



Full Length Article

Vegetation Cover Estimation Based on in-suit Hyperspectral Data: A Case Study for Meadow Steppe Vegetation in Inner Mongolia, China

Xiao-Bing Li¹, Hui-Ling Long^{2*} and Hong Wang¹

¹State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Resources Science and Technology, Beijing Normal University, Beijing, 100875, China

² Beijing Research Center for Information Technology in Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing, 100097, China

*For correspondence: hllong53@gmail.com

Abstract

Changes of key parameters of vegetation are essential indicators of ecosystem and global change. Hyperspectral data, as a powerful tool to estimate vegetation parameters, needs to be used more efficiently and effectively, especially in the aspect of massive information extraction. The objectives of the present study were to provide guidance on how to select the optimal subset of hyperspectral data to improve the accuracy of estimating vegetation cover using hyperspectral data measured in the field, and to compare the predictive ability of several estimation models. Based on the field-measured hyperspectral curves for completely covered land, bare soil, and the vegetation canopy, we used vegetation cover data obtained by analyzing digital camera photos and different vegetation indices to calculate the accuracy of estimation of vegetation cover by the different models and we discuss differences among the models. We found the most accurate estimate of vegetation cover in our study area using a single optimal combination of wavelengths based on *MSAVI2* indices and the semi-empirical model proposed by Gutman and Ignatov. © 2013 Friends Science Publishers

Keywords: Vegetation cover; Hyperspectrum; Field measurement; Wavelength selection

Introduction

Vegetation is an important indicator in global change research because it is the main component of most ecosystems and represents a natural link among soil, atmosphere, and water. Vegetation plays an essential role in biogeochemical and hydrological cycles, as well as in energy exchange at the Earth's surface (Li *et al.*, 2003). Vegetation cover is a key parameter that represents the amount of vegetation and its growing conditions (Chaichi *et al.*, 2005; Jiang *et al.*, 2006). Studies of vegetation cover can provide information about land cover change and about the vegetation's ability to maintain the global carbon balance, control CO₂ emissions, and stabilize the global environment against climate change (Wang *et al.*, 2008; Louhaichi *et al.*, 2010).

Remote sensing is an important tool for estimating vegetation cover. The spectral signal of vegetation is characterized by a low reflectance value in the red band, a sudden increase in the 700- to 740 nm band, and a high value in the green band (Collins, 1978; Horler *et al.*, 1983a and b). There are three main methods to monitor vegetation cover: empirical or semi-empirical models, vegetation indices and sub-pixel decomposition (Graetz, 1988; Dymond *et al.*, 1992; Shoshany *et al.*, 1996; Gutman

and Ignatov, 1998). The empirical models have been widely used for a long time due to their convenience and ease of use. The most widely used index has been the normalized-difference vegetation index (*NDVI*) (; Tucker, 1979; Huete and Liu, 1994; Ustin *et al.*, 2004; Jiang *et al.*, 2006), combining various models like dense-vegetation mosaic-pixel model established by Gutman and Ignatov (1998), the genetic semi-empirical relationship between vertical gap fraction and vegetation index (Baret *et al.*, 1995), and the semi-empirical relationship proposed by Carlson and Ripley (1997).

Data sets from remote-sensing platforms and sensors are widely used for the analysis of vegetation conditions and provide significant information about vast areas, thereby permitting an overview of land cover at regional scales. However, these systems have limited accuracy because of their coarse spectral resolution (Roberts *et al.*, 1993; van Leeuwen and Huete, 1996; Numata *et al.*, 2008). The development of hyperspectral sensors has provided continuous data across a large number of spectral bands. This detailed information improves the accuracy of monitoring, quantifying, and estimating vegetation chemical and physical properties (Kokaly and Clark, 1999; Curran *et al.*, 2001; Pu *et al.*, 2003; Mutanga *et al.*, 2004). At the same time, hyperspectral data provide more opportunities to

estimate vegetation fractional coverage based on semi-empirical models derived from vegetation indices (Rouse *et al.*, 1973; Richardson and Wiegand, 1977). These methods of band selection regard hyperspectral data as just another type of broadband data, and therefore fail to take advantage of the full potential offered by the high resolution of hyperspectral data.

