Korean Journal of Remote Sensing, Vol.28, No.6, 2012, pp.661~670
http://dx.doi.org/10.7780/kjrs.2012.28.6.6

Mapping Vegetation Volume in Urban Environments by Fusing LiDAR and Multispectral Data

Jinha Jung*† and Bryan Pijanowski**

*Institute for Environmental Science and Policy, University of Illinois at Chicago
**Forestry and Natural Resources Department, Purdue University

Abstract : Urban forests provide great ecosystem services to population in metropolitan areas even though they occupy little green space in a huge gray landscape. Unfortunately, urbanization inherently results in threatening the green infrastructure, and the recent urbanization trends drew great attention of scientists and policy makers on how to preserve or restore green infrastructure in metropolitan area. For this reason, mapping the spatial distribution of the green infrastructure is important in urban environments since the resulting map helps us identify hot green spots and set up long term plan on how to preserve or restore green infrastructure in urban environments. As a preliminary step for mapping green infrastructure utilizing multi-source remote sensing data in urban environments, the objective of this study is to map vegetation volume by fusing LiDAR and multispectral data in urban environments. Multispectral imageries are used to identify the two dimensional distribution of green infrastructure, while LiDAR data are utilized to characterize the vertical structure of the identified green structure. Vegetation volume was calculated over the metropolitan Chicago city area, and the vegetation volume was summarized over 16 NLCD classes. The experimental results indicated that vegetation volume varies greatly even in the same land cover class, and traditional land cover map based above ground biomass estimation approach may introduce bias in the estimation results.

Key Words : Vegetation volume mapping, Data fusion, LIDAR, Urban forest, Green Infrastructure, Carbon sequestration, Above-ground biomass

1. Introduction

Urban forests provide great ecosystem services to population in metropolitan areas even though they occupy little green space in a huge gray landscape. Unfortunately, urbanization inherently results in threatening the green infrastructure, and the recent urbanization trends drew great attention of scientists and policy makers on how to preserve or restore green infrastructure in metropolitan area. In addition, measuring above-ground biomass (AGB) in large scales recently gained great interests among scientists

Received October 19, 2012; Revised November 8, 2012, Revised November 30, 2012; Accepted December 2, 2012.
† Corresponding Author: Jinha Jung (jinha@purdue.edu)

due to the international efforts to reduce greenhouse emissions associated with deforestation and forest degradation. Initiatives such as the United Nations' REDD (Reducing Emissions from Deforestation and Degradation) must depend on accurate map of forest carbon storage in large scales (Houghton *et al*., 2010). For these reasons, mapping the spatial distribution of the green infrastructure and their ecosystem services is important in urban environments since the resulting map not only helps us identify hot green spots and set up long term plan on how to preserve or restore green infrastructure in urban environments, but also contribute to understand large scale carbon cycle studies as the urban areas are expanding due to ongoing urbanization trends.

Earlier efforts to map spatial distribution of green infrastructure have been limited to field work based approach especially in urban environments (McPherson *et al*., 1997). More recently, passive optical remote sensing data have been widely utilized to map the spatial extent of green infrastructure in urban environments (Myeong *et al*., 2001; Walker and Briggs, 2007; Xiao *et al*., 2010). A land cover map is generated from the passive remote sensing data, then factors for each land cover is adopted from field measurements to calculate overall vegetation amount in a landscape scale (ICF International, 2012). Although the traditional land cover map based estimation approach may provide overall amount of vegetation in large scales and may be appropriate for simple analysis such as balancing carbon cycle in large scales, the traditional approach fails to provide detailed information about spatial distribution of the green infrastructure. Recently, LiDAR (Light Detection And Ranging) technology has gained significant attention of scientists due to its unique ability to capture vertical structural information objects. LiDAR is an active remote sensing technique that enables three dimensional measurements of objects on the Earth's surface based on range measurements. The range measurements are accomplished by measuring time difference between outgoing laser pulse and return laser pulse. The calculated range measurements are combined with location of the sensor computed from the GPS (Global Positioning System) and attitude of the sensor estimated from the IMU (Inertial Measurement Unit) sensor, so that three dimensional coordinates of the objects are calculated through a vector addition operation.

