Korean Journal of Remote Sensing, Vol.26, No.6, 2010, pp.665~673

Vegetation Classification Using Seasonal Variation MODIS Data

Hyun-Ah Choi*,, Woo-Kyun Lee** †, Yowhan Son**, Toshiharu Kojima***, and Hiroyuki Muraoka*****

*Korea Adaptation Center for Climate Change, Korea Environment Institute, 290 Jinheungno, Eunpyeong-Gu, Seoul, 122-706, Korea

**Department of Environmental Science and Ecological Engineering, Korea University, Seoul, 136-713, Korea

***Institute for Basin Ecosystem Studies, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan

Abstract : The role of remote sensing in phenological studies is increasingly regarded as a key in understanding large area seasonal phenomena. This paper describes the application of Moderate Resolution Imaging Spectroradiometer (MODIS) time series data for vegetation classification using seasonal variation patterns. The vegetation seasonal variation phase of Seoul and provinces in Korea was inferred using 8 day composite MODIS NDVI (Normalized Difference Vegetation Index) dataset of 2006. The seasonal vegetation classification approach is performed with reclassification of 4 categories as urban, crop land, broad-leaf and needle-leaf forest area. The BISE (Best Index Slope Extraction) filtering algorithm was applied for a smoothing processing of MODIS NDVI time series data and fuzzy classification method was used for vegetation classification. The overall accuracy of classification was 77.5% and the kappa coefficient was 0.61%, thus suggesting overall high classification accuracy.

Key Words : Vegetation classification, seasonal variation, MODIS, Best Index Slope Extraction (BISE), fuzzy classification.

1. Introduction

Vegetations are major part of land coverage, and their changes have important influences on energy and mass biochemical cycles. Moreover, they are a key indicator of regional ecological environment change (Sun *et al*., 1998). It is of great significance in acquiring surface vegetation cover and its variation to analyze the effect of human activities on ecological environment and the feedback of nature (Townshend, 1994). In addition, Normalized Difference Vegetation Index (NDVI) is the most widely used vegetation index due to their simplicity, ease of application, and wide-spread familiarity (Gu *et al*., 2008). Time-series Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI products provide consistent, spatial and temporal comparisons of global vegetation conditions, which are used to monitor the Earth's

Received November 24, 2010; Revised December 21, 2010; Accepted December 26, 2010.

† Corresponding Author: Woo-Kyun Lee (leewk@korea.ac.kr)

terrestrial photosynthetic vegetation activity in support of phenologic, change detection, and biophysical interpretations (Van Leeuwen *et al*., 1999).

NDVI time series, however, suffer from numerous limitations: the calculated NDVI is not only a function of vegetation density and type, but it is also influenced by the atmosphere and illumination as well as observation geometry, which result in noisy NDVI time series (Holben, 1986; Gutman, 1991). In order to extract meaningful information on vegetation dynamics regardless of these distortions, various methods for the elimination of spurious data were developed, such as the Maximum Value Composite (MVC; Holben, 1986), Best Index Slope Extraction (BISE; Viovy *et al*., 1992), Fourier adjustment (Sellers *et al*., 1996), Savitzky-Golay filter (Chen *et al*., 2004), and asymmetric Gaussian function fitting (Jönsson *et al*., 2002). All the methods aim at approaching an upper NDVI envelope, based on the assumption that NDVI values are depressed by any of the above-mentioned effects (Holben, 1986). However, smoothing algorithms yet hold the danger of introducing artifacts and suppressing natural variations in the NDVI time series (Fisher *et al*., 2007).

Wide area vegetation map is important for the integrated eco-meteorological model. However, it is difficult to get a clear wide-area classification map due to low accuracy of existing vegetation maps, such as U.S. Geological Survey (USGS), and quality variation of traditional remote sensing image classification methods with difference of acquired date, photo-interpreter and difficulties of image acquisition all at once (Kojima *et al*., 2008). Although vegetative time variation is also affected by acquisition aspect, such as sun elevation, atmospheric, and seasonal vegetative variation, its influence is less severe than the influence of spectral data. Thus, we examined the vegetation classification using phenology data by MODIS NDVI instead of multi-channel data.

2. Material and methods

1) Study site description and reference data

This study was carried out in Seoul and provinces around Seoul in Korea (Fig. 1). The climate of the selected region is humid continental with warm, humid summers and cool, dry winters. The average annual precipitation of 1108mm is mainly distributed in summer. The average temperatures in January and July are -3.0 and 25.5℃, respectively.

