International Journal of Business and Applied Social Science (IJBASS)

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

 www.cpernet.org

LIDAR-derived Biomass and Height Inventory of *Avicennia marina* in Eastern Mangrove Lagoon National Park, Abu Dhabi

Tareefa S. Alsumaiti
Geography and Urban Planning Department
College of Humanities and Social Sciences
United Arab Emirates University
PO Box 15551 Alain, UAE
E-mail: tareefa@uaeu.ac.ae
UAE

Jason A. Tullis
Department of Geosciences
Center for Advanced Spatial Technologies
Fulbright College of Arts and Sciences
University of Arkansas
Fayetteville, AR 72701, USA

Bowei Xue
Department of Geosciences
Center for Advanced Spatial Technologies
Fulbright College of Arts and Sciences
University of Arkansas
Fayetteville, AR 72701, USA

Abstract

Aboveground biomass (AGB) of Avicennia marina forest was mapped using airborne LIDAR-derived metrics in Eastern Mangrove Lagoon National Park,Abu Dhabi. From small footprint near-infrared LIDAR, multiple percentile heights were calculated using a neighbourhood algorithm. Multiresolution image segmentation was then employed to transform the 2D LIDAR-derived image into structurally homogeneous units. Given that neighbourhood size affects the calculation of the height metrics, neighbourhood sizes with 3 and 5 m radii were tested to evaluate workflow performance. AGB measurements from twenty field sample plots, each 154 m², were incorporated into a machine learning regression tree to map per-segment biomass throughout the study area. Based on in situ reference, the larger 5 m neighbourhood resulted in higher accuracy (RMSE= 10.17, R= 0.87, R²= 0.76) and was thus selected for biomass estimation. Final segmentsizesranged from 42 to 20,000 m² with an average of 2,445± 118 m², whereas biomass density per segment ranged from 0.23 to 13.18 (kg/m²) with an average of 5.16 ± 0.14 (kg/m²) and a total of 14,850 kg. Additionally, about 49% of the study area had a low biomass density (≤ 4.15 kg/m²), 23% had a moderate biomass density (4.16 to 8.01 kg/m²), and 28% had a high biomass density (8.02 to 13.17 kg/m²). Based on aper-pixel canopy height model generated from the same LIDAR mass points, minimum, average, and maximum height was0.1, 3.0, and 7.9 m, respectively. While this study represents the first LIDAR-derived AGB and height inventory of Avicennia marina in the UAE, it also combines in situ reference, segmentation, and multiple percentile height statistics in an innovative machine learning approach that is replicable for other mangrove inventories. The study demonstrates encouraging results in biomass mapping of mangroves in the UAE.

Keyword: *Avicennia marina*, aboveground biomass, LIDAR, percentile heights, segments.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Introduction

Mangrove forests are national treasures of the United Arab Emirates (UAE) and other arid countries with limited forested areas. Mangroves form a crucial part of the coastal ecosystem and provide numerous benefits to society, economy, and the environment (Alsumaiti, 2017). Detailed knowledge of the various mangroves' physical and biological distinguishing properties such as height, trunk diameter, crown spread, and biomass is important for proper coordination, preservation, conservation, and change monitoring (Wannasiri et al. 2013; White et al. 2013). In fact, one of the most vital measures of mangrove forest ecosystem structure and function is aboveground biomass (AGB). Mangrove AGB is important for evaluation of carbon sequestration and forest response to climate change and anthropogenic disturbances (Bombelli et al. 2009; Houghton, 2009). According to Jackowski et al. (2013), mangrove store up to four times more carbon than any other tropical forests.

Estimation of AGB from forest inventory plots generally includes both the measurement of diameter at breast height (DBH) and ideally tree height. AGB can be estimated from these two measurements using algometric equations (White et al. 2013). However, it is very unlikely that enough plots can be inventoried in this way to characterize a large forested area due to restrictions in resources, time and access, and it is impractical to monitor the forests using *in situ* data alone (Rasolofoharinoro et al.1998; Laba et al. 1997; Ramsey and Jensen, 1996; Aschbacher et al. 1995).

Several remote sensing technologies can be used to estimate AGB of a large forested area (Song, 2013; Mitchard et al. 2009). First, there is the use of optical data and subsequent estimates of spatially-averaged biomass values. Optical remote sensing typically makes use of visible and near-infrared reflectance from the surface of the earth to produce images. This lays a foundation for contemporary global-scale vegetation monitoring through numerous sensor systems such as Lands at, ASTER, IKONOS, MODIS, World View and others (Jensen, 2005). Such tools find wide application in research aimed at linking forest biomass measurements from *in situ* data to those obtained from aerial or satellite images. Major obstacles to this approach include persistent cloud cover and the presence of ubiquitous sun shadows. Furthermore, most satellite images obtained through optical sensors do not provide important vegetation characteristics such as canopy height. While this information can be derived via stereo analysis of overlapping imagery, the stereo imagery may be costly, have a limited spatial coverage, and demand a huge allocation of time to analyze (Geotz et al. 2009; Lucas et al. 2000).

Second, researchers may determine values of characteristics such as volume or height using laser-based airborne light detection and ranging or LIDAR which captures large numbers of unevenly arranged x, y and z positions within the volume of the forest structure (Fatoyinbo& Armstrong, 2010). The sensor in a LIDAR system continuously records the amplitude of the pulse registered as initiated by reflectance from laser targets of the forest canopy. Through this system, the vertical structure of vegetation can be estimated in remarkable detail, thus providing a clear advantage compared with optical imagery. Small-footprint airborne LIDAR systems are capable of detecting objects with a horizontal accuracy of less than one meter and a vertical accuracy of a few centimetres (Aguilar & Mills, 2008; Simardet al. 2006). The structural form of mangrove forests (in terms of height and density) may be extracted from LIDAR and is very applicable in the derivation of forest AGB (Wannasiri et al. 2013). To use LIDAR data for forest biomass estimation, machine learning approach is used to understand the relationship between AGB calculated from *in situ* data (based on allometric equations) and LIDAR height metrics (White et al. 2013).

