LiDAR Above Ground Biomass Calculation Tool Project

GISC 3200k - Programming for Geospatial Science & Technology

Author:

Jeff Robertson

Prepared For:

Dr. Cho

October 28, 2020

Abstract

The ability to efficiently monitor forest carbon is essential to the management and sequestration of carbon. With the development of smaller sensor packages such as those attached to drones, data can be collected and processed more efficiently and cost effectively. This allows decision makers to make informed decisions regarding the effects of anthropogenic change. This project looks at ways to automate the processing of the multispectral data collected by sensors to estimate above ground vegetation biomass. Other spatial and quartile data is collected in the process, which can be utilized for further resource management.

Large data sets can be processed in a more consistent and efficient way through automation. The data processing step was automated by creating two ArcPro Python Toolboxes. This could also be accomplished using open source software. The first toolbox was used to process LiDAR point cloud data. This data provides accurate elevation information, which produces raster tiles smaller than a foot, depending on cloud density. The LiDAR processing tool creates a digital elevation map (DEM), digital surface map (DSM), and normalized digital surface map (nDSM). Each one of these layers can be used for environmental monitoring calculations. The nDSM is important for biomass calculation, which was the focus of this project. This layer allows for the calculation of forest canopy height. The second toolbox was used to clip a study area, create a normalized difference vegetation index (NDVI), and estimate the above ground biomass.

Both the NDVI and biomass estimates are important for understanding the health of the forest and its potential concentration of carbon. This tool uses an algorithm that was developed by a graduate student at Duke University. It uses canopy height data and NDVI to output megagrams per hectare (Mg/ha) of yield. This allows for the calculation of stored carbon by species. The goal is to use these tools to develop a more attractive and updated forest monitoring system which utilizes drones and ground sensors.

Table of Contents

Introduction

The ability to efficiently measure aboveground biomass (AGB) is beneficial for a better understanding of carbon sequestration. (Zhao et al., 2012) Above ground biomass accounts for all the living vegetation above the soil. This has a great effect on forest watershed, which is crucial for freshwater supply. In the United States, forest watersheds account for 80% of our fresh water. (Mwamba et al.2016) So, the ability to calculate AGB efficiently and accurately is important for freshwater supply and its anthropogenic response. The total amount of potential carbon emissions from AGB is not certain due to the lack of suitable spatial metrics and algorithms, and inconsistent elevation data. (Lu et al., 2012) Some of this inconsistent elevation data is due to spatial resolution and the sensing system used for collection.

Allomeric equations are regression models for estimating biomass based on height data. This is the primary method of estimating biomass in the United States. In 2003, a case study to estimate forest biomass in the Sierra National Forest was published. The study's author provided ten equations for estimating total aboveground biomass. (Hoover & Smith, 2016) These equations calculated and categorized the parts of the trees being studied. (Figure 1) This included stems, wood, stem bark, foliage, and below ground coarse roots, which are all combined for total AGB. (Zhou & Hemstrom, 2009) The initial allomeric equation breaks the species into ten species groups consisting of five softwood, four hardwood and one woodland group. (Jenkins, Chojnacky, Heath, & Birdsey, 2004) This gave a total of ten possible equations for each of the ten tree species. (Figure 2) The objectives of the Sierra National Forest study were to develop a consistent generalized biomass regression equation for use in large scale inventory, based on carbon inventory. The goal was that the carbon budget should take into consideration several ecosystem components including live biomass, debris, and soil. (Jenkins et al.2003)

The allomeric equation was updated in 2013 by Chojnack. In 2002, he published a paper to show the inconsistency of the calculation on a regional scale. He posed that the allometric scaling theory could improve the reliability of the Jenkins allometric equations. (Chojnacky, 2002) Chojnacky later expanded the original ten equations to thirty-six generalized equations consisting of thirteen conifers, eighteen hardwood, and four woodland. These equations resulted in a 20% higher biomass than the original model estimated. (Chojnacky, Heath, & Jenkins, 2013)

There are many studies showings the use of LIDAR and Landsat imagery to estimate AGB. (Badreldin & Sanchez-Azofeifa, 2015) The purpose of this study will be to make a python script to process LIDAR data in order to find tree height, diameter of base height, and canopy width. The script should make it easier to reproduce this project. Finally, this project will build an ArcPro tool that will calculate the AGB for the different species of tree.