The objective of this paper was to explore the performance of different band-selection methods used to calculate vegetation indices and vegetation cover based on three semi-empirical models. We estimated the vegetation cover using coverage data obtained with a digital camera, and then compared this with the results produced by three models and four vegetation indices. Our study was done at a community scale, and all the hyperspectral data was obtained from field measurements using a portable spectroradiometer.

Materials and Methods

Experimental Design

We conducted our experiment in a meadow steppe environment in Ewenke County of Hulunbeier Banner of China's Inner Mongolia autonomous region (Fig. 1). The study area is flat, with a relatively homogeneous land surface and open topography that is favorable for the collection of vegetation cover and hyperspectral data. The dominant species in the investigation plots are *Leymus chinensis*, *Stipa baicalensis* and *Filifolium apetalum*. For our field study, we chose a representative area of 90×90 m. We established 90 m² sample plots at intervals of 15 m within the study area (i.e., n = 25 plots), excluding points at the edges of the overall area to avoid edge effects (Fig. 2). We examined vegetation cover and the associated hyperspectral curve data for each study plot.

Hyperspectral Measurements

We observed the study site from 10:00 to 14:30 on 8 August 2009, on a sunny day. To obtain the hyperspectral curves, we used an ASD FieldSpec 3 portable spectroradiometer (<http://www.asdi.com/products/fieldspec-3-hi-res-portable-spectroradiometer>) with a band range of 325 to 1110 nm. The viewing angle of the sensor was 25°. During measurements, we held the sensor's probe vertically, facing down from a height of 1 m above the vegetation canopy, thereby forming a field of view of a circular field of view about 0.5 m in diameter projected on the land surface. The sampling was first calibrated using a white board to reduce subsequent spectral measurement errors as much as possible. At every sampling point, we captured five consecutive samples and used the average value as the spectral curve at the sampling point.

We used the spectroradiometer to measure the

spectrum of the original vegetation canopy in every sample plot, of land with full vegetation cover and of bare soil in the same plot. The spectral of the original vegetation canopy in a plot refers to the spectral curve measured above the canopy under natural conditions. To obtain the spectral curve with full cover of the same kinds of plants at each sampling point, we cut the aboveground parts of all plants in the measured area after measuring the spectrum of the canopy, then laid this material in the plot to cover all the ground in the plot; the spectrum measured by the spectroradiometer was then used to represent the same plot with full vegetation coverage. We then obtained a spectrum for bare soil in this plot by removing all of the aboveground plant parts.

Vegetation Cover from Digital Photos

We obtained photos of the vegetation in every plot under natural conditions using a digital camera. From these images, we manually traced sketched "vegetation polygons" as accurately as possible by means of visual interpretation, and then calculated the percentage of the total area covered by these polygons.

Vegetation Cover Based on Vegetation Indices

To compare different methods of using hyperspectral information, we examined different hyperspectral bands and different vegetation indices. We also compared three frequently used semi-empirical estimation models to estimate vegetation cover based on *NDVI*. Detailed model are referred to Baret *et al.* (1995), Carlson and Ripley (1997) and Gutman and Ignatov (1998).

Optimal Hyperspectral Bands Selection

In this study, we compared the field-measured vegetation cover (calculated from the digital photos) with the value calculated from the corresponding hyperspectral curve for each image. We used the following three methods to select wavelengths and calculate vegetation indices:

Method 1: For each pair of red and NIR bands in the hyperspectral data (71×201 pairs) and calculated each vegetation index (see section 2.4) for images with full vegetation cover and exposed soil for each sampling spot for each of these pairs. We then calculated the vegetation cover using each of the three semi-empirical models. We verified all vegetation cover values obtained using this method using the field-measured vegetation cover determined from the digital photos as the assumed "actual" value. Based on the results of this comparison, we selected the optimal combination of bands for the vegetation cover of the study area.

