Evaluating Urban Vegetation Cover Using LiDAR and High Resolution Imagery

Y.A. Ayad and D. C. Mendez
Clarion University of Pennsylvania

Abstract

One of the key planning factors in urban and built up environments is to identify the percentage of natural versus man made covers. Vegetation contributes to the health and the quality of its surrounding land covers. Its existence in the same fabric as the built environment adds to the health and well being of its inhabitants, creating a self sustaining environment and contributing to reducing its carbon footprint. This project is an attempt to assess and to quantify urban vegetation of the town of Clarion and to relate vegetation and tree cover density and percent coverage to the surrounding built and man made land covers using LiDAR data in conjunction with high resolution infrared aerial photography of the National Agricultural Image Program (NAIP). Two classes (tree density/trees and structures) were extracted using LiDAR LAS point cloud data directly by filtering different returns for classes 2 and 12. The remaining three classes (bare ground, roads, and other pavement) resulted from the non-supervised classification of the NAIP 4-band aerial photograph. The final classification resulted in an overall 77% accuracy. 77% of the study area was barely covered with any trees (0-10% density), while 11.71% were 70-100% covered. On the other hand, non-permeable surfaces constituted 50.65% of the study area while permeable surfaces occupied 49.35%. The project outcomes are summarized in the form of maps for different stages and outputs as well as a variety of tables reflecting the spatial distribution as well as the comparisons of different land cover percentages of the identified permeable and non-permeable surfaces within the study area.

Evaluating Urban Vegetation Cover Using LiDAR and High Resolution Imagery

Y.A. Ayad and D. C. Mendez
Clarion University of Pennsylvania

June 2014

Contents

Figure List

Table List

Introduction

In the past decades, many studies explored remote sensing technologies in studying forest ecosystems. Most recently, with the widespread use of LiDAR technology, many applications have emerged. Studies have extended to cover the possibilities of automating land cover extraction and the delineation of landscape features using hybrid methods that involve traditional image processing as well as innovative techniques for the extraction of features from LiDAR point cloud data (e.g. Chen *et al.,* (), and Syed *et al.,* 2005).

In this study, methods of LiDAR point cloud data manipulation combined with traditional multi-spectral image classification are applied in order to investigate the natural versus non-natural elements in the town of Clarion. Extracting those features would help quantifying permeable and non-permeable surfaces and assess the current built environment in Clarion. In addition, those methods would also help the estimation of the tree density cover in town.

Data

The following table summarizes the two main datasets that were needed in order to accomplish this project:

Data	Year	Format	Data Originator	Notes
LiDAR Elevation Points	2006	LAS	DCNR (PAMAP Program)	Classified
Infrared Digital Ortho Photos	2010	TIF	USDA-FSA Aerial Photography Field Office	Consists of 4 spectral bands (R, G, B and near IR)

It is important to note that there is a 4 year time difference between the two datasets which could lead to some discrepancies due to land cover changes between the two dates. Also, the LiDAR dataset was collected during leaf-off season (early spring) which would result in some difficulties calculating an accurate tree canopy cover in some areas of mostly deciduous cover.

Figure 1 The selected study area. The selection was based on the mostly built environment in Clarion.

Both datasets were downloaded in separate tiles. Those tiles were merged to cover the area of interest. All of the processes were restricted to the study area boundaries. The following figure shows the study area and the corresponding dataset tiles.

