Estimation of forest carbon stock using satellite imagery and NFI data - comparing *k*NN algorithm and regression model

Jaehoon Jung[1], Honggyoo Sohn[2], Miseon Hong[3],

[1,2,3]*GRS Lab., University of Yonsei, Seoul (120-749), South Korea; Tel: + 82-2-2123-2809; Fax: +82-2-2123-8209*
[1]*Ph.D student, E-mail: lionheart_kr@yonsei.ac.kr*
[2] *Professor, E-mail: sohn1@yonsei.ac.kr*
[3]*Master student, E-mail: mshong@yonsei.ac.kr*

KEY WORDS: *k*NN, regression, NFI, Landsat TM, ratio image, carbon stock

ABSTRACT: *k*-Nearest Neighbor (*k*NN) algorithm and regression model have been widely used for a variety of forest parameter estimation and mapping application due to its intuitiveness and ease of use. The objective of this study is to comparing both algorithms for estimation of aboveground carbon stock in Danyang-Gun, South Korea. Field data from 5th NFI and Landsat TM satellite image were used as dataset. Additionally, various ratio images, such as vegetation indices, topographic effect correction indices, and spectral angle indices, were generated and compared to the Landsat TM original bands. As a result, *k*NN algorithm and Landsat TM original bands were determined to be a suitable method and dataset for forest carbon stock estimation in Danyang-Gun, respectively.

1. INTRODUCTION

To monitoring the amount and spatial distribution of aboveground forest biomass carbon stock, satellite remote sensing has become an important tool for forest parameter estimation with finer temporal resolution and its broad coverage. Also, satellite remote sensing can be applicable to estimate forest parameters of uninvestigated area where no field data exists. For the interpretation of remotely sensed data, *k*-Nearest Neighbor (*k*NN) algorithm and regression model have been widely used for a variety of forest parameter estimation and mapping application due to its intuitiveness and ease of use.

In this study, aboveground forest biomass carbon stock of Danyang-Gun, in South Korea, was estimated using both *k*NN algorithm and regression model, and the results were compared. Field data from the 5th National Forest Inventory (NFI) and Landsat TM satellite image were used as a dataset. Additionally, ratio images, such as vegetation indices, topographic effect correction indices, and spectral angle indices were generated from the Landsat original bands and the results were compared each other.

2. MATERIALS AND METHODS

2.1 Study area

Danyang-Gun, 78,067 ha in size, is located in South Korea. Forest stands in the study area covers about 80% of total area (62,745 ha). Forest stands in the western part of study area are dominated by pine tree and pitch pine (51.93%). Deciduous species (31.03%) and mixed forest (12.70%) are found in eastern and southern part, respectively. Its mean annual temperature is about 11.6℃, and mean annual precipitation 1,702mm.

2.2 Data sets

Forest inventory data were available from Korea Forest Research Institute (KFRI). It is 5th NFI data conducted from 2006 to 2009. Measurements of field sampling are location, diameter at breast height, tree species, age, aboveground biomass, etc. (Korea Forest Research Institute, 2009). Final carbon stock was obtained from aboveground biomass using the conversion factor of 0.5. Total 130 plots are included in the study area.
As a remotely sensed data, four seasonal Landsat TM satellite images, acquired on January 20, April 9, August 31 in 2004 and November 22 in 2005 were used. Additionally, various ratio images such as vegetation indices (case 2), topographic effect correction indices (case 3), and spectral angle indices (case 4), which are expected to be related to the better estimation of forest carbon stock, were generated and compared to the Landsat TM original bands (case 1).

2.3 Data processing

*k*NN algorithm and regression model were selected in this study. The set of points for which both the independent and dependent variables are known is referred to as the reference set, whereas the set of points for which only independent variables are known is referred to as the target set (Fuchs et al. 2009). Using the NFI plot as the reference set, the *k*NN algorithm estimates forest carbon stock by calculating a weighted mean value of the reference plots' forest parameter of the k nearest samples in spectral space. On the other hand, regression model was derived from the relation between independent and dependent variables of reference set. Finally, using leave-one-out cross validation, the accuracy was tested on each datasets.

3. RESULTS

In table 1, estimated total forest carbon stock and RMSE of Danyang-Gun are presented. Firstly, we compared RMSE results between *k*NN algorithm and regression model, and it was found that *k*NN algorithm has shown better RMSE (1~2 tonC/ha) than regression model. The total carbon stocks estimated from both methods were also different each other but a specific pattern was not found. In the comparison of dataset, there were no significant difference but case 1, Landsat TM original bands, shows slight better result than other cases. Based on the RMSE results, we found the best estimate from the dataset of 08/31 in case 2, kNN algorithm. Consequently, the total forest carbon stock in Danyang-Gun was determined to be about 3,310,000 tonC with RMSE of 23.97 tonC/ha.

Table 1. Total carbon stock and RMSE of Danyang Gun

Method	Date	Case 1		Case 2		Case 3		Case 4	
		Total Carbon	RMSE	Total Carbon	RMSE	Total Carbon	RMSE	Total Carbon	RMSE
*k*NN	04/09	3,452	25.58	3,516	25.95	3,455	27.11	3,472	26.80
	08/31	3,517	25.40	3,310	23.97	3,308	25.26	3,352	26.28
	11/22	3,190	24.87	3,204	26.10	3,140	25.91	3,183	25.96
	01/20	3,205	25.94	3,272	26.56	3,166	25.92	3,245	26.16
Regression	04/09	3,361	27.36	3,344	26.57	3,377	27.50	3,338	27.32
	08/31	3,353	26.02	3,382	26.44	3,351	26.06	3,372	25.98
	11/22	3,559	27.99	3,193	27.69	3,303	27.60	3,334	26.51
	01/20	3,480	27.28	2,999	29.49	3,236	27.61	3,266	26.60

*Total Carbon Unit : thousand tonC, RMSE Unit: tonC/ha

4. CONCLUSIONS

In this study, aboveground forest carbon stock of Danyang-Gun was estimated using *k*NN algorithm and regression model. Field data from 5th NFI and Landsat TM satellite image were used as datasets. Additionally, various ratio images were generated from the Landsat original bands, and results were compared each other. Consequently, *k*NN algorithm and Landsat TM original bands are determined to be a suitable method and dataset to estimate the forest carbon stock in Danyang-Gun. Further study will focus on the test in various study areas with different forest conditions.

ACKNOWLEDGEMENT

This work was supported by a grant from Evaluation of CO2 storage capacity on geological stratum and search for storage site by GIS method (No. 16-2008-04-007-00) funded by the Korean Ministry of Education, Science and Technology.

REFERENCES

Fuchs, H., Magdon, P., Kleinn, C., and Flessa, H., 2009, Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory, Remote Sensing of Environment, 113, pp. 518–531.

Korea Forest Research Institute, 2009, The fifth national forest inventory surveying: field surveying guide book ver. 1.3, Korea Forest Service, pp. 37.