# Technical Details on the

# Riparian Classification from LiDAR (RCL) Tool

Authors: Sky Jones1; Racha El Kadiri2, PhD; Henrique Momm3, PhD;

1Middle Tennessee State University, [rj3h@mtmail.mtsu.edu](mailto:rj3h@mtmail.mtsu.edu)

2Assistant Professor, Middle Tennessee State University, [racha.elkadiri@mtsu.edu](mailto:racha.elkadiri@mtsu.edu)

3Asssociate Professor, Middle Tennessee State University, [henrique.momm@mtsu.edu](mailto:henrique.momm@mtsu.edu)





## Abstract

*A vast catalogue of LiDAR point cloud data is available online in government and private databases, and the amount of data available is constantly increasing as new missions are contracted and completed. Aerially-collected LiDAR point clouds are useful for many analytical and visualization purposes, including the delineation and characterization of vegetation. Under contract with the NRCS, the MTSU Department of Geosciences has produced a model and associated ArcGIS tool to classify riparian landcover called the Riparian Classification from LiDAR (RCL) model. This document details the design decisions and workflow used to generate this model, the scope of the model’s applicability and future plans for modifications to the model.*

## Goals

Goal of NRCS project

## An Introduction to LiDAR

Light Detection and Ranging, or LiDAR, uses laser pulses to measure the distance between a sensor and a target. Upon striking the target, the laser pulse is reflected and the sensor records the time between the pulse leaving the LiDAR apparatus and returning to the sensor. This information can be used to determine the distance between the sensor and object. In coordination with a GPS system and when mounted on aircraft, LiDAR systems can be used to generate detailed representations of the ground below the flightpath. These data, called *point clouds*, typically contain millions of points each representing a *return* (reflected laser pulse). Each point is associated with geographic information (northing, easting, elevation) as well as ancillary information, such as the *intensity* of the return (a measure of how reflective the target is) and the *return number*. Individual laser pulses from the LiDAR apparatus spread as they travel, resulting in an ever-widening *footprint*. If the footprint strikes an object with multiple elevation levels (such as a building edge or a tree), then the originating pulse will return in multiple pulses rather than one discrete pulse; the order that the split pulses return determines the return number for each returned pulse. LiDAR vendors will occasionally classify landcover using proprietary models, but the classification methods vary in quality and detail.

Traditionally three types of elevation models are generated from aerially-collected LiDAR point clouds. A Digital Surface Model (DSM) is created by interpolating only the first returns in the point cloud, and roughly represents the elevation of the ground or, if present, above-ground structures such as trees and power lines. A Digital Elevation Model (DEM) is created by interpolating only the last returns in a point cloud, and roughly represents the elevation of the ground without any above-ground structures. Occasionally points classified as buildings will be excluded when generating a DEM, but vendor classifications cannot always be relied on for this. The third elevation product, the Digital Height Model, is the difference between the DSM and the DEM. This roughly represents the height of above-ground structures. Because the DHM is often used as a proxy for canopy height it is sometimes called the Canopy Height Model (CHM). These data products, as well as those derived from them such as slope models, can be used to train machine learning algorithms to identify the unique morphometric signatures associated with different landcover types.

## Methods

A watershed in western Tennessee (HUC 080102040304) was the primary focus for this study. Data for a 2012 LiDAR mission entirely covering the watershed was obtained and used to create multiple data products, including but not limited to a DEM, DSM, DHM, and derived slope models. Approximately 6.4% of the watershed’s landcover was manually classified into one of 19 categories **(Table 1)**, which was then used to train a decision tree to classify landcover. This process was repeated for six additional watersheds across the continental US **(Table 2)** in order to evaluate the effects of physiography and LiDAR vendor on model validity; training coverage accounted for approximately 10.0% of all land in the study areas. Though 18 landcover types were manually classified, model output was restricted to 2 (trees and all other) or 3 (trees, herbaceous vegetation and all other) classes due to insufficient differences in LiDAR signature between most classes.

Though this model is primarily intended to classify riparian landcover, landcover throughout the entirety of the watersheds was classified. Doing so increases the amount of training data available to the model and diversifies the LiDAR-signatures encountered during training.

## Results

The decision trees for a ternary and binary, western Tennessee-specific (HUC 080102040304) ArcGIS-friendly models are shown in **Figures 1 and** **2**. More general models **(Figures 3 and 4)** are provided as well. Quality metrics for each model are available in **Table 3**.

## Discussion

Discussion, including model limitations. Packaged model is the TN model.

## Future Plans

Allow RCL tool to either use the pre-trained decision tree or accept training input and use SVM model.

## Acknowledgements

Many thanks to the National Resources Conservation Service for providing funding for this project.