Although LiDAR data have been widely utilized to characterize vegetation structure even in heavily forested area (Harding and Carabajal, 2005; Lefsky *et al*., 2007; Sun *et al*., 2008; Miller *et al*., 2010; Muss *et al*., 2010; Hwang and Lee, 2011; Jung and Crawford, 2012; Park *et al*., 2012) due to its unique ability to capture vertical structural information, little research has focused on the utilization of LiDAR data for mapping urban green infrastructure in urban environments. As a preliminary step for mapping green infrastructure utilizing multi-source remote sensing data in urban environments, the objective of this study is to map vegetation volume by fusing LiDAR and multispectral data in urban environments. Multispectral imageries are used to identify the two dimensional distribution of green infrastructure, while LiDAR data are utilized to characterize the vertical structure of the identified green structure.

2. Methods

1) Study site

This study is conducted in the Cook County that is located in the north eastern part of Illinois, USA (Fig. 1). The Cook County encompasses the metropolitan Chicago city area which has a population of 2,695,598 and the Chicago city is the third largest city

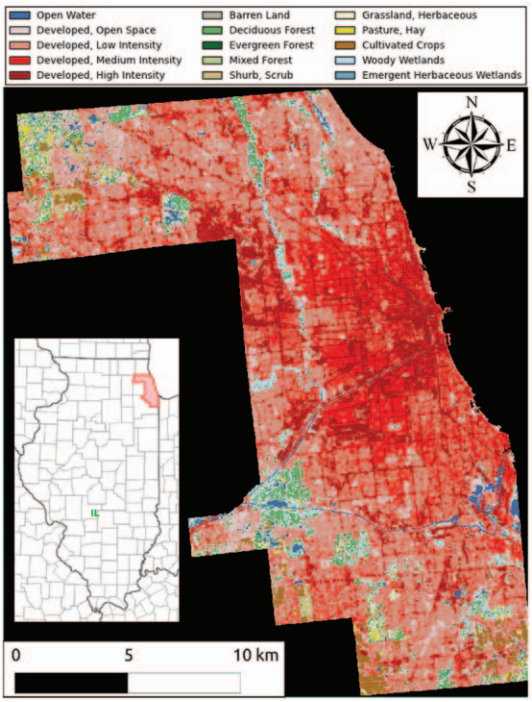

Fig. 1. NLCD 2006 data over the study area.

Table 1. Land cover area percent distribution over Cook, IL calculated from the NLCD (National Land Cover Dataset) 2006 data

Land cover type	Area (%)
Open Water (OW)	1.75
Developed, Open Space (DOS)	9.34
Developed, Low Intensity (DLI)	36.00
Developed, Medium Intensity (DMI)	26.16
Developed, High Intensity (DHI)	13.65
Barren Land (BL)	0.19
Deciduous Forest (DF)	3.91
Evergreen Forest (EF)	0.05
Mixed Forest (MF)	0.75
Shrub, Scrub (SS)	0.87
Grassland, Herbaceous (GH)	1.15
Pasture, Hay (PH)	0.51
Cultivated Crops (CC)	1.96
Woody Wetlands (WW)	3.50
Emergent Herbaceous Wetlands (EHW)	0.21

in the United States (U.S. Bureau Census, 2011). National Land Cover Database 2006 (NLCD2006) is a national scale land cover map of whole United States that has 16-class land cover classification scheme. The NLCD2006 was mainly generated by applying unsupervised classification algorithms to Landsat Enhanced Mapper+ (ETM+), and it has 30 m spatial resolution (Fry *et al*., 2011). Fig. 1 shows the NLCD2006 data over the study area, and area distribution of the 16 classes was calculated from the NLCD2006 data and summarized in the Table 1. Table 1 indicates that the study area consists of more than 85 percent of developed area, approximately 4.7 percent of forest, 3.7 percent wetlands, and 2.5 percent of grasslands.

2) Multispectral imageries and LiDAR data

Aerial digital ortho imageries were taken in 2010 during the agricultural growing seasons over the Cook County, Illinois as a part of NAIP (National Agriculture Imagery Program). The ortho imageries were acquired in four bands (Red, Green, Blue, and Near Infrared). The ortho imageries were projected onto UTM (Universal Transverse Mercator) 16N coordinate system with 1 m spatial resolution when provided. The existence of the fourth (Near Infrared) band in the ortho imageries enables calculation of NDVI (Normalized Difference Vegetation Index), and the NDVI can be potentially utilized as an indicator to differentiate vegetation from urban objects such as buildings and roads (Carlson and Ripley, 1997). The NDVI values ranges between -1 and 1, and higher NDVI value indicates higher chance of vegetation existence at the pixel. Using red and near infrared band from the ortho imageries, NDVI values were calculated using Equation 1, where a_{nir} and a_{red} represent surface reflectance in the near infrared ($\lambda \sim 0.8\ \mu$m) and red ($\lambda \sim 0.6\ \mu$m) regions of the spectrum, respectively.

$$\mathrm{NDVI} = \frac{a_{nir} - a_{red}}{a_{nir} + a_{red}} \quad (1)$$

A topographic LIDAR mapping system with the Leica Geosystem ALS50 (Airborne Laser Scanner

50) sensors were flown over the study area to collect LIDAR data during two different time periods: November 21-29, 2012 and April 11-14 2009, and this campaign consists of a total of 167 flight lines. The ALS50 is a discrete return LiDAR system, and can operate up to 150 KHz pulse repetition frequency (PRF) rate and 75 degree field of view (FOV). Operational altitude of the system ranges from 400 m to 4,700 m (Cook County Board of Commissioners, 2010). The system can record up to 5 returns from a return pulse, but it was configured to record only three returns for this data acquisition. An average LiDAR point density for the acquisition is approximately 2.6 points/m^2. The LiDAR point cloud data were projected onto Illinois East State Plane Coordinate System (SPCS) when provided. Spatial grid structure with 5 foot resolution was adopted for LiDAR data processing since the 5 foot spatial grid structure ensures approximately 5.85 points for every pixel (2.6 points/$m^2 \times 25$ ft^2 / pixel=5.85 points/pixel). LiDAR point cloud data were classified into ground and non-ground classes using LAStools (Isenburg, http://lastools.org), and the ground points were used to generate 5 foot digital terrain model (DTM) using the natural neighbor interpolation algorithm. Digital surface model (DSM) was calculated by calculating maximum elevation within each grid cell. Digital height model (DHM) was generated by subtracting the DTM from the DSM. The DHM generation process eliminates topographic effects from the DSM and converts the elevation unit from elevation above sea level to elevation above ground. The DHM only contains structural information and it is challenging to discriminate vegetation from urban objects from the DHM layer.

3) Vegetation volume calculation

Fig. 2 illustrates flow charts for calculating vegetation volume by fusing LiDAR and multispectral data. First, the DHM layer was generated from the LiDAR data, and the NDVI layer was calculated from the ortho imageries. Second, the DHM and the NDVI layers were fused at the pixel level. Due to different coordinate frames and spatial resolution used in the DHM (SPSC IL East) and NDVI (UTM 16N) layers, the NDVI layer was projected onto the spatial grid structure of the DHM layer, then the NDVI layer was resampled at 5 foot spatial resolution. Third, the NDVI layer was used to mask out urban objects such as buildings and roads. Threshold value of 0.2 was used as a break point in the masking process, i.e., *I* value was set to 1 when the NDVI value was greater than the threshold value, otherwise *I* value was set to 0 in Eq. 2. After masking out urban pixels using the NDVI layer, vegetation volume (V_{veg}) was computed by multiplying DHM value (*DHM*) with corresponding pixel area (*A*) as a last step (Eq. 2). Fig. 3 illustrates resulting vegetation volume layer over the study area generated by the methods described in this section by fusing the multispectral imageries and LiDAR data.

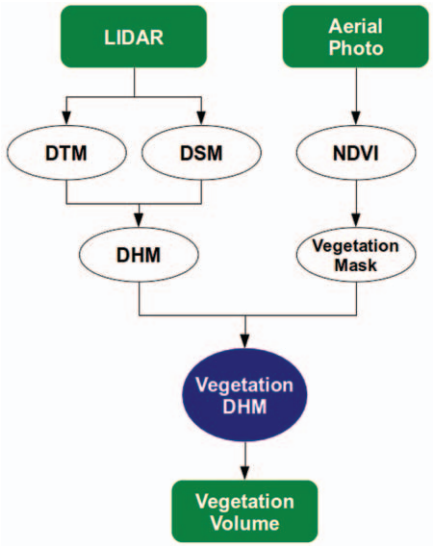

Fig. 2. Flow chart for vegetation volume mapping using LiDAR and multispectral data.

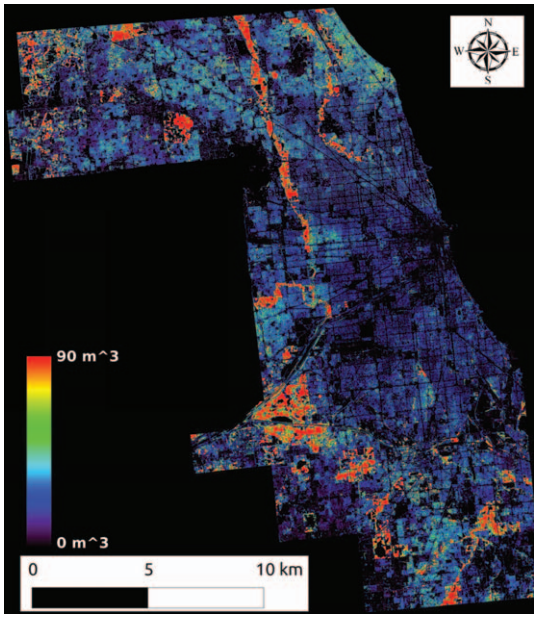

Fig. 3. Vegetation volume map generated from LiDAR and multispectral data.

$$V_{veg} = I \times DHM \times A \quad (2)$$

4) Summarizing vegetation volume for individual land cover class

The vegetation volume layer was generated for the 5 foot spatial grid which uses the SPSC IL East projection. However, the NLCD 2006 data uses the Albers Conical Equal Area projection with NAD83 (North American Datum 1983) datum with 1 arc-second (approximately 30 m) spatial resolution. Due to the different projection used in the vegetation volume layer and the NLCD 2006 data, coordinates of the vegetation volume layer is converted into the Albers Conical Equal Area projection to co-register both dataset. Then, the 1 arc-second grid structure of the NLCD 2006 data was placed on top of the converted vegetation volume layer and vegetation volume of each land cover was calculated by summing up the vegetation volume layer pixels within each NLCD 2006 grid cell.

Table 2. Mean vegetation volume for individual land covers

Land cover type	Mean vegetation volume (m^3)
Open Water (OW)	1355.07
Developed, Open Space (DOS)	4297.82
Developed, Low Intensity (DLI)	4286.08
Developed, Medium Intensity (DMI)	1616.47
Developed, High Intensity (DHI)	289.95
Barren Land (BL)	239.88
Deciduous Forest (DF)	12444.62
Evergreen Forest (EF)	13579.77
Mixed Forest (MF)	11811.37
Shrub, Scrub (SS)	4593.43
Grassland, Herbaceous (GH)	1698.15
Pasture, Hay (PH)	1617.56
Cultivated Crops (CC)	614.09
Woody Wetlands (WW)	14704.97
Emergent Herbaceous Wetlands (EHW)	1345.19

3. Results and Discussion

Vegetation volume for individual land covers was summarized using the procedure described in the Methods section. Table 2 shows an average vegetation volume for individual land covers, and this table indicates that land covers can be divided into four different groups based on the calculated average vegetation volume. First, a forested land cover group such as Deciduous Forest (DF), Evergreen Forest (EF), Mixed Forest (MF), and Woody Wetlands (WW) has the highest average vegetation volume (> 10,000 m^3/pixel). These four land covers are expected to contain tall canopies within their boundaries, which results in a group of land covers with the highest vegetation volume. Second, Shrub/Scrub (SS), Developed Open Space (DOS) and Developed Low Intensity (DLI) land covers form a group that has second highest vegetation volume (~ 4,000 m^3/pixel). This group of land covers contains smaller vegetation volume than the forested land covers since the SS land cover is expected to contain mostly shrubs that are much shorter than trees and much

