Leaf area index retrieval using gap fractions obtained from high resolution satellite data: Comparisons of approaches, scales and atmospheric effects

Alemu Gonsamo*

Department of Geography, University of Helsinki, Gustaf Hällströminkatu 2, P.O. Box 64, FIN-00014 Helsinki, Finland

A R T I C L E   I N F O

Article history:
Received 13 August 2009
Accepted 16 March 2010

Keywords:
Atmospheric correction
High resolution satellite data
Large scale LAI inversion
Scale effect
Vegetation index

A B S T R A C T

This study is aimed at demonstrating the feasibility of the large scale LAI inversion algorithms using red and near infrared reflectance obtained from high resolution satellite imagery. Radiances in digital counts were obtained in 10 m resolution acquired on cloud free day of August 23, 2007, by the SPOT 5 high resolution geometric (HRG) instrument on mostly temperate hardwood forest located in the Great Lakes – St. Lawrence forest in Southern Quebec. Normalized difference vegetation index (NDVI), scaled difference vegetation index (SDVI) and modified soil-adjusted vegetation index (MSAVI) were applied to calculate gap fractions. LAI was inverted from the gap fraction using the common Beer–Lambert’s law of light extinction under forest canopy. The robustness of the algorithm was evaluated using the ground-based LAI measurements and by applying the methods for the independently simulated reflectance data using PROSPECT + SAIL coupled radiative transfer models. Furthermore, the high resolution LAI was compared with MODIS LAI product. The effects of atmospheric corrections and scales were investigated for all of the LAI retrieval methods. NDVI was found to be not suitable index for large scale LAI inversion due to the sensitivity to scale and atmospheric effects. SDVI was virtually scale and atmospheric correction invariant. MSAVI was also scale invariant. Considering all sensitivity analysis, MSAVI performed best followed by SDVI for robust LAI inversion from high resolution imagery.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Land surface parameters retrieved from remote sensing are required for several applications such as: forestry, agriculture, landscape studies, land management, hydrological, meteorological and meso-scale weather circulation forecasts (e.g., Stern and Donald, 1961; Bounoua et al., 2006; GCOS, 2008; Gobron and Verstraete, 2008; Gonsamo Gosa, 2009). Among several land surface parameters, leaf area index (LAI) is one of the principal biophysical parameters in climate, weather, and ecological studies, and has been routinely estimated from remote sensing measurements (Asner et al., 2003). LAI is a part of the essential climate variables identified by the Global Climate Observing System (GCOS, 2006). LAI is a dimensionless value of the amount of foliage area of a vegetation canopy and is defined as one half the total radiation intercepting leaf area per unit ground horizontal surface area (Chen and Black, 1992; Gonsamo and Pellikka, 2008). The major physiological processes of vegetation including photosynthesis and evapotranspiration are determined by the vegetation biophysical parameters that describe the canopy structure as an exchange medium for energy and matter between the terrestrial ecosystem and atmosphere. Remotely sensed data offers a wide range of spectral, spatial, temporal, and angular resolutions to estimate LAI. Several numerical models require a continuous field of high spatial and temporal resolution LAI measurements due to heterogeneity and size of natural vegetation or agricultural patches, and the large seasonal dynamics of vegetation. To fulfill these needs, the LAI retrieval methods including the processing of the remotely sensed data are expected to be efficient and convenient for the modelers and end users.

There are several global initiatives for ground-based measurement validation networks, and global satellite products and intercomparisons for mapping LAI on continuous bases (Weiss et al., 2004; Myneni et al., 2002; Morissette et al., 2006; Garrigues et al., 2008). The retrieval of LAI from remote sensing data is usually done based on statistical/empirical relationships between spectral vegetation indices (SVI) and ground-based measurements (Huete, 1988; Baret et al., 1989; Myneni et al., 1995). While these relationships work well under a particular viewing and illumination geometry and for specific vegetation phenology, they produce inaccurate results when applied over the broad range of land cover types and optical and geometrical conditions normally encountered in satellite images. Another approach for canopy LAI retrieval is to invert a physically based canopy reflectance models as they describe the transfer and interaction of radiation inside the canopy based on physical laws to provide explicit connection between the biophys-
tical parameters and the canopy reflectance (e.g., Vohland et al., 2010). However, there are no ideal canopy reflectance models that can be directly used for LAI inversion over large areas as the inversion process is ill-posed by nature due to measurements and model uncertainties since different combinations of model parameters may correspond to almost identical spectra (Combal et al., 2002). Additionally, the selection of the parameters for the inversion of canopy reflectance models is complicated, and parameterization technique typically relies on the existence of experimental data collected at the site of interest. In general, the success of LAI estimation from remotely sensed data remains cumbersome and there is always a need to calibrate remotely retrieved parameters with ground-based measurements. This led to the cascades of methodological studies for LAI retrievals from remote sensing observation (Richardson and Wiegand, 1977; Baret et al., 1989; Myneni et al., 1995; Gutman and Ignatov, 1998; Combal et al., 2002; Asner et al., 2003; Bréda, 2003; Xiao and Moody, 2005; Jiang et al., 2006; Sprintsin et al., 2007).