## Tables and Figures

|  |  |
| --- | --- |
| **Category** | **Reclassification** |
| Forest | Trees |
| Linear Trees | Trees |
| Individual Trees/Small Clusters | Trees |
| Building Tops | Other |
| Building Edges | Other |
| Dirt/Bare Field | Other |
| Crops | Herbaceous Vegetation |
| Rough Vegetation | Herbaceous Vegetation |
| Other Impervious Surfaces | Other |
| Water | Other |
| Snow | Other |
| Bare Rock | Other |
| Sand | Other |
| Wetlands | Herbaceous Vegetation |
| Power Lines | Other |
| Charred Trees and Vegetation | (excluded) |
| Utility Easement | Herbaceous Vegetation |
| Large-Scale Urban | (excluded) |
| Canyon | Other |

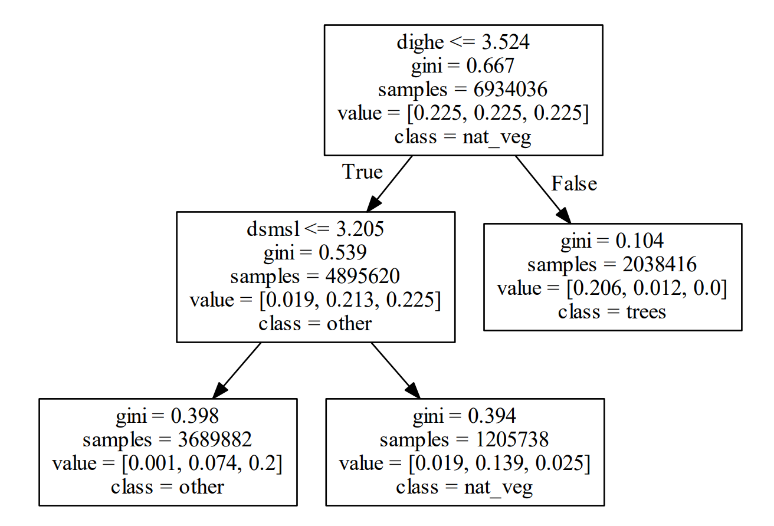
**Table 1.** Landcover classes used to train the model. The more specific categories (left column) were grouped together (right column) in order to improve model quality.



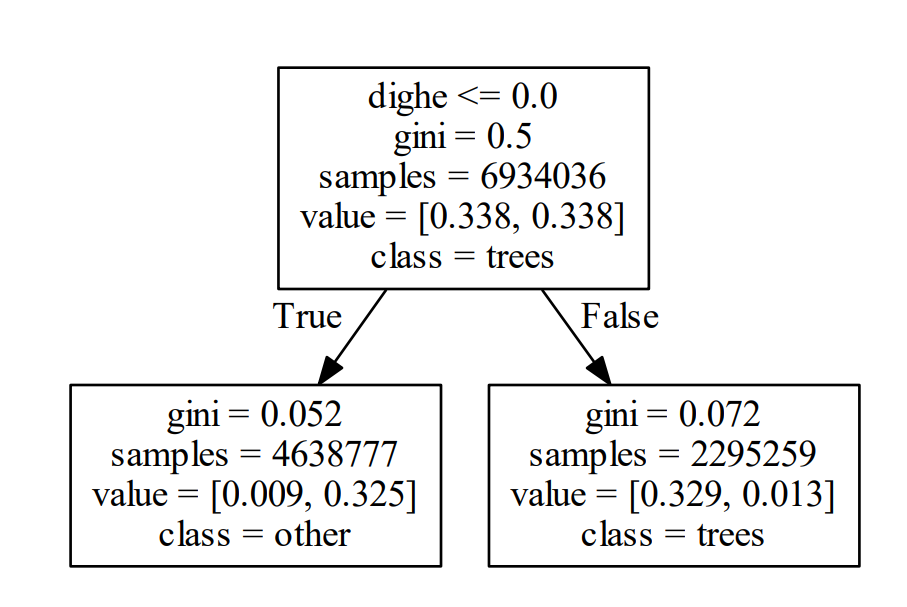
**Table 2.** Locations used to generate the general model.

|  |  |  |  |
| --- | --- | --- | --- |
| **TN, Ternary** | | | |
|  | **Precision** | **Recall** | **F1-Score** |
| Trees | 97.3% | 91.4% | 94.3% |
| Herb. Veg. | 49.5% | 61.7% | 55.0% |
| Other | 91.2% | 88.7% | 89.9% |
| *Accuracy* | 85.8% | | |
| *Macro Avg.* | 79.4% | 80.6% | 79.7% |
| *Weighted Avg.* | 87.3% | 85.8% | 86.4% |
| **TN, Binary** | | | |
|  | **Precision** | **Recall** | **F1-Score** |
| Trees | 92.1% | 97.3% | 94.6% |
| Other | 98.8% | 96.2% | 97.4% |
| Accuracy | 96.5% | | |
| Macro Avg. | 95.4% | 96.8% | 96.0% |
| Weighted Avg. | 96.7% | 96.5% | 96.6% |
| **General, Ternary** | | | |
|  | **Precision** | **Recall** | **F1-Score** |
| Trees | 96.4% | 93.8% | 95.1% |
| Herb. Veg. | 17.5% | 60.0% | 27.1% |
| Other | 93.7% | 77.6% | 84.9% |
| *Accuracy* | 86.9% | | |
| *Macro Avg.* | 69.2% | 77.1% | 69.0% |
| *Weighted Avg.* | 92.9% | 86.9% | 89.2% |
| **General, Binary** | | | |
|  | **Precision** | **Recall** | **F1-Score** |
| Trees | 95.4% | 96.8% | 96.1% |
| Other | 94.8% | 92.7% | 93.7% |
| Accuracy | 95.2% | | |
| Macro Avg. | 95.1% | 94.7% | 94.9% |
| Weighted Avg. | 95.2% | 95.2% | 95.1% |

