ESTIMATING ABOVE-GROUND BIOMASS OF CORN BY COMBINING SATELLITE IMAGERY AND FIELD INVENTORY

Nguyen Cong Hieu[1], Jeahoon Jung[2], Jeonghyun Kim[3], Shuhe Zhao[4], Joon Heo[5]

[1]Dept. of Civil and Environmental Engineering, Yonsei University, Seoul, South Korea, *hieunguyen@yonsei.ac.kr*
[2]Dept. of Civil and Environmental Engineering, Yonsei University, Seoul, South Korea, *lionheart_kr@yonsei.ac.kr*
[3]Dept. of Civil and Environmental Engineering, Yonsei University, Seoul, South Korea, *jhkim2014@yonsei.ac.kr*
[4]Geographical Information Science, Nanjing University, China, *zhaosh@nju.edu.cn*
[5]Dept. of Civil and Environmental Engineering, Yonsei University, Seoul, South Korea, *jheo@yonsei.ac.kr*

Corresponding to Dr. Shuhe Zhao via zhaosh@nju.edu.cn and Dr. Joon Heo via jheo@yonsei.ac.kr.

Abstract: Recently, crop growth monitoring has been become urgently needed through the growing season in order to estimate crop yields. Above-ground biomass is one of important crop parameters that can be monitored and assessed by remote sensing technology. In this paper, we aim to estimate above-ground biomass of corn using a non parametric method, *k*-nearest Neighbors algorithm (*k*NN) which to extract information about spatial distribution and total carbon stock. A case study area was selected in Shandong Province of China. The satellite image HJ-1 data and Corn biomass were acquired by the synchronous date. Fifteen band ratios were produced and Pearson correlation between spectral properties and corresponding ground data was tested. Based on the analysis, five optimal band ratios were chosen as independent variables. Also, number of k in *k*NN algorithm was tested; it is from 5 to 15. As a result, the best RMSE is estimated at 118.213 g/m2 when *k* is 15. Though the number of samples is limited, a model combining HJ-1 and *k*NN algorithm could be used to create map of above-ground corn biomass.

Keyword: corn biomass, *k*NN, HJ-1, mapping.

Note: this paper for poster section.

I. Introduction

Crop biochemical and biophysical variables are very important in precision agriculture, which improve the ability to accurately monitor crop-growth status and forecast final crop yield. Crop growth status such as above ground biomass can be quantified spatially and timely by using satellite imagery based method and field survey (Liu et al., 2010). However, for agricultural purposes, spatial and temporal resolutions of satellite systems are of concern, because a crop season is normally from 3 months to 6 months and size of crop field in some countries such as China, Japan and Korea is small compared to North America's. Among optical satellite systems, HJ-1 system presents some advantages, it's resolution is 30m and revisit is 4 days for single HJ-1 satellite and 2 days for HJ-1A and HJ-1B. Furthermore, a high correlation between band ratios achieved from HJ-1 and response ground biomass was verified (Gao, Niu, Huang, & Hou, 2013). In addition, *k*-Nearest Neighbours (*k*NN) algorithm is widely used to estimate above biomass over the world (Gjertsen, 2007; Magnussen, McRoberts, & Tomppo, 2009; Soenen, Peddle, Hall, Coburn, & Hall, 2010; Tomppo, Gagliano, Natale, Katila, & McRobert, 2009; Zheng et al., 2004). The objective of this paper is to aim to assess feasibility of *k*NN to estimate corn above-ground biomass by using HJ-1 imagery on a study site in Shangdong province, China.

II. Materials and methods

Study area: Yucheng commune, Shangdong province, Eastern of China was chosen as studied site, its location is between $36^o32'39''$N – $37^o4'57''$N and $116^o27'16''$E – $117^o5'2''$E. Yucheng belongs to continental climate with distinct seasons even Shangdong has a long coastal line. Studied site is flat and is an agricultural area with full of corn at the time of study. The corn crop covered 138817.75 ha (62% of total area).

Data sets: Ground truth data of corn above-ground biomass was collected by field survey from 24th August to 8th September 2011, when corn tree was full developed. Field data was performed by Remote Sensing group, Department of Geographical Information Science, Nanjing University, China. There are eighteen samples of corn, in each sample there are 9 subsamples distributing within a rectangular. Corn fresh biomass was collected and dried for all 9 subsamples, and then it was averaged and assigned to the central subsample.

One HJ-1 scene with cloud free was acquired in 24th August 2011. Several image pre-processing steps were conducted: 1) geo-rectify and projecting to UTM-WGS 84 as same as field work data; 2) geometry correction using RST model and Nearest Neighbours in ENVI software; and 3) atmospheric correction using FLAASH model in ENVI. Fifteen band ratios including four HJ-1 raw bands were then created.

Data processing: Basically, *k*NN algorithm coding by MATLAB program language is to estimate corn above-ground biomass for Yucheng. There is assumption about strong relationship between spectral properties and responding corn above-ground biomass. It means that the algorithm will estimate biomass of target pixel based on the similar of spectral radiation in determined radius (Labrecquea, Fourniera, Lutherb, & Piercey, 2006). In term of accuracy assessment, ten fold validations were conducted and Root Mean Square Error (RMSE) and Relative Root Mean Square Error (%RMSE) were calculated. In this study, two cases based on band ratios were of concern: 1) all band ratios and 2) five band ratios with the highest correlation (Pearson's correlation).