Method 2: We selected pairs of bands and calculated the corresponding vegetation index using the same approach as in Method 1. However, we then calculated the average

NDVI for all combinations of bands and used that to obtain the average vegetation index for images with full vegetation coverage and exposed soil for each sampling spot, and used the resulting overall mean to calculate vegetation cover using the three models.

Method 3: We averaged the reflectance of all wavelengths in the red and NIR bands to obtain the average reflectance at each sample spot, then used this average to calculate the vegetation index and vegetation cover of land with full vegetation coverage and exposed soil for each sampling spot.

Vegetation Indices Selection

We compared the suitability of several vegetation indices: the ratio vegetation index (*RVI*; Tucker, 1979), the optimization of soil-adjusted vegetation index (*OSAVI*; Rondeaux *et al.*, 1996), the modified soil-adjusted vegetation index (*MSAVI2*; Pu and Gong, 2000), and the scaled-difference vegetation index (*SDVI*; Jiang *et al.*, 2006).

Results

We calculated the values of the four vegetation indices (*NDVI*, *RVI*, *OSAVI*, *MSAVI2*) using $71 \times 201 = 14\ 271$ combinations (pairs) of reflectance values in the red and NIR bands. We then estimated the vegetation cover based on the four vegetation indices using the three models described in section 2.2. The results were compared with the actual vegetation cover estimated using the digital photos and selected the best models and their optimal combination of bands for estimation of the vegetation cover. Table 1 shows the results.

Using the optimal red and NIR bands, to analyze the relation between actual vegetation cover and vegetation cover calculated with different vegetation indices and estimation methods, the maximum correlation coefficient between the estimated result and the actual value ranged from 0.67 to 0.79. Overall, the vegetation cover estimated with *OSAVI* provided the weakest correlation, and for that index, the lowest correlation was for the Carlson and Ripley model. The vegetation cover estimated with *NDVI* was relatively stable among the three models (with correlations ranging from 0.76 to 0.77), and the optimal combination of bands was also stable. When estimating the vegetation fractional coverage with *RVI* and *MSAVI2*, the correlation was lowest with the Carlson and Ripley model. The Gutman and Ignatov model produced the strongest correlations with all four vegetation indices, followed by the Baret *et al.* model.

The correlation coefficient between the vegetation fractional coverage estimated based on *NDVI* (the mostly widely applied index) using the Gutman and Ignatov (1998) model was about 0.77, which was only slightly lower than the correlation coefficient (0.79) obtained by that

model using *MSAVI2*. This result can be explained by the fact that the vegetation cover was fairly high for some plots. However, there is an acknowledged limitation in the application of *NDVI* when vegetation cover was greater than a certain threshold value. The *NDVI* cannot reflect the actual vegetation condition. *MSAVI2* was designed to reduce the influence of soil in the image and to enhance the spectral sensitivity for concentrated vegetation. The success of this effort has been confirmed by the present results. To determine the applicability of this index in the field, we combined the index with the model that produced the best result, the Gutman and Ignatov model ($R_{\max} = 0.79$), and compared its predictions with the field-measured values. The resulting regression was $y = 1.0137x - 8 \times 10^{-5}$ ($p < 0.01$), where y represents the estimated value and x represents the field-measured value. The predictions were reasonably accurate, with optimal wavelengths of 687 nm (red) and 952 nm (NIR).

The optimal combination of bands (the one that produced the maximum correlation coefficient between the estimated and actual vegetation fractional coverage values for steppe vegetation in Inner Mongolia) for all three models using *NDVI* and *RVI* were 671 nm (red) and 951 nm (NIR). These values are similar to the traditional single band combination used to calculate *NDVI*. When estimating vegetation fractional coverage with *OSAVI* and *MSAVI2*, the optimal combination of bands was 687 or 690 nm (red), depending on the model, and 952 nm (NIR), which is slightly different from the optimal combination for the other two vegetation indices and clearly different from the frequently used combination of 670 nm (red) and 800 nm (NIR). Hence, when using hyperspectral data, selection of the optimal bands will greatly influence the final results. Therefore, it may not always be appropriate to follow standard practice and select a frequently used combination of bands without confirming that no other combination is more suitable for the vegetation type being studied.

