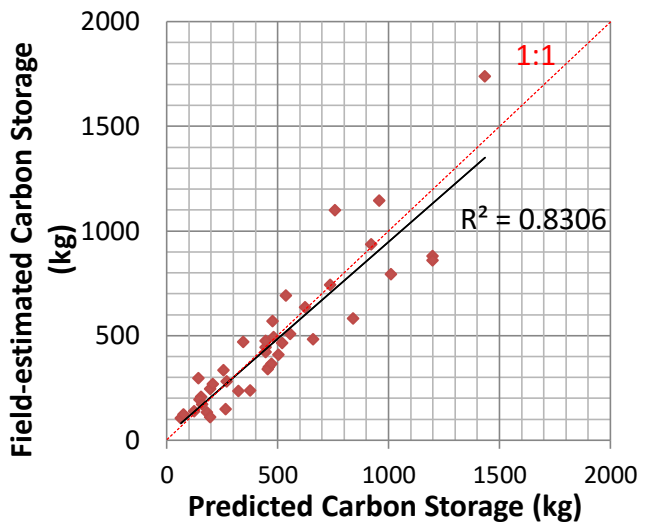


Figure 5.9 Scatterplot of the ALS-predicted carbon storage vs. field-estimated carbon storage

Parameter	RMSE	RMSE%	Bias	Bias%
DBH (cm)	6.39	13.1%	0.44	0.1%
Carbon (kg)	142.0	28.6%	14.4	2.9%

5.6 Analysis of the ALS-estimated Results

The accuracy of tree-height measurements using ALS data has previously been studied by Yu et al. (2004), Kaartinen et al. (2012), and Hadaś & Estornell (2016). Yu et al. (2004) found that as flying height increased from 400 m to 1500 m the accuracy of tree heights lowered from 0.76 to 1.16 m for single tree species. Kaartinen et al. (2012) reported that the best methods which utilized the local maxima finding with a point density of 8 points / m² can obtain a RMSE of 60 cm to 80 cm for tree heights. Hadaś & Estornell (2016) showed that the bias of tree height measurements can be decreased from -1.48 m to -0.72 m if the point density increased from 3.5 points / m² to 9 points / m². The RMSE achieved in the present study is in line with these studies and is potentially affected by the errors generated during the field measurements. Both overestimations of the crown size and underestimation of the tree heights are likely a result of the overlaying crown covers of the dominant tree and the suppressed trees but could be mitigated if the resolution of CHM is at the sub-meter level.

The results of ALS-modeled DBH are in line with the finding in Hauglin et al. (2014) and Popescu et al. (2007) regarding the tree height and crown width as good predictors to predict DBH using linear regression. The accuracy found in the present study is higher than these two studies. Hauglin et al. (2014) reported a RMSE of 35% for ALS-estimated DBH of Norway spruce. Popescu et al. (2007) reported a lower RMSE (4.9 cm) and higher R² compared to the present study. However, considering the accuracies in Popescu et al. (2007) was for single tree species which had an average DBH as 29.55 cm meanwhile they used all 43 sampled trees to construct the model and validate the model using the same dataset; the RMSE% and R² in their study would be higher than this study. For aboveground biomass Popescu et al. (2007) reported a RMSE of 47% for single tree species. Kankare et al. (2013)

reported a RMSE of 26.3% and 36.8% for pines and spruce, respectively. Huaglin et al. (2014) achieved a RMSE of 35.1% for biomass estimation. Though the achieved accuracy of carbon estimation is higher in the present study, these studies are not entirely comparable because some studies used field-destructive measurements as reference data which is not available in this study. And because genus information is also not available, the estimation of carbon stocks was done by the allometry equations for all species in Lambert et al. (2005) so that the derived carbon mainly depended on the ALS-derived DBH and height with little considerations given to the differences in species.