The reference data for classes were established from 2001 Ministry of Environment land cover map data. The land cover classification scheme includes

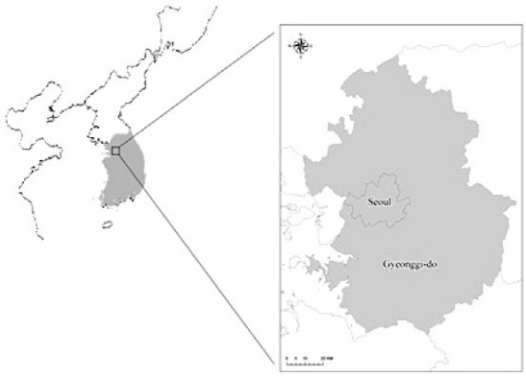


Fig. 1. Study area: Gyeonggi province and Seoul.

Table 1. Land cover classification scheme

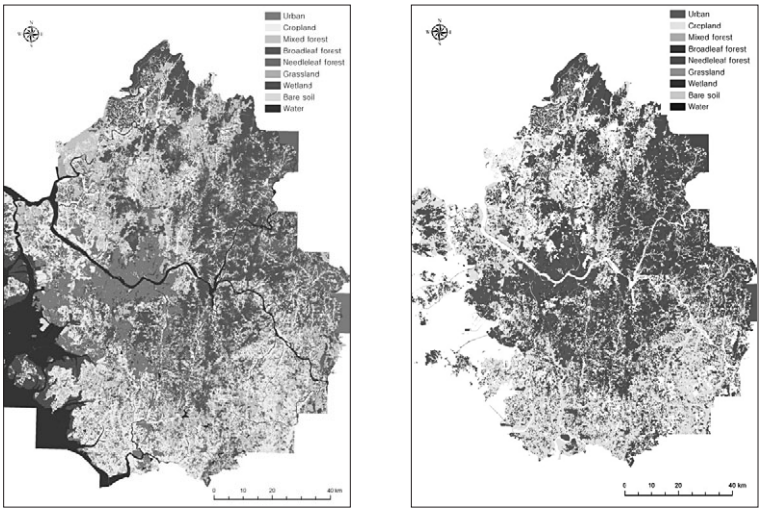
Class	Description
Urban	Commercial, industrial, residential and traffic area
Cropland	Paddy , crop field, orchard
Forest area	Needle-leaf, broad-leaf, mixed forest types
Grass land	Natural grassland, golf courses, lawns
Wetland	Wetland types
Bare land	Bare areas, not cropped in summer or fall
Water	Permanent open water; lakes and rivers

the categories listed in Table 1.

2) Methods

The image series used in this study consisted of 8 day composite 250m resolution Terra/MODIS data acquired during January to December of 2006 and this area covered 1 MODIS title (h28v5). Annual NDVI data were calculated using these data. The seasonal classification approach was performed with the following steps: i) We reclassified 7 categories of land cover data, which is the national actual vegetation map by Ministry of Environment, to 4 categories (Fig. 2). ii) Next, we extracted annual NDVI data from highly occupancy ratio pixels of each category and investigated the phonological properties of the categories. iii) Finally, we classified using a fuzzy classification method.

Changes in the NDVI derived from satellite image data are usually indicative if changes in the surface conditions, most predominantly changes in vegetation. However, there are other extrinsic factors that cause changes in the overall NDVI profile, among which are cloud contamination, atmospheric variability and bi-directional effects (Gutman 1991); these changes are usually considered as undesirable noise in vegetation studies. A widely used technique to reduce this noise is the Best Index Slope Extraction (BISE) (Viovy *et al*, 1992). Prior to the application of the algorithm, missing data were set to NDVI=0. The BISE algorithm introduced by Viovy *et al*. (1992) approximates an upper NDVI envelope by excluding low values and accepting high values within a predefined time period (here: 30 days), based on user defined thresholds. A data point is only accepted, if it is greater than the previous value. Decreases from one day to the next are only accepted, if there is no regrowth within the predefined period greater than 20 percent of the difference between the preceding high and the decreased value. This threshold is applied because it is assumed that changes in vegetation remain visible for a number of days, during which regrowth is slow (Jose *et al*., 2002). Viovy *et al*. (1992) excluded all points with an increase greater than 0.1 from one point to the next to eliminate unrealistic fluctuations.