To demonstrate the importance of selecting the most appropriate bands, we estimated the vegetation cover using the above mentioned frequently used methods and compared the result with the estimate produced using the optimal combination of bands (Table 2). Comparing the correlation coefficients (Table 2) with the correlation coefficients based on the optimal combination (Table 1) shows that the strength of the correlation decreased for both methods when the optimal wavelengths were not used. The most obvious decrease was for *NDVI*, for which the correlation coefficient decreased from 0.77 to 0.70 for the Gutman and Ignatov model. *RVI* showed the next strongest decrease, with the correlation coefficient falling from 0.76 to 0.68 for the Gutman and Ignatov model. The *MSAVI2* estimate changed least, and also produced the best correlation with the field-measured value, but the correlation was still weaker than the result obtained with the optimal combination of bands. These results demonstrate that to take maximum advantage of the hyperspectral data, it

Table 1: Selection of the optimal combination of bands and the best model for the estimation of vegetation cover (f) using hyperspectral image data

Model	Correlation coefficient (R_{max}) and optimal red and NIR bands								
	Gutman and Ignatov (1998)			Carlson and Ripley (1997)			Baret et al. (1995)		
	R_{max}	Red	NIR	R_{max}	Red	NIR	R_{max}	Red	NIR
$f(NDVI)$	0.767	671	951	0.764	671	951	0.762	671	951
$f(RVI)$	0.758	671	951	0.739	671	951	0.751	671	951
$f(OSAVI)$	0.747	687	952	0.668	690	795	0.728	690	952
$f(MSAVI2)$	0.792	687	952	0.736	690	952	0.778	687	952

Notes: The optimal bands are those for which the vegetation cover was closest to the actual value based on analysis of the digital photos. R_{max} is the maximum Pearson's correlation coefficient; Red and NIR refer to the optimal red and near-infrared bands (i.e., the combination that produced the maximum correlation coefficient). All correlations were statistically significant ($P < 0.05$)

Table 2: Correlation coefficients between the field-measured vegetation cover and the value estimated using methods 2 and 3

Methods	Model	Gutman and Ignatov (1998)	Carlson and Ripley (1997)	Baret et al. (1995)
Method 2	$f(NDVI)$	0.700	0.693	0.691
	$f(RVI)$	0.688	0.665	0.668
	$f(OSAVI)$	0.738	0.654	0.717
	$f(MSAVI2)$	0.771	0.723	0.758
Method 3	$f(NDVI)$	0.700	0.696	0.697
	$f(RVI)$	0.688	0.676	0.685
	$f(OSAVI)$	0.725	0.643	0.705
	$f(MSAVI2)$	0.762	0.716	0.750

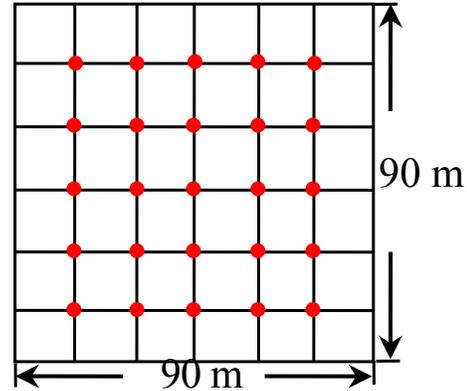


Fig. 2: Sample plots setting in the study area (Dots at the grid intersections represent the positions of the sample plots)

