MAPPING ABOVEGROUND FOREST BIOMASS CARBON STOCK BY USING SATELLITE IMAGE AND NFI DATA – A COMPARISON BETWEEN *k*NN AND REGRESSION TREE MODEL

Hieu Cong NGUYEN[1a*], Jaehoon JUNG[2a], Doyeon KIM[3a], Joon HEO[4a], Kyoungmin KIM[5b]

[a]Dept. of Civil and Envi Eng., Yonsei University, Seoul, South Korea;
Tel: + 82-2-2123-2809; Fax: +82-2-2123-8209
[1a*]Ph.D student, E-mail: conghieu.ig@gmail.com
[2a]Ph.D student, E-mail: lionheart_kr@yonsei.ac.kr
[3a]Research fellow, E-mail: edkim@yonsei.ac.kr
[4a]Professor, Email: jheo@yonsei.ac.kr

[b]Dept. of Forest & Climate Change, KFRI, Seoul, South Korea;
Tel: +82-2-961-2862; Fax: +82-2-961-2869
[5]Research fellow, E-mail: greenann@forest.go.kr

KEY WORDS: *k*NN, regression tree, Landsat TM, NFI, carbon mapping.

Abstract: To achieve quantitative information of aboveground forest biomass carbon stock by remote sensing approach, a number of methods have been conducted. This study is to examine application of regression tree (tree decision) and *k*-Nearest Neighbor (*k*NN) algorithm for forest carbon stock estimation of Chungnam province in South Korea. Dataset used for this research includes Landsat Thematic Mapper (TM) images and field data from 5 th National Forest Inventory (NFI). As a result, total above forest carbon stock estimated by *k*NN (20,467,652.900 tonC) model is closer to the number given by Korean Institutes of Forest than regression tree's number (20,239,247.239 tonC), however RMSE from latter algorithm is less than the former, 19.168 tonC/ha and 20.063 tonC/ha, respectively.

INTRODUCTION

Forest is an important part of terrestrial ecosystems, which stores a large part of the total organic carbon. After Kyoto Protocol, measuring forest carbon stock is increasing considerably after by different methods such as field m easurement, geospatial information system and remote sensing applications (Lu 2006; Fuchs, Magdon et al. 2009). Among remote sensing algorithms used to estimate aboveground forest carbon stock, *k*-Nearest Neighbor (*k*NN) and regression models have been widely applied, and many comparisons have been also conducted. However, *k*NN and regression model (especially regression tree model) has not been widely used for forest parameter estimation in South Korea, thus intensive investigation is required for both algorithms.
In this research, we aim to examine *k*NN and regression tree algorithms in estimating aboveground forest carbon stock and making a comparison. Chungnam province in South Korea was selected as a study area. Used data comes from two sources, Landsat Thematic Mapper (TM) image and 5th Korean National Forest Inventory (NFI).

MATERIALS AND METHODS

Study area: Chungnam province is located in the middle west of South Korea (between 35°58'N to 37°7'N and 125°31'E to 127°28'E). It belongs to continental climate with distinct for seasons, and average temperature ranges 11 to 13 ^{o}C. Average height above mean sea level is approximately 100 m so, Chungnam province is considered as flat area. Its total area is 860,096 ha with four types of forestry land: conifer (179,089 ha), deciduous (128,890 ha), mixed forest (111,292) and bamboo (255 ha) respectively.

Data sets: NFI sample plot consists of one subplot at center and three others forming a circle around the center subplot. Information included in each plot is location, diameter at breast height, tree height, species, age, and biomass etc. Aboveground forest biomass carbon stock is obtained from forest biomass using the conversion factor of 0.5. In this study, the 5th NFI data, which was established from 2006 to 2010, was used as field measurements.

To cover the whole Chungnam province, total four Landsat TM images are used. Because of cloud condition over the period that collects field data, images with different acquisition dates were used in the current study: two of them were obtained in July 2010 and others were in May 2009. All images were geometrically calibrated using 1:5,000 numerical map of Chungnam province, then reflectance and topographic corrections were also performed.

In addition, non-forest areas were removed from study site using the forest type map provided from Korea Forest Research Institute.

Data processing: Both estimation models, *k*NN algorithm and regression tree assumes that there is a strong relationship between forest statuses and spectral properties at the same location. It means that both algorithms will estimate carbon stock of target pixel based on the similar of spectral radiation in determined radius(Labrecquea, Fourniera et al. 2006). For the accuracy assessment of both algorithms, Root Mean Square Error (RMSE) was calculated. All processes were conducted using MATLAB and Cubist commercial software.