Zhao et al. (2012) note that several researchers have efficiently utilized LIDAR data to estimate forest biomass based on relationships between LIDAR canopy heights metrics, such as mean canopy height and canopy height percentiles. However, recent studies attempt to predict AGB using image segmentation of LIDAR canopy height metrics, which allows the estimate of AGB to be conducted on structurally homogeneous units of forested

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

areas.The image segmentation process reduces the variability of subsequent AGB estimations. For example, Riggins et al. (2009) reports that the coefficient of determination values (R^2) of LIDAR-derived AGB estimates analyzed on a per-segment scale is higher than corresponding values from a plot level study. Biomass estimation of various forest environments, especially complex, heterogeneous ones, should benefit by machine learning analysis of multiple LIDAR derived canopy height metrics and image segmentation approaches.

Various biomass models have been published for different types of forests worldwide including mangroves (Poudel &Temesgen, 2016; Lu et al. 2012; Zhao et al. 2009; Simard et al. 2006). However, none have been created to estimate biomass of *Avicennia marina* forests in Abu Dhabi. The main objective of this study is to predict and characterize AGB of mangrove forest in the study area through innovative processing of small footprint LIDAR. To this end, a combination of LIDAR-derived height percentiles statistics, structurally homogenous forest segments, and *in situ* reference data were organized in a machine learning environment. Through this approach, per-segment estimates of both a) total AGB per unit area, and b) biomass density classes were extracted for the study area.

Materials and Methods

1. Study Area

The study area is found in Eastern Mangrove Lagoon National Park, which is the first mangrove protected area to be designated in the United Arab Emirates (UAE). The main vegetation cover in the study area includes Avicennia marina, also known as grey mangrove, and salt marshes dominated by Arthrocnemummacrostachyum. The study area is characterized by hot arid climate conditions, nearly flat topography, and several interconnected tidal creeks. The mangrove stands in the forest are usually inundated by tides twice daily. The study area is contained within a 1.7 km²rectangle with an upper left corner at 24°27'20"N and 54°25'18"E, and a lower right corner at 24°26'52"N and 54°26'24"E (Figure 1).

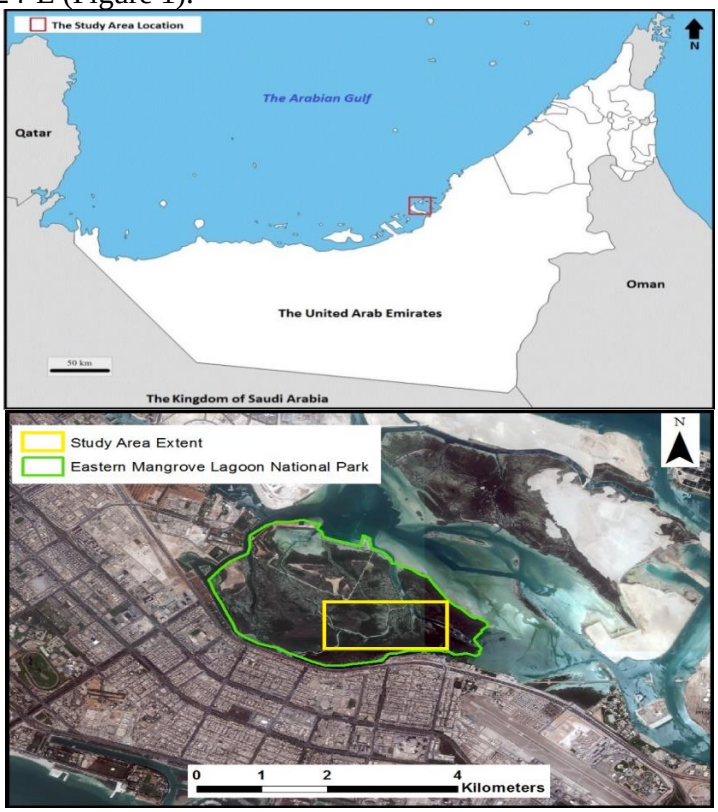

Figure1.Location of Abu Dhabi Island (top), and location of the study area (bottom; yellow box) within Eastern Mangrove Lagoon National Park (green polygon) in Abu Dhabi, UAE.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

2. Plot Design and In SituSurvey

An in situ survey was conducted throughout the study area from January to April 2014 using fixed area circular plots. Circular plots are easier to establish in situ compared to square plots because only the centers of circular plots are registered unlike those of square plots that have four corners. Furthermore, circular plots have approximately 11% less perimeter compared to square plots of an equal area, which minimizes the negative impact of edge effects such as the error in metric calculation (White et al. 2013; Wulder et al. 2012).

Several studies report that circular ground sampling plot sizes range from 50 m² (8 m diameter) to 2500 m² (56 m diameter)(Thomas et al. 2008; Naesset&Okland, 2002). However, these studies do not clarify the choice of sampling plot size. Even though there is no universal optimum plot size to study forest attributes (Frazer et al. 2011; Gobakken&Næsset, 2009), White et al. (2013) recommend that a ground sampling plot should have an area of at least 200 m² (16 m diameter). They recommend such plot size to minimize errors in modelled outcomes associated with ground data that does not capture a full range of the forest structural variability as captured by LIDAR data. Furthermore, this plot size increases both the efficiency of sampling and the accuracy of target and explanatory variables. In the present study, the ground sampling plots covered an area of 154 m2 (14 m diameter). Even though they were slightly smaller in size than those recommended by White et al. (2013), the high-density trees in the study area meant that a large number of trees were sampled.

Based on the Natural Resource Information System (NRIS) of the United States Department of Agriculture (USDA), it is recommended to have at least one sampling plot per every 0.04 km². Since the Abu Dhabi Island study area covers 1.7 km², ideally 42 plots should be established. In the present study, only 20 plots were established due to a number of environmental factors: 1) in situ access in mangrove swamps was extremely difficult due to the presence of mud, large root-like mangrove structures on the marshy ground, and the rise and fall of tidal waters; 2) foot travel through the very dense forest to collect tree attributes was impractical; and 3) collection of tree attributes was limited to the winter season (January-April) due to dangerously high temperatures during the summer. Fortunately, the 20 sampling plots covered a clear range of forest structural variability present in the study area. In order to capture variability within the study area, the locations of the sampling plots were randomly distributed within very accessible forested areas close to the main water creek (Figure2).