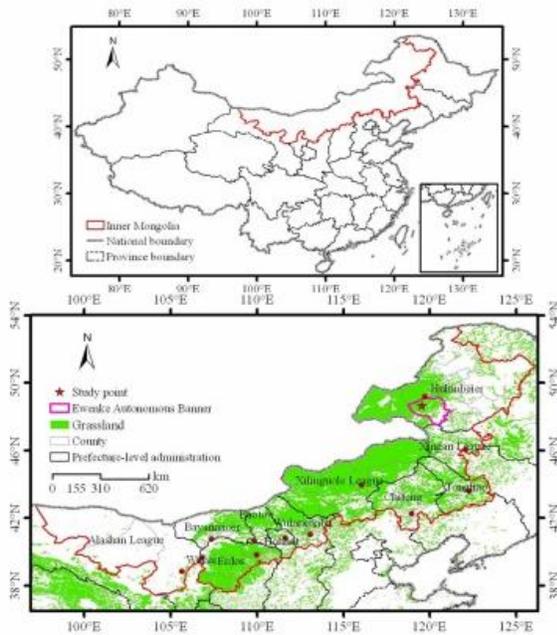


Fig. 1: Location of the study area

is necessary to identify the optimal wavelengths for vegetation being studied.

Discussion

Hyperspectral remote sensing technique development has significantly promoted earth studies, especially in key parameters identification and modeling of terrestrial

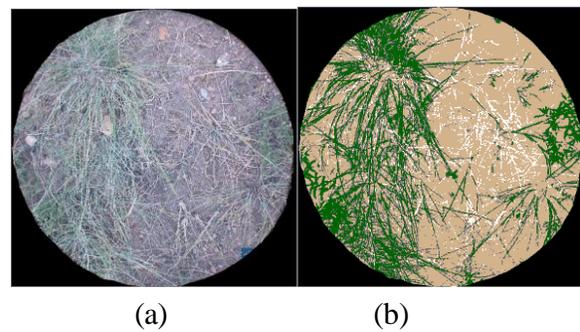


Fig. 3: Vegetation cover estimated by visual interpretation of digital photos. (a) The original digital photos. (b) The digital photos after visual interpretation and tracing of vegetation polygons

ecosystem due to its spectral measurement ability (Sykioti et al., 2012). Meanwhile, we are also challenged more and more in the aspect of valuable information extraction from massive data (Clevers et al., 2010). In order to validate simulation results from space remote sensing data, it is better to discover the relationship of these parameters with in-suit hyperspectral data to eliminate the influence of atmosphere (Delegido et al., 2010).

In this study, we identified the optimal wavelengths to estimate vegetation cover when using hyperspectral data,

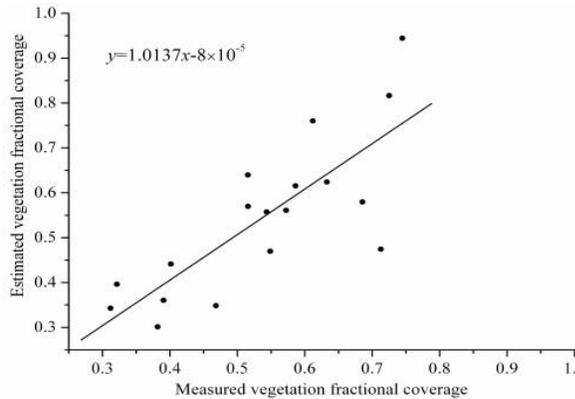


Fig. 4: Comparison between the actual field-measured vegetation cover and the result estimated using the *MSAVI2* vegetation index and the Gutman and Ignatov (1998) model. For this combination, the optimal bands were 687 nm (red) and 952 nm (NIR)

based on the in-suit hyperspectral curves and vegetation cover for samples with full vegetation coverage and exposed soil. We also evaluate the performance of different vegetation indices and models when dealing with vegetation cover estimation. The results that vegetation cover estimated using a single optimal combination of bands is closer to the actual value than the result estimated with an average value (e.g., the average for the full red or full NIR band). This reveals the greater potential of narrow-band spectrum in grassland characteristics extraction than broad-band ones. This is similar to existing researches about other ecosystems (Botha *et al.*, 2007; Meggio *et al.*, 2010). The performance of three common vegetation cover estimation models and the vegetation indices is different. As a result, it is not only important to select optimal bands but also essential to select indices and models. For grassland vegetation types, the optimal combination of bands was 687 nm and 952 nm, which is different with the frequently used combination of 680 nm and 800 nm (Rouse *et al.*, 1973).

