

International Society for Environmental Information Sciences 2010 Annual Conference (ISEIS)

An assessment of correlation on MODIS-NDVI and EVI with natural vegetation coverage in Northern Hebei Province, China

Zhongfei Li^{a,b}, Xiaobing Li^{a*}, Dandan Wei^a, Xu Xu^b, Hong Wang^a

^a State Key Laboratory of Earth Surface Process and Resource Ecology, College of Resources Science and Technology, Beijing Normal University, Beijing 100875, China

^b Department of Environment Science and Engineering, Southwest Forest University, Kunming 650224, China

Abstract

Regional vegetation cover plays an important role in modeling ecosystem change and conservation. In this study, MODIS-NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index) image of 32 days composite were made in August 2008. 39 samples of grassland, 40 samples of shrub and 55 samples of forest were established in Northern Hebei Province in August 2008. The MODIS image we used was taken in the same period as the time of field sampling. The MODIS-NDVI and EVI values of field samples were extracted from MODIS image. The optimal regression equation was developed between the NDVI values and the vegetation cover of field samples in grassland, shrub and forest respectively. The similar process was conducted for EVI values. The simulation precision of the equation was tested by independent field data. Lastly, the predictive validity of natural vegetation cover by using NDVI and EVI values was compared and assessed. The results show that all selected optimal regression equation pass significance testing of 0.05. As for forest, shrub and grassland, the coefficient Correlation (R²) of selected optimal regression equations based on NDVI values is higher than that based on EVI (from 0.488 to 0.644), and Mean Absolute Error (MAE) based on NDVI values is lower (from 0.0617 to 0.0916) than that based on EVI values. There is a conclusion that MODIS-NDVI are more correlation with field data of vegetation cover and have obvious advantages for predicting natural vegetation coverage than MODIS-EVI in our study area.

© 2010 Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

Keywords: MODIS, NDVI, EVI, Vegetation Coverage, Correlation, Regression

1. Introduction

Presently ecological informatics provides an important support for development of modern ecology, and it serves for ecosystem management directly [1]. Regional vegetation coverage is an important integral part of spatial

*Corresponding author Tel.: +86-10-58808460
E-mail address: xbli@ires.cn

database of terrestrial eco-information. It plays an important role in modeling ecosystem change and conservation. Natural vegetation coverage can be surveyed by using satellite remote sensing technology, which have become an important mean of obtaining eco-information on regional or global scale [2]. Contrasted with NOAA/AVHRR, the terra-MODIS is a kind of new-style and important satellite remote sensor, and has higher spatial and spectral resolutions. MODIS-derived Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) are the most widely used Vegetation indices in satellite monitoring of vegetation information [3, 4, 5, 6]. Gao et al. found that the NDVI was more saturated at high biomass levels [7], however EVI can eliminate impact of soil background and atmospheric aerosols [3, 8], and can had improved sensitivity over high biomass areas [9]. Research has shown that when developed a linear relationship with photosynthetically active radiation, leaf area index and biomass of vegetation, EVI is superior to NDVI [10]. Furthermore, Huete et al. also reported that the EVI can provide greater dynamic range than the normalized difference vegetation index [3]. However, there is a lack of understanding of correlation on MODIS-NDVI and EVI respectively with natural vegetation cover in China. Thus, the objective of this study is to assess the correlation on MODIS-NDVI and EVI with natural vegetation coverage in Northern Hebei Province, China. Based on their correlation, optimal regression equation between the MODIS-NDVI/EVI values and the natural vegetation coverage of field samples in forest, shrub and grassland respectively were developed and selected. Finally, the simulation precision of regression formula was tested by using field data of vegetation coverage.

2. Study Area and Site Description

Hebei province (36°05' N-42°37' N, 113°27' E-119°45' E) is located in Northeast China. The study area in this paper is located in Northern Hebei Province (39°6'-42°37'N, 113°48'-119°45'E) (fig.1), including four cities (Zhangjiakou; Chengde; Qinhuangdao; Tangshan) in the administrative regionalization. The landforms are complex and diverse in the study area. The geomorphological types are mountains and hills, and the altitude is about 1000 m and in northern study area and the geomorphological type is plain and the altitude is less than 50 m in southeast study area. The study area is dominated by continental monsoon climate of middle temperate zone. The mean annual temperature was 0-13°C and the mean annual rainfall was 300-700 mm in Northern Hebei Province. Most of the annual rainfall occurred during the rainy season from June to August. Temperate grassland and temperate deciduous broad-leaved forest are the main zonal vegetation types in our study area. In the ecotone from forest zone to the grassland, many shrub-grassland distributes [11].