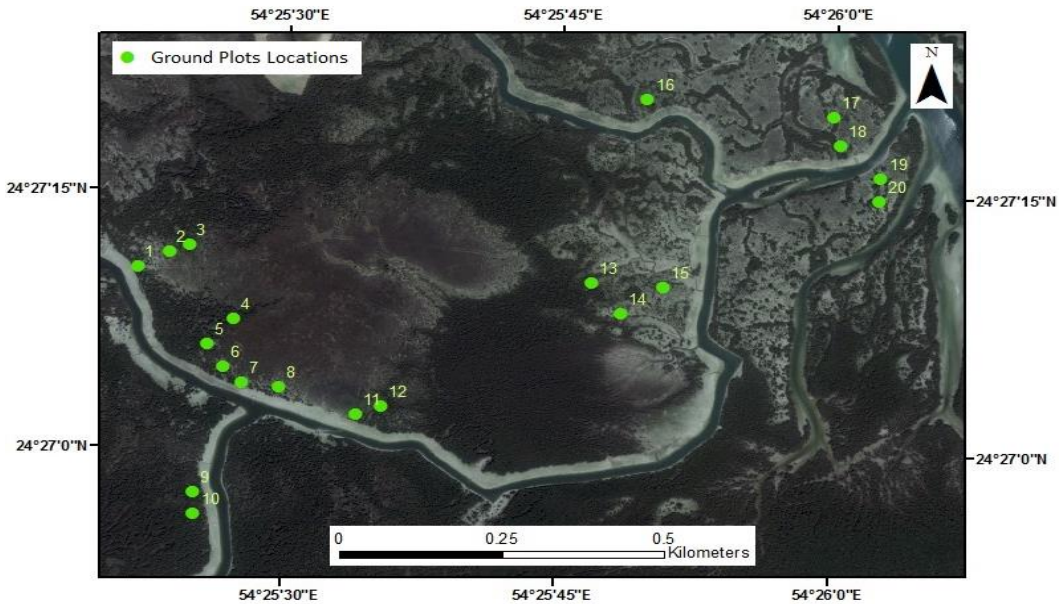

Figure 2. Locations of 20 circular ground sample plots in the study area.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

A GNSS receiver was used to determine the coordinates of each sampling plot. Several factors were taken into consideration to record precise positions including atmospheric conditions, satellites' geometry, number of location measurements at each point, multipath problems, and quality of GNSS equipment. Thus, a Leica Viva GS14 GNSS RTK receiver was used to obtain the highest level of accuracy, and real-time differential correction was applied to improve the precision of location data collected. The average horizontal accuracy of the location data is estimated to be 7 ± 4 cm.

Mangrove trees' structural attributes were collected by direct measurements at the ground plot level, including height and diameter at breast height (DBH). Using a height stick, the height of individual trees was measured. While using a diameter tape, the DBH was measured at 1.3m from the ground. Generally, a minimum threshold for measurement is specified between 5 to 10 cm and all trees with a DBH larger than this threshold should be measured (McGarrigle et al. 2011). Nevertheless, most of Avicennia marina trees in the sampling plots have a DBH value of ≤ 5 cm (approximately 69%). Thus, all trees with DBH ≥ 1 cm were measured. For trees with multiple trunks, each trunk was measured separately. A total of 2,216 mangrove trees were measured, and their structural attributes were recorded.

3. Plot-Level AGB Estimation

AGB, which is the quantity of vegetative matter per unit of area, is calculated as the dry weight of tree elements above ground including stems, branches, and leaves (Houghton, 2005). AGB was estimated for each tree, and then for each plot. In situ AGB was estimated from DBH ground measurements using the following allometric equation:

$$M = aD^b$$

Where M is the total aboveground tree dry biomass (kg), D is diameter at breast height (cm), and a and b are constants, which are estimated to be 0.5317 and 1.7476, respectively (Kirui, 2006).

AGB can be estimated using several variables such as the tree height, DBH, or both. However, in Kirui's study (2006) which was conducted in Kenya, the estimate of Avicennia marina aboveground biomass was best determined using only DBH values. Kirui found a significant relationship between the DBH and the dry weight of Avicennia marina, where r2 = 0.96. Thus, Kirui'sallometric equation was used to estimate AGB since no allometric equations have been developed for Avicennia marina in the study area. Nevertheless, a better estimation of AGB in the study area can be obtained by measuring the oven-dry-weight of Avicennia marina of all sizes, which can be used to develop refined allometric equations for precise AGB estimation.

4. LIDAR Data

Airborne LIDAR mass points (Figure 3) were collected in January 2014 using an Airborne Laser Terrain Mapper (ALTM) 3100 EA system mounted in a Beech craft King Air 350 aircraft. The aircraft was operated by Bayanat Company, which provides national geospatial products and services in UAE, with a flight altitude of approximately 1,520 m above ground level (AGL), a flight speed of 90 m per second (175 knots), and a scan angle varying from -25° to 25° from nadir. The LIDAR system was configured with a pulse rate of 100 kHz and recorded first and last returns of an emitted laser pulse. LIDAR swaths measuring 1,420 m in width were acquired with laser footprints spaced approximately every 2 m beneath the flight path. LIDAR data was received in raw LAS format (version 1.2), which is an American Society for Photogrammetric and Remote Sensing (ASPRS) standard binary format for the interchange of LIDAR data and associated metadata. Most proprietary GIS software, such as ArcGIS Pro and LP360, supports LIDAR data that is provided in LAS file format.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