Our results suggest that in simulation and monitoring of vegetation parameters based on hyperspectral data, it is important to identify the optimal wavelengths and models for the specific type of vegetation being studied instead of relying on average values. In this paper, we describe a simple and practical method for identifying the optimal bands and model. Although our approach provided good results for the estimation of vegetation cover, it's important to note that our results should be replicated under different environmental conditions (e.g., under drought or high temperature) to determine what factors might affect the optimal bands.

Acknowledgements

The project was supported by National Natural Science Foundation of China (No. 41030535 and No. 31000229), Special Fund for Scientific and Technological Innovation

Ability Construction in Beijing Academy of Agriculture and Forestry Sciences (KJCX201104012) and Beijing Nova program (2010B024).

References

- Baret, F., J.G.P.W. Clevers and M.D. Steven, 1995. The robustness of canopy gap fraction estimations from red and near-infrared reflectances: A comparison of approaches. *Remote Sensing Environ.*, 54: 141–151
- Botha, E.J., B. Leblon, B. Zebarth and J. Watmough, 2007. Non-destructive estimation of potato leaf chlorophyll from canopy hyperspectral reflectance using the inverted PROSAIL model. *Int. J. Appl. Earth Observ. Geoinform.*, 9: 360–374
- Carlson, T.N. and D.A. Ripley, 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing Environ.*, 62: 241–252
- Chaichi M.R., M.M. Saravi and A. Malekian, 2005. Effects of livestock trampling on soil physical properties and vegetation cover (Case study: Lar rangeland, Iran). *Int. J. Agric. Biol.*, 7: 904–908
- Clevers, J.G.P.W., L. Kooistra and M.E. Schaepman, 2010. Estimating canopy water content using hyperspectral remote sensing data. *Int. J. Appl. Earth Observ. Geoinform.*, 12: 119–125
- Collins, W., 1978. Remote sensing of crop type and maturity. *Photogram. Eng. Remote Sensing*, 26: 43–55
- Curran, P.J., J.L. Dungan and D.L. Peterson, 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: testing the Kokaly and Clark methodologies. *Remote Sensing Environ.*, 76, 349–359
- Delegido, J., L. Alonso, G. Gonzalez and J. Moreno, 2010. Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (NAOC). *Int. J. Appl. Earth Observ. Geoinform.*, 12: 165–174
- Dymond, J.R., P.R. Stephens, P.F. Newsome and R.H. Wilde, 1992. Percent vegetation cover of a degrading rangeland from SPOT. *Int. J. Remote Sensing*, 13: 1999–2007
- Graetz, R.D., R.P. Pech and A.W. Davis, 1988. The assessment and monitoring of sparsely vegetated rangelands using calibrated Landsat data. *Int. J. Remote Sensing*, 9: 237–251.
- Gutman, G and A. Ignatov, 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sensing*, 19: 1533–1543
- Horler, D.N.H., M. Dockray and J. Barber, 1983a. The red edge of plant leaf reflectance. *Int. J. Remote Sensing*, 4: 273–288
- Horler, D.N.H., M. Dockray, J. Barber and A. R. Barringer, 1983b. Red edge measurements for remotely sensing plant chlorophyll content. *Adv. Space Res.*, 3: 273–277
- Huete, A.R. and H. Liu, 1994. An error and sensitivity analysis of the atmospheric- and soil-correcting variants of the NDVI for the MODIS-EOS. *IEEE Trans. Geosci. Remote Sensing*, 32: 897–905
- Jiang, Z.Y., A.R. Huete, J. Chen, Y.H. Chen, J. Li, G.J. Yan and X.Y. Zhang, 2006. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing Environ.*, 101: 366–378
- Kokaly, R.F. and R.N. Clark, 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing Environ.*, 67: 267–287
- Li, X.B., Y.H. Chen, P.J. Shi and J. Chen, 2003. Detecting vegetation fractional coverage of typical steppe in Northern China based on multi-scale remotely sensed data. *Acta Bot. Sin.*, 45: 1146–1156
- Louhaichi, M., M.D. Johnson, A.L. Woertz, A.W. Jasra and D.E. Johnson, 2010. Digital charting technique for monitoring rangeland vegetation cover at local scale. *Int. J. Agric. Biol.*, 12: 406–410
- Meggio, F., P.J. Zarco-Tejada, L.C. Núñez, G. Sepulcre-Cantó, M.R. González and P. Martín, 2010. Grape quality assessment in vineyards affected by iron deficiency chlorosis using narrow-band physiological remote sensing indices. *Remote Sensing Environ.*, 114: 1968–1986

