Remote Sensing of Environment 139 (2013) 257–270

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Retrieval of leaf area index from MODIS surface reflectance by model inversion using different minimization criteria

G. Leonenko [1], S.O. Los *, P.R.J. North [2]

Department of Geography, Swansea University, Singleton Park, Swansea SA2 8PP, UK

ARTICLE INFO

Article history:
Received 23 April 2012
Received in revised form 25 April 2013
Accepted 6 July 2013
Available online 6 September 2013

Keywords:
Cost function
Statistical distance
LAI
Look-up table inversion
Biophysical parameter estimation
MODIS

ABSTRACT

Leaf area index (LAI) is one of the key parameters for the calculation of the energy budget, photosynthesis, and the interception of precipitation in land-surface models at local to global scales. Estimation of LAI from satellite data is a challenging and difficult problem. Studies over the past decades have focused predominantly on the improvement of forward modeling of the radiative transfer problem and on the application of more realistic numerical inversion schemes. Little or no attention has been paid to alternatives for the least squares method as a statistical distance measure or cost function, used to minimize the distance between observations and model predictions. The least-squares method has properties that assume noise with a Gaussian distribution and zero mean, an assumption often violated when LAI is estimated from satellite reflectance data. Here, we test the use of alternative statistical distance measures or cost functions to estimate LAI. We combine a look-up table (LUT)-inversion method based on the FLIGHT radiative transfer model and test how well it estimates LAI from MODIS reflectance data for a large set of alternative cost functions. We consider three classes of statistical distance measures or cost functions: information divergence measures, M-estimates, and minimum contrast methods. We estimate LAI from the Moderate Resolution Imaging Spectrometer (MODIS) surface reflectance product (MOD09GA) for 11 VALERI and BigFoot sites around the globe. These sites consist of a wide range of tree-cover types that include conifer, broadleaf and mixed (conifer, broadleaf, grassland) forest sites. We develop LUTs with FLIGHT for conifer and broadleaf forests and we show that improvements can be obtained for the estimation of LAI by choosing a cost function appropriate for a particular problem. Results show error reductions of 20% compared with the MODIS LAI retrieval (MOD15A2).

1. Introduction

The one-sided leaf area index (LAI) is an important parameter in agricultural, ecological, hydrological and land-surface models. LAI is used to quantify key biophysical processes such as photosynthesis and is used to model the exchange between the land and atmosphere of carbon dioxide, water, energy (latent and sensible heat) and momentum (Bounoua et al., 2000; Doraiswamy et al., 2004; Potter et al., 1993; Sellers et al., 1996). LAI depends on factors such as species composition, developmental and phenological stage, prevailing site conditions, environmental stresses, disturbance and management practices. LAI changes rapidly at the start and end of the growing season and shows variations from year to year dependent on variations in land management, disturbance, precipitation and temperature.

There is considerable uncertainty in measuring LAI both from the ground and from space (Baret & Buis, 2008; Gower, Kucharik, & Norman, 1999; Kussner & Mosandl, 2000; Weiss, Baret, Smith, Jonckheere, & Coppin, 2004). Moreover, the collection of ground-based measurements of LAI is costly and time consuming. Satellite based techniques provide a relatively cost-effective means to obtain frequent updates of LAI estimates for large areas. The effective time resolution of global satellite based LAI estimates and associated biophysical parameters is 8 days to a month and the spatial resolution is between 1 and 8 km (Gobron et al., 2005; Myneni et al., 2002).