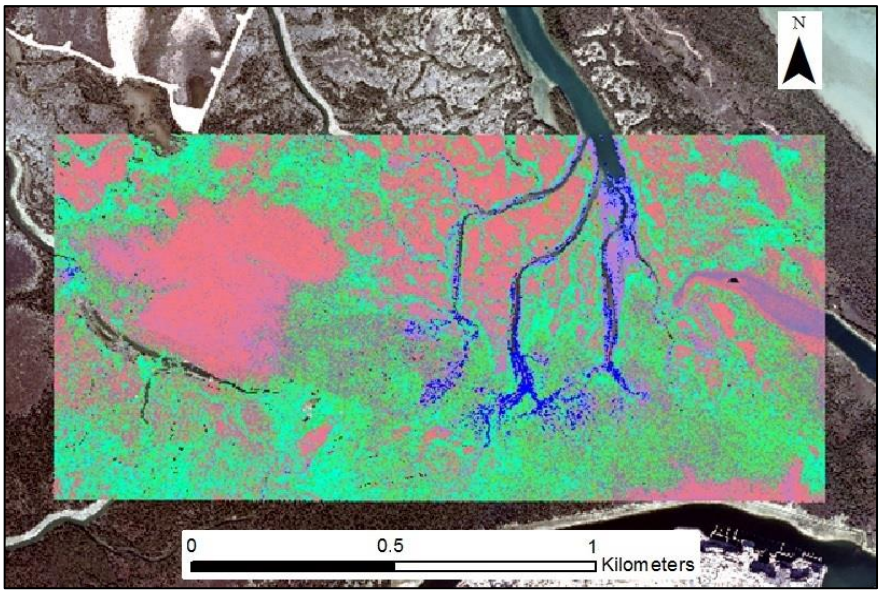

Figure 3. Digital surface model created using LIDAR data acquired in Jan 2014 by Bayanat Company. Blue dots represent ground points with lower elevation values. Pink dots represent moderate elevations, and green dots represent relatively higher ground point elevations.

5. LIDAR-derived Statistics

The first step in LIDAR data processing was to classify LIDAR points into three main categories: noise points, ground points, and high vegetation points. The noise points, identified as the very low or high points compared to the actual elevations of the study area, were identified and ignored. Usually, low points can be a result of laser multipath. On the other hand, high points can be caused by aircraft flying at low levels, birds, and/or atmospheric aerosols (McGaughey, 2014). Several software programs can be used to filter or classify LIDAR points; LAS tools were used in the present study for this purpose. Out of a total of 862,016 points, 79 noise points were ignored. Furthermore, the classification results show that the total number of ground points is 208,430; while the total number of vegetation points is 653,506. After classification, the modified LAS file was transformed to a single text file for convenient use in R statistical software. This text file contains information about x, y, z values for each LIDAR point, as well as the return number and class number. With such information, LIDAR-derived metrics can be generated, and various statistics can be calculated.

The next LIDAR processing step was to create grid points for the study area, separated by pre-defined grid spacing, using a neighbourhood technique. A neighbourhood cantered on each grid point was generated, and LIDAR points within the neighbourhood were extracted. This allowed statistics to be calculated based on the distribution of the points' z (or elevation) value and then assigned as attributes to the grid points. This technique transforms discrete LIDAR points into 2D image layers showing the spatial distributions of LIDAR height metrics with user-specified resolutions and generalization extents. The user specified resolutions, equal to the grid spacing, control the amount of horizontal detail that could be revealed. In contrast, the generalization extents, or spatial area where LIDAR points are queried, are determined by the neighbourhood size and shape. The neighbourhood size should be large enough to include a sufficient number of LIDAR points to calculate statistics precisely. It should also be larger than the grid spacing to allow overlap between adjacent neighbourhoods. However, neighbourhoods that are too large may result in masking spatial details and creating very similar adjacent grid points. Additionally, too large of a neighbourhood may include more LIDAR points falling outside the plots' boundaries, presenting unwanted information to the plots and making subsequent biomass regression models inaccurate. Furthermore, the

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

pre-defined grid spacing affects the performance of the neighbourhood method. For instance, large grid spacing results in coarse image resolutions but decreases the computing time. On the other hand, small grid spacing results in fine image resolutions but increases the computing time (Miyazaki et. al., 2014).

In the present study, grid points were created for the study area, including the 20 circular plots, and were separated by 0.5 m. The advantage of using such fine spatial resolution is to capture more spatial details and not to increase computing cost significantly due to the relatively small plot area. The spatial extents of the sampling plots were buffered by half of the neighbourhood radius. In this way, a sufficient number of LIDAR points is included to calculate statistics for grid points located at the plot boundaries. Two neighbourhood sizes, 3 and 5 m, were used as the neighbourhood radius in order to test the effect of different neighbourhood sizes and to select the most feasible size to be used to calculate LIDAR-derived height percentile statistics for the whole study area. For the neighbourhood with a radius of 3 m, a buffer of 1.5 m was used to clip the entire point cloud. In contrast, for the neighbourhood with a radius of 5 m, a buffer of 2.5 m was used. The combination of grid spacing and neighbourhood size contributes to the performance of the final regression models. Thus, the performance of the models was considered to decide the best neighbourhood size to be applied for the entire study area.

An algorithm was prepared to create the circular neighbourhood windows cantered on each of the grid points to extract LIDAR points falling within the neighbourhood. This means that the window stopped at every point on the half meter grid and selected all LIDAR returns within its boundary. LIDAR-derived statistics, including total number of points, ratio of high vegetation points, minimum elevation, and height above the minimum elevation on the 5th, 15th, 25th, 35th, 45th, 55th, 65th, 75th, 85th, 95th, and 100thpercentiles, were calculated and assigned as grid attributes to describe the distribution of the z-value within a given neighbourhood. The calculations resulted in 0.5 × 0.5 m 2D raster images with 13 layers, each representing a continuous statistical surface. However, when the numbers of LIDAR points was not sufficient to calculate statistics of some pixels within the raster images, an Inverse Distance Weighted (IDW) interpolation method was used to estimate the values of these pixels. For visualization purposes, the grid was transformed to multilayer images showing spatial distributions of LIDAR height metrics after they were clipped by the plot polygons.

6. Biomass Modelling

Forest AGB can be extracted in at least four spatial scales: 1) individual tree level, 2) pixel level, 3) plot level, 4) and segment level. In this study, the segment level was chosen largely because this approach avoids difficulties associated with predicting AGB at the individual tree or pixel scales. Image segmentation was used to group individual pixels, with similar attributes (color, texture, contextual information and other image features), into image objects that correspond to real objects (Skurikhin, 2009). Different algorithms have been developed and used to find objects from given image layers including the multi-resolution image segmentation. This algorithm is widely used for segmenting canopy height models using eCognition software (Baatz&Schape, 2000). It employs a bottom-up process to connect similar pixels that represent a homogeneous forest structure in the study area. This process is a spatial clustering technique that would place each pixel in the image into a segment based on the degree of similarity to the neighboring pixels, so it becomes more spatially meaningful than an arbitrary 1 × 1 m pixel.