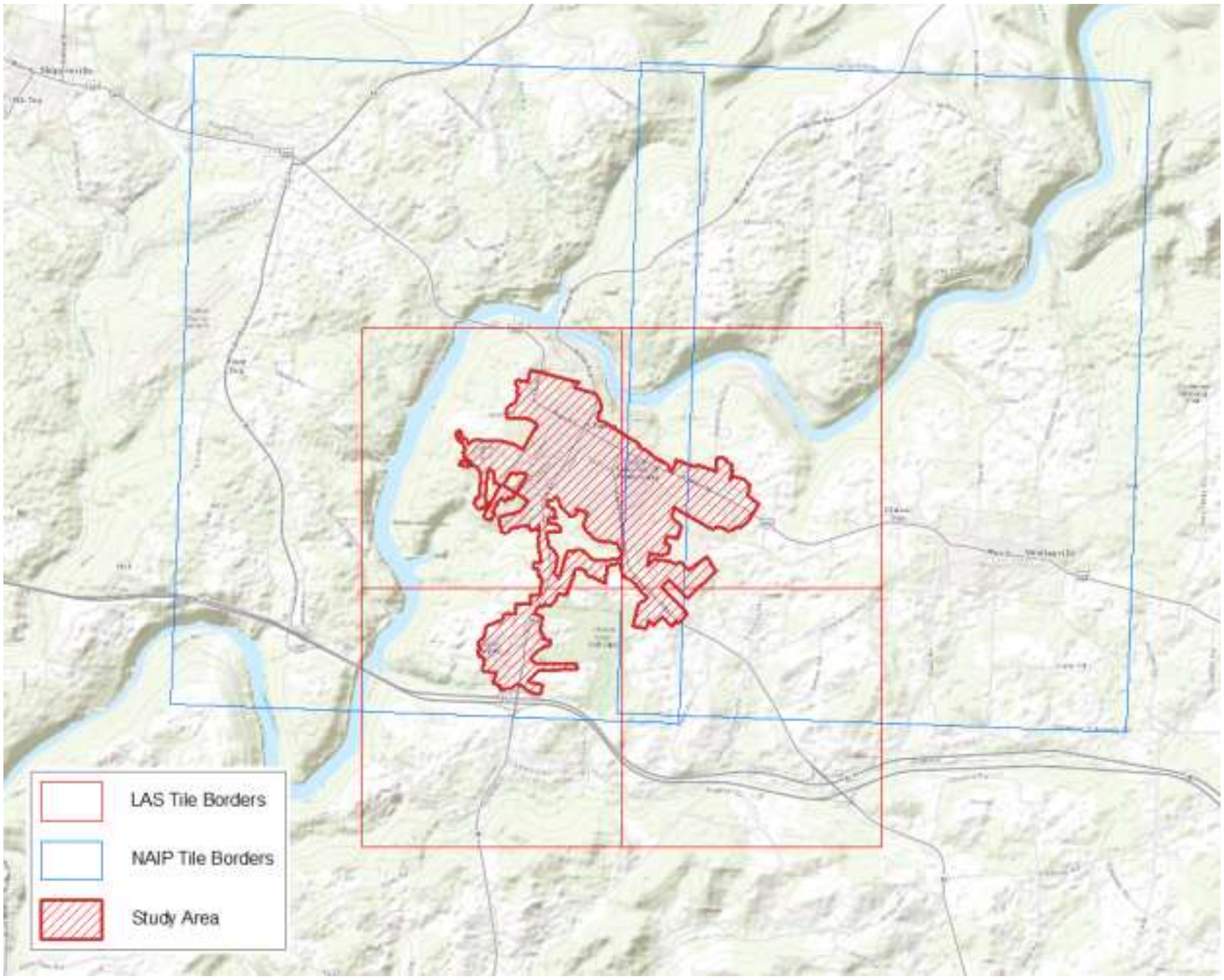

Figure 2 Borders of the different dataset tiles that cover the selected study area

Methods

The LAS dataset are already classified according to the following table (information extracted from file's metadata):

Class	Description
1	Default - These are the points that are a mixture of the remaining points after the ground classification. These would contain bridges, overpasses, buildings, cars, parts of vegetation, etc.
2	Ground - These are points on the bare earth surface. Th8ey are from the automated processing, as well as the manual surface review.
8	Model Key - These are the educated, thinned points to represent the final bare earth surface. This is from our automated processing. These are the points that we have used to generate the final contours.
9	Water - These are points inside of hydrographic features, as collected by photogrammetric methods. These are from automated processing, as well as the manual surface review.
12	Non-Ground - These are points that are identified as first of many return or intermediate of many returns from the LIDAR pulse. These are points that are most likely vegetation returns or points identified to be not on the ground surface.
15	Road Edges - These are the points that fall within +/- 1.5' of road break lines.