3. Data Source and Data Process Methods

3.1 Field survey of vegetation coverage

In study area, we selected 55 forest samples, 40 shrub samples and 39 grassland samples randomly in August, 2008. The spatial distance between samples of same vegetation type is about from 5 to 10 km (Fig. 1). Every sample was positioned by GPS based on topographic map. These samples were representative of vegetation coverage in Northern Hebei Province. The size of every sample was 250 m × 250 m, which can match with the resolution of chosen MODIS image perfectly. In every sample, we made 5 plots based on gradient of vegetation coverage of this sample. The size of plot is 1 m² for grassland, 16 m² for shrub and 100 m² for forest respectively. The vegetation coverage of every plot was measured. The mean of 5 plots was representative of the vegetation coverage of the sample.

3.2 Remotely sensed data source

MODIS-NDVI and EVI products with spatial resolution of 250 m × 250 m were provided by the NASA. The data in August 2008 were used for this study. Firstly, we obtained the MODIS-NDVI and EVI production of 16 day composite. To avoid the presence of cloud cover and match with time of filed survey of vegetation coverage, the MODIS production of 16 day composite were converted to 32-day composites based on the Maximum Value Composite [12], which can also eliminate the impact resulting from difference of the time series of MODIS image

and field survey. Lastly, the NDVI and EVI values of field sample localized by GPS were extracted from the MODIS images.

3.3 Development of regression equation

Statistical analyses were performed to evaluate the applicability of MODIS 250 m VI for vegetation coverage in Northern Hebei Province. We developed linear and nonlinear (including quadric, cubic, logarithmic, and so on) regression equations between NDVI values and vegetation cover of field samples of grassland, shrub and forest respectively. Based on the correlation coefficient of these regression equations, the optimal regression equation was selected for each vegetation type. Thus, we obtained three optimal regression models corresponding to grassland, shrub and forest respectively. The similar process was conducted for EVI values.