In order to perform the multi-resolution image segmentation, eCognition requires some parameters to be configured manually (Table 1). The estimation of the scale parameter tool (ESP), developed by Lucian et al. (2010), and was employed to automatically determine the scale parameter. Two sets of raster images, each with 13 bands, were created using the multi-resolution image segmentation algorithm. The first was generated using a neighbourhood radius of 3 m, while the second using a radius of 5 m. The resulting image objects, or segment polygons, of connected pixels that represent homogeneous forest units, were created; and the mean values of all the height metric image pixels falling within the image object were taken as its attribute values. After assigning height

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

metric attributes to all the image objects of the plots, they were stored as ESRI Shape files to be processed in R to build AGB regression models. The known aboveground biomass of each plot, calculated based on in situ data and the allometric equation, was included in the model as the dependent variable. Decision tree regression equations, each employing rule-managed linear regression to predict the dependent variable, were developed from the LIDAR percentile height values of each image object. Using Cubist software, regression models to estimate AGB per square meter in the image segments were developed.

Table 1. Image segmentation parameters

Parameter	Value
Use of Hierarchy	1
Starting scale level1	1
Step size level1	1
Starting scale level2	1
Step size level2	3
Starting scale level3	1
Step seize level3	5
Shape	0.5
Compactness	0.5
Number of loops	100

The size of the neighbourhood affects the calculation of LIDAR derived height metrics and statistics, and therefore it affects the performance of the decision tree regression model. It is important to determine which of the two neighbourhood sizes (3 m or 5 m) is the most appropriate to be used to calculate LIDAR height metrics statistics in the study area. Thus, a 10-fold cross validation approach was applied to train models to evaluate the model performance. For each fold (10% of the data), regression models were generated using data outside this fold (90% of the data). Thus, a total of 10 models for each neighbourhood size were built. As seen in Table 2 and 3, the following statistics were used to evaluate the model performance with the different neighbourhood sizes: root mean square error (RMSE), correlation coefficient (R) between predicted target values and reference target values, and adjusted coefficient of determination (R^2). These statistics were calculated on the training set as well as the validation set.Tables4 and 5 show the statistics of the model performance using all of the data with the two neighbourhood sizes.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Table 2. The statistics of the 10-fold cross validation approach using a neighbourhood size of 3 m to evaluate the model performance.

Fold#	Training Set			Validation Set		
	RMSE	R	R^2	RMSE	R	R^2
1	13.84	0.71	0.50	15.52	0.59	0.33
2	14.39	0.64	0.40	17.98	0.35	0.08
3	13.71	0.73	0.52	12.44	0.42	0.14
4	14.73	0.66	0.43	8.23	0.81	0.64
5	13.39	0.69	0.47	19.18	0.57	0.30
6	15.13	0.61	0.37	11.70	0.51	0.23
7	11.44	0.78	0.61	25.48	0.54	0.26
8	11.27	0.83	0.68	16.00	0.50	0.21
9	13.09	0.74	0.55	14.71	0.60	0.33
10	15.16	0.62	0.38	9.89	0.54	0.26
Average	**13.61**	**0.70**	**0.49**	**15.11**	**0.54**	**0.28**

Table 3. The statistics of the 10-fold cross validation approach using a neighbourhood size of 5 m to evaluate the model performance.

Fold#	Training Set			Validation Set		
	RMSE	R	R^2	RMSE	R	R^2
1	9.88	0.89	0.79	12.21	0.68	0.44
2	9.88	0.87	0.75	18.43	0.76	0.56
3	9.99	0.88	0.77	15.96	0.58	0.30
4	9.33	0.89	0.79	13.04	0.66	0.41
5	9.92	0.89	0.79	10.38	0.69	0.46
6	9.77	0.89	0.79	14.77	0.64	0.39
7	10.76	0.84	0.71	13.18	0.66	0.41
8	9.60	0.90	0.80	12.05	0.81	0.64
9	9.95	0.87	0.76	17.86	0.83	0.67
10	10.45	0.87	0.75	14.79	0.64	0.38
Average	**9.95**	**0.88**	**0.77**	**14.27**	**0.70**	**0.47**

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Table 4. The statistics to evaluate model performance using all data with the 3 m neighbourhood size.

RMSE	R	R^2
14.10	0.66	0.44

Table 5. The statistics to evaluate model performance using all data with a neighbourhood size of 5 m.

RMSE	R	R^2
10.17	0.87	0.76

Results

As seen in Tables2 and 3, higher average R and R^2 values of the training and validation sets are observed with a 5 m neighbourhood (training R= 0.88, training R^2= 0.77, validation R= 0.70, validation R^2= 0.47) compared to a 3 m neighbourhood (training R= 0.70, training R^2= 0.49, validation R= 0.54, validation R^2= 0.28). The average RMSE is lower with the 5 m neighbourhood (training RMSE= 9.95, validation RMSR=14.27) compared to the 3 m neighbourhood (training RMSE= 13.61, validation RMSR=15.11). These statistics indicate that Cubist models performed much better on the 10-fold training and validation data (of the cross validation approach) generated using the 5 m neighbourhood radius. It is noted that the model performance of the validation data is lower than the training data even with the 5 m neighbourhood radius. In order to increase the performance on the validation data, a larger neighbourhood size can be used. However, due to the relatively small plot size (7 m radius), it will not be reliable to used neighbourhood size larger than the size of the training plots. This would present unwanted information outside the plots, and would affect the accuracy of the model.

Additionally, the model performance of all data using the 5 m neighbourhood (RMSE= 10.16, R= 0.87, R^2= 0.75) is significantly better than the 3 m neighbourhood (RMSE= 14.09, R= 0.66, R^2= 0.43). The R and R^2 values of the 5 m neighbourhood are high enough to ensure the algorithms can provide reliable and accurate predictions on the data set. Since the 5 m neighbourhood corresponded to the best model performance, it was selected to calculate LIDAR statistics in order to estimate biomass for the entire study area. The image of the study area was segmented and the model was applied to each segment. Plots of known AGB, calculated from ground reference DBH measurements using allometric equations, were then utilized as training data to generate a model for predicting AGB in the remainder of the image. Biomass density of the 20 sampling plots ranges from 0.10 to 4.28 (kg/m^2), with an average of 1.33 ± 0.24 (Table 6).The use of the Cubist model to utilize LIDAR percentile height statistics to image segmentation resulted in a segment-level aboveground biomass map of the study area as seen in Figure 4. The total number of the image segments is 715. While the size of the segments ranges from 42 m^2 to 20,005 m^2, with an average of 2,445 ± 117.88 m^2. The aboveground biomass density per segment ranges from 0.23 to 13.18 (kg/m^2), with an average of 5.16 ± 0.14. While the total aboveground biomass per segment ranges from 0.90 to 76.89 kg, with an average of 29.88 and a total of 14,850.26 kg.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Table 6. Above ground biomass of the sampling plots estimated using *in situ* measurements and allometric equations.