Source: LAS files metadata (http://www.pasda.psu.edu/uci/FullMetadataDisplay.aspx?file=pamap_lidar_LAS.xml)

This classification helped in most cases in the extraction of different features based on their assigned description. Nonetheless, it was challenging to extract those exact classes within the current study area due to the complexity of features as well as the season of the LAS dataset collection (leaf off). Therefore, the process was divided into four major tasks: extraction of the tree canopy, building and other structures, and ground only areas, as well as the classification of those ground only surfaces into natural and non-natural (man-made) classes. The following diagram summarizes the major steps for this project:

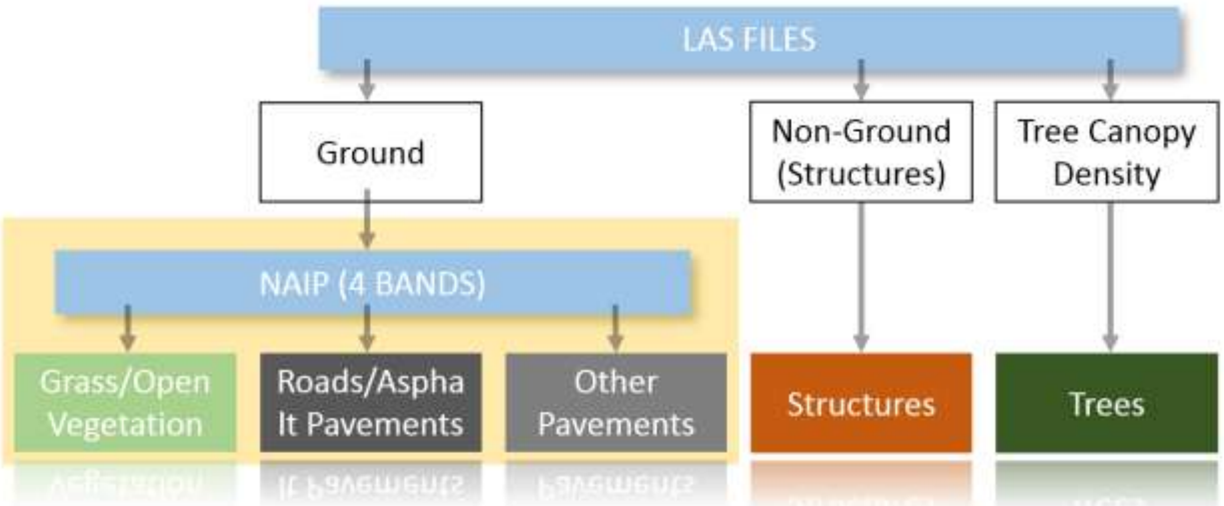

Figure 3 Flow chart summarizing the main inputs and outputs extracted from the project procedures

A model was constructed in order to help in automating the process and to apply the same process in other study areas with similar datasets (Figure 12).

The process was applied to the predefined area of interest in Clarion town. The model could be applied to any other area that is covered by the same datasets (PAMAP LiDAR LAS files as well as NAIP 4-band aerial photography).

Tree Canopy Cover Calculation

For tree canopy calculation, class number 2 with the second and third returns from the LAS point cloud data were selected. This filtered out any other possible classes that were mainly ground and most likely not trees. Also class number 12 was also added since most of it was inspected and proved to be mostly trees as well, all returns from this class were used. For both, point counts were calculated for a 10ftx10ft cell size and stored in a raster dataset. The same was done to the ground LAS class number 2 with all returns excluding the second and the third. The tree cover point counts were added then divided by the total of all counts (trees and ground). This resulted in a raster dataset containing the ratio between tree return count and all the count of all returns for each cell.