5.7 Analysis of Carbon Storage

There are a total of 2555 dominant trees in the study area. The trees were located along the roadsides, in the backyard, and around the lakes. The average tree carbon was 484.3 kg and resulted in total 1.24 kt C (10^3 tons Carbon). The study area can be generally divided into four land use types: residential, park and recreational, open area, and water. The open area occupied the largest portion of the study area which had 47.0 ha. The residential area occupied 25.8 ha. And the park had a small area of 2.5 ha. The extracted carbons in trees were therefore grouped based on the land use type and the amount of carbon stored in each land cover type was calculated.

Within the study area, open areas contained the largest tree carbon stocks (682.7 t), followed by residential areas (362.6 t), and last parks and recreational areas (29.2 t). The tree carbon storage for the open area, residential, and parks on a per-unit-area basis were 14.54 t C/ha, 14.08 t C/ha, and 11.57 t C/ha, respectively. In the citywide, open area occupied 83.7% (191.5 km^2) of the total city area and contained the largest carbon storage as 278.4 kt C. Residential area covered 8.3% (19.4 km^2) of the total city area and contained 27.3 kt C tree carbons. Parks covered 12.8 km^2 with a total carbon storage estimated as 14.8 kt C. The carbon storage map shows that large tree carbon stocks are accumulated in urban environment and are distributed heterogeneously among land use types (see Figure 5.10).

The estimated tree carbons in the study area are in line with the estimation of carbon storage in Canadian urban trees conducted by researchers at Environment Canada (Pasher et

al., 2014). Pasher et al. (2014) estimated the carbon stocks in trees by applying the crown cover area of urban trees and a Canadian-specific area-based growth rate for urban trees. Pasher et al. (2014) reported a total urban area of 5317 km² in Ontario Mixedwood Plains with an estimation of carbon storage as 9177.6 kt C resulting in a carbon storage per unit urban area as 17 kt C/ha. This is slightly more than what has been predicted in this study (around 14 kt C/ha). The main source of the difference comes from the approaches to estimate the biomass amounts between the present study and that used in Pasher et al. (2014).