Plot #	AGB (kg/m²)	Plot #	AGB (kg/m²)
1	2.301	**11**	2.262
2	1.559	**12**	2.527
3	2.254	**13**	0.459
4	0.610	**14**	0.407
5	1.495	**15**	0.306
6	0.882	**16**	0.417
7	4.279	**17**	0.218
8	1.475	**18**	0.103
9	2.053	**19**	0.253
10	2.061	**20**	0.645

Using Esri's ArcGIS for Desktop 10.2, biomass estimates were classified into 10 classes from the lowest to the highest. The results indicate that 49% of the study area has a relatively low biomass density with values less than 4.15 (kg/m²), 23% of the study area has a relatively moderate biomass density with values ranging from 4.16 to 8.01, and 28% of the study area has a relatively high biomass density with values ranging from 8.02 to 13.17kg/m². Furthermore, a canopy height model (CHM) was created of the whole study area using the classified LIDAR points as seen in Figure 5. The maximum height value of the canopy model pixels is 7.86 m, which means that no trees in the study area exceed this height. The average height value of the pixels is 3.03 m while the minimum height value is 0.12m.

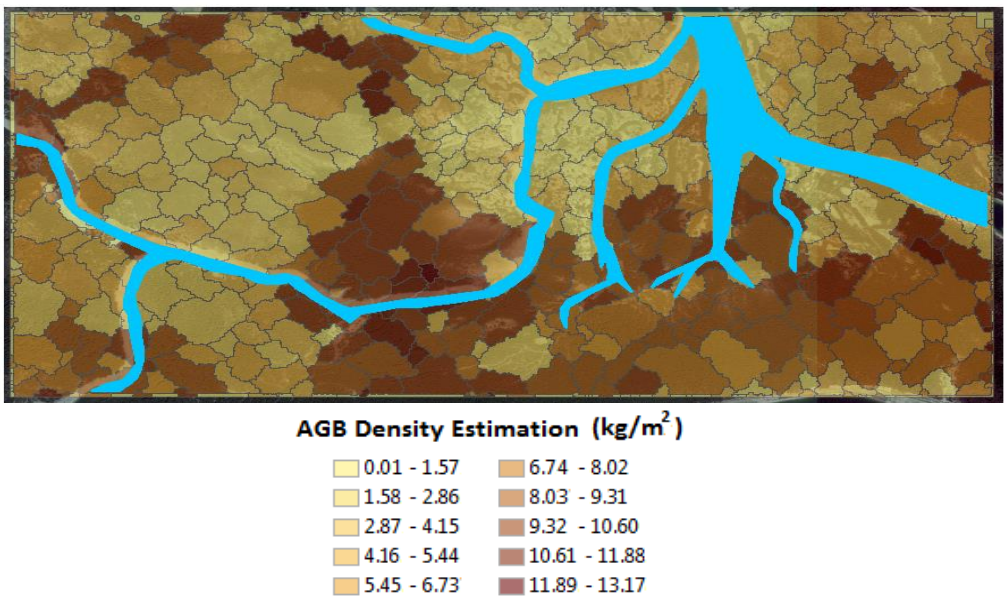

Figure 4. The map represents aboveground biomass density (kg/m²) of each segment in the entire study area. Biomass estimates are differentiated into 10 classes from the lowest to the highest.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