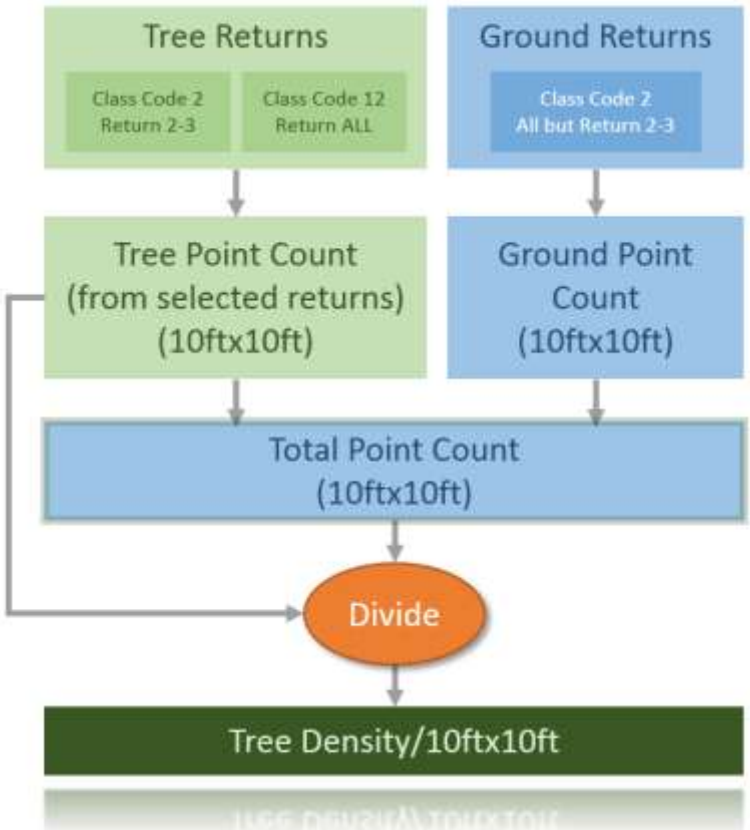

Figure 4 Flowchart of the tree canopy calculation using LiDAR's different classes and returns

The final result was then reclassified to 5 different classes and percentages were added to the attribute table.

Table 1 Tree density reclassification values

Value Range	Class
0-0.10	10
0.10-0.20	20
0.20-0.40	40
0.40-0.60	60
0.60-0.80	80
0.80-1.0	100

The areas between 5th avenue, almost at the vertical center of the study area and the Clarion University campus in the east, as well as the far west of the study area had the most concentration of tree cover in the study area, other locations scattered all over the study area contained less tree cover and more of the open areas and built up environment.

Figure 5 Tree density distribution in the study area.

Extraction of Buildings and Other Structures

The extracted ground areas from the ground point count in the previous step included a non-ground class that was reclassified to represent the structures which included any buildings or large enough object.