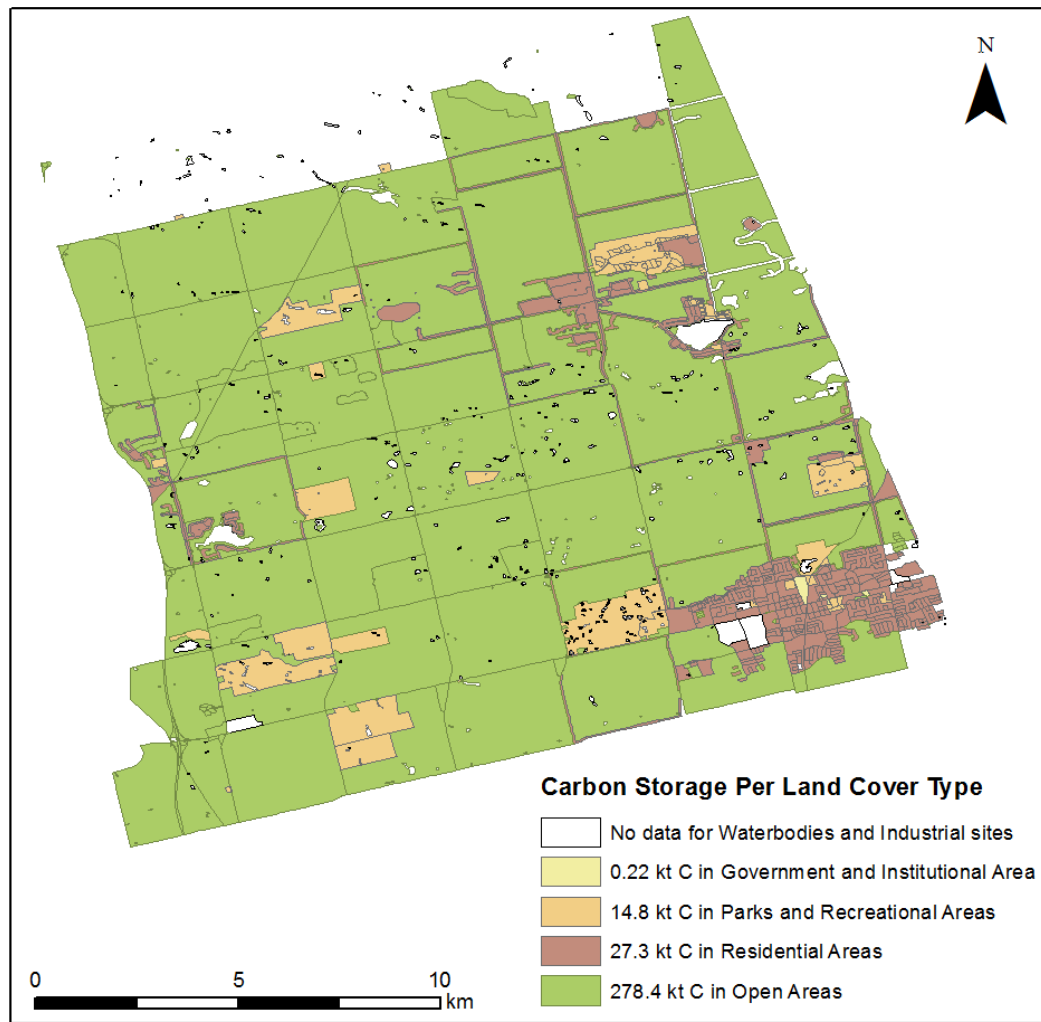


Figure 5.10 Carbon storage map for the Town of Whitchurch-Stouffville, Ontario

5.8 Chapter Summary

This chapter provided the results generated by the stepwise approaches proposed in Chapter 4. A classification map was created for the study area by using the SVM classifier with the multispectral ALS datasets. An analysis of the spectral behaviors of land covers in multispectral ALS channels was presented. The carbon storage in trees was estimated at the individual-tree level by applying a modified marker-controlled watershed segmentation algorithm and the allometry-based ALS-DBH modeling. The accuracies of ground-point filtering, the land cover classification, segmentation results, and the ALS-derived dendrometric parameters were provided. The parameterization process for the ALS-DBH modeling and the rationale behind parameter selection were presented. The tree carbon in the study area was estimated and used to extrapolate the carbon storage in the whole city. A carbon storage map for the Town of Whitchurch-Stouffville, Ontario was presented.

Chapter 6

Conclusions and Recommendations

In this chapter, the key findings and contributions of this study are summarized. The limitations of the proposed methodology are also discussed. And recommendations for further studies in terms of the applications of the multispectral ALS data are provided at the end.

6.1 Conclusions

This study proposed a workflow to map land covers and estimate aboveground carbon storage in trees at the spatial resolution of 1 m using the multispectral ALS data. This study shows that good classification result can be obtained solely from the multispectral ALS datasets. It achieved an overall accuracy of 90% which was 11% higher than that obtained using a single-wavelength ALS data. Spectral patterns for imperious surfaces (road, building, rooftops) and single-return vegetation (grass) are observed to have similar patterns as in the optical imagery. The dendrometric parameters at single-tree level can be derived directly from the multispectral ALS data. The study also shows the use of both spectral and geometric properties of the multispectral ALS data can improve the detection of treetops. The improvement could be more significant if the resolution of the CHM gets finer.

The study presented the feasibility of applying forest based allometric methods to assess carbon stocks in urban environment. Dominant trees with less underneath or nearby trees were better detected and analyzed in the proposed study. Though DBH cannot be directly measured from the ALS data, the ALS-predicted DBH remains to be a power predictor to estimate tree carbon at the individual-tree level. More accurate tree carbon could be obtained if genus information and crown base heights are further investigated. An improvement of derivation of the crown width would also help in better prediction of tree carbon stocks. This study derived similar carbon amounts per unit area on both residential area and open area within the study area because the open area has twice the size of the residential area but the density of the canopy covers is less than it in the residential area. Citywide carbon storage

estimation is derived in this study by extrapolating the values within the study area to the entire city based on the specific proportion of each land cover type in the entire city. This approach is applicable in the present study because the Town of Whitchurch-Stouffville has a relatively simple city structures and the study area has included the major components of the city and those land covers excluded in the study area only occupied a little proportion of the entire city.