Most of the LAI retrieval limitations associated with the empirical approaches can be addressed by using SVIs which normalize out the atmospheric, topographic, and scale effects. The normalized and scaled difference vegetation indices such as normalized difference vegetation index (NDVI, Rouse et al., 1974) and scaled difference vegetation index (SDVI, Jiang et al., 2006) minimize the atmospheric and topographic effects, and the scale effect, respectively. In view of soil background reflectance effect, Qi et al. (1994) put forward modified soil-adjusted vegetation index (MSAVI) with correction parameters which can be retrieved from the image scene. Thus, the general outcome of interpretation is that normalized difference SVIs remain the only practical approach to the global analysis of vegetation using remotely sensed data owing to partial cancellation of the bi-directional, atmospheric and other interfering effects on observed radiances (Gutman and Ignatov, 1998). Based on these premises, a number of studies analysed SVI for estimation of fractional vegetation cover (fF) aiming particularly to relate with LAI from satellite remote sensing with arguable convergence of results (Richardson and Wiegand, 1977; Baret et al., 1995; Gutman and Ignatov, 1998; Xiao and Moody, 2005; Jiang et al., 2006; Sprintsin et al., 2007). These studies ignored the sensitivity analysis of atmospheric, soil background and scale effects opposed to the major aim of finding LAI methodology with a lesser amount of parameterization. Moreover, the use of SVIs for gap extraction is limited often by the coarse resolution of satellite observations, which results in mixed pixels made up of more than one land cover types.

Therefore, it was hypothesised that the robustness of canopy gap fraction estimations can be improved using high spatial resolution remote sensing imagery. This paper first presents the comparative analysis of the algorithms used to retrieve LAI and its sensitivity for soil background, atmospheric and scale effects. The second part of this paper describes the validation of the retrieved LAI and interactive correction to optimize the algorithms. It should be noted that the LAI retrieved solely from SVIs is the ‘green leaf area index’ compared to the LAI reported in literature which is derived from optical field instruments as a structural parameter (Bréda, 2003). The green leaf area index derived from SVI is a photosynthetically active LAI which is strongly related with the canopy optical thickness, whereas LAI as a structural parameter from ground measurement represents ‘plant area index’ including both photosynthetic and non-photosynthetic canopy materials. In spite of this difference, ground-based LAI estimated from optical field instruments is being used as a ground-truth validation dataset even in recent studies (e.g., Ganguly et al., 2008; Houborg and Boegh, 2008).

In this regard, ground-based measurements are used as comparison dataset rather than as the ground-truth. Besides, the effectiveness assessment was extended by comparing the retrieved LAI with MODIS LAI product, and 1D Radiative Transfer Model inversion.

2. Site description and datasets

2.1. Site description

The study site is located in the Great Lakes–St. Lawrence forest in Southern Quebec, Canada. It is part of the Gatineau Park (Fig. 1), which is managed by the National Capital Commission (NCC) of Canada and centred at 45°30′N, 75°52′W. Gatineau Park is a vital ecological and conservation reserve, consisting of many endangered and rare wildlife and plant species, and is one of the largest and oldest Parks managed by NCC (NCC, 2005). More than 80% of the park’s approximately 36,000 ha is forested. The park is about 10 km by 50 km and is mostly temperate hardwood forest with a dominant overstorey of sugar maple (Acer saccharum Marsh.), and small patches dominated by American beech (Fagus grandifolia Ehrh.), trembling aspen (Populus tremuloides Michx.), and red oak (Quercus rubra L.). Small numbers of red maple (Acer rubrum L.), American basswood (Tilia americana L.), ironwood (Ostrya virginiana (Mill.) K. Koch), white ash (Fraxinus americana L.), black ash (Fraxinus nigra Marsh.), white birch (Betula papyrifera Marsh.), and black cherry (Prunus serotina Ehrh.) are also present.