Material and Methods

For this study, a 1-meter LiDAR was used from the study area of Walton County. The first step was to extract the digital elevation model (DEM), digital surface model (DSM), and canopy height model (CHM). For the CHM, a normalized digital elevation model (nDSM) was used. An algorithm developed by a graduate student at Duke University was used. (Riegel, 2012) The algorithm uses average canopy height and NDVI to output Mg/ha of yield.

Two ArcPro Python Toolboxes were created to process National Agriculture Imagery Program (NAIP) data and US Geological Survey (USGS) 3D point cloud data. The data was then clipped to the desired area and used for calculations. The goal was to come up with an efficient automated process for getting data that can be used for analyzing forest health and carbon concentration, which could be a building block for future research.

While this study specifically looked at biomass calculations, other data sets which could be useful in monitoring natural resources were considered. For instance, elevation data could be utilized for slope, aspect, and watershed modeling. The NAIP only looks for vegetation health, but eventually could be developed for vegetation classification. This would allow for a much more accurate calculation of above ground biomass.

The ability to have a standard and proficient way of processing temporal data is important, especially with the increasing developments in remote sensing technology like drones. This type of data processing is going to be crucial.

Data and Software

- ESRI ArcPro ModelBuilder and Python Toolbox
- USGS 3D Elevation LiDAR data
- USGS NAIP Imagery on AGOL
- NDVI Formula
  - NDVI = (NIR – RED)/(NIR + RED)
- Algorithm developed at Duke University for estimating vegetation biomass
  - (exp[4.33 + 0.28 * Log10(nDSM_Mean)-1.05 * Log10(NDVI_Mean) + 3.96 *NDVI_Max]1)/10

Process and Results

These tools were developed in the ArcGIS ModelBuilder. The approach was to work out all the details in ModelBuilder and then use the Python toolbox to build the script. ModelBuilder allowed for much quicker results then starting a script from scratch in the toolbox. The first step was to process the LiDAR data and produce an nDSM. This was needed to find the mean of the forest canopy height. Processing the point cloud return values was difficult in the ModelBuilder.

The LAS to Raster Tool does not allow for the selection of return values. Normally, when processing LiDAR in ArcGIS, the user must go to the properties or the quick access toolbar to change the settings. For this tool, the properties were changed to select the desired return values. This allows the user to turn off all but the ground returns to create a DEM. To classify these returns and create an accurate DEM in the model, the Classify LAS Ground Tool was used. Then, the Make LAS Dataset Layer tool was used twice to create ground and vegetation layers. The LAS Dataset to Raster tool was then used to create the DEM and DSM. The Minus tool was used to subtract the ground from the surface to get the canopy height or nDSM. The final output raster did not have an attribute table. This may be important for future research, so the Build Raster Attribute Table tool was used.

The LiDAR processing script for this tool was then exported directly to a Python file. A new toolbox was created, and the script was inserted. This tool only has two parameters: input and output location. The tool was designed so that it does not require the .las file to be loaded into the map. This was limiting the amount of memory required for processing multiple .las files. The output is also automatically placed in the project folder for the same reason. The tool produces an .las dataset, DEM, DSM, nDSM and a nDSM with attributes. If multiple. las inputs and outputs were used, this would allow the data to run in the background without affecting the active map project. This allows the user to select the desired elevation data without overloading system.

The Biomass Algorithm Toolbox was also designed using ModelBuilder. This also allowed a way for the process to be worked out before converting the model to script. Unlike the last tool, this tool can use maps that are imported into the project map. The output also automatically adds the files to the map. This is important for visualization and selection of an area. The toolbox has two tools: one to clip an area of homogeneous ground cover and another to calculate the NVDI, mean height, mean NDVI, and max NDVI. A process was then scripted to use the calculated data and apply it to the estimating vegetation biomass algorithm.