(a) (b)

Fig. 2. Land cover map prepared by the Ministry of Environment: (a) National actual land cover map, (b) Land cover map with 4 categories.

Because vegetation often occurs as a mosaic, it is often difficult or erroneous to classify the vegetation or land cover of an areal unit into only one class (Gopal and Woodcock, 1994). For this reason, we also used the accuracy assessment method based on fuzzy classification. Fuzzy classification recognizes uncertainty in the mapping process and allows for an areal unit to correctly fall into more than one category. The benefit of using fuzzy set theory is that it provides additional tools for analyzing map error (Muller *et al*., 1998). A fuzzy set A is characterized by a membership function (MF), $\mu_A(x)$ that assigns to each element x a grade of membership ranging from zero to one. X is a space of points (objects), with a generic element of X denoted by x. Thus, $X = \{x\}$.

$$A = \{x, \mu_A(x)\} \text{ for each } x \in X \; \mu_A(x) \rightarrow [0, 1] \qquad (1)$$

where, $\mu_A(x) = 0$ means that x does not belong to the subset A, $\mu_A(x) = 1$ indicates that x fully belongs, and $0 < \mu_A(x) < 1$ means that x belongs to the degree μ_A. Two different but complementary approaches exist for deriving MFs: the similarity relation model and the semantic import model. The similarity relation model is based on a cluster analysis such as fuzzy c-means or on fuzzy neural networks. The semantic import model an expert or an empirical model is applied to specify a formula for the membership function without reference to data (Zhao *et al*., 2008).

In this study the use of fuzzy classification method is to map land cover and to better represent the vegetation landscape.

3) Accuracy Assessment

In general, an accuracy assessment of a map classification is performed by comparing it to a more detailed, independently sampled data set (Muller *et al*., 1998). A considerable amount of research has focused on various methods of assessing the accuracy of maps derived from remotely sensed data; most of these are designed for and applicable to categorical data. Hord and Brooner (1976), van Genderen and Lock (1977), and Hay (1979) presented and discussed methods for determining appropriate sample size. Others presented discussions, research, and empirical experiments on various sampling designs (Congalton, 1988; Gong and Howarth, 1990; Stehman, 1992). Story and Congalton (1986) discussed the use of error matrices for the analysis of reference data, including calculation of descriptive statistics such as producer's and user's accuracy. Congalton *et al*. (1983), Rosenfield and Fitzpatrick-Lins (1986), Hudson and Ramm (1987), and Foody (1992) discussed various uses of the Kappa coefficient of agreement, which is derived from an error matrix. In this study, we used producer's accuracy, user's accuracy, and Kappa value for comparative accuracy.

Kappa value is a chance-corrected measure of agreement between two raters, each of whom independently classifies each of a sample of subjects into one of a set of mutually exclusive and exhaustive categories (King, 2004). It is computed as the following:

$$K = \frac{P_o - P_e}{1 - P_e} \qquad (2)$$

where p is the proportion of ratings by two raters on a scale having k categories, P_o is the proportion of observed agreement and P_e the proportion of expected agreement in the hypothesis of independence between observers. P_o and P_e are calculated according to Eq. 3.

$$P_o = \sum_{i=1}^{k} P_{ii} \text{ and } P_e = \sum_{i=1}^{k} P_{iT} P_{Ti} \qquad (3)$$

where P_{ii} is the number of objects placed in category. The quantity P_{iT} is the proportion of cells and P_{Ti} is the proportion of cells of category.

3. Results and discussion

The vegetation classification was assessed with the land cover map generated from the NDVI time series using the BISE algorithm and the fuzzy classification. In order to scale the computed NDVI results to byte data range the NDVI data range of 0 to 1 is scaled to the range of 0 to 2000, where computed 0 equals 0, computed 1 equals 2000. Fig. 3 is the vegetation type classification with the proposed fuzzy classification.

After labeling the 4 classes, we verified the accuracy of urban area, crop land and forest area and sample points are selected randomly. The results of image classification can be summarized as Fig. 4.