2.2. Ground-based LAI measurements

20 m by 20 m plots were established in the Gatineau Park (n = 61) along two north-south oriented transects (Fig. 1), which their size was set by a semivariance analysis conducted on field and remote sensing data of two Quebec Ministry of Natural Resources (QMNR) permanent sample plots (Butson and King, 1999). Both transects have been part of ongoing research on monitoring forest damage, structure, health and succession following the ice storm of

Fig. 1. Location map of study area, a SPOT image (colour infrared composite) taken July 23, 2007, showing the Gatineau Park, Canada.
1998, and were sites of LAI determination from airborne imagery (Gonsamo et al., in press). A subset \((n = 54)\) of previously inventoried forest plots that could be easily found in 2007 were selected for LAI measurements. Plot corners were surveyed using differential GPS to provide positional accuracy on the order of \(\pm 1\ m\).

The ground LAI measurements were collected using digital hemispherical photography between August 10th and 20th, 2007. The photographs were acquired using a high resolution (8 mega pixels) Nikon Coolpix 8800 VR digital camera equipped with a fish-eye Nikon FC-E9 lens adapter (Nikon Inc., Japan). In total, five photographs were acquired in each study plot, one at each corner and one at the centre. All photographic procedures are described in Gonsamo et al. (in press). LAI was computed for the range of 0–60° view zenith angle to reduce the growing effects of mixed pixels near the horizon, which result from light scattering and coarse resolution. All photographs were analysed using CAN_EYE software (http://www.avignon.inra.fr/can_eyes). CAN_EYE allows interactively classifying RGB colours into vegetation elements and background. At the end of this process a binary image, background vs. vegetation elements (including both green and non-green elements) is obtained to retrieve gap fraction. LAI is inverted from gap fraction assuming an ellipsoidal distribution of leaf inclination (Campbell, 1986), and using the lookup table (LUT) from zenithal variation of gap fraction (Weiss et al., 2004). These LAI estimation procedures from hemispherical photography using CAN_EYE software were proven to be a robust approach (Demarez et al. 2008; Garrigues et al., 2008). Further details on LAI derivation from gap fraction using hemispherical photographs can be found in Weiss et al. (2004). For simplicity, the effective LAI, which assumes random foliage distribution, was used in this study and is hereafter referred to as LAI. Effective LAI is a part which directly contributes to the canopy interception of incident radiation and therefore well related with SVIs. Plotwise LAI ranged from 2.6 to 5.7 with an average value of 4 and was close to normal distribution along the two sampling transects of Gatineau Park.

2.3. Remote sensing data pre-processing

Radiance in digital counts in the green (500–590 nm), red (610–680 nm), near infrared (NIR) (780–890 nm) and short-wave infrared (SWIR) (1580–1750 nm) wavelength regions were obtained in 10 m resolution acquired on cloud free day of July 23, 2007, 10:49 local time by the SPOT 5 High Resolution Geometric (HRG) instrument for 60 km x 60 km image swath. The sun zenith (\(\theta_s\)) and azimuth (\(\phi_s\)) angles were 30° and 140°, respectively. The view zenith angle (\(\theta_v\)) was 20°. The SWIR band was originally acquired at 20 m pixel size and resampled into 10 m in order to match the other bands. The SWIR Band was used only for classification of land cover types (Section 3.2).

2.3.1. Geometric correction

The image was orthorectified using a 3 in. digital elevation model (DEM) obtained from Canadian Digital Elevation Data (CDED) (www.geobase.ca). The digital source data for CDED at scales of 1:50,000 were extracted from the hypsographic and hydrographic elements of the National Topographic Data Base and various scaled positional data acquired from the provinces and territories. The DEM was resampled based on the nearest-neighbour interpolation into 10 m to match the SPOT HRG image resolution. The geometric correction model was based on the SPOT XS sensor colinearity equation. The SPOT image was registered using 139 ground control points collected from orthorectified digital aerial photography and national road network of Canada (www.geobase.ca) into UTM zone 18 with a NAD83 datum. The X and Y root mean square deviation was below 6 m (0.6 pixels).