- Mutanga, O., A.K. Skidmore and H.H.T. Prins, 2004. Predicting *in situ* pasture quality in the Kruger National Park, South Africa, using continuum removed absorption features. *Remote Sensing Environ.*, 89: 393–408
- Numata, I., D.A. Roberts, O.A. Chadwick, J.P. Schimel, L.S. Galvao and J.V. Soares, 2008. Evaluation of hyperspectral data for pasture estimation in the Brazilian Amazon using field and imaging spectrometers. *Remote Sensing Environ.*, 112: 1569–1583
- Pu, R.L. and P. Gong, 2000. *Hyperspectral Remote Sensing and its Application*, p: 205. Beijing: Higher Education Press
- Pu, R.L., P. Gong, G.S. Biging and M.R. Larrieu, 2003. Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index. *IEEE Trans. Geosci. Remote Sensing*, 41: 916–921
- Richardson, A.J. and C.L. Wiegand, 1977. Distinguishing vegetation from soil background information. *Photogram. Eng. Remote Sensing*, 43: 1541–1552
- Roberts, D.A., M.O. Smith and J.B. Adams, 1993. Green vegetation, nonphotosynthetic vegetation, and soils in Aviris data. *Remote Sensing Environ.*, 44: 255–269
- Rondeaux, G., M. Steven and F. Baret, 1996. Optimization of soil-adjusted vegetation indices. *Remote Sensing Environ.*, 55: 95–107
- Rouse, J.W., R.H. Haas, J.A. Shell and D.W. Deering, 1973. *Monitoring Vegetation Systems in the Great Plains with ERTS-1, Third Earth Resources Technology Satellite Symposium*, pp: 309–317. Goddard Space Flight Center, Washington DC, USA
- Shoshany, M., P. Kutiel and H. Lavee, 1996. Monitoring temporal vegetation cover changes in Mediterranean and arid ecosystems using a remote sensing technique: case study of the Judean Mountain and the Judean Desert. *J. Arid Environ.*, 33: 9–21
- Olga Sykioti, O., D. Paronis, S. Stagakis and A. Kyparissis, 2012. Band depth analysis of CHRIS/PROBA data for the study of a Mediterranean natural ecosystem. Correlations with leaf optical properties and ecophysiological parameters. *Remote Sensing Environ.*, 115: 752–766
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing Environ.*, 8: 127–150
- Ustin, S.L., D.A. Roberts, J.A. Gamon, G.P. Asner and R.O. Green, 2004. Using imaging spectroscopy to study ecosystem processes and properties. *Bioscience*, 54, 523–534
- Van Leeuwen, W.J.D. and A.R. Huete, 1996. Effects of standing litter on the biophysical interpretation of plant canopies with spectral indices. *Remote Sensing Environ.*, 55: 123–138
- Wang, H., X.B. Li, H.L. Long, X. Xu and C. Zhang, 2008. Simulating vegetation fractional coverage for temperate grassland in Northern China combining NDVI with precipitation time series from 1982 to 1999. *J. Basic Sci. Eng.*, 16: 525–536

(Received 07 June 2012; Accepted 07 December 2012)