RESULTS

Results from two methods were compared in terms of basic parameters such as RMSE, min, max and total carbon stock as well as pattern of carbon stock distribution. From table 1, it is shown that minimum, maximum and total carbon stock from *k*NN are higher than those from regression tree, however RMSE from regression tree model is slightly less than the one from *k*NN model, they are 19,168 tonC/ha and 20,063 tonC/ha, respectively. It reveals that regression tree is more suitable for carbon stock estimation of Chungnam province.

Table 1: Comparison between *k*NN and regression tree algorithm.

category	*k*NN	Regression tree
Minimum carbon stock	17.348 tonC/ha	1.000 tonC/ha
Maximum carbon stock	123.220 tonC/ha	110.000 tonC/ha
RMSE	20.063 tonC/ha	19.168 tonC/ha
Total carbon stock	20,467,652.900 tonC	20,239,247.239 tonC

A comparison of distribution of carbon stock form both algorithms was also conducted. Figure 1 denotes the difference map of both results in which difference is compared by two maps of forestry carbon stock by two methods at the same pixel and the difference ranges from -70.264 tonC/ha to 73.220 tonC/ha. From the visual inspection, it can be seen that difference between two results is negligible because the yellow part, which denotes the low difference, is occupying the most areas in the difference map. Additionally, table 2 demonstrate that difference distributes mostly from -10 to 10 tonC/ha, accounts for 85.55% out of total forest area, whereas, variation more than 20 tonC/ha only occupy 0.8% of total forest area.

Table 2: Variance of Carbon stock in the difference map of *k*NN and regression tree algorithms.

Variance (tonC/ha)	Area (ha)	Percentage (%)
-70.264 - -20	1343.97	0.32
-20 - -10	17786.07	4.22
-10 - -2	113843.60	27.03
-2 - 2	96407.73	22.89
2 - 10	150074.10	35.63
10 - 20	39714.48	9.43
20 - 73.22	2006.46	0.48
Total	421176.40	100.00

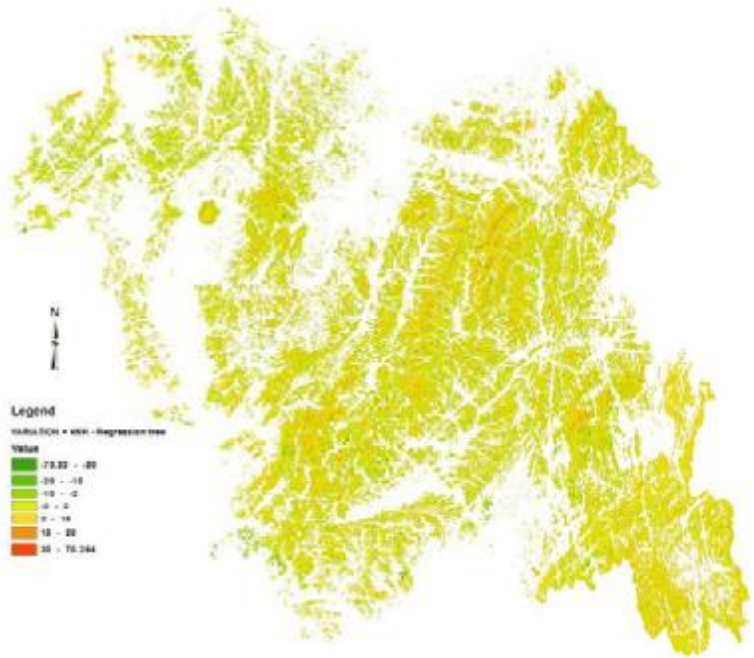

Figure 1: Difference map of forest carbon stock of *k*NN and regression tree model

CONCLUSIONS

To estimate above ground biomass in Chungnam province, *k*NN and regression tree methods a long with NFI data and Landsat TM imagery are sufficient, in which RMSE from regression tree is improved. Results from both are compared to each other in term of map pattern and mass of carbon stock within each pixel at the same location. Some further experiments need to be conducted are: using vegetation indices, energy bands and ratio bands in order to find optimal bands for both algorithms, on the other hand, same approached will be apply in other areas with different natural conditions.

ACKNOWLEDGEMENT

This work was supported by a grant from "Spatiotemporal patterns of forest carbon stocks by urban land use change" (No. 20120811-043-00) project funded by the Korean Forest Research Institute and Brain Korea21, Center for Future Infrastructure System Yonsei University.

REFERENCES:

References from Journals:

Fuchs, H., P. Magdon, et al. (2009). Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory. Remote Sensing of Environment, **113**(3), pp. 518-531.
Labrecquea, S., R. A. Fourniera, et al. (2006). A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland. Forest Ecology and Management, **226**(1-3), pp.129-144.
Lu, D. (2006). The potential and challenge of remote sensing-based biomass estimation. International Journal of Remote Sensing, **27**(7): pp. 1297-1328.