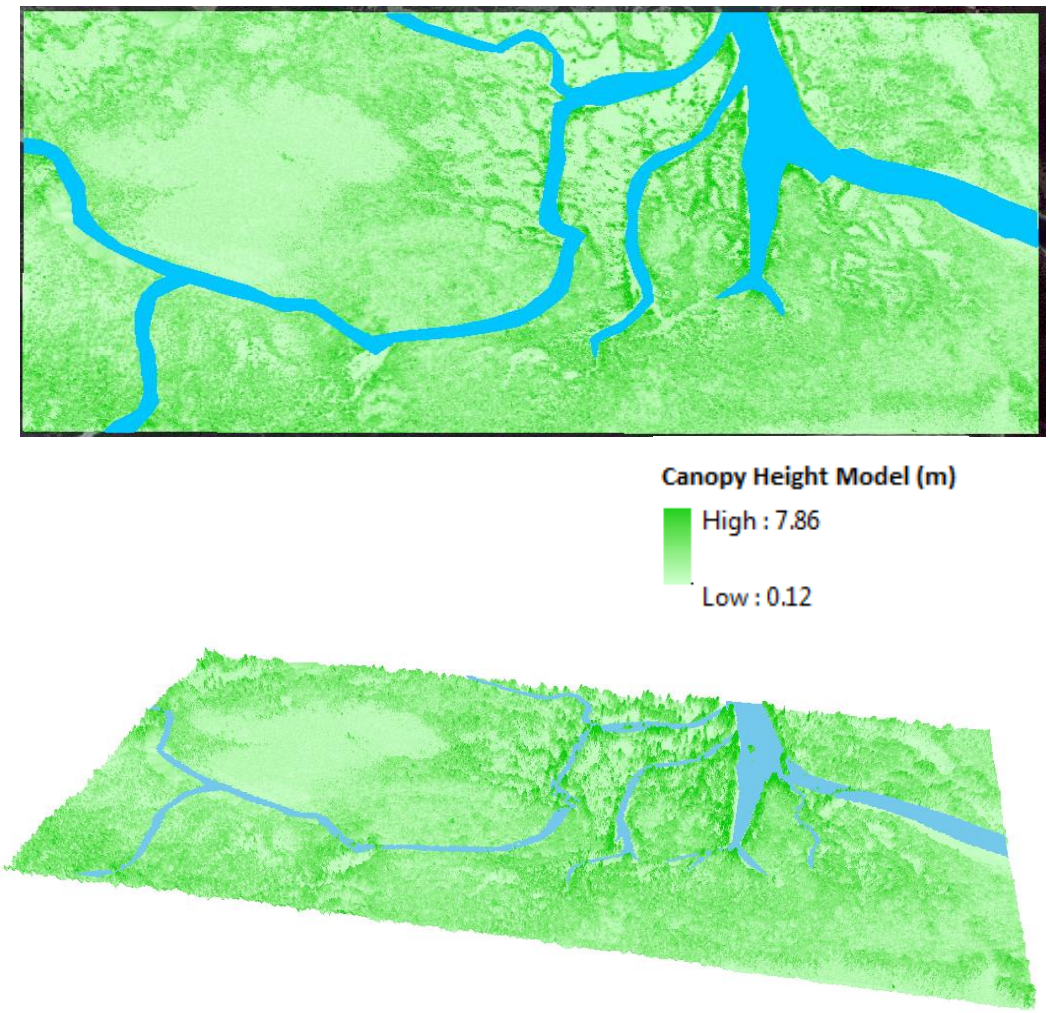

Figure 5. Canopy height model of the study area created in ArcMap and ArcScene (10.2). The darker green color indicates higher pixel values while the lighter green color indicates lower pixel values.

Conclusion

Several studies attempted to estimate aboveground biomass for various types of forests using an object-based segmentation approach, while other studies attempted to estimate aboveground biomass using multiple percentile height statistics. The current study combined the two approaches with a machine learning environment and ground reference data. Riggins et al., (2009) research is considered one of the first attempts to combine these two methods together. As in Riggins paper, combing the two methods together provide a powerful approach for deriving forest biophysical variables from LIDAR data. Additionally, this study used spatial aggregation methods to estimate aboveground biomass from small foot-print airborne LIDAR data. Similar methods can be used to estimate other forest biophysical characteristics such as basal area or leaf area index of the whole study area. However, the collection of a limited number of ground sampling plots for training purposes is still required.

The present study is considered as the first attempt to estimate and map aboveground biomass of *Avicennia marina* forest in Eastern Mangrove Lagoon National Park, Abu Dhabi. By utilizing LIDAR percentile canopy

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

height and image segmentation in a machine-learning algorithm and *in situ* reference data, a map of aboveground biomass density per segments was successfully produced. Segments-level processing divides the data into homogenous forested units according to the values of the neighboring pixels. Therefore, selecting the best neighbourhood size is critical for calculating LIDAR height metrics, and it influences the accuracy of the regression model. Due to the small size of the circular sampling plots (7 m radius each), the choice of neighbourhood size is limited. In this study two neighbourhood radii were tested (3 m and 5 m). The latter resulted in the highest model accuracy (RMSE= 10.17, R= 0.87, R^2= 0.76). However, if the training plots were larger in size, larger neighbourhood sizes could be tested in an attempt to increase the model performance.

The number of the percentiles height layers affects the accuracy of the model performance. The fewer number of percentile heights means that less information about a canopy structure is taken into account. Some studies used 5 or 6 percentiles heights (Lim & Treitz, 2004; Mariappan et. al. 2012); however, this study used 11 percentiles heights (0, 5, 15, 25, 35, 45, 55, 65, 75, 85,100th) to provide better information about the canopy structure at more elevations. Although creating more percentiles layers increase the computing time, the small study area and the small LIDAR point cloud data (less than a million points) did not lead to any computing disadvantages.

The biomass model presented in this study can be very useful in monitoring mangrove status and biomass conditions, leading to a better management of the forested areas in Abu Dhabi. Landscape mangrove biomass estimates are needed because of the importance of mangrove in the carbon cycle. By quantifying the amount of forest biomass, carbon emission and storage can be easily estimated. Thus, measuring mangrove extent, structure and biomass is vital for addressing climate change mitigation and adaptation.

Acknowledgments

We would like to thank the Marine Biodiversity Sector and the Environment Information Management Sector at the Environment Agency in Abu Dhabi (EAD) for facilitating forest survey and providing water transportation. We extend our sincere thanks to Mr. Yousef Al-Marzouqi from Bayanat Company and Mr. Jorge Merrit from Abu Dhabi Systems and Information Company (ADSIC) for providing remote sensing data including LIDAR and WorldView-2 satellite image of the study area. We would further like to thank General Enterprises Co. Engineering Division (GECO) in Abu Dhabi for providing GPS receivers and assisting in collecting sampling plots' locations. We are very grateful for Mr. Khalfan Al-Meherbi from EAD, Mr. Zia Rehman from GECO, and Mr. Abdulla Al-Marzouqi from ADSIC for their great assistance in gathering *in situ* data.

References

Aguilar, F.J., Mills J.P. (2008). *Accuracy Assessment of LiDAR-derived Digital Elevation Models*. Photogrammetric Record, 23 (122), 148-169.

Alsumaiti, Tareefa S., Hussein, Khalid, Al-Sumaiti, Ameena.(2017).*Mangrove of Abu Dhabi Emirate, UAE, in a Global Context*. International Journal of Environmental Sciences. CRDEEP. 6(3), 110- 121. ISSN: 2277-1948

Aschbacher, J. K., Ofren, R. B., Delsol, J. P., Suselo, T. B., Vibulsresth, S. K. and T. K. Charrupat. (1995). *An Integrated Comparative Approach to Mangrove Vegetation Mapping Using Advanced Remote Sensing and GIS Technologies*; Preliminary Results. Hydrobiologia, 295:285-294.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Baatz M., Schape A. (2000). Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. *In: Strobl, J., Blaschke T., Griesebner G. (eds.), AngewandteGeographischeInformationsverarbeitung XII*. Wichmann-Verlag, Heidelberg, pp 12-23.

Bombelli A., Avitabile V., Balzter H., Belelli Marchesini L., Bernoux M., Brady M., Hall. R., Hansen M., Henry M., Herold M., Janetos A., Law B.E,Manlay R., Marklund L.G., Olsson. H., Pandey D., Saket M., Schmullius C., Sessa R., Shimabukuro Y.E., ValentiniR.,Wulder M. (2009). Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables: BIOMASS. GTOS 67, AFO, Global Terrestrial Observing System. Rome.