Urban ecosystems are an important component in the global carbon cycle. In the context of urban sprawl, quantifying the carbon storage for urban areas is very important in terms of getting reliable estimation of carbon sequestration rate and magnitude. But it is a difficult and complex task that requires advanced analysis techniques and data sources to achieve fine-scale estimation. The methods developed here provide an accurate and detailed estimate of how urban trees in a Canada's city plays the role as a carbon sink. The presented approach of estimating carbon stocks in urban trees takes the advantages of the available Canada-wide allometry relationship between biomass and the tree DBH and height, and also the power of the ALS system in providing the estimation of dendrometric parameters. The methodology proposed in the present study does not require destructive sampling or large-scale field works. It is applicable to other urban areas and is beneficial to better understand urban carbon budgets and urban heat island effects. It also provides valuable information on the impact of climate change to city planners.

In conclusion, the present study has developed a detailed workflow to estimate tree carbon stocks from the multispectral ALS data by using a series of techniques including SVM classification, watershed segmentation, and allometry-based linear regression modeling. This study also demonstrated the strong capability of multispectral ALS data in land cover mapping and tree-level inventory in urban environment. The Titan scanner showed a trend for the development of airborne LiDAR technology. This scanner is able to realize a series of environmental and topographic applications that have been done previously by photogrammetric data.

6.2 Recommendations

Nevertheless, the 3D information of ALS technology has been widely applied in vegetation analysis. The information of ALS intensity was less frequently used in previous studies and only played as an ancillary role in land cover mapping and forest inventory. The findings in the present study suggest that the use of the multispectral ALS intensities may be undervalued and should be investigated further. The understanding of multiple-return intensities will be more complex and confused than single-channel ALS or optical imagery. Hence, the usefulness of the multispectral ALS intensity depends on the calibration process which requires in-depth investigation in the future. Though superior object-based classification methods have been proposed in recent years, the analysis of the multispectral ALS intensity data showed that individual land cover types cannot be distinguished clearly in single ALS channels. Hence, the classification method applied in this study remains to be a pixel based method. However, it is strongly recommended that researchers conduct in-depth analysis of the NIR and SWIR intensity bands and examine the applicability of performing object-based classification on individual bands in order to classify tree genus. The present study reveals some clues of the potential of the multispectral ALS data in providing fine-scale and class-diverse land cover mapping. The classification of hardwood and softwood is especially important for vegetation analysis. Though this study did not achieve a tree-genus level of classification, it is still highly possible to be accomplished by taking full usage of the 3D geometry and intensity information of the multispectral ALS datasets. The proposed work should be extended further in terms of estimating the urban vegetation conditions for specific species, estimating the carbon sequestration and emission rate.