Two common types of methods to estimate LAI from satellite data are widely used. The first type estimates LAI from an empirical relationship with a vegetation index (VI); the second uses a canopy radiative transfer (RT) model to link LAI to reflectance values collected at multiple solar and view angles. The VI-based methods use simple empirical relationships that vary with vegetation type (Los et al., 2000; Sellers et al., 1996). Other factors such as spatial variations in soil background reflectance are assumed dependent on land-cover type. The RT-based inversions are more complex and put a high demand on computing resources. The first canopy RT models were developed in the late 1970s and early 1980s (Kimes, Norman, & Walthall, 1985; Ross, 1981; Verhoef, 1984). Further complexity and realism to these canopy reflectance models or bidirectional reflectance factor (BRF) models was added over the past three decades; a suite of recent models is compared in Widlowski et al. (2007). LAI estimates based on inversion of RT models require a priori assumptions about for example leaf optical

* Corresponding author. Tel.: +44 1792 295144.
E-mail addresses: g.leonenko@swansea.ac.uk (G. Leonenko), s.o.los@swansea.ac.uk (S.O. Los), p.r.j.north@swansea.ac.uk (P.R.J. North).
[1] Tel.: +44 1792602810.
[2] Tel.: +44 1792295955.

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.07.012

properties, soil background reflectance and canopy structure to solve what is in essence an underdetermined set of equations. These properties are commonly derived from land-cover maps or plant functional type maps (Myneni et al., 2002). Hybrid methods are used as well; these methods use either a RT-based inversion if sufficient observations are available or a VI-based empirical relationship where observations are few and a RT-based inversion solution cannot be determined (Myneni et al., 2002).

A number of methods are used to invert RT models: conventional numeric optimization methods (Jacquemoud et al., 2000), inversions based on look up tables (LUTs; Combal et al., 2002; Darvishzadeh, Skidmore, Schierf, & Atzberger, 2008; North, 2002; Richter, Atzberger, Vuolo, Weihs, & D'Urso, 2009; Weiss, Baret, Myneni, Pragnere, & Knyazikhin, 2000), and methods based on machine learning such as neural networks and support vector machines (Atzberger, 2004; Bacour, Baret, Beal, Weiss, & Pavageau, 2006; Danson & Rowland, 2003; Fang & Liang, 2003; Gong, Wang, & Liang, 1999; Verrelst et al., 2012). While much research has been carried out to improve RT models and inversion techniques, only few studies exist that investigate the properties of the residual error and how these affect LAI retrievals. In particular, little attention is paid to the development of robust methods that perform well when errors or residuals between models and data are non-linear, not normal or asymmetrically distributed (Susaki, Hara, Kajiwara, & Honda, 2004). Other statistical estimation methods used for parameter estimation (but not necessarily biophysical parameter estimation) in large geophysical data sets are found in the environmental literature. These methods can be based on a hierarchical Bayesian approach; see for example Cressie and Johannesson (2008) for references to further work or can employ dimension reduction techniques and regularization of inverse estimation problems (Ruiz-Medina & Espejo, 2012, 2013).

Finding appropriate metrics for non-linear parameter estimation is essential in many statistical problems. Numerous diagnostic tests exist in a wide variety of research areas to examine model residuals such as independency, reversibility and other properties. The most commonly used measures of dependence and test statistics are convenient functions of correlation that are motivated by linear relations involving Gaussian processes. These measures tend to fail when variables are discrete, or when they face non-linear or non-Gaussian properties. In particular, the residuals of RT models are likely to have non-linearities and heterogeneities and therefore assumptions of Gaussian properties underlying the use of least squares estimation (LSE) are violated.

RT models describe the interactions of solar radiation with the vegetation canopy based on physical laws and probabilities of light being scattered, reflected and absorbed by optical elements in the canopy (leafs, twigs, branches and stems), understory and soil and thus provide a cause–effect relationship between scattering elements, their biochemical constituents, structure and reflectance (Jacquemoud et al., 2009). RT model inversions are based on the adjustment of input values of biophysical variables such as LAI, cover fraction, canopy shape, leaf area distribution through the canopy and leaf angle distribution (Prieto-Blanco, North, Barnsley, & Fox, 2009). The best match, usually in terms of the least-squares criterion between simulated BRF and measured BRF for various wave bands, leads to the most likely set of biophysical parameters. The inverse problem can be solved only if it is well posed (if a unique solution exists). In general the inverse problem is by nature an ill-posed problem, i.e. a set of solutions can lead to the same BRF. Measurement and model uncertainties will exacerbate this problem (Atzberger, 2004). The measurement uncertainties come from the noise associated with the sensor and the data processing required to transform the sensor raw output signal into BRFs (atmospheric effects, residual cloud effects and calibration errors; see e.g. Holben, 1986; Los et al., 2000). The model uncertainties come from the type of canopy architecture or the soil background reflectance assumed, and these may not be consistent with the actual surface conditions. Furthermore, the computation of the RT requires approximations yielding additional uncertainties in the model simulations.

