

# ESTIMATION OF LAND-USE IN AN URBANIZED LANDSCAPE USING LIDAR INTENSITY DATA: A REGIONAL SCALE APPROACH

Kunwar K. Singh, John B. Vogler and Ross K. Meentemeyer

Center for Applied Geographic Information Sciences,  
University of North Carolina at Charlotte, Charlotte, NC 28223, USA – ksingh9@uncc.edu

**KEYWORDS:** LiDAR data, intensity, normalization, high resolution, urban landscape, heterogeneity

## ABSTRACT:

Spatially heterogeneous patterns of land use in urban environments have long posed a challenge to remote sensing. High spatial resolution passive sensors provide detailed data of urban regions at sub meter level but are frequently limited by shadows of the built environment. Moderate resolution data can provide synoptic perspectives of such landscapes but tend to obscure information of spectrally similar objects. Due to its height-above-ground component, which is unaffected by shadows, Light Detection and Ranging (LiDAR) data are increasingly being used as an alternative to passive sensors. However, LiDAR's intensity component is infrequently utilized in urban studies presumably because its range of digital number values is similar between urban impervious and tree canopy covers. Previous investigations have concentrated on mapping either tree canopy or buildings using local-scale normalization procedures but the use of normalized intensity to map multiple land-use types in a heterogeneous urban landscape at a regional scale has received little attention. Our approach uniquely utilizes normalized intensity data in combination with structural components derived from LiDAR masspoints using maximum likelihood estimation of land use classes. Preliminary results show that our approach accurately distinguishes impervious surfaces and tree canopy over broad metropolitan contexts, with an overall accuracy of 96.7% for the ML classification of integrated LiDAR. In summary, we found that normalized LiDAR intensity data can be integrated with LiDAR surface models improving our ability to map heterogeneous urban geographies.

## 1. INTRODUCTION

Accurate mapping of land-use patterns in urbanized landscapes is important for efficient urban planning, preserving the aesthetic value of the urban landscape, and for monitoring local climate variability. High and moderate resolution passive remote sensing datasets are common sources for mapping land use; map accuracies, however, have been hampered due to the effects of shadow in high-resolution data and the inability to distinguish urban land use types in moderate resolution data (Dare, 2005). Light Detection and Ranging (LiDAR), an airborne laser scanning, active sensor, has emerged as a standard tool for collecting very high-resolution topographic data for describing the Earth's surface (Wang and Glenn, 2009; Kaasalainen et al., 2005; Kaasalainen et al., 2009). The height-above-ground component of LiDAR data has been used extensively in previous investigations, for example, to create 3D surfaces of urban environments, estimate characteristics of forest stands (Yu et al., 2004; Popescu, Wynne, and Nelson 2002; Popescu and Wynne, 2004; Hudak et al., 2002; Anderson et al. 2008), and conduct flood mapping and modeling (Raber et al., 2007). However, the radiometric property of LiDAR data (intensity) is often overlooked or underutilized in mapping land use in urbanized landscapes (Yoon, Shin, and Lee, 2008). Previous studies utilized LiDAR intensity data either through integration with other remote sensing data (e.g., elevation, hyperspectral) (Wang and Glenn, 2009; Dalponte, Bruzzone, and Gianelle, 2008) or to estimate a particular characteristic of the urban landscape, such as forest cover (Yoon, Shin, and

Lee, 2008) or identifying residential properties in an urban area (Jutzi, 2009).

Intensity is a radiometric component of LiDAR data that is recorded by the sensor as the amount of energy backscattered from objects on the Earth's surface. Intensity measurements are affected by several factors including surface reflectance, atmospheric transmission, local incidence angle, and the distance between the sensor and Earth objects (Wagner et al., 2006; Wehr and Lohr, 1999; Mazzarini et al., 2007;). Based on these factors, theoretical formulas of laser backscattering are simplified into three variables: backscattering coefficients related to the reflectance of objects and incidence angles, atmospheric attenuation, and the range between the sensor and objects (Yoon, Shin, and Lee, 2008; Baltsavias, 1999). Previous experiments indicate the effect of incidence angle is negligible at small angles, which is essential for the sensor collecting backscattered energy, and that laser intensity is weaker at longer wavelengths than at shorter wavelengths (Kaasalainen et al., 2009). Additionally, atmospheric absorption has an insignificant effect on airborne near-infrared LiDAR intensity data due to the higher wavelength laser pulses (Wang and Glenn, 2009). Relative to other factors, the distance between the sensor and objects plays the most significant role, in determining the amount of energy backscattered from objects. Normalizing LiDAR intensity data with respect to sensor-to-object distance produces an 8-bit panchromatic raster data that when integrated with other LiDAR derivatives provide a clear distinction between impervious surfaces and forest cover.

This study demonstrates a novel approach to mapping land use patterns in the rapidly urbanizing region of Mecklenburg County, NC, by combining normalized LiDAR intensity data with canopy height model (CHM) and a surface model, generated by subtracting CHM from normalized digital surface (nDSM) and applying a maximum likelihood (ML) classifier. Significant improvements in urban land use map accuracies are realized, particularly with respect to impervious surfaces, forest cover, bare earth, and water categories.

## 2. METHOD

LiDAR point cloud data for a sub-region ( $90\text{km}^2$ ) of Mecklenburg County, NC, were acquired and processed using LiDAR Analyst image processing software to generate intensity, first return (FR) and last return (LR) raster data. Further intensity data were normalized by sensor-to-object distance. FR and LR raster data were used to generate CHM (FR - LR) (Figure 1), nDSM and then a surface model was generated by subtracting CHM from nDSM raster data (Figure 2). Intensity of LiDAR data was rasterized into an 8-bit panchromatic raster data (Figure 3.). These rasters data were combined into a 3-band composite image (Figure 4). A maximum likelihood (ML) classifier using a supervised statistical approach to pattern recognition was applied to the integrated dataset.

ML is based on a Bayesian probability function calculated from training site data collected for each land use class. A total of 47 training sites were selected across four land-use classes: 1) impervious surfaces (e.g., asphalt and concrete), 2) forest (includes coniferous, deciduous trees and managed forest), 3) bare earth (includes farmland, open space, median, golf course and transmission line, rock, soil, and other non-forest pervious surfaces), and 4) water bodies. These training sites were used to estimate the parameters of the ML classifier and assess the accuracy of the final classification. Integrating LiDAR intensity with CHM and nDSM produces a  $d$  dimensional feature space  $x$ , where  $d$  is the number of features in the feature vector  $x$ . Assuming there are  $C$  classes, the probability of a data sample  $x$  belonging to a particular class  $i$  can be computed as:

$$P(i|x) = p(x|i)P(i)/p(x) \quad (1)$$

where  $p(x|i) = \sum p(x|j)P(j, P(i))$  is the prior probability of class  $i$ .

With no prior information about  $P(i)$ , it is usually safe to assume that  $P(i)$ 's for all the classes are equal ( $1/C$ ). Therefore, in order to determine the posterior probability  $P(i|x)$ , the class conditional densities  $p(x|i)$ . Finally, the data sample  $x$  is assigned to the class  $i$  for which  $P(i|x)$  is maximized (Tso, 2001; Charaniya, 2004; Bartels, 2006).

Based on the 153 randomly distributed ground truth points, a confusion matrix was produced to assess the final classification accuracy. Overall accuracy is estimated by the ratio of the sum of its main diagonal and the total number of classified pixels. In addition, the Kappa coefficient, a measure of overall agreement of a matrix, was calculated based on the following equations:

$$\kappa = (p_o - p_e)/(1 - p_e) \quad (2)$$

$$p_o = \text{accuracy of observed agreement}, \frac{\sum X_{ii}}{N} \quad (3)$$

$$p_e = \text{estimate of chance agreement}, \frac{\sum X_{i+}X_{+i}}{N^2} \quad (4)$$

Where  $X_{ii}$  = represents observation in row  $i$  and column  $i$ ,  $N$  = the total number of observation,  $X_{i+}$  and  $X_{+i}$  = are the sum of row  $i$  and column  $i$ , respectively (Conglton, 1991).

## 3. RESULTS AND DISCUSSION

Figure 5 shows the final ML classification derived from the integrated LiDAR dataset. Impervious surfaces are clearly distinguished from forest canopy, which has historically been a challenge. The addition of forest canopy and nDSM with

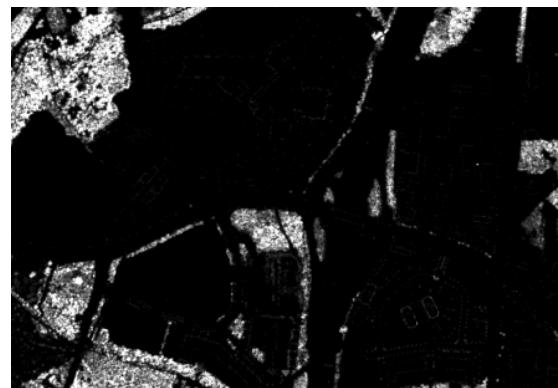


Figure 1. Canopy Height Model

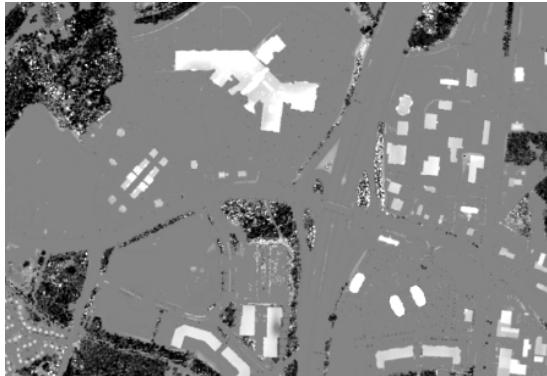


Figure 2. Surface Model



Figure 3. LiDAR Intensity

LiDAR intensity data allows for distinguishing these land use types with a very high degree of accuracy as shown in Table 1. Overall accuracy for the ML classification of integrated LiDAR was 96.7%. Producer's accuracies for bare earth, forest, impervious surfaces, and water were 100%, 92.98%, 96.42%, and 100% respectively. The Kappa coefficient for the final land use map was 0.95.

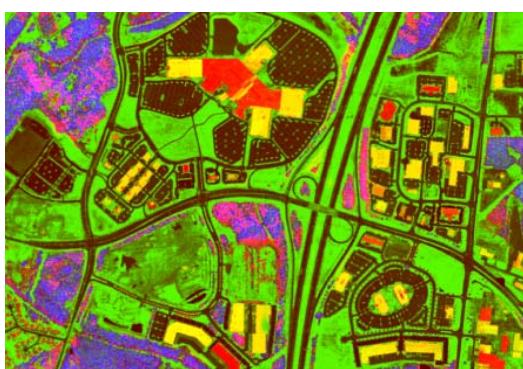


Figure 4. Composite image of CHM, Surface Model, and Intensity



Figure 5. Final land use classification map.

Land Use	Bare Earth	Forest	Impervious	Water	Sum
Bare Earth	<b>45</b>	2	0	0	47
Forest	0	<b>53</b>	0	0	53
Impervious	0	2	<b>27</b>	0	29
Water	0	0	1	<b>23</b>	24
Sum	45	57	28	23	<b>153</b>
Producer's accuracy	100	92.98	96.42	100	
Kappa Hat	Varia nce	Z	P	95% CI Upper	95% CI Lower
0.95	0.000 3959	47.981	< 0.00001	0.915757	0.9937 58
Overall Accuracy: (148 / 153) = 0.967320261					

Table 1. Confusion matrix and Kappa coefficient

#### 4. CONCLUSIONS

Using this methodology, forest cover can be classified into sub categories, such as managed forest and different aged stands. Sub classifying the bare earth category remains a challenge; however, using LiDAR data acquired during peak growing season for area crops and a proper normalization method can provide desirable results. We conclude that the integration of normalized LiDAR intensity data with LiDAR surface models significantly improves class discrimination among impervious surfaces and forest cover, and this methodology can be utilized to successfully map land use patterns at regional scales in urbanizing landscapes.

## REFERENCES

- Anderson, J. E., L. C. Plourde, M. E. Martin, B. H. Braswell, M. L. Smith, R. O. Dubayah, M. A. Hofton, and J. B. Blair., 2008. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment* 112 (4), pp.1856-1870.
- Baltsavias, E. P., 1999. Airborne laser scanning: basic relations and formulas. *Isprs Journal of Photogrammetry and Remote Sensing* 54 (2-3), pp. 199-214.
- Bartels, M., and H. Wei, 2006. Rule-based Improvement of Maximum Likelihood Classified LiDAR Data Fused with Co-Registered Bands: The Remote Sensing and Photogrammetry Society, United Kingdom, pp 1-9.
- Charaniya, A. P., R. Manduchi, and S. K. Lodha, 2004, "Supervised Parametric Classification of Aerial LiDAR Data" in Real Time 3D Sensor and Their Use: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop, Washington, pp. 30-38.
- Conglton, R., 19991. A review of assessing the accuracy of classifications of remote sensing data. *Remote Sensing of Environment* 37, pp. 35-46.
- Dalponte, M., L. Bruzzone, and D. Gianelle., 2008. Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. *Ieee Transactions on Geoscience and Remote Sensing* 46 (5), pp. 1416-1427.
- Dare, P. M., 2005. Shadow analysis in high-resolution satellite imagery of urban areas. *Photogrammetric Engineering and Remote Sensing* 71 (2), pp. 169-177.
- Hudak, A. T., M. A. Lefsky, W. B. Cohen, and M. Berterretche., 2002. Integration of lidar and Landsat ETM plus data for estimating and mapping forest canopy height. *Remote Sensing of Environment* 82 (2-3), pp. 397-416.
- Jutzi, B., 2009. Normalization of LIDAR intensity data based on Range and surface incidence angle. Laserscanning09, Paris, France, pp. 1-6.
- Kaasalainen, S., E. Ahokas, J. Hyypa, and J. Suomalainen., 2005. Study of surface brightness from backscattered laser intensity: Calibration of laser data. *Ieee Geoscience and Remote Sensing Letters* 2 (3), pp. 255-259.
- Kaasalainen, S., H. Hyypa, A. Kukko, P. Litkey, E. Ahokas, J. Hyypa, H. Lehner, A. Jaakkola, J. Suomalainen, A. Akujarvi, M. Kaasalainen, and U. Pyysalo., 2009. Radiometric Calibration of LIDAR Intensity With Commercially Available Reference Targets. *Ieee Transactions on Geoscience and Remote Sensing* 47 (2), pp. 588-598.
- Mazzarini, F., M. T. Pareschi, M. Favalli, I. Isola, S. Tarquini, and E. Boschi., 2007. Lava flow identification and aging by means of lidar intensity: Mount Etna case. *Journal of Geophysical Research-Solid Earth* 112 (B2), pp.1-43.
- Popescu, S. C., and R. H. Wynne., 2004. Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering and Remote Sensing* 70 (5), pp. 589-604.
- Popescu, S. C., R. H. Wynne, and R. F. Nelson., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture* 37 (1-3), pp. 71-95.
- Raber, G. T., J. R. Jensen, M. E. Hodgson, J. A. Tullis, B. A. Davis, and J. Berglund., 2007. Impact of lidar nominal post-spacing on DEM accuracy and flood zone delineation. *Photogrammetric Engineering and Remote Sensing* 73 (7), pp.793-804.
- Tso, B. and P. M. Mather, 2001. *Classification Methods for Remotely Sensed Data*. Taylor & Francis, pp.41-61.
- Wagner, W., A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *Isprs Journal of Photogrammetry and Remote Sensing* 60 (2), pp. 100-112.
- Wang, C., and N. F. Glenn., 2009. Integrating LiDAR Intensity and Elevation Data for Terrain Characterization in a Forested Area. *Ieee Geoscience and Remote Sensing Letters* 6 (3), pp. 463-466.
- Wehr, A., and U. Lohr., 1999. Airborne laser scanning - an introduction and overview. *Isprs Journal of Photogrammetry and Remote Sensing* 54 (2-3), pp. 68-82.
- Yoon, J. S., J. I. Shin, and K. S. Lee. 2008. Land Cover Characteristics of Airborne LiDAR Intensity Data: A Case Study. *Ieee Geoscience and Remote Sensing Letters* 5 (4), pp. 801-805.
- Yu, X. W., J. Hyypa, H. Kaartinen, and M. Maltamo., 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment* 90 (4), pp. 451-462.

## ACKNOWLEDGEMENTS

Financial and logistic support of this research was provided by the Renaissance Computing Institute engagement site at University of North Carolina at Charlotte, North Carolina. The authors would like to express sincere thanks to Amy Rockwell, Sr. Business Analyst, Storm water Services, and David B. Weekly, Division Manager, Engineering and Property Management, Land Development Services, City of Charlotte for providing high resolution 2007 LiDAR point cloud data.