References

- Akbari H, Matthews HD, Seto D, 2012. The long-term effect of increasing the albedo of urban areas. *Environmental Research Letters* 7(2), 024004.
- Aldred H, Bonnor M, 1985. Application of airborne lasers to forest surveys. Petawawa National Forestry Institution, Information Report, NO. PI-X-51, pp.1-62.
- Axelsson, P, 2000. DEM generation from laser scanner data using adaptive TIN models. *ISPRS Archives* 33(B4), 111-118.
- Bakuła K, 2015. Multispectral airborne laser scanning—a new trend in the development of LiDAR technology. *Archiwum Fotogrametrii, Kartografii i Teledetekcji* 27, 25-44.
- Ben-Arie JR, Hay GJ, Powers RP, Castilla G, St-Onge B, 2009. Development of a pit filling algorithm for LiDAR canopy height models. *Computers & Geosciences* 35(9), 1940-1049.
- Beraldin J, Blais F, Lohr U, 2010. Laser scanning technology, In: Vosselman G and Maas H (eds.), *Airborne and Terrestrial Laser Scanning*. Whittles Publishing, 1-42.
- Beucher S, Lantuejoul C, 1979. Use of watersheds in contour detection. *International Workshop on Image Processing: Real-Time Edge and Motion Detection/Estimation 1979*, Rennes, France, 12–21.
- Blaschke T, 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65(1), 2-16.
- Brook McIlroy, 2002. Community of Stouffville urban design guidelines. http://www.townofws.ca/en/residents/resources/Documents/Planning/Stouffville_UrbanDesignGuidelines.pdf, Accessed 20 February 2016.
- Carlson TN, Arthur ST, 2000. The impact of land use—land cover changes due to urbanization on surface microclimate and hydrology: a satellite perspective. *Global and Planetary Change* 25(1), 49-65.
- Casas Á, García M, Siegel RB, Koltunov A, Ramírez C, Ustin S, 2016. Burned forest characterization at single-tree level with airborne laser scanning for assessing wildlife habitat. *Remote Sensing of Environment* 175, 231-241.
- Chen Q, Baldocchi D, Gong P, Kelly M, 2006. Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering & Remote Sensing* 72(8), 923-932.

Chen Y, Su W, Li J, Sun Z, 2009. Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas. *Advances in Space Research* 43(7), 1101-1110.

Corfee-Morlot J, Kamal-Chaoui L, Donovan MG, Cochran I, Robert A, Teasdale PJ, 2009. Cities, climate change and multilevel governance. *OECD Environment Working Papers*, No. 14. doi: 10.1787/220062444715.

Culvenor DS, 2002. TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Computers & Geosciences* 28(1), 33-44.

Dalponte M, Bruzzone L, Gianelle D, 2011. A system for the estimation of single-tree stem diameter and volume using multireturn LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing* 49(7), 2479-2490. doi: 10.1109/TGRS.2011.2107744.

Davies ZG, Edmondson JL, Heinemeyer A, Leake JR, Gaston KJ, 2011. Mapping an urban ecosystem service: quantifying above-ground carbon storage at a city-wide scale. *Journal of Applied Ecology* 48(5), 1125-1134.

DMTI Spatial Inc., 2015. Land cover region. Geospatial Center at University of Waterloo. http://geol.scholarsportal.info.proxy.lib.uwaterloo.ca/#r/details/_uri@=4259707280, Accessed 10 January 2016.

Dodman D, Bicknell J, Satterthwaite D (eds.), 2012. *Adapting Cities to Climate Change: Understanding and Addressing the Development Challenges*. Earthcan, London.

Donovan GH, Butry DT, 2011. Trees in the city: valuing street trees in Portland, Oregon. *Landscape and Urban Planning* 94(2), 77-83.

Elmqvist M, Jungert E, Lantz F, Persson A, Soderman U, 2001. Terrain modelling and analysis using laser scanner data. *ISPRS Archives* 34(3-W4), 219-226.

Environment Canada, 2015. National inventory report 1990-2013: greenhouse gas sources and sinks in Canada - executive summary. <http://www.ec.gc.ca/GES-GHG/default.asp?lang=En&n=5B59470C-1&offset=7&toc=show>, Accessed 20 March 2016

Estornell J, Ruiz LA, Velázquez-Martí B, López-Cortés I, Salazar D, Fernández-Sarría A, 2015. Estimation of pruning biomass of olive trees using airborne discrete-return LiDAR data. *Biomass and Bioenergy* 81, 315-321.

Google Earth 7.1.2.2041, 2004. Town of Whitchurch-Stouffville. 44°01'29.7"N 79°16'03.6"W, Eye alt 582 m. First Base Solutions 2016. <http://www.earth.google.com>, Accessed February 26 2016.

Google Earth 7.1.2.2041, 2015. Town of Whitchurch-Stouffville.44°01'29.7"N 79°16'03.6"W, Eye alt 1.41 km. First Base Solutions 2016.
<http://www.earth.google.com>, Accessed February 26 2016.

Gougeon FA, 1995. A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. *Canadian Journal of Remote Sensing* 21(3), 274-284.