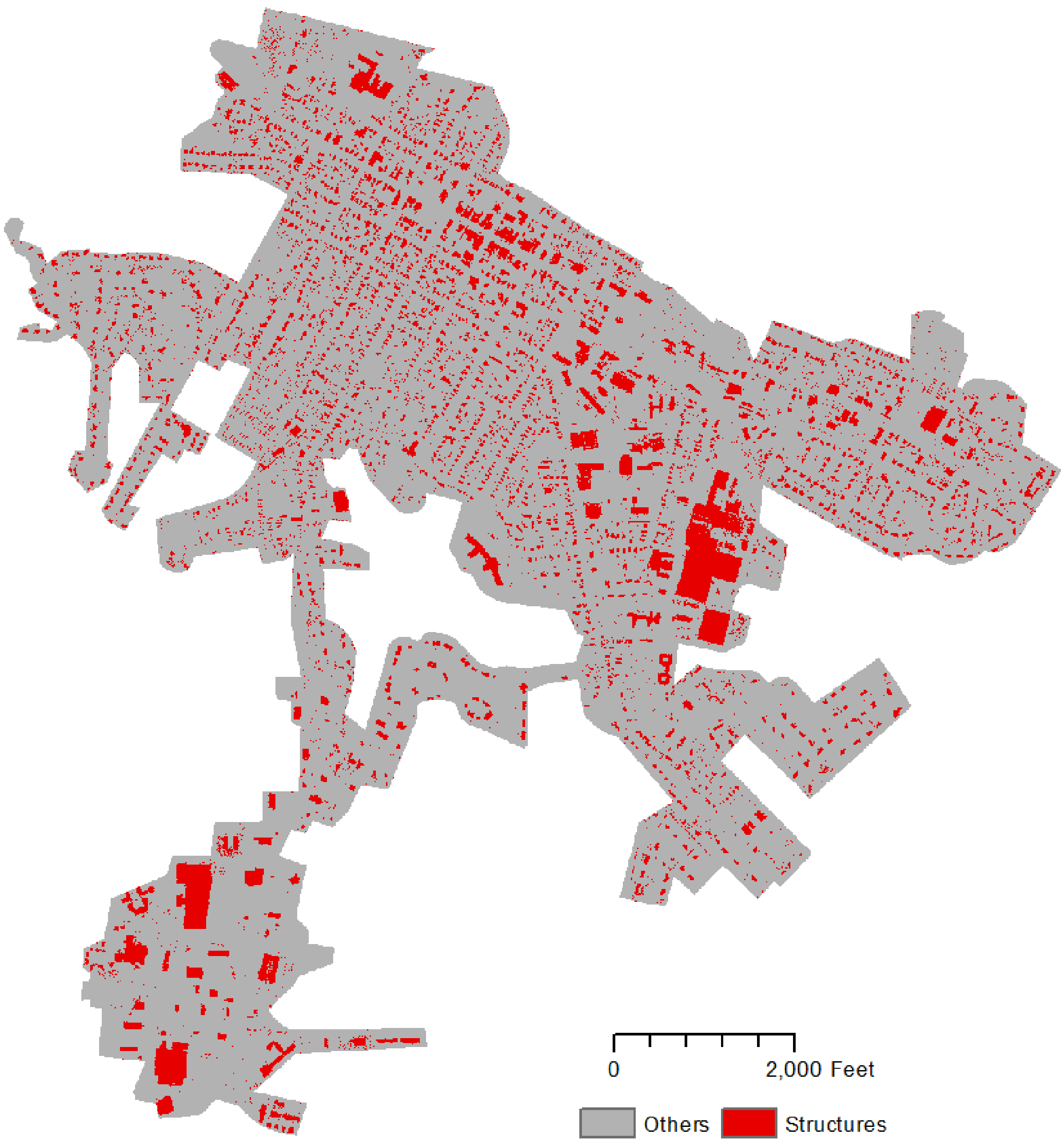

Figure 6 Extracted structures

Ground Only Extraction

The point count of the ground returns was reclassified in order to create a mask for ground areas. The areas that were previously classified as trees were then masked out from this ground dataset in order to extract the bare ground with no structures nor tree cover. Those were used as a mask for the 4-band NAIP aerial photograph for land cover classification.

Figure 7 Extracted bare ground areas

Ground Only Classification

The extracted bare ground areas were used to extract the multi-band images of the NAIP aerial photograph. Those were then classified using non-supervised method (ISO Cluster) to extract 8 different bare ground land cover classes. Those classes were then aggregated based on observations to three classes: 1-Open green areas, 2-Paved surfaces (mainly roads), and 3-Other paved surfaces (including gravel, artificial turfs, and non-vegetated ground.