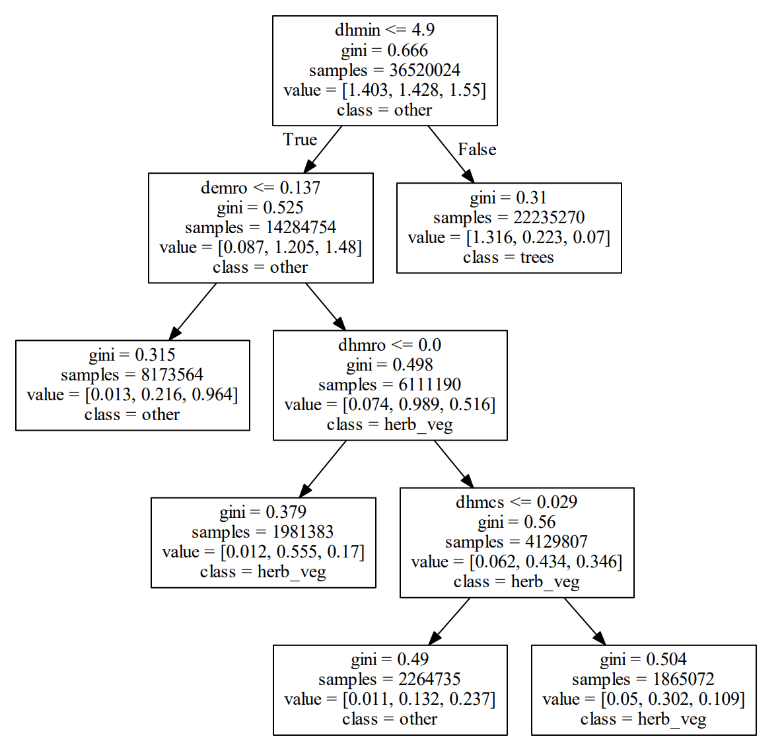
**Table 3.** Quality metrics for each model.



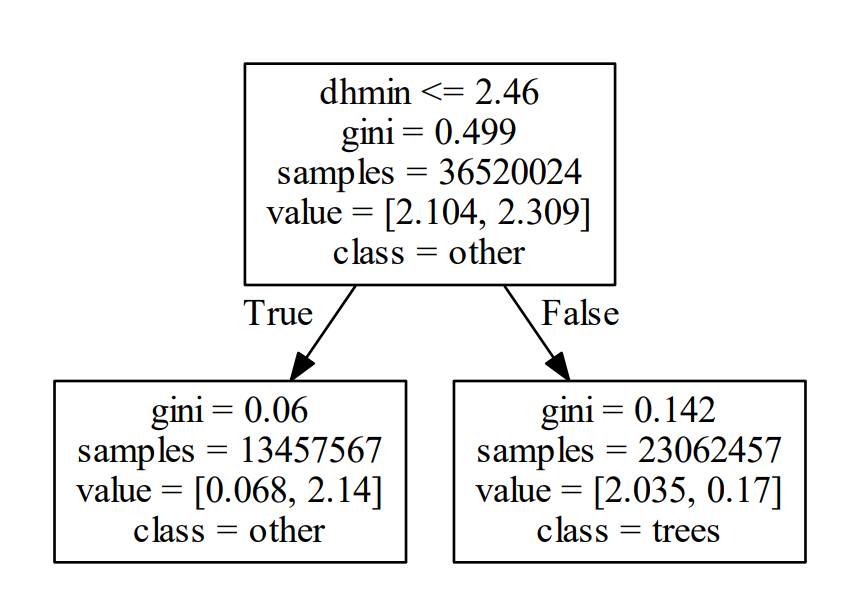
**Figure 1.** The decision tree for the ternary. western Tennessee specific, ArcGIS friendly model.



**Figure 2.** The decision tree for the binary, western Tennessee specific, ArcGIS friendly model.



**Figure 3.** The decision tree for the ternary general model.



**Figure 4.** The decision tree for the binary general model.

## References

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621. doi: 10.1109/tsmc.1973.4309314

Kashani, A., Olsen, M., Parrish, C., & Wilson, N. (2015). A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. *Sensors*, 15(11), 28099–28128. doi: 10.3390/s151128099

Momm, H., Easson, G., & Kuszmaul, J. (2009). Evaluation of the use of spectral and textural information by an evolutionary algorithm for multi-spectral imagery classification. *Computers, Environment and Urban Systems*, 33(6), 463–471. doi: 10.1016/j.compenvurbsys.2009.07.007