Guo L, Chehata N, Mallet C, Boukir S, 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(1), 56-66.

Hadaś E, Estornell J, 2016. Accuracy of tree geometric parameters depending on the LiDAR data density. *European Journal of Remote Sensing* 49, 73-92.

Hartfield K A, Landau K I, Van Leeuwen W J, 2011. Fusion of high resolution aerial multispectral and LiDAR data: land cover in the context of urban mosquito habitat. *Remote Sensing* 3(11), 2364-2383.

Hauglin M, Gobakken T, Astrup R, Ene L, Næsset E, 2014. Estimating single-tree crown biomass of Norway spruce by airborne laser scanning: a comparison of methods with and without the use of terrestrial laser scanning to obtain the ground reference data. *Forests* 5(3), 384-403.

Hebel M, Stilla U, 2012. Simultaneous calibration of ALS systems and alignment of multiview LiDAR scans of urban areas. *IEEE Transactions on Geoscience and Remote Sensing* 50(6), 2364-2379. doi: 10.1109/TGRS.2011.2171974.

Heidemann H, Stoker J, Brown D, Olsen M, Singh R, Williams K, Chin A, Karlin A, McClung G, Janke J, Shan J, Kim K, Sampath A, Ural S, Parrish C, Waters K, Wozencraft J, Macon C, Brock J, Wright C, Hopkinson C, Pietroniro A, Madin I, Conner J, 2012. Applications, In: Renslow MS (eds.), *Manual of Airborne Topographic Lidar*. American Society for Photogrammetry and Remote Sensing, pp. 283-423.

Heurich M, 2008. Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the Bavarian Forest National Park. *Forest Ecology and Management* 255(7), 2416-2433.

Huang Y, Yu B, Zhou J, Hu C, Tan W, Hu Z, Wu J, 2013. Toward automatic estimation of urban green volume using airborne LiDAR data and high resolution remote sensing images. *Frontiers of Earth Science* 7(1), 43-54.

Hyypä J, Hyypä H, Leckie D, Gougeon F, Yu X, Maltamo M, 2008. Review of methods of small - footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing* 29(5), 1339-1366.

Johnson, T, 2009. Deforestation and greenhouse-gas emissions. <http://www.cfr.org/forests-and-land-management/deforestation-greenhouse-gas-emissions/p-14919>, Accessed February 01 2016.

Kaartinen H, Hyypä J, Yu X, Vastaranta M, Hyypä H, Kukko A, Holopainen M, Heipke C, Hirschmugl M, Morsdorf F, Næsset E, 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing* 4(4), 950-974.

Kankare V, Rätty M, Yu X, Holopainen M, Vastaranta M, Kantola T, Hyypä J, Hyypä H, Alho P, Viitala R, 2013. Single tree biomass modelling using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 85, 66-73.

Kavzoglu T, Colkesen I, 2009. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation* 11(5), 352-359.

Ke Y, Quackenbush LJ, 2011. A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. *International Journal of Remote Sensing* 32(17), 4725-4747.

Koch B, Kattenborn T, Straub C, Vauhkonen J, 2014. Segmentation of forest to tree objects, In: Maltamo M, Naesset E, Vauhkonen J (eds.), *Forestry Applications of Airborne Laser Scanning*. Springer, pp. 89-112.

Korpela I, Ørka HO, Hyypä J, Heikkinen V, Tokola T, 2010. Range and AGC normalization in airborne discrete-return LiDAR intensity data for forest canopies. *ISPRS Journal of Photogrammetry and Remote Sensing* 65(4), 369-379.

Lambert M C, Ung C H, Raulier F, 2005. Canadian national tree aboveground biomass equations. *Canadian Journal of Forest Research* 35(8), 1996-2018.

Leeuwen M, Nieuwenhuis M, 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research* 129(4), 749-770.

Lieth, H, 1963. The role of vegetation in the carbon dioxide content of the atmosphere. *Journal of Geophysical Research* 68, 3887-3898.