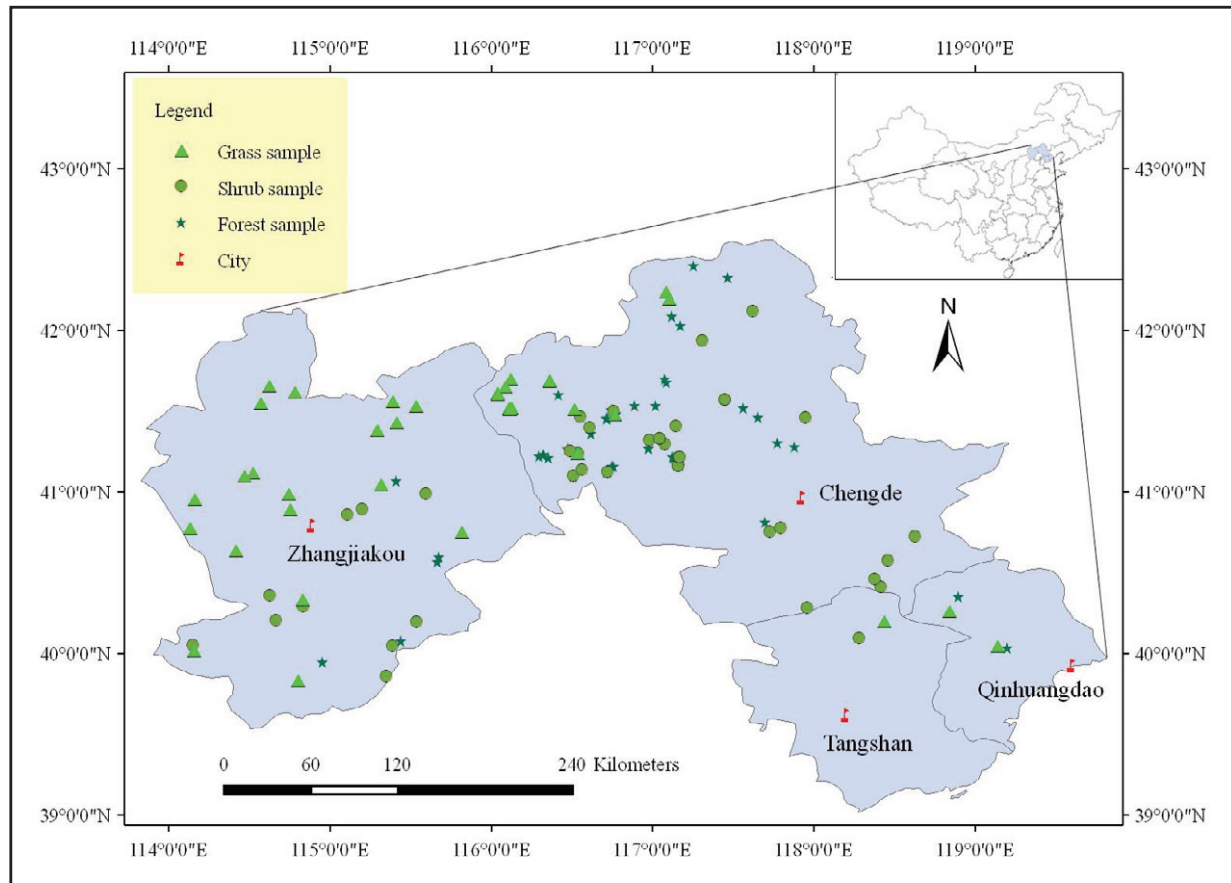


Fig.1 The study area and space distribution of field samples

3.4 Verifying accuracy of regression equation

In order to verify prediction accuracy of the optimal regression equation, we tested the predicting results with independent field data. Namely, before developing the regression model, we reserved the field data of ten samples of each vegetation type. Then, Mean Absolute Error (MAE) between the predicted values of vegetation cover based on the optimal regression equation and corresponding field data was calculated. Thus, based on the mean absolute errors, the accuracy of selected optimal regression equation can be evaluated. Lastly, the optimal regression equations for different vegetation type based on NDVI value were compared with that based on EVI value.

4. Results and Analysis

Table 1 presents a series of optimal regression equation based on NDVI and EVI values for each vegetation type. Firstly, based on the Correlation Coefficient (R^2), the optimal regression equation between NDVI and vegetation coverage of forest, a quadric curve equation, were selected, which was $y = -5.0770 + 14.2646 * NDVI - 8.4566 * NDVI^2$ ($R^2 = 0.522$, $N = 45$). where y was the vegetation coverage (%). Similarly, a liner regression equation for between NDVI and vegetation coverage of shrub and other liner regression equation for between NDVI and vegetation coverage of grassland were selected respectively, which were respectively $y = 0.0008 + 1.0163 * NDVI$ ($R^2 = 0.522$, $N = 30$) and $y = 0.0824 + 0.9688 * NDVI$ ($R^2 = 0.522$, $N = 29$). Similarly, the EVI-based optimal regressions equation for three vegetation type of forest, shrub and grassland were respectively $y = 0.6039 + 0.6491 * EVI$ ($R^2 = 0.259$, $N = 45$), $y = -0.4520 + 5.1863 * EVI - 5.3051 * EVI^2$ ($R^2 = 0.315$, $N = 30$) and $y = 0.4150 + 0.8076 * EVI$ ($R^2 = 0.301$, $N = 29$). The results show that all optimal regression equations pass significance testing of 0.05, which meet the statistical demand ($p < 0.05$). Furthermore, judged from the R^2 , NDVI values reveal a stronger correlation with vegetation coverage of three kinds of type which are forest, shrub and grassland respectively than EVI values. Especially, the Correlation between NDVI and vegetation coverage of grassland was strongest, and the R^2 reach to 0.644.

As for each vegetation type, verifying results to optimal regression equations using independent field data show that MAE based on NDVI was lower than that based on EVI respectively, which varies from 0.06 to 0.09 (table 1). However, MAE based on EVI varies from 0.08 to 0.13 for three kinds of vegetation types. The optimal regression equation based on NDVI hold more prediction accuracy to vegetation coverage. So the statistic summary showed that NDVI presented much better results than EVI in the study area.