The vegetation classification revealed that urban area, crop land and forest area were clearly changeable. Urban class always indicates low NDVI (Fig. 4a) and NDVI of crop land highly arise during

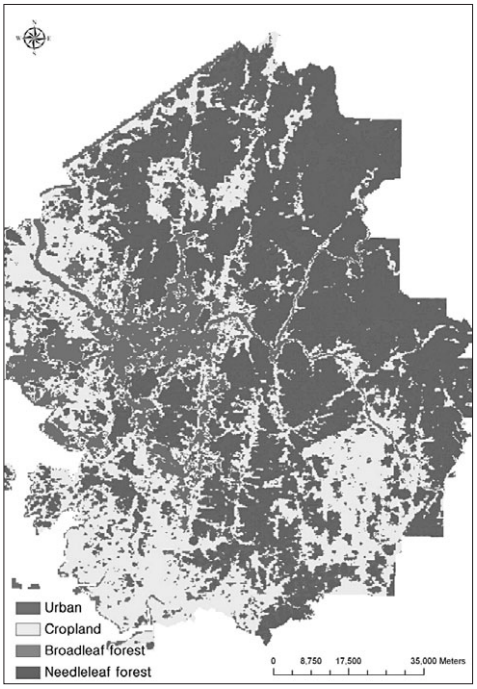


Fig. 3. Vegetation type classification with phenological classification method.

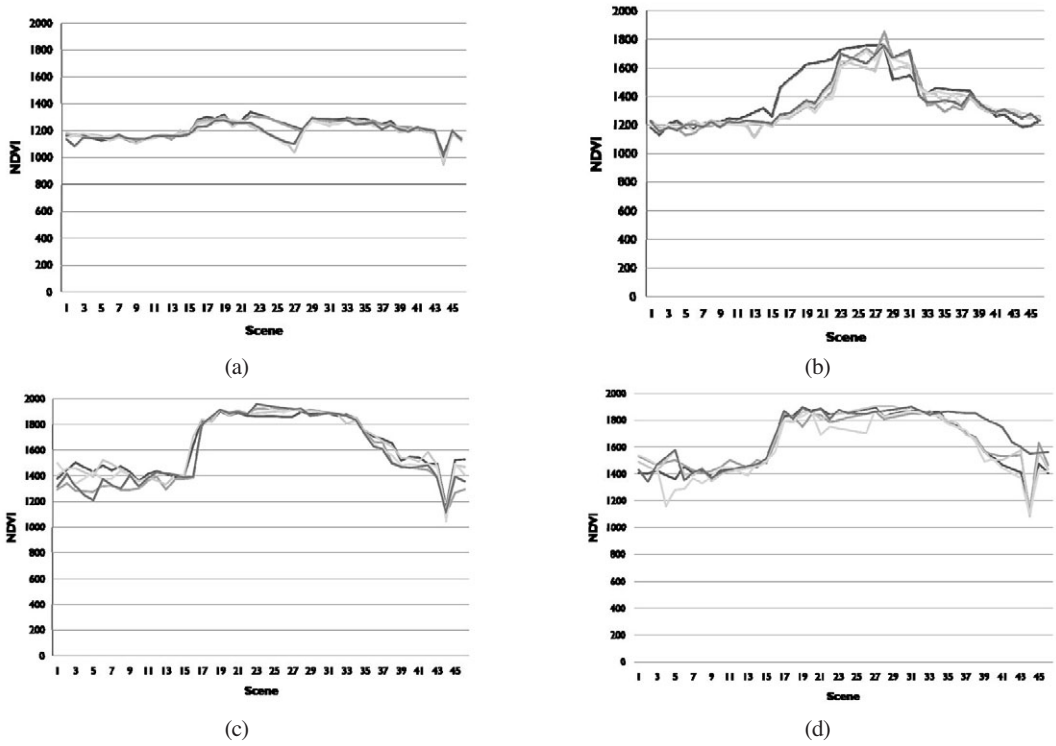


Fig. 4. Seasonal variation of each category from sample points: (a) urban, (b) crop land, (c) broad-leaf and (d) needle-leaf.

summer season (Fig. 4b). However, it is difficult to divide between deciduous broad-leaf and deciduous coniferous, also between ever green broad-leaf and ever green coniferous. Thus, two categories such as deciduous tree and ever green tree are combined. Broad-leaf forest area clearly shows seasonal differences (Fig. 4c). Needle-leaf forest area shows similar pattern with that of broad-leaf forest, but still display high value during winter seasons (Fig. 4d).

The kappa coefficient was 0.61% and the overall accuracy of classification was 77.5%, thus suggesting overall high classification accuracy (Table 2). Producer's accuracy of classification in each class was generally over 80%, whereas low classification accuracy was obtained for classes of needle-leaf forest and crop. Moreover, user's accuracy of classification in each class was generally around 90% except for broad-leaf forest areas.