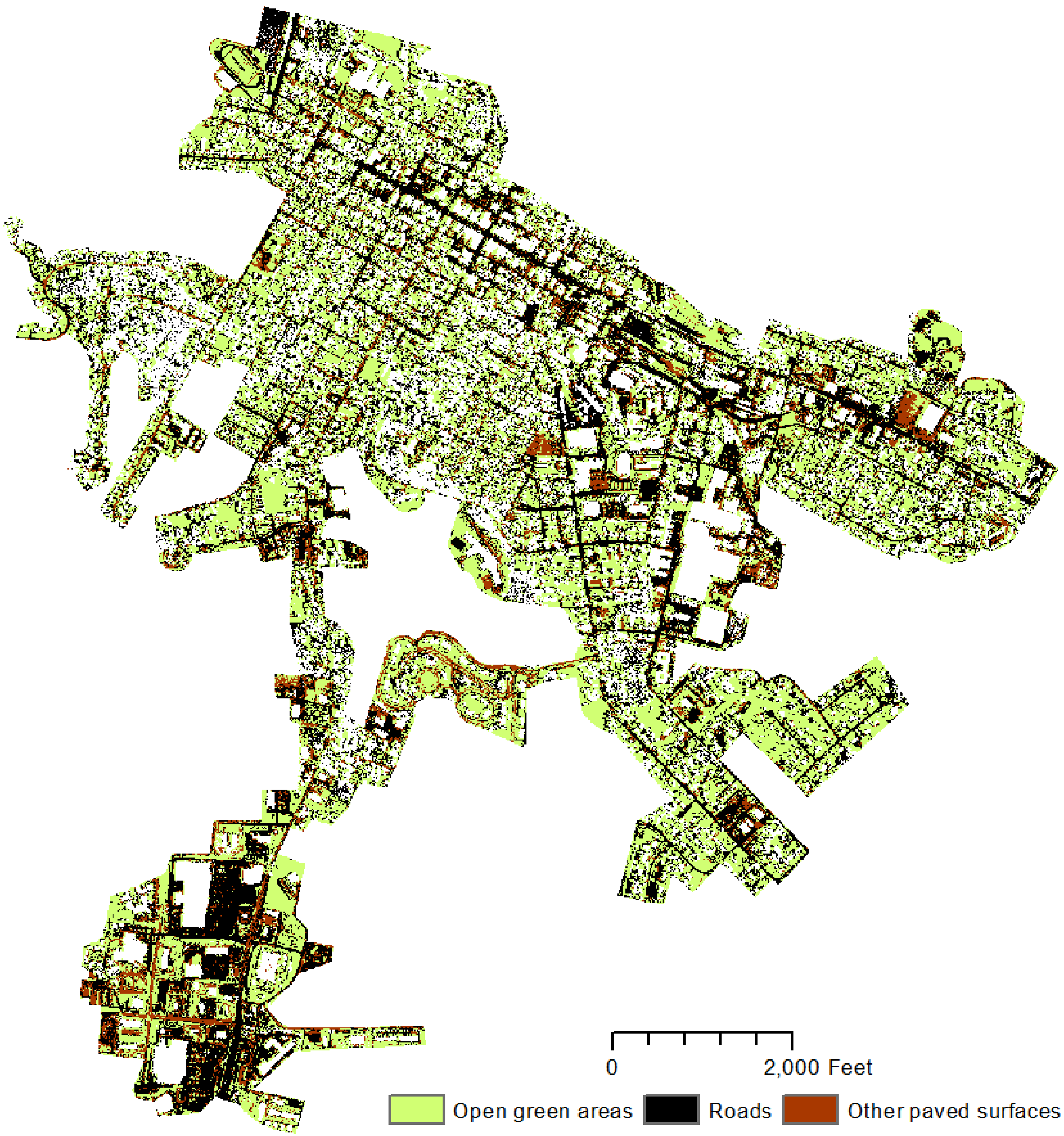

Figure 8 Classified bare ground areas

Assembling land cover Data

All three layers of land cover; trees, structures and ground (classified into three classes) were combined together into one raster dataset in order to facilitate their manipulation as well as the calculation of different cover percentages and to enable the process of accuracy assessment.

Calculation of the total area of each of the final five classes was added to the attribute table of the resulting land cover dataset.

Figure 9 Classified bare ground areas

Accuracy Assessment

300 random points were created within the study area and each was assigned the classified value as well as checked for accuracy from the NAIP aerial photograph. All true and classified values were recorded and summarized in a transition matrix (Table 2). The percent omitted from and committed to each class was calculated which revealed an overall 77% accuracy. The most accurately classified class was the open areas (class 1) with a total of 80.34% accuracy, and the lowest was the structures with 70% accuracy.

Table 2 Classification accuracy assessment: Transition Matrix showing counts as well as percentage of correctly classified points (total 300).

		Classified					
		1	2	3	4	5	
Truth	1	94 (80.34%)	16 (13.68%)	3 (2.56%)	2 (1.71%)	2 (1.71%)	117
	2	5 (7.58%)	52 (78.79%)	7 (10.61%)	0 (0%)	2 (3.03%)	66
	3	2 (10.53%)	2 (10.53%)	15 (78.95%)	0 (0%)	0 (0%)	19
	4	3 (7.5%)	4 (10%)	0 (0%)	28 (70%)	5 (12.5%)	40
	5	4 (6.9%)	11 (18.97%)	1 (1.72%)	0 (0%)	42 (72.41%)	58
		108	85	26	30	51	**300**

Results

Tree Canopy Evaluation

The tree canopy calculations revealed that 75% of the study area is either not covered with any tree or has less than 10% tree cover, 7.15% has 60-80% cover while 6.97% has a 40-60% tree cover. The remaining values are mostly distributed between 20-40% and 80-100% coverages (Table 2). For the study area, more than 18% of the study area was 40% or higher in tree cover.

