A comparison of random forest and kNN for FRI mapping in northern Canada

* Craig MAHONEY¹, Chris HOPKINSON¹, Ron HALL², and Michelle FILIATRAULT²

1. Department of Geography, University of Lethbridge, Lethbridge AB, Canada, T1K 3M4
craig.mahoney@uleth.ca
2. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton AB, Canada, T6H 3S5

* Corresponding Author

ABSTRACT

Large-scale contemporary forest resource inventories (FRI) are of paramount importance in enabling policy makers to pursue informed decisions with respect to sustainable forest management strategies. The consequences of climate change have been reported as most severe and rapid in the northern regions of the globe, thus driving the need for such forest inventories. The current study follows on from previous work that investigated the use of the historically favored k-nearest neighbor (kNN) imputation technique for predicting FRI. The current study follows the exact same methods as previously documented, but substitute’s random forest (RF) imputation in place of kNN to predict stand height and crown closure over a 200,000 km² region in northern Canada. That is, field plot data are utilized to develop models of stand height and crown closure for airborne LiDAR data, which in turn are utilized to infer attributes to spaceborne Geoscience Laser Altimeter system (GLAS) footprints. Quality controlled GLAS data are then submitted to the RF algorithm in order to produce regional predictions of stand height and crown closure. Predictions from both methods are compared with spatially coincident values from an independent ALS dataset, suggesting that RF (R²=0.45, RMSE=4.44 m) better characterizes stand height relative to kNN (R²=0.39, RMSE=4.66 m), whereas crown closure exhibits a slight decrease in accuracy (kNN: R²=0.41, RMSE=7.00 %, RF: R²=0.41, RMSE=7.09 %). A brief overview of methodological pros and cons are presented with the view of evaluating which method is better suited to FRI mapping. The performance of RF highlights its use as a means to quantify FRI in dynamic landscapes over large geographies, however, the current results do not support the replacement of kNN by RF for FRI mapping. Further analysis is require to conclusively indicate if one method should be favoured, and if particular methods are better suited to certain landscape characteristics.