Table 1. Correlation on NDVI and EVI with natural vegetation cover

Vegetation Index	Vegetation Type	Optimal Regression Equation	Correlation Coefficient (R^2)	Sig.	Mean Absolute Error
NDVI	Forest	$y = -5.0770 + 14.2646 * NDVI - 8.4566 * NDVI^2$	0.522	0.000	0.06
	Shrub	$y = 0.0008 + 1.0163 * NDVI$	0.448	0.000	0.07
	Grassland	$y = 0.0824 + 0.9688 * NDVI$	0.644	0.000	0.09
EVI	Forest	$y = 0.6039 + 0.6491 * EVI$	0.259	0.001	0.08
	Shrub	$y = -0.4520 + 5.1863 * EVI - 5.3051 * EVI^2$	0.315	0.006	0.14
	Grassland	$y = 0.4150 + 0.8076 * EVI$	0.301	0.002	0.13

** significant at $p < 0.01$.

5. Discussion and Conclusions

Methods and data for assessment and validation of natural vegetation coverage obtained from remote sensing is a critical issue. Over the last two decades, remotely sensed data has offered a means of measuring vegetation properties at regional to global scales [13]. MODIS-NDVI and EVI are closely tied to vegetation coverage, therefore NDVI and EVI will change closely natural vegetation coverage. Previous studies have used NDVI and EVI in semi-arid climates [14, 15]. But few studies have compared the two indices for natural vegetation coverage of forest, shrub and grassland in our study area.

This study used MODIS NDVI and EVI, enabling comparison of the two indexes in a continental monsoon climate. In this study, EVI value would be expected to most closely correlate with natural vegetation coverage in Northern Hebei Province, because studies show that EVI was developed to improve sensitivity in high biomass regions, where NDVI tended to saturate [16], and the NDVI can result identifies the loss of dense forest cover canopies, regardless of forest type [17]. However, the smaller MAE and higher correlation coefficient indicate that

NDVI can predict more reasonable data than EVI. So, among the two vegetation indices in this study, NDVI is more perfect for predicting natural vegetation coverage. Many factors can cause this result. Because estimating vegetation parameters from NDVI or EVI is not straightforward, these remotely sensed indices contain combined information affected by many factors, such as vegetation phenology, disturbance events, signal contamination, sensor conditions, and so on [13, 18, 19]. Furthermore, in our study area, the climate is continental monsoon climate of middle temperate zone. Annual precipitation was lacking. Therefore, this region do not has high biomass, and NDVI is sufficient for our study area. All these causes can explain perhaps why the NDVI-vegetation coverage regression model is more perfect than EVI-vegetation coverage regression model.

The results also show that the correlation coefficient between NDVI and vegetation coverage of forest, shrub and grassland are not very high, and are also not available for all other regions. Perhaps, we can find out some reasons. For example, owing to the limit of geographic extent and ecological conditions, these collected filed data are clearly relatively insufficient, which can not provide robust characterization of error in the MODIS product and more adequate checking to this correlation. Nevertheless, more profound reasons are unclear. Many important questions need to be study continually in order to both address weaknesses identified in this paper and meet the needs of scientists who wish to use these results better.

In any rate, our study results exhibited a true situation of correlation on MODIS-NDVI and EVI with natural vegetation coverage in study area. The conclusions can be drawn that MODIS-NDVI values are more correlation with field data of vegetation cover and have obvious advantages than MODIS-EVI values in our study area when predicting the vegetation cover. Moreover, the optimal regression equation based on NDVI stands for better simulation precision than that based on EVI. This study also suggests that selected optimal NDVI-vegetation coverage regression mode can be applied to estimate the natural vegetation coverage within the growing season in our study area.

Acknowledgment

Financial supports from The National High Technology Development Plan 863 in China (No:2006AA120108) are highly appreciated. We also thank highly Dr. Bin Zhang and Dr. Kejian Shen at College of Resources Science and Technology of Beijing Normal University for his great help and valuable suggestions in the process of writing this paper.