Lodha SK, Kreps EJ, Helmbold DP, Fitzpatrick DN, 2006. Aerial LiDAR data classification using Support Vector Machines (SVM). *3DPVT 2006*, pp. 567-574. doi: 10.1109/3DPV-T.2006.23.

- Mather PM, Koch M, 2011. *Computer Processing of Remotely-sensed Images: An Introduction*. 4th ed. John Wiley & Sons, London.
- Meng X, Currit N, Zhao K, 2010. Ground filtering algorithms for airborne LiDAR data: a review of critical issues. *Remote Sensing* 2(3), 833-860.
- Meyer F, Beucher S, 1990. Morphological segmentation. *Journal of Visual Communication and Image Representation* 1(1), 21-46.
- Najman L, Schmitt M, 1996. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12), 1163-1173. doi: 10.1109/34.546254.
- Nowak, D. J., 1993. Atmospheric carbon reduction by urban trees. *Journal of Environmental Management* 37, 207-217.
- Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA, 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment* 148, 42-57.
- Pasher J, McGovern M, Khoury M, Duffe J, 2014. Assessing carbon storage and sequestration by Canada's urban forests using high resolution earth observation data. *Urban Forestry & Urban Greening* 13(3), 484-494.
- Pollock RJ, 1996. *The Automatic Recognition of Individual Trees in Aerial Images of Forests Based on A Synthetic Tree Crown Image Model*, PhD Thesis, Department of Computer Science, University of British Columbia.
- Pontius Jr RG, Millones M, 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32(15), 4407-4429.
- Popescu SC, Wynne RH, Nelson RF, 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing* 29(5), 564-577.
- Popescu SC, 2007. Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy* 31(9), 646-655.
- Raciti SM, Hutya LR, Newell JD, 2014. Mapping carbon storage in urban trees with multi-source remote sensing data: relationships between biomass, land use, and demographics in Boston neighborhoods. *Science of the Total Environment* 500, 72-83.

Roy S, Byrne J, Pickering C, 2012. A systematic quantitative review of urban tree benefits, costs, and assessment methods across cities in different climatic zones. *Urban Forestry & Urban Greening* 11(4), 351-363.

Sawka M, Millward AA, Mckay J, Sarkovich M, 2013. Growing summer energy conservation through residential tree planting. *Landscape and Urban Planning* 113, 1-9.

Schreyer J, Tigges J, Lakes T, Churkina G, 2014. Using airborne LiDAR and QuickBird data for modelling urban tree carbon storage and its distribution—a case study of Berlin. *Remote Sensing* 6(11), 10636-1055.

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

Sithole G, Vosselman G, 2001. Filtering of laser altimetry data using a slope adaptive filter. *ISPRS Archives* 34(3-W4), 203-210.

Sithole G, Vosselman G, 2005. Filtering of airborne laser scanner data based on segmented point clouds. *ISPRS Archives* 36(3-W19), 66-71.

Skaloud J and Schwartz K P, 2000. Accurate orientation for airborne mapping systems. *Photogrammetry Engineering and Remote Sensing* 66(4), 393–401.

Statistics Canada, 2015. Census profile Town of Whitchurch-Stouffville. <http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=3519044&Geo2=CD&Code2=3519&Data=Count&SearchText=Whitchurch-Stouffville&SearchType=Begins&SearchPR=01&B1=All&Custom=&TABID=1>, Accessed 20 February 2016.

Solberg S, Naesset E, Bollandsas OM, 2006. Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. *Photogrammetric Engineering & Remote Sensing* 72(12), 1369-1378.

Sohn G, Dowman IJ, 2002. Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion. *ISPRS Archives* 34(3-A), 336-344. doi:10.1016/j.isprsjprs.2007.01.001.

Sousa AM, Gonçalves AC, Mesquita P, da Silva JR, 2015. Biomass estimation with high resolution satellite images: a case study of *Quercus rotundifolia*. *ISPRS Journal of Photogrammetry and Remote Sensing* 101, 69-79.

Teledyne Optech Titan, 2015. Multispectral LiDAR system: high precision environmental mapping, <http://www.teledyneoptech.com/wp-content/uploads/Titan-Specsheet-150515-WEB.pdf>, Accessed June 05 2015.

Tong X, Li X, Xu X, Xie H, Feng T, Sun T, Jin Y, Liu X, 2014. A two-phase classification of urban vegetation using airborne LiDAR data and aerial photography. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7(10), 4153-4166.

Tovari D, Pfeifer N, 2005. Segmentation based robust interpolation-a new approach to laser data filtering. *ISPRS Archives* 36(3-W19), 79-84.

United Nations, 2014. World urbanization prospects 2014: highlights. <http://esa.un.org/unpd/wup/highlights/wup2014-highlights.pdf>, Accessed February 01 2016.

Vauhkonen J, Ene L, Gupta S, Heinzel J, Holmgren J, Pitkänen J, Lien V, 2012. Comparative testing of single-tree detection algorithms under different types of forest. *Forestry* 85(1), 27-40.

Vauhkonen J, Maltamo M, McRoberts RE, Næsset E, 2014. Introduction to forestry applications of airborne laser scanning, In: Maltamo M, Naesset E, Vauhkonen J (eds.), *Forestry Applications of Airborne Laser Scanning*. Springer, 1-16.

van den Berg M, Wendel-Vos W, van Poppel M, Kemper H, van Mechelen W, Maas J, 2015. Health benefits of green spaces in the living environment: a systematic review of epidemiological studies. *Urban Forestry & Urban Greening* 14(4), 806-816.

Vosselman G, 2000. Slope based filtering of laser altimetry data. *ISPRS Archives* 33(B3), 935-942. doi: 10.5194/isprsannals-II-3-W5-113-2015.

Wichmann V, Bremer M, Lindenberger J, Rutzinger M, Georges C, Petrini-Monteferri F, 2015. Evaluating the potential of multispectral airborne LIDAR for topographic mapping and land cover classification. *ISPRS Annals II-3*, 113-119. doi:10.5194/isprsannals-II-3-W5-113-2015.

Whittaker R H, Likens G E, 1973. Carbon in the biota, In: G M Woodell and E V Pecans (eds), *Carbon and the Biosphere*. Springfield/Virginia: U.S. National Technical Information Service, pp. 281-302.

Wulder M, Niemann KO, Goodenough DG, 2000. Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery. *Remote Sensing of Environment* 73(1), 103-114.

Xu Q, Hou Z, Maltamo M, Tokola T, 2014. Calibration of area based diameter distribution with individual tree based diameter estimates using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 93, 65-75.

Xu T, Sathaye J, Akbari H, Garg V, Tetali S, 2012. Quantifying the direct benefits of cool roofs in an urban setting: reduced cooling energy use and lowered greenhouse gas emissions. *Building and Environment* 48, 1-6.

Yan WY, Shaker A, Habib A, Kersting AP, 2012. Improving classification accuracy of airborne LiDAR intensity data by geometric calibration and radiometric correction. *ISPRS Journal of Photogrammetry and Remote Sensing* 67, 35-44.

Yan W Y, Shaker A, El-Ashmawy N, 2015. Urban land cover classification using airborne LiDAR data: a review. *Remote Sensing of Environment* 158, 295-310.

Yu X, Hyypä J, Kaartinen H, Maltamo M, 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment* 90(4), 451-462.

Zhao K, Popescu S, Nelson R, 2009. Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment* 113(1), 182-196.

Zhang C, Zhou Y, Qiu F, 2015. Individual tree segmentation from LiDAR point clouds for urban forest inventory. *Remote Sensing* 7(6),7892-7913.

Zhang J, 2015. Single Tree Detection from Airborne Laser Scanning Data: A Stochastic Approach, PhD Thesis, Department of Earth and Space Science, York University.

Zhou W, 2013. An object-based approach for urban land cover classification: integrating LiDAR height and intensity data. *IEEE Geoscience and Remote Sensing Letters*, 10(4), 928-931.