III. Results and conclusions

As a result from Pearson's Correlation analysis, the five band ratios which are the highest correlation are TVI, band 2 (green), MSAVI2, DVI and SAVI (table 1), these five band ratios are used as independent variables in *k*NN model. In both cases, the number of *k* is from 5 to 15 and radius parameter is 50km to cover the whole studied area. From table 2, it is shown that at the same number of k, all band ratio's has better accuracy than 5 band ratios, and the more number of k, the higher accuracy will be. As a consequence, the lowest RMSE and relative RMSE were achieved when k is 15 in all cases, they are 118.214 g/m^2, 14.435% (all band ratios); 127.958 g/m^2 and 15.610%, respectively. Thus, it can conclude that *k*NN with all band ratios present better accuracy than selected one's.

Table 1. List of variables

Variable set		Correlation *r*	Correlation *p*
Raw Data	Band 1	0.165	0.512
	Band 2	-0.274	0.270
	Band 3	0.011	0.965
	Band 4	-0.178	0.479
Spectral vegetation indices	DVI	-0.189	0.454
	SR	0.147	0.560
	NDVI	-0.148	0.558
	GEMI	-0.116	0.645
	SAVI	-0.189	0.451
	OSAVI	-0.181	0.473
	SAVI2	-0.166	0.510
	MSAVI2	-0.190	0.451
	EVI	-0.126	0.619
	PVI	-0.159	0.529
	TVI	-0.299	0.227

Table 2. Results of *k*NN accuracy assessment

k	All band ratios				5 band ratios			
	RMSE	%RMSE	BIAS	%Bias	RMSE	%RMSE	BIAS	%Bias
5	139.218	16.947	-15.374	-1.808	144.578	17.590	-23.810	-2.838
6	131.452	16.018	-7.601	-0.876	139.927	17.041	-18.339	-2.183
7	127.162	15.500	-6.177	-0.708	136.356	16.611	-16.539	-1.970
8	125.735	15.337	-2.908	-0.306	135.013	16.457	-14.228	-1.685
9	123.895	15.118	-2.867	-0.299	132.967	16.213	-14.442	-1.711
10	122.995	15.009	-4.746	-0.518	131.816	16.072	-15.995	-1.891
11	120.506	14.704	-4.580	-0.502	130.215	15.876	-15.510	-1.835
12	120.496	14.706	-5.659	-0.626	130.097	15.865	-16.400	-1.938
13	119.905	14.637	-6.513	-0.730	129.490	15.793	-17.079	-2.021
14	119.175	14.549	-7.011	-0.791	128.799	15.709	-17.395	-2.060
15	118.213	14.435	-4.445	-0.480	127.958	15.610	-15.129	-1.786

From the accuracy assessment results, the case which has the best result was chosen to create a map of above-ground corn biomass, it is 15 nearest neighbours and all band ratios. The map presents spatial distribution of corn biomass over Yucheng area (Figure 1). Total dry biomass is 114,933.2487kg, maximum is 1057.790 g/m^2 and minimum is 617.262 g/m^2. This results could be used to improve corn crop yield as well as corn field management.

Figure 1: Above-ground corn dry biomass map.

ACKNOWLEDGEMENT

This work was supported by a grant from “Integration and implementation of Satellite image processing and biophysical parameter estimation algorithm” (No PJ009978) project funded by the National Academy of Agricultural Science, Rural Development Administration.

References

Gao, S., Niu, Z., Huang, N., & Hou, X. H. (2013). Estimating the Leaf Area Index, height and biomass of maize using HJ-1 and RADARSAT-2. International Journal of Applied Earth Observation and Geoinformation, 24, 1-8. doi: DOI 10.1016/j.jag.2013.02.002

Gjertsen, Arnt Kristian. (2007). Accuracy of forest mapping based on Landsat TM data and a kNN-based method. Remote Sensing of Environment, 110(4).

Labrecquea, S., Fourniera, R.A., Lutherb, J.E., & Piercey, D. (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), 129-144.

Liu, J. G., Pattey, E., Miller, J. R., McNairn, H., Smith, A., & Hu, B. X. (2010). Estimating crop stresses, aboveground dry biomass and yield of corn using multi-temporal optical data combined with a radiation use efficiency model. Remote Sensing of Environment, 114(6), 1167-1177. doi: DOI 10.1016/j.rse.2010.01.004

Magnussen, Steen, McRoberts, Ronald E, & Tomppo, Erkki O. (2009). Model-based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. Remote Sensing of Environment, 113(476-488.).

Soenen, Scott A., Peddle, Derek R., Hall, Ronald J., Coburn, Craig A., & Hall, Forrest G. (2010). Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain. Remote Sensing of Environment, 114(7), 1325–1337.

Tomppo, Erkki O., Gagliano, Caterina, Natale, Flora De, Katila, Matti, & McRobert, Ronald E. (2009). Predicting categorical forest variables using an improved k-Nearest Neighbour estimator and Landsat imagery. Remote Sensing of Environment, 113(3), 500–517.

Zheng, Daolan, Rademacher, John, Crow, Jiquan ChenaThomas, Bresee, Mary, Moine, James Le, & Ryu, Soung-Ryoul. (2004). Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. Remote Sensing of Environment, 94(3), 402–411.