Previous studies show that Kim *et al*. (2007) investigated the optimal land cover classification algorithm for the monitoring of North Korea with MODIS multi-temporal data based on monthly phenological characteristics. The overall classification accuracy of ISODATA (Isodata stands for Iterative Self-Organizing Data Analysis Techniques), SMA (Spectral Mixture Analysis), and SOM (The self-organizing map) was 69.03%, 64.28%, and 73.57%, respectively. In addition, Jang and Kim (2003) analyzed high-resolution IKONOS image and classify land cover map using fuzzy classification method with minimum operator used as a tool for joint membership functions. It showed that the kappa value was 0.94%, the overall accuracy of classification was 95.0%, and accuracy of classification in each class was generally over 90%.

It is difficult to get a good wide area classification map related with ground truth vegetation map and low accuracy of existing vegetation map. Thus, this study proposed the new vegetation type classification method based on BISE filtering. As yet, no studies have applied BISE filtering for reducing noise in Korea. It is seemed that the new method has a possibility to produce the equivalent quality classification and accurate performance for wide area vegetation types. And more accurate results would be expected if this method is applied in future studies.

4. Conclusion

With high temporal resolution and spectral resolution, MODIS data may supply high-quality data of vegetation index, NDVI, surface temperature, and provide conditions for analyzing the regional vegetation classification mapping (Lu *et al*., 2008). In this study, the vegetation classification mapping in Seoul and provinces around Seoul in Korea were analyzed using MODIS NDVI seasonal data. The

Table 2. Summary statistics for 4 class land cover classification

Classified data	Reference (m^2)					User's Accuracy (%)
	Broad-leaf	Needle-leaf	Crop	Urban	Sum	
Broad-leaf	1657	1140	17	8	2822	59
Needle-leaf	309	3200	47	0	3556	90
Crop	9	23	292	0	324	90
Urban	0	0	59	399	458	87
Sum	1975	4363	415	407	7160	
Producer's Accuracy (%)	84	73	70	98		

Overall classification accuracy=77.5 %. Kappa value=0.61

BISE filtering algorithm was applied for a smoothing processing of MODIS NDVI time series of 2006. The curves of different classes were characterized by an obvious differentiation that created a better result for classifying vegetation. The NDVI time series curves of broad-leaf forest and crop land appear peak during the plant growing season. However, urban areas always indicate low NDVI due to industrial, residential and traffic area. The accuracy test illustrates that the overall accuracy of classification was 77.5% and the kappa coefficient was 0.61%, thus suggesting overall high classification accuracy.

Although this study specifically analyzed vegetation classification with MODIS data using fuzzy classification, it would be possible to apply the proposed method on similar data for vegetation classification. In addition, this research results may better describe information on vegetation coverage and be helpful for studying ecology, climate and environment.

Acknowledgment

This study was supported by the National Research Foundation of Korea (A3 Foresight Program: Grant No. A307-K001) and partly supported by Japanese Association of University Women (JAUW) International Fellowship program.

Reference

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., and Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter, Remote Sensing of Environment, 91: 332-344.

Congalton, R. G., 1988. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data, Photogrammetric Engineering and Remote Sensing, 54(5): 593-600.

Congalton, R. G., Oderwald, R. G., and Mead, R. A., 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques, Photogrammetric Engineering & Remote Sensing, 49(12): 1671-1678.

Fisher, J. I. and Mustard, J. F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data, Remote Sensing of Environment, 109: 261-273.

Foody, G. M., 1992. On the compensation for chance agreement in image classification accuracy assessment, Photogrammetric Engineering & Remote Sensing, 58(10): 1459-1460.

Gong, P. and Howarth, P. J., 1990. An assessment of some factors influencing multi-spectral land-cover classification, Photogrammetric Engineering & Remote Sensing, 56(5): 597-603.

Gopal, S. and Woodcock, C., 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets, Photogrammetric Engineering and Remote Sensing, 60(2): 181-188.

Gu, J., Li, X., and Huang, C., 2008. Spatio-Temporal Reconstruction of MODIS NDVI Data Sets Based on Data Assimilation Methods, Proc. of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Shanghai, P. R. China, Jun. 25-27, 2008, 242-247.

Gutman, G. G., 1991. Vegetation indices from AVHRR: An update and future prospects,

Remote Sensing of Environment, 35: 121-136.