Fatoyinbo, T., Armstrong, A. (2010). *Remote Characterization of Biomass Measurements: Case Study Of Mangrove Forests Biomass.*Sciyo Publishing.

Goetz, S.J., Baccini, A., Laporte, N.T., Johns,T., Walker,W., Kellndorfer, J., Houghton, R. A., and Sun, M.(2009). *Mapping and monitoring carbon stocks with satellite observations: a comparison of methods*. Carbon Balance and Management.

Houghton, A. (2005). Aboveground forest biomass and the global carbon balance. *Global Change Biology* 11(6):945–58.

Houngton, R.A. Hall, F.G. Goetz, S.J. (2009). The importance of biomass in the global carbon cycle. *Journal of Geophysical Research* – Biogeosciences. 114(G2).

Jachowski N. R. A., Quak M. S. Y. ,Friess D. A., Duangnamon D., Webb E. L., Ziegler A. D. (2013). Mangrove biomass estimation in Southwest Thailand using machine learning. *Applied Geography* (45): 311-321

Jensen, John R. (2005). Introductory Digital Image Processing, 3rd Ed. New Jersey. Prentice Hall: Upper Saddle River.

Kirui, B. (2006). Allometric relations for estimating aboveground biomass of naturally growing mangroves, Avicennia marina FORSK (VIERH). and Rhizophora mucronata LAM. along the Kenya coast. Egerton University.

Laba, M., S.D. Smith, and S.D. Degloria (1997). *Landsat-based land cover mapping in the lower Yuna River watershed in the Dominican Republic*, International Journal of Remote Sensing, 18(14):3011–3025.

Lim K. S., Treitz P. M. (2004). Estimation of Above Ground Forest Biomass From Airborne Discrete Return Laser Scanner Data Using Canopy-based Quintile Estimators. *Scandinavian Journal of Forest Research*, 19(6):558-570.

Lu D., Chen Q., Wang G., Moran E., Batistella M., Zhang M., Laurin G.V., and Saah D. (2012). Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *International Journal of Forestry Research.* Volume 2012, Article ID 436537, 16 pages.

Lucas, R.M. ; Sch. of Geogr. Sci., Bristol Univ., UK ; Milne, A.K. ; Mitchell, A. ; Donnelly, B. (2000). Use of stereo aerial photography for assessing changes in the extent and height of mangrove canopies in tropical Australia. *IGRASS Proceedings, 5: 1880 - 1882*

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Lucian D., Tiede D., Levick S. R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Sciences.* Vol. 24 No. 6. pp 859- 871.

McGarrigle, E.; Kershaw, J.A.; Lavigne, M.B.; Weiskittel, A.; Ducey, M. (2011). Predicting the Number of Trees in Small Diameter Classes Using Predictions from a Two-parameter Weibull Distribution. *Forestry* 84: 431-439.

Mariappan, M. Lingava, S. Murugaiyan, R. Kolanuvada, S.R. Thirumeni, R.S.L.(2012). Carbon Accounting of Urban Forest in Chennai City using Lidar Data. *European Journal Of Scientific Research.* 81(3):314-328.

McGaughey R. J. (2014). FUSION/LDV: Software for LIDAR Data Analysis and Visualization. United States Department of Agriculture (USDA). Forest Service. Pacific Northwest Research Station. FUSION Version 3.42.

Mitchard E. T. A., Saatchi S. S.,Woodhouse I. H., Nangendo G., Ribeiro N.S., Williams M., Ryan C.M., Lewis S.L., Feldpausch T.R. , Meir P. (2009). Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different African landscapes. *Geophysical Research Letters*, vol. 36.

Miyazaki R., Yamamoto M., Hanamoto E., Izumi H., Harada K. (2014). A line-based approach for precise extraction of road and curb region from mobile mapping data. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume II-5, 2014, pp.243-250.

Poudel, K.P.,Temesgen, H. (2016)Methods for estimating aboveground biomass and its components for Douglas-fir and lodgepole pine trees. Canadian Journal of Forest Research. 46(1): 77-87.

Ramsey III, E., and J. Jensen (1996). Remote sensing of mangroves: relating canopy spectra to site-specific data. *Photogrammetric Engineering and Remote Sensing*, 62(8):939-948

Rasolofoharinoro et al. (1998). A remote sensing based method for mangrove studies in Madagascar. *Int. J. of Remote Sensing* . 19 (10) : 1873-1886.

Riggins J.J., Tullis J.A. & Stephen F.M. (2009) Per-segment aboveground forest biomass estimation using LIDAR-derived height percentile statistics. *GIScience&RemoteSensing* 46: 232-24.

Simard, M., Zhang, K., Rivera-Monroy, V.H., Ross, M., Ruiz, P., Castaneda-Moya, E., Rodriguez, E., Twilley, R. (2006). Mapping mangrove height and estimate biomass in the Everglades using SRTM elevation data. *Photogrammetric Engineering and Remote Sensing*, 72 (3):299-311.

Skurikhin A. N. (2009). Hierarchal Image Feature Extraction by an Irregular Pyramid of Polygonal Partitions. ASPRS Annual Conference, Baltimore, MD.

Song C.(2013). Optical remote sensing of forest leaf area index and biomass. Progress in *Physical Geography*, vol. 37, 1: pp. 98-113.

VOL: 5, ISSUE: 1
January/2019
https://ijbassnet.com/
E-ISSN: 2469-6501

Wannasiri W., Nagai M., Honda K., Santitamnont P., Miphokasap P. (2013). *Extraction of Mangrove Biophysical Parameters Using Airborne LiDAR.* Remote Sensing 04/2013; 5(4):1787-1808. DOI:10.3390/rs5041787.

White, J.C.; Wulder, M.A.; Varhola, A.; Vastaranta, M.; Coops, N.C.; Cook, B.D.; Pitt, D.; Woods, M.(2013). Best Practices Guide for Generating Forest Inventory Attributes from Airborne Laser Scanning Data Using an Area-based Approach. *Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre, Victoria, BC.*

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

Zhao F., Guob Q., Kelly M. (2012). Allometric equation choice impacts lidar-based forest biomass estimates: A case study from the Sierra National Forest, CA. *Agricultural and Forest Meteorology*165:64-72.