References

- [1] Yu, G.R., Niu, D., and He, H.L. (2003). Ecosystem management, eco-informatics and data resource management. *Resources Science*, 25(1), 48-53. (in Chinese).
- [2] Liu, J.Y., Yue, T.X., Zhang, R.H., Zhang, R.H., and Shao, Q.Q. (2006). The infomation technology based ecosystem assessment. *Resources Science*, 28(4), 5-7. (in Chinese).
- [3] Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.*, 83(1-2), 195-213.
- [4] Mildrexler, D.J., Zhao, M.S., and Running, S.W. (2009). Testing a MODIS global disturbance index across North America. *Remote Sens. Environ.*, 113, 2103-2117.
- [5] Peckham, S.D., Ahl, D.E., Serbin, S.P., and Gower, S.T. (2008). Fire-induced changes in green-up and leaf maturity of the Canadian boreal forest. *Remote Sens. Environ.*, 112, 3594-3603.
- [6] Wang, Z.X., Liu, C., Chen, W.B., and Lin, X. (2006). Preliminary comparison of MODIS-NDVI and MODIS-EVI in Eastern Asia. *Geomatics and Information Science of Wuhan University*, 31(5), 407-427. (in Chinese).
- [7] Gao, X., Huete, A.R., Ni, W., and Miura, T. (2000). Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sens. Environ.*, 74, 609-620.
- [8] Liu, H.Q. and Huete, A.R. (1995). A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.*, 33(2), 457-465.
- [9] Wardlow, B.D., Egbert, S.L., and Kastens, J.H. (2007). Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. central great plains. *Remote Sens. Environ.*, 108, 290-310.

- [10] Wim, J.D., Huete, A.R., and Laing, T.W. (2002). MODIS vegetation index composition approach: a prototype with AVHRR data. *Remote Sens. Environ.*, 69(3), 264-280.
- [11] Liu, L. (1981). The distribution of the forest steppe region in Hebei Province. *Acta Phytocologia Et Geobotanica Sinica*, 5(4), 308-312. (in Chinese).
- [12] Holben, B.N. (1986). Characteristics of maximum-value composite images for temporal AVHRR data. *Int. J. Remote Sens.*, 7(11), 1417-1434.
- [13] Lu, H., Raupach, M.R., McVicar, T.R., Barrett, D.J. (2003). Decomposition of vegetation cover into woody and herbaceous components using AVHRR NDVI time series. *Remote Sens. Environ.*, 86, 1-18.
- [14] Laneve, G., and Castronuovo, M.M., (2005). Comparison between vegetation change analysis in Kenya based on AVHRR and SeaWiFS images. *Int. J. Remote Sens.*, 26 (12), 2549-2559.
- [15] Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., and Wang, S. (2005). Comparing MODIS vegetation indices with AVHRR NDVI for monitoring the forage quantity and quality in Inner Mongolia grassland, China. *Grassland Science*, 51, 33-40.
- [16] Terrestrial Biophysics and Remote Sensing Lab (TBRS). Theoretical basis for the enhanced vegetation index (United States). <http://tbrs.arizona.edu/cdrom/Index.html>. (accessed Mar. 2008)
- [17] Hansen, M.C., Shimabukuro, Y.E., Potapov, P., and Pittman, K. (2008). Comparing annual MODIS and PRODES forest cover change data for advancing monitoring of Brazilian forest cover. *Remote Sens. Environ.*, 112, 3784-3793.
- [18] Sims, D.A., Luo, H., Hastings, S., Oechel, W.C., Rahman, A.F., and Gamon, J.A. (2006). Parallel adjustments in vegetation greenness and ecosystem CO₂ exchange in response to drought in a Southern California chaparral ecosystem. *Remote Sens. Environ.*, 103, 289-303.
- [19] Zhao, B., Yan, Y.N., Guo, H.Q., He, M.M., Gu, Y.J., and Li B., (2009). Monitoring rapid vegetation succession in estuarine wetland using time series MODIS-based indicators: An application in the Yangtze River Delta area. *Ecol. Indic.*, 9, 346-356.
-

Xiao-Bing Li is presently a research scientist at College of Resources Science and Technology, Beijing Normal University, and is also the dean of this college. Professor Li has always been engaged in study on global change and terrestrial ecosystem, remote sensing for natural resources and environment, natural resources assessment and ecosystem management. He has published more than 50 journal papers and several books in these research fields.

Zhong-Fei Li is studying on College of Resources Science and Technology, Beijing Normal University for his Ph.D. During the last several years, he have spend a considerable amount of time in tropical forest research, and obtained his master's degree in ecological field in July 2005. Presently, his major is land resources management. He engage mainly in variety ecological effects resulted from land use change.