Hay, A. M., 1979. Sampling designs to test land-use map accuracy, Photogrammetric Engineering b Remote Sensing, 45(4): 529-533.

Holben, B. N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data, International Journal of Remote Sensing, 7: 1417-1434.

Hord, R. M. and Brooner, W., 1976. Land-use map accuracy criteria, Photogrammetric Engineering b Remote Sensing, 42(5): 671-677.

Hudson, W. and Ramm, C., 1987. Correct formulation of the Kappa coefficient of agreement, Photogrammetric Engineering & Remote Sensing, 53(4): 421-422.

Jang, D. H. and Kim, M. K., 2003. Improving of land-cover map using IKONOS image data, The Journal of GIS Association of Korea, 11(2): 101-117.

Jönsson, P. and Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data, IEEE Transactions on Geoscience and Remote Sensing, 40: 1824-1832.

Jose, E., Ryutaro, T., and Renchin, T., 2002. The polynomial least squares operation (PoLeS): a method for reducing noise in NDVI time series data, Proc. of the 23rd Asian Conference on Remote Sensing, Section of Data Processing, Algorithm and Modelling, Kathmandu, Nepal, November.

Kim, D. H., Jeong, S. G., and Park, C. H., 2007. Comparison of Three Land Cover Classification Algorithms - ISODATA, SMA, and SOM - for the Monitoring of North Korea with MODIS Multi-temporal Data, Korean Journal of Remote Sensing, 23(3): 181-188.

King, J., 2004. Software solutions for obtaining a kappa-type statistic for use with multiple raters. Southwest Educational Research Association.

Kojima, T., Nagai, S., Ishihara, M., and Akiyama, T., 2008. Vegetation type classification with MODIS phenology data, SATECO 2008 Symposium, Gifu, Japan, July. 30 - August. 1, 2008.

Lu, Y., Yan, Y., and Tao, H., 2008. Classification of vegetation in North Tibet Plateau based on MODIS time-series data, Wuhan University Journal of Natural Sciences 13(3): 273-278.

Muller, S. V., Walker, D. A., Nelson, F. E., Auerback, N. A., Bockheim, J. G., Guyer, S., and Sherba, D., 1998. Accuracy assessment of a land-cover map of the Kuparuk river basin, Alaska: considerations for remote regions, Photogrammetric Engineering and Remote Sensing, 64(6): 619-628.

Rosenfield, G. H. and Fitzpatrick-Lins, K., 1986. A coefficient of agreement as a measure of thematic classification accuracy, Photogrammetric Engineering & Remote Sensing, 52(2): 223-227.

Sellers, P., Los, S. O., Tucker, C. J., Justice, C. O., Dazlich, D. A., Collatz, G. J., and Randall, D. A., 1996. A revised land surface parametrization (SiB2) for atmospheric GCMs. Part II. The generation of global fields of terrestrial biophysical parameters from satellite data, Journal of Climate, 9: 706-737.

Stehman, S. V., 1992. Comparison of systematic and random sampling for estimating the accuracy of maps generated from remotely sensed data, Photogrammetric Engineering & Remote Sensing, 58(9): 1343-1350.

Story, M. and Congalton, R. G., 1986. Accuracy assessment: A user's perspective, Photogrammetric Engineering & Remote

Sensing, 52: 397-399.

Sun, H., Wang, C., Niu, Z., and Bukhosor, L. B., 1998. Analysis of the Vegetation Cover Change and the Relationship between NDVI and Environmental Factors by Using NOAA Time-Series Data, Journal of Remote Sensing - BEIJING, 2(3): 210-216.

Townshend, J. R. G., 1994. Global Data Sets for Land Applications from the Advance very High Resolution Radiometer: An Introduction, Int J Remote Sensing, 15: 3319-3332.

van Genderen, J. L. and Lock, B. F., 1977. Testing land-use map accuracy, Photogrammetric Engineering & Remote Sensing, 43(9): 1135-1137.

Van Leeuwen, W. J. D., Huete, A. R., and Laing, T. W., 1999. MODIS vegetation index compositing approach: Aprototype with AVHRR data, Rem. Sens. Environ., 69: 264-280.

Viovy, N., Arino, O., and Belward, A. S., 1992. The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time series, International Journal of Remote Sensing, 13:1585-1590.

Zhao, X., Chen, X., and Stein, A., 2008. Integrating multi-source information via fuzzy classification method for wetland grass mapping. Proc. of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing.