International Journal of Applied Earth Observation and Geoinformation 50 (2016) 26–38

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Linking in situ LAI and fine resolution remote sensing data to map reference LAI over cropland and grassland using geostatistical regression method

Yaqian He[a,b,c], Yanchen Bo[a,b,*], Leilei Chai[a,b], Xiaolong Liu[d], Aihua Li[e]

[a] State Key Laboratory of Remote Sensing Science, Research Center for Remote Sensing and GIS, School of Geography, Beijing Normal University, Beijing 100875, China
[b] Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Beijing 100875, China
[c] Department of Geology and Geography, West Virginia University, Morgantown, WV 26506, USA
[d] College of Tourism & Geography Science, Yunnan Normal University, Kunming, Yunnan Province 650500, China
[e] Department of Geoscience, Boise State University, Boise, ID 83725, USA

ARTICLE INFO

Article history:
Received 8 December 2015
Received in revised form 22 February 2016
Accepted 26 February 2016
Available online 11 March 2016

Keywords:
Leaf area index
Up-scaling
Geostatistical regression
Reduced major axis
Vegetation index

ABSTRACT

Leaf Area Index (LAI) is an important parameter of vegetation structure. A number of moderate resolution LAI products have been produced in urgent need of large scale vegetation monitoring. High resolution LAI reference maps are necessary to validate these LAI products. This study used a geostatistical regression (GR) method to estimate LAI reference maps by linking in situ LAI and Landsat TM/ETM+ and SPOT-HRV data over two cropland and two grassland sites. To explore the discrepancies of employing different vegetation indices (VIs) on estimating LAI reference maps, this study established the GR models for different VIs, including difference vegetation index (DVI), normalized difference vegetation index (NDVI), and ratio vegetation index (RVI). To further assess the performance of the GR model, the results from the GR and Reduced Major Axis (RMA) models were compared. The results show that the performance of the GR model varies between the cropland and grassland sites. At the cropland sites, the GR model based on DVI provides the best estimation, while at the grassland sites, the GR model based on DVI performs poorly. Compared to the RMA model, the GR model improves the accuracy of reference LAI maps in terms of root mean square errors (RMSE) and bias.

1. Introduction

Leaf Area Index (LAI), defined as half the total leaf area per unit ground surface areas (Chen and Black, 1992), is an important parameter of vegetation structure and function (Abuelgasim et al., 2006). LAI provides substantial information on the exchange of energy, mass, and momentum flux between the Earth's surface and its atmosphere (Morisette et al., 2006; Myneni et al., 1997). LAI has been widely used as an input in climate, hydrology, and biogeochemistry models (Berterretche et al., 2005; Knyazikhin et al., 1998; Morisette et al., 2006). To date, a number of global and regional moderate-resolution LAI products have been produced, including Moderate Resolution Imaging Spectroradiometer (MODIS), Carbon Cycle and Change in Land Observational Products from and Ensemble of Satellites (CYCLOPES), Canada Centre for Remote Sensing (CCRS), and Global Land Surface Satellite (GLASS) (Chen et al., 2002; Tian et al., 2000; Weiss et al., 2007; Xiao et al., 2014). Owing to the influence of model algorithms, vegetation heterogeneity, and observation conditions, these LAI products inevitably have inherent uncertainties (Chen et al., 2002), which subsequently may impact the accuracy of any resulting modeling activities. Specifying the uncertainties of these coarse spatial resolution LAI products is essential for users to determine the most appropriate dataset for their applications, and for producers to improve methodological algorithms. However, a direct comparison between in situ LAI measurements and these corresponding moderate resolution LAI products is not recommended because of scale-mismatch, geolocation errors, and land surface heterogeneity (Huang et al., 2006; Yang et al., 2006). The proposed way to validate coarse resolution remote

* Corresponding author at: State Key Laboratory of Remote Sensing Science, Research Center for Remote Sensing and GIS, School of Geography, Beijing Normal University, Beijing 100875, China.
E-mail address: boyc@bnu.edu.cn (Y. Bo).

http://dx.doi.org/10.1016/j.jag.2016.02.010
0303-2434/

sensing products is using fine reference maps derived from upscaling in situ measurements (Fernandes et al., 2014; Iiames et al., 2015; Kang et al., 2015; Morisette et al., 2006; Wang et al., 2014). Previous studies have generated fine resolution LAI reference maps through fusing in situ LAI measurements and fine resolution remote sensing images (e.g. TM, ETM+, ASTER, SPOT) (Baret et al., 2005; Chen et al., 2002; Cohen and Justice, 1999; Garrigues et al., 2008; Li et al., 2013a; Martinez et al., 2009; Morisette et al., 2006; Pisek and Chen, 2007).

There are three categories of methods for estimating reference LAI maps using in situ LAI observations and fine spatial resolution remote sensing data, including regression, vegetation radiation transfer equation inversion, and geostatistical methods (Cohen et al., 2003; Martinez et al., 2010; Yang et al., 2006). Of these, the radiation transfer equation inversion method is not used widely due to the difficulty in collecting certain model parameters (e.g., canopy structure) and the fact that the solution of the model is not unique (Yang et al., 2006). Geostatistical methods have become popular in linking field data to image data, and been applied to estimate forest parameters (basal area, height, health conditions, etc.), detect land use and land cover change, and map vegetation index (e.g., normalized difference vegetation index: NDVI and LAI) (Van der Meer, 2012). Traditional geostatistical methods, such as Kriging, predict unknown points through spatially interpolating surrounding field observations (Berterretche et al., 2005; Li et al., 2013a,b). The limited number of field observations and the spatial non-stationarity of in situ observations distribution could lead to uncertainty of predicting results. Regression methods, such as ordinary least squares regression, attempt to improve the predicting accuracy through accounting for high resolution remote sensing data (e.g., reflectance or vegetation indices (VIs) derived from Landsat ETM+). Cohen et al. (2003) compared three regression methods (i.e., traditional ordinary least squares regression, inverse ordinary least square regression, and reduced major axis: RMA) over the BigFoot AGRO and NOBS sites. They reported that the performance of RMA method was superior to the other two. However, none of the regression methods consider the spatial/temporal correlation of in situ observations and high resolution reflectance or VI data, which may lead to an underestimation of the uncertainty along with the regression coefficients (Chatfield, 2003).

Geostatistical regression (GR) method conserves merits from both traditional geostatistical methods and regression methods. It has been used in examining the relationships between terrestrial carbon dioxide flux and its primary environmental drivers (Mueller et al., 2010), and estimating snow cover and gross primary productivity (Erickson et al., 2005; Yadav et al., 2010). Compared to traditional regression methods, the GR method is improved in one distinct way, which is the ability to account for the spatial/temporal correlation of the residuals from in situ observations (such as field LAI measurements) and auxiliary data (such as NDVI) (Erickson et al., 2005; Mueller et al., 2010; Yadav et al., 2010). Unlike traditional geostatistcal methods (e.g., Kriging), the GR method attempts to provide better estimating of unknown points by exploring the correlation between high resolution remote sensing data and field observations. To our knowledge, no attempts have been made to use the GR method to estimate LAI reference maps. This study applied the GR method to estimate high resolution LAI reference maps over cropland and grassland sites through fusing in situ LAI measurements and high resolution remote sensing images (i.e., Landsat TM/ETM+ and SPOT). To investigate the discrepancy of employing different VIs on estimating LAI reference maps, this study established the GR models for the following VIs: difference vegetation index (DVI), NDVI, and ratio vegetation index (RVI). To robustly assess the performance of the GR model, the results from GR and RMA models were compared.

2. Methodology

2.1. Geostatistical regression method

The GR method not only models the relationships between auxiliary variables (DVI, NDVI, and RVI in this study) and field measurements (in situ LAI measurements in this study), but also accounts for the spatial/temporal correlation of the regression residuals (Erickson et al., 2005). As with the linear regression method, the GR method decomposes LAI into a deterministic and a stochastic component:

$$LAI = X\beta + \varepsilon \tag{1}$$

where $X\,(n \times P)$ is the DVI, NDVI, and RVI, respectively, $\beta\,(P \times 1)$ is the corresponding regression coefficient, and $\varepsilon\,(n \times 1)$ is assumed to be second-order stationary and zero-mean residual for DVI, NDVI, and RVI (Leung and Cooley, 2014; Mueller et al., 2010; Yadav et al., 2010). Unlike the traditional linear regression approach, which regards ε as white noise, the GR method uses spatial covariance to recognize the spatial autocorrelation structure of the regression residuals ε. The experimental covariance of residuals ε for DVI, NDVI and RVI, respectively is:

$$Q(h) = E(\varepsilon(X)\,\varepsilon(X+h)) \tag{2}$$

where h is the spatial and/or temporal distance, $Q(h)$ is the covariance of residual at separation distance h (Erickson et al., 2005). Many theoretical covariance functions (such as nugget, exponential, spherical, and Gaussian functions) can be used to model the experimental covariance (Schabenberger and Pierce, 2001). In this study, a linear combination of nugget and exponential functions is used following the previous studies (Erickson et al., 2005; Li et al., 2013a; Mueller et al., 2010). This function is defined as:

$$Q(h) = \begin{cases} \sigma_N^2 + \sigma_S^2, h = 0 \\ \sigma_S^2 \exp\left(-\dfrac{h}{l}\right), h > 0 \end{cases} \tag{3}$$

σ_N^2 is the measurement error or the variability at small scale that is uncorrelated in space and/or time, σ_S^2 is the variance of the variability correlated in space and/or time, and l is the correlation range parameters (Leung and Cooley, 2014). The Restricted Maximum Likelihood (RML), which maximizes the marginal distribution of the covariance function parameters, is used to estimate the parameters (σ_N, σ_s, l) (Kitanidis and Shen, 1996).

The best linear unbiased estimator of β on the basis of Aitken (1935) is the generalized-least-squares estimator, that is, the value of β that minimizes $\left(LAI - X\beta\right)^T Q^{-1}\left(LAI - X\beta\right)$. Thus,

$$\hat{\beta} = \left(X^T Q^{-1} X\right)^{-1} X^T Q^{-1} LAI \tag{4}$$

2.2. Reduced major axis method

To robustly assess the performance of the GR model, we compare the results from GR and RMA models. We choose RMA method because it is regarded as the 'standard' method for estimating LAI reference map in BigFoot project (Berterretche et al., 2005; Cohen et al., 2003), which is a well known project linking in situ measurements, remote sensing and models to validate MODIS products including LAI product. The form of RMA is identical to a simple linear regression method:

$$LAI = \beta_0 + \beta_1 X + \varepsilon \tag{5}$$

where X is DIV, NDVI, and RVI, respectively. ε is white noise residual.

RMA method is superior to traditional ordinary least squares regression when both dependent (LAI in this study) and independent variables (DVI, NDVI, and RVI in this study) are measured with