smaller number of canopies than the forested land cover group. The experimental results that the DOS and DLI land covers belongs to this group and they contain significant amount of vegetation volume may be explained by the fact that 1) residential area is one of the major components of the DLI land cover and trees are usually planted around the residential area, and 2) the DOS land cover is usually co-located around the developed areas and used as ecosystem service area with vegetation where residents can relax and entertain themselves. Third, a group of land covers including Open Water (OW), Developed Medium Intensity (DMI), Grassland Herbaceous (GH), Pasture Hay (PH), and Emergent Herbaceous Wetlands (EHW) forms a cluster which has third highest vegetation volume (~ 1,000 m^3/pixel). Last, a group of land covers including Developed High Intensity (DHI), Barren Land (BL), and Cultivated Crops (CC) has the smallest vegetation volume (< 1,000 m^3/pixel).

Histograms of vegetation volume of the first two groups were generated from the vegetation volume layer. Fig. 4 is the histogram of vegetation volume for a group that has the highest vegetation volume. The figure indicates that quite large variance of vegetation volume exists in the first group. The EF and MF land covers have peak around mean vegetation volume value, while the DF and WW land covers show right skew distribution indicating that the use of representative factor for the forested land cover group

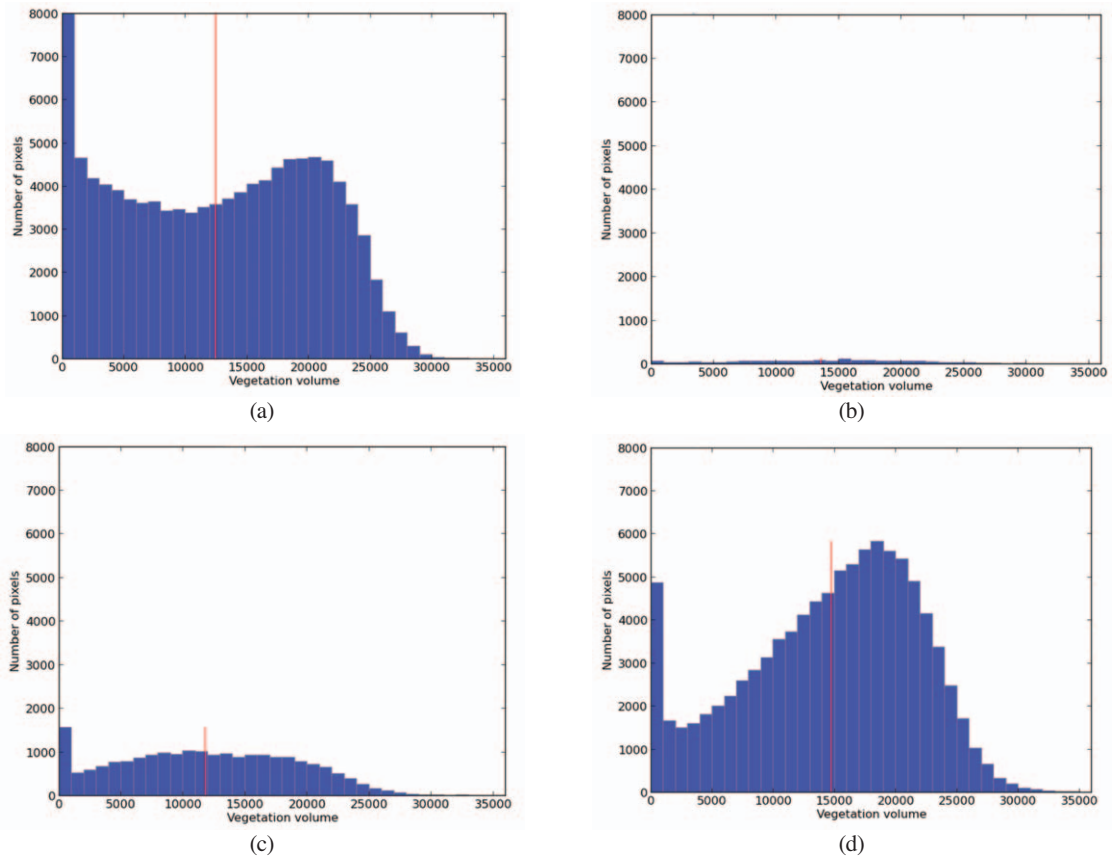

Fig. 4. Histogram of vegetation volume histogram for forest land covers: (a) Deciduous Forest, (b) Evergreen Forest, (c) Mixed Forest, and (d) Woody Wetlands.

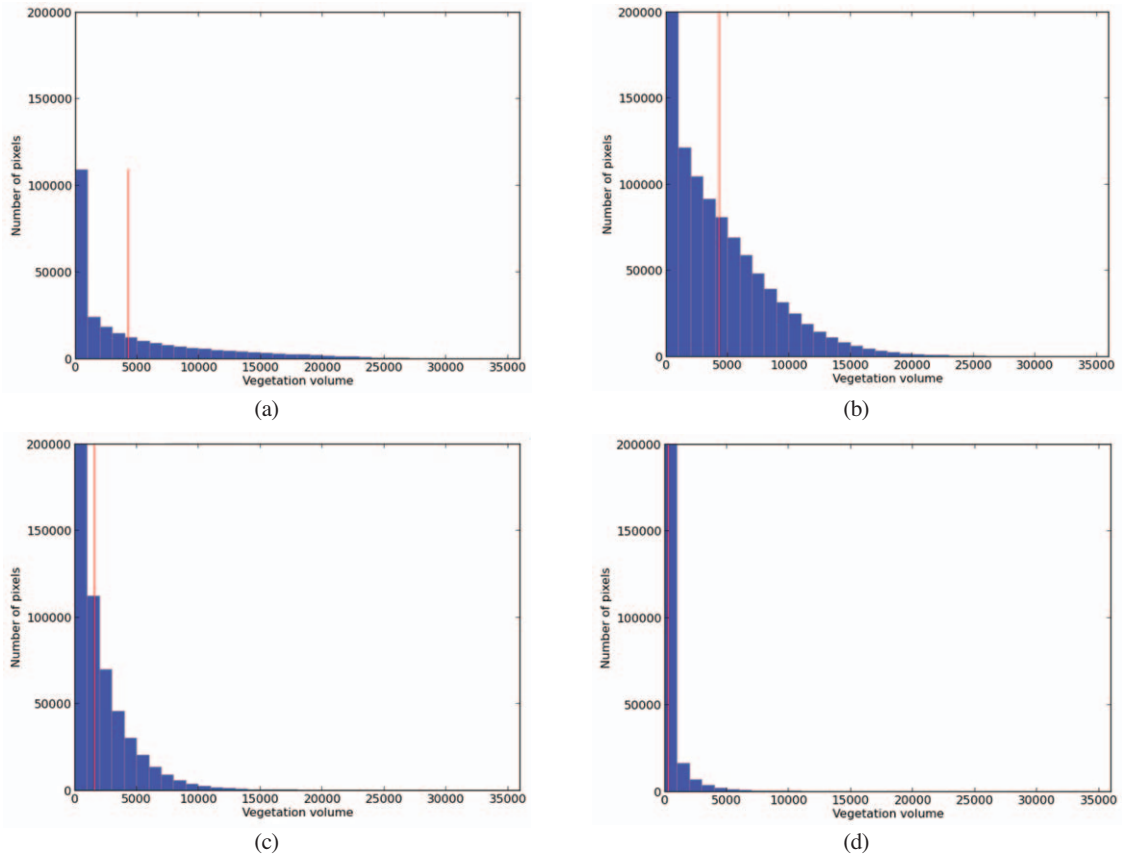

Fig. 5. Histogram of vegetation volume for developed land covers: (a) Developed, Open Space, (b) Developed, Low Intensity, (c) Developed, Medium Intensity, and (d) Developed, High Intensity.

- which is often adopted in the traditional approach - may cause great bias in the estimation results. Fig. 5 is the histogram of vegetation volume for a group that has the second highest vegetation volume. All the land covers in this group show similar histogram pattern. Most areas contain little vegetation, and the number of pixels with higher vegetation volume decreases exponentially. While the DOS and DLI land covers have similar average vegetation volume (Table 2) of approximately 4,200 m^3/pixel, the DOS land cover has more pixels with higher vegetation volume and the DLI land cover has more pixels with medium low vegetation volume. The DMI land cover has fewer pixels with medium low vegetation volume, and the DHI land cover has most pixels with little vegetation volume.

Total vegetation volume and vegetation volume percentage of each land cover was calculated from the vegetation volume layer. Fig. 6 shows the vegetation volume percentage of each land cover. The DLI land cover contains almost half (42.7%) of total vegetation volume, while area percentage of the DLI land cover is about a third (36.0%) of total area of the study area. Even though the DLI belongs to a group that has the second highest vegetation volume, the amount of vegetation volume contained in the DLI land cover is the greatest due to its dominant coverage within the study area. The WW, DF, DMI, and DOS land covers contains 14.3%, 13.5%, 11.7%, and 11.1% of total vegetation volume, respectively.

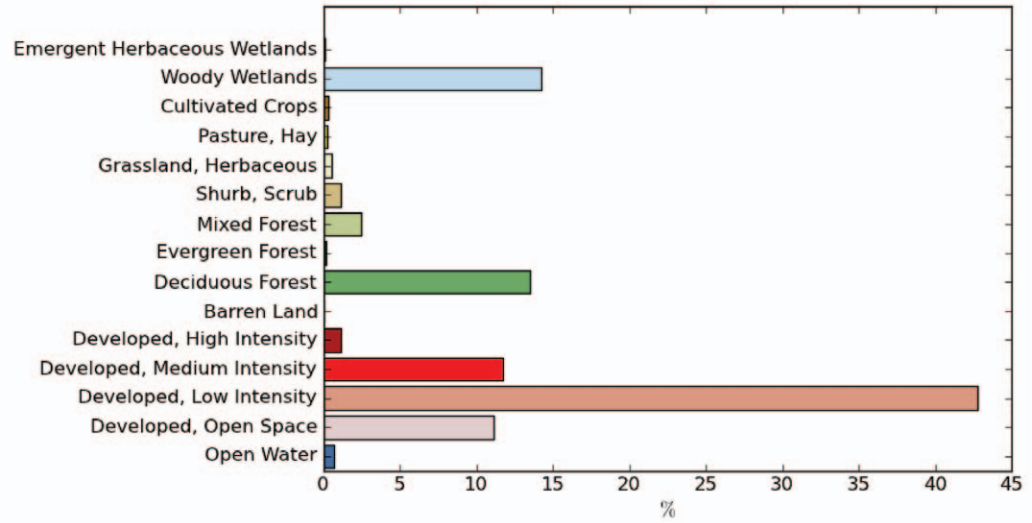

Fig. 6. Vegetation volume for individual land use covers.

The WW and DF land covers belongs to a group with the highest vegetation volume, but their area percentages are relatively small, 3.5% and 3.9% of total area, respectively. Although area percentages of the WW and DF land covers are less than 8%, they contain about 28% of total vegetation volume. The experimental results also indicate that quite amount of vegetation volume (11.1%) is contained in the DHI land cover class due to its larger coverage (13.7% of total area) even though the DHI land cover belongs to a group with the smallest vegetation volume. In the traditional approach to map urban forest using optical remote sensing data, developed land cover classes are usually not included in the analysis since little research has been done in urban environments and no factor is available for these land cover classes. However, the experimental results indicate that developed land cover classes (DOS, DLI, DMI, DHI) contain major portion (approximately 76.6%) of total vegetation volume in urban environments and the developed land cover classes should not be ignored in the carbon storage analysis.

4. Conclusion

Mapping green infrastructure in urban environments is critical since the resulting map not only helps us identify hot green spots and set up long term plan on how to preserve or restore green infrastructure in urban environments, but also contribute to understand large scale carbon cycle studies. Among various remote sensing data types, LiDAR and multispectral imageries were utilized to map vegetation volume in metropolitan area located within Cook County, IL. Multispectral imageries were used as mask to filter out urban objects such as buildings and roads, and LiDAR data were used to calculate volume of vegetated area. Vegetation volumes of the 16 classes were calculated by overlaying the NLCD2006 grid structure over the vegetation volume layer and summing up vegetation volume within the each grid cell. The experimental results indicated that Forested classes (DF, EF, MF, and WW) forms a group which has the highest average vegetation volume, and they also showed highest vegetation volume variation within them. The experimental results also showed that DOS and DLI

classes belongs to a group which has second highest average vegetation volume, but they contains the most vegetation volume among all classes due to their dominant coverage within the study area. The experimental results indicated that vegetation volume varies greatly even in the same land use cover, and the traditional land cover map based above ground biomass estimation approach may introduce bias in the estimation results since the traditional approach lacks the ability to capture this variation that exists especially in the developed land cover classes. Unfortunately, this study adopted the simplest vegetation volume estimation model with an assumption that the amount of vegetation is linearly correlated with the vegetation height, and this model does not take into account vertical structural differences that exist even in the same land cover class. Future work will include developing better estimation models that are based on field measurement data and taking advantages of vertical structural information extracted from the LiDAR data.

Acknowledgement

This research was supported by the National Science Foundation and the U.S. Forest Service under the UNLTRA-Ex program (grant number 0948484).

References

Carlson, T.N. and D.A. Ripley, 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index, *Remote Sensing of Environment*, 62(3): 241-252.

Cook County Board of Commissioners, 2010. Cook County 2008 LiDAR and topographic data services (Contract No. 08-41-342), digital terrain model "bare earth" for Cook County, Illinois. DTM, Version 1.0, Published May 2010.: Cook County Board of Commissioners, Chicago, Illinois.

Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham, 2011. Completion of the 2006 national land cover database for the conterminous United States, *Photogrammetric Engineering and Remote Sensing*, 77(9): 858-864.

Harding, D.J. and C.C. Carabajal, 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure, *Geophysical Research Letters*, 32, L21S10.

Houghton, R.A., N. Greenglass, A. Baccini, A. Cattaneo, S. Goetz, and J. Kellndorfer, 2010. The role of science in REDD, *Carbon Management*, 1: 253-259.

Hwang, S. and I. Lee, 2011. Current status of tree height estimation from airborne LiDAR data, *Korean Journal of Remote Sensing*, 27(3): 389-401.

ICF International, 2012. Chicago 2010 regional greenhouse gas emission inventory, (ICF 112831.0.001).

Jung, J. and M.M. Crawford, 2012. Extraction of features from LiDAR waveform data for characterizing forest structure, *IEEE Geoscience and Remote Sensing Letters*, 9(3): 492-497.

Lefsky, M.A., M. Keller, Y. Pang, P.B. De Camargo, and M.O. Hunter, 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms, *Journal of Applied Remote Sensing*, 1: 1-18.

McPherson, E.J., D. Nowak, G. Heisler, S.

Grimmond, C. Souch, R. Grant, and R. Rowntree, 1997. Quantifying urban forest structure, function, and value: the Chicago Urban Forest Climate Project, *Urban Ecosystems*, 1: 49-61.

Miller, M.M., M. Lefsky, and P. Yong, 2010. Optimization of Geoscience Laser Altimeter System waveform metrics to support vegetation measurements, *Remote Sensing of Environment*, 115(2): 298-305.

Muss, J.D., D.J. Mladenoff, and P.A. Townsend, 2010. A pseudo-waveform technique to assess forest structure using discrete LiDAR data, *Remote Sensing of Environment*, 115(3): 824-835.

Myeong, S., D.J. Nowak, P.F. Hopkins, and R.H. Brock, 2001. Urban cover mapping using digital, high-spatial resolution aerial imagery, *Urban Ecosystems*, 5(4): 243-256.

Park, T., W.K. Lee, J.Y. Lee, M. Hayashi, Y. Tang, D.A. Kwak, H. Kwak, M.I. Kim, G. Cui, and K. Nam, 2012. Maximum canopy height estimation using ICESat GLAS laser altimetry, *Korean Journal of Remote Sensing*, 28(3): 307-318.

Sun, G., K.J. Ranson, D.S. Kimes, J.B. Blair, and K. Kovacs, 2008. Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data, *Remote Sensing of Environment*, 112(1): 107-117.

Walker, J.S. and J.M. Briggs, 2007. An object-oriented approach to urban forest mapping in Phoenix, *Photogrammetric Engineering and Remote Sensing*, 73(5): 577-583.

Xiao, Q., L. Ustin, and E.G. McPherson, 2010. Using AVIRIS data and multiple-masking techniques to map urban forest tree species, *International Journal of Remote Sensing*, 25(24): 5637-5654.