The first tool then clips NAIP imagery and the nDSM. This was done by scripting two Extract by Mask tools into the process. The ability to select the area of interest polygon was accomplished by defining the datatype on the input to GPFeatureRecordSetlayer. This input allows the creation of a point, line, or polygon feature layer. This is what is used by the Extract by Mask script. It allows the user to label the outputs as desired.

The final tool uses the clipped NAIP, nDSM and polygon produced in the last tool as the inputs. The first process in the tool calculates the NDVI by using the red and NIR band of the NAIP imagery. This is accomplished by using the "arcpy.sa.NDVI" script. This script allows for the input of the NAIP and then requires the band numbers to be entered. The band 4 is entered first because it is the first band in the calculation. Then, using a scripted Zonal Statistics tool, the NVDI, mean height, mean NDVI, and max NDVI is calculated. The final process is the biomass algorithm calculation. This is done by using the spatial analysis math tool set syntax.

This tool worked as desired. The final output of the tool is yield in megagrams per hectare (Mg/ha). However, it is important to remember this is just an estimate of above ground biomass. The tool still needs some modifying and calibrating. To get a more accurate assessment, each species should be considered individually. This tool has the potential be useful for observing change, growth, and health.

Conclusion

Although this tool was created for a python class project, it has potential for further development. The additional use of high-resolution data could introduce further usefulness for this tool. The use of drones would allow for even higher resolution imagery. There are limitations in using the ESRI software to run these tools. The software can be expensive and is not available to everyone. Also, there are limitations in using the toolboxes ap opposed to a standalone script. The software requires processing power and produces residual data.

The future goal for this tool is to adapt the scripts to work in the command line and to work with drone data. There is currently some interesting open source software available, like

Open Drone Map. This would be a good starting point for further development. Being able to process forestry and vegetation data is import for proper management. Manned aircraft data collection and processing is expensive. It is also limited. The current commercial drone and processing software on the market is also expensive, so it wouldn’t likely be an option for the smaller producer.

Therefore, I think it is important that opensource data that can work for applications using affordable drones is produced. This allows for affordability on a scale that can be widely used and applied. The software should be user friendly so that in-depth technical experience is not necessary. The end user can then utilize the tool to analyze crops and natural resources at the local level. Empowering individuals with affordable user-friendly tools in this manner could lead to a much broader level of understanding. This knowledge informs decision making at all levels and could lead to a much greater impact on our environment.

Calculations

Mg/ha = (Exp (4.33 + 0.28 * Log10(nDSM_Mean)- 1.05 * Log10(NDVI_Mean) + 3.96 * NDVI_Max)-1)/10

NDVI = (NIR - RED) / (NIR + RED)

dbh = 2 x {square root (basal area/3.142)}

V 10 = P 10 × (P 10 × Eq. (5))2

Jenkins allometric equation

B = exp (b0 + b1 ln (DBH))

Where:

basal = diameter at breast height

P10 = percent of height

V10 = Volumetric metrics at each percent of hight

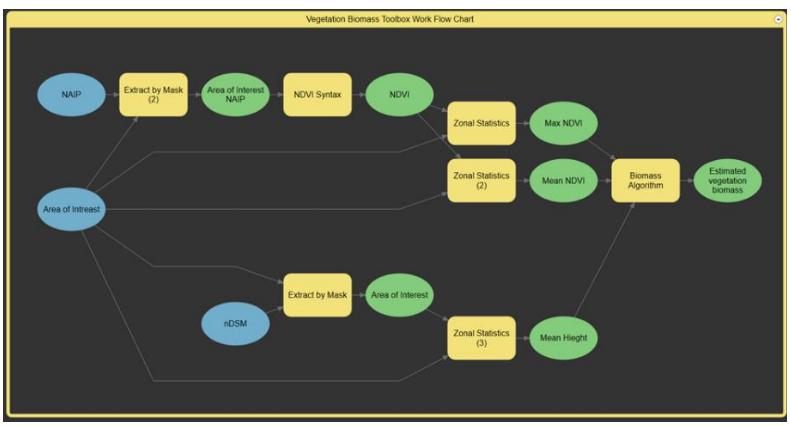
Bm = total aboveground biomass (kg) for tress 2.5 cm and larger in d.b.h

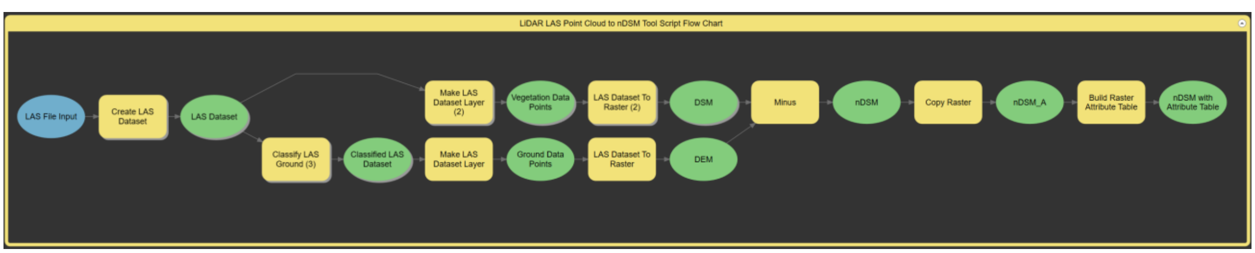
Dbh = diameter at breast height (cm)

Exp = exponential function

In = natural log base "e" (2.718282)

Tables and Charts





References

Jenkins JC, Chojnacky DC, Heath LS, Birdsey RA. National-scale biomass estimators for United States tree 2016 billon ton-Ton Report | 183 06 Water Quality Response to Forest Biomass Utilization Benjamin Rau [Press release]. (n.d.).

Allometric equation choice impacts LiDAR-based forest biomass estimates: A case study from the Sierra National Forest, CA [Press release]. (n.d.).

Atgamosa, R., Paringit, E., Quinton, K., Tandoc, F., Faelga, R., Ibañez, C., . . . Zaragosa, G. (2016). FULLY AUTOMATED GIS-BASED INDIVIDUAL TREE CROWN DELINEATION BASED ON CURVATURE VALUES FROM A LIDAR DERIVED CANOPY HEIGHT MODEL IN A CONIFEROUS PLANTATION. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI*(B), 8th ser., 563-569.

Badreldin, N., & Sanchez-Azofeifa, A. (2015). Estimating Forest Biomass Dynamics by Integrating Multi-Temporal Landsat Satellite Images with Ground and Airborne LiDAR Data in the Coal Valley Mine, Alberta, Canada. *Remote Sensing, 7*(3), 2832-2849. doi:10.3390/rs70302832

Chojnacky, D. C. (2002). ALLOMETRIC SCALING THEORY APPLIED TO FIA BIOMASS ESTIMATION. *Forester, Forest Inventory Research, Enterprise Unit*.

Chojnacky, D., Heath, L., & Jenkins, J. (2013, December 29). Updated generalized biomass equations for North American tree species. Retrieved September 16, 2020, from https://academic.oup.com/forestry/article/87/1/129/602137

Hoover, C. M., & Smith, J. E. (2016). Evaluating revised biomass equations: Are some forest types more equivalent than others? *Carbon Balance and Management, 11*(1). doi:10.1186/s13021-015-0042-5

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2004). Comprehensive database of diameter-based biomass regressions for North American tree species. doi:10.2737/ne-gtr-319

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., . . . Saah, D. (2012). Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *International Journal of Forestry Research, 2012*, 1-16. doi:10.1155/2012/436537

Muwamba, A., Trentin, C., Panda, S., Amatya, D. M., Tollner, E. W., & B. R. (n.d.). Water Quality Response to Forest Biomass Utilization. *2016 Billion-Ton Report,* 184-209.

Riegel, J.B., 2012. A Comparison of Remote Sensing Methods for Estimating Above-Ground Carbon Biomass at a Wetland Restoration Area in the Southeastern Coastal Plain.

Zhao, F., Guo, 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. doi:10.1016/j.agrformet.2012.05.019

Zhou, X., & Hemstrom, M. A. (2009). Estimating aboveground tree biomass on forest land in the Pacific Northwest: A comparison of approaches. doi:10.2737/pnw-rp-584