It is important to note that the tree canopy and tree cover calculation depends mainly on the quality of the LiDAR point cloud data and the season of acquisition. In the case of the present study, the LiDAR data were collected during leaf-off season and therefore would not be the best dataset for such an application, but it was the only readily available data at the time of this study. Nevertheless, acceptable accuracies have been achieved.

Table 3 summary of percent tree cover in the study area

Class	Pixel Count	Percent From Total
10	355,107	**75.03**
20	5,236	1.11
40	24,539	5.18
60	33,005	6.97
80	33,821	**7.15**
100	21,582	4.56
Totals	473,290	100.00

It would be prudent to conduct a further study to compare those results with other similarly sized towns in a different location and then compare the percentage of tree canopy cover calculated from each, in order to have a bench mark for evaluation.

Land Cover and Surface Permeability

Non permeable surfaces, which include Roads/Paved, Other Paved Surfaces, and Structures (50.65%) were slightly more than the permeable surfaces including the Open Areas and Trees (49.35%) (Table 3). The roads and the other paved surfaces constituted the majority of all surfaces in the study area (37.82%), while open areas and trees occupied 49.35%.

Table 4 Final result of land cover classification and the corresponding summary areas/percentages

Class	Description	Pixel Count	Area (SqFt)	Percent From Total
1	Open Areas	191,481	19,148,100	35.27
2	Roads/Paved	168,195	16,819,500	30.98
3	Other Paved Surfaces	37,151	3,715,100	6.84
4	Structure	69,646	6,964,600	12.83
5	Trees	76,463	7,646,300	14.08
	Totals	542,936	54,293,600	100.00

General

The results of accuracy assessment was around what was expected since the LiDAR point cloud data, which most of the procedures relied upon, was acquired during early spring when most coniferous trees bared no leafs. LiDAR technology relies heavily on the reflected portions of a light ray and its returns, and since trees were uncovered, this might have contributed to the relatively low accuracy (72.41%) of tree cover classification. Structures identification accuracy was the lowest too, although most large structures were properly classified, some areas had missing or displaced buildings due to the time difference between the LiDAR data (2006) and the NAIP aerial photo acquisition (2010), especially around the Clarion University campus were obvious changes between the two dates can be easily discerned (Figure 10 and Figure 11). For conducting similar studies, it would be recommended to use synchronized LiDAR and multi-band aerial photograph in order to avoid similar discrepancies.

Figure 10 PAMAP aerial photo of 2004

Figure 11 NAIP aerial photo of 2010

Also, the accuracy assessment compared the resulting land cover with the NAIP aerial photo. It might be of value if the same accuracy assessment was repeated using the 2004 PAMAP image instead, since it is leaf off and was acquired at the same time as the LiDAR data, some features, including roads, hidden driveways, and even some structures might show up in a better fashion on those images.

For future reference, the same model would be applied to other study areas, with similar datasets as inputs in Pennsylvania. The results will be compared in order to evaluate Clarion's tree density as well as the land cover composition of permeable versus non-permeable surfaces in relation to other similar towns and neighborhoods.

The model that was created during this study (Figure 12) is designed to accept a LiDAR dataset, a NAIP aerial photograph (or any multi-band high resolution image) and a study area polygon feature class as inputs while it'll output a classified tree canopy raster dataset as well as a final land cover of the study area.

References

Chen, L., Chiang, T., and Teo, T. (?). *Fusion of LIDAR Data and High Resolution Images for Forest Canopy Modeling*. Center for Space and Remote Sensing Research, National Center University, Jhong-Li, Taiwan.

Syed, S., Dare, P. and Jones, S. (2005). *Automatic Classification of Land Cover Features with High Resolution imagery and LIDAR data: An Object-Oriented Approach*. Proceedings of SSC2005 Soatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute. September 2005. Melbourne: Spatial Sciences Institute.

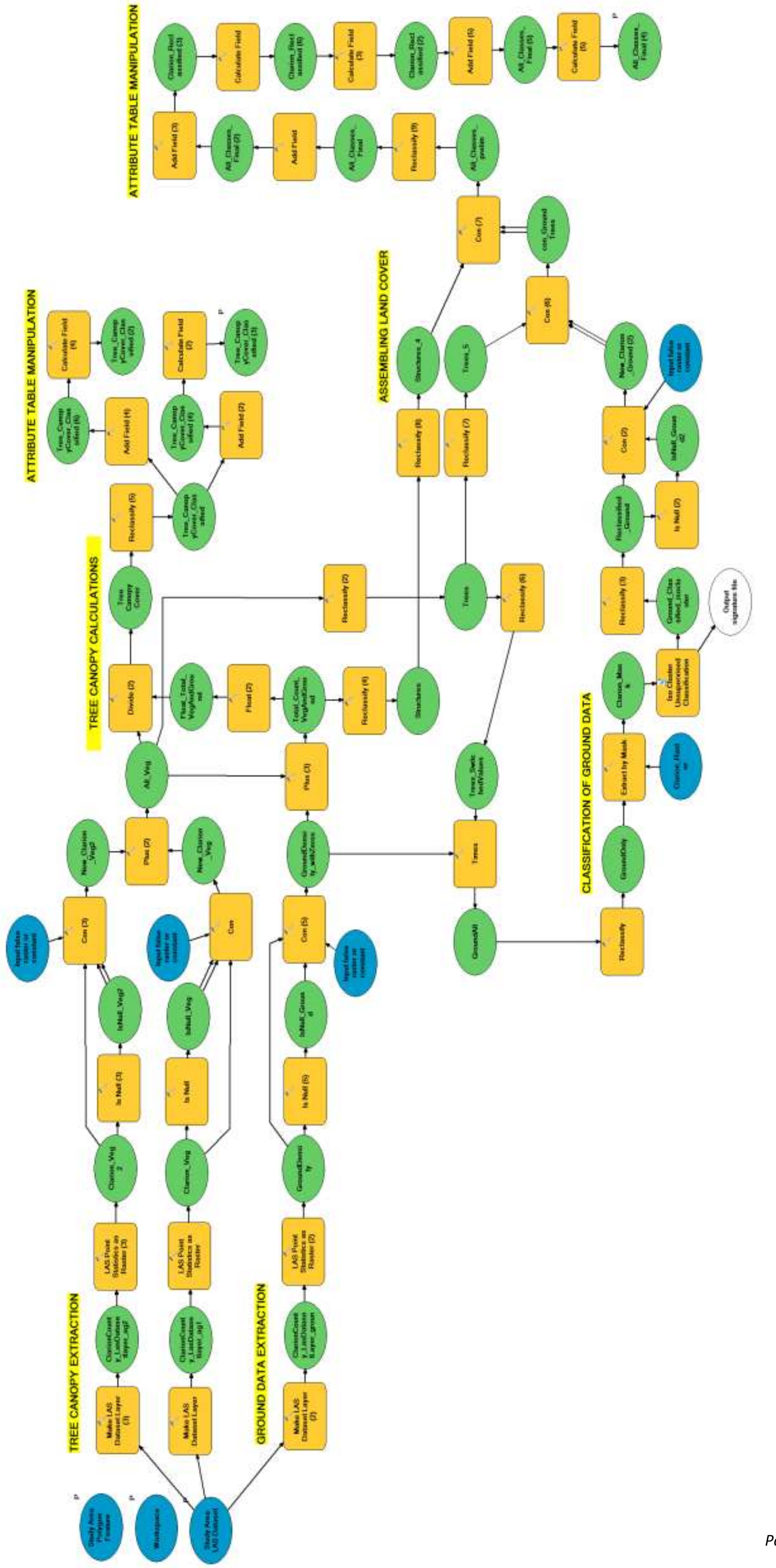

Figure 12 Overall model for tree density calculation as well as land cover classification