In a previous paper (Leonenko, North, & Los, 2013), it was explored how the use of different cost functions, or statistical distance measures, could improve the estimation of biophysical parameters from simulated reflectance data. Over 60 statistical distances were obtained from the literature; these distances can be divided in three classes: information measures, M-estimates and minimum contrast approach. Numerical results, based on both simulated LUTs and simulated observations, showed that a number of distances measures work better than the commonly used LSE. Here, we use the same approach to retrieve LAI from MODIS surface reflectance data (MOD09GA). The LUT was derived with FLIGHT (North, 1996; North, Rosette, Suarez, & Los, 2010). FLIGHT simulates the scattering and absorption of solar radiation in vegetation canopies. It is based on Monte Carlo simulations of photon transport and it allows representation of complex vegetation structures as well as angular geometry. FLIGHT has been compared with other three-dimensional radiative transfer codes under the Radiative Transfer Model Inter-comparison (RAMI) framework (Widlowski et al., 2007, 2008).

The retrieval method involves three steps: 1) construction of the LUT (Prieto-Blanco et al., 2009), i.e. simulation of BRFs associated with a set of solar and view zenith angle configurations, biophysical (leaf and canopy) parameters and soil background reflectance values 2) minimization of the different cost functions between MODIS reflectance values and the LUT and 3) comparing the estimated and ground-measured LAI. Cost functions are evaluated in terms of how well the ground measured and satellite estimated LAI agree.

The LUT based retrievals with different cost functions are evaluated for a wide range of land-cover types that included conifer, broadleaf and mixed (conifer with broadleaf) forest sites (Section 2 and Appendix A). The retrieved LAI was also compared with the MODIS LAI product (MOD15A2).

The paper is organized as follows: In Section 2, the MODIS data and ground data from 11 sites are discussed. In Section 3, an explanation of the construction of the LUTs is provided. In Section 4, a brief summary of alternative distances is given, and in Section 5, the alternative distances are applied and the numerical results presented. Guidance for application of the alternative distances and conclusions are presented in Section 6. An evaluation of all cost functions in terms of estimating LAI can be found in Appendix A.

2. Data description

2.1. MODIS data

MOD09 (MODIS Surface Reflectance; (Vermote et al., 1997)) is a seven-band product computed from the MODIS Level 1B land bands 1 (620–670 nm), 2 (841–876 nm), 3 (459–479), 4 (545–565 nm), 5 (1230–1250 nm), 6 (1628–1652 nm), and 7 (2105–2155 nm). The product corrects top-of-the-atmosphere reflectance for atmospheric scattering and absorption and provides an estimate of the surface spectral reflectance for each of the seven bands as measured at ground level. We use the MOD09GA product, for the main part collected by the Terra instrument. The spatial resolution of this MODIS product is 1 km.

We used the MOD15A2 level-4 MODIS global leaf area index (LAI) product (Myneni et al., 2002). The temporal resolution of this product is 8 days and the spatial resolution is 1 km; data are projected on a Sinusoidal grid. The MODIS LAI estimation algorithm uses reflectance values in up to 7 spectral bands. The product was developed jointly by personnel at Boston University and the University of Montana under contract with the National Aeronautic and Space Administration.

2.2. Forest sites

Ground data from 11 forests, collected at the BigFoot and VALERI sites, were used to test the retrieval of LAI from MODIS reflectance