2.3.2. Radiometric calibration

Satellite radiances were obtained for SPOT HRG quantised calibrated pixel values in 8-bit digital numbers (DNs) in four spectral bands. The DNs were converted to at-satellite radiance based on:

\[
L_{\text{sat}} = \frac{\text{DN}}{\gamma}\sin(\theta_v)
\]

where \(L_{\text{sat}}\) is the satellite radiance in W m\(^{-2}\) sr\(^{-1}\) m\(^{-1}\) for band \(\lambda\), and \(\gamma\) is band specific absolute gain. Satellite radiances were converted to top of the atmosphere (TOA) reflectance following:

\[
\rho_{\text{TOA}} = \frac{\pi L_{\text{sat}}}{(\text{ESUN} / \lambda)^2 \cos(\theta_v)}
\]

where \(\rho_{\text{TOA}}\) is TOA reflectance for band \(\lambda\), \(d\) is the earth–sun distance in astronomical units, and \(\text{ESUN}\) is the mean solar exo-atmospheric irradianc in W m\(^{-2}\) sr\(^{-1}\) m\(^{-1}\) for band \(\lambda\).

2.4. MODIS LAI product

The MODIS Collection 5 LAI product was acquired in a form of HDF subset from the Warehouse Inventory Search Tool (WIST) client for searching and ordering earth science data from various NASA and affiliated centres (https://wist.echo.nasa.gov/). The subset LAI product was in 1 km resolution with 8 days composite based on the maximum FPAR measurement. The HDF subset has been re-projected and they matched the study area layout. The available LAI was downloaded from date 23/07/2007 (20–27/07/2007) to cover the same period of ground and SPOT data acquisitions. The final MODIS LAI and associated quality layers are clipped to preferred study area covering 40 km by 40 km (1600 MODIS pixels). MODIS LAI product quality flags were extracted for the subset as the time period of the data acquisition coincided with the peak of growing summer period. Under optimal circumstances, a look-up-table (LUT) method is used to achieve inversion of a three-dimensional radiative transfer model for MODIS LAI retrieval (Myneni et al., 2002). When this method fails to optimize a solution, a back-up algorithm based on a relationship between the normalized difference vegetation index (NDVI) and LAI is utilised (Myneni et al., 1995). The Quality Flags serve to determine the origin of the calculated value or mark pixels where no retrievals were made. In the subset of 1600 MODIS pixels used in this study, 87% were retrieved using the main LUT algorithm with or without saturation, 1% were retrieved with main LUT algorithm containing significant cloud, 6% were retrieved using the back-up algorithm and 6% were fill values containing non-vegetated land covers. Only pixels retrieved with main algorithm without cloud contamination were used in further analysis.

2.5. Simulated data using PROSPECT and SAIL model

The combined PROSPECT leaf model (Jacquemoud and Baret, 1990) and the SAIL canopy model (Verhoef, 1984) were used to simulate canopy reflectance data in order to test the robustness of the methodologies developed in this study and to derive correction factors as described in the latter sections. The PROSPECT model simulates the leaf optical properties from visible to mid infrared as a function of leaf characteristics. The SAIL model is a turbid medium model with the assumption of homogeneous semi-infinite medium crop canopy with Lambertian reflecting leaves. The canopy reflectance has been simulated for two wavelength bands centred around 645 and 834 nm, corresponding to red (R) and NIR bands of SPOT HRG, respectively, for a range of plausible input parameters. PROSPECT uses the following input parameters: chlorophyll content \(a\) and \(b\) content (\(C_{CHL}, \mu g\ cm^{-2}\)), dry biomass content (\(C_{BM}, g\ cm^{-2}\)),
equivalent leaf water content \((C_w, \text{g cm}^{-2})\), and mesophyll structure parameter \((N_i)\); and SAIL uses: LAI, leaf angle distribution (LAD), soil reflectance \((R_s)\) and external parameters like View, Sun zenith \((\theta_v, \theta_s)\) and azimuth \((\phi_v, \phi_s)\) angles, hot-spot size parameter \((H_s)\), and ratio of diffuse to total incident radiation \((D_F)\). For this analysis, the range of input LAI was between 0.25 and 9 by interval of 0.25, and 12 as the largest LAI value. Three different soil backgrounds ranging from dark to bright were specified from Bowker et al. (1985). The other input parameters were kept constant: \(N = 1.55, C_{b0} = 34.24, C_w = 0.0137, C_m = 0.0045, \text{spherical LAD, } \theta_v = 0, \theta_s = 30, \phi_v = 0, \phi_s = 0, \text{HS} = 0, \text{and } D_F = 0.\) The input parameters which have been used in the analysis are within the realistic range and adopted in other simulation studies (Jacquemoud and Baret, 1990; Jacquemoud et al., 2000). Several studies have investigated the effect of changing one input variable at a time for vegetation characterization using PROSPECT and SAIL models (e.g., Jacquemoud et al., 2009). Therefore, in this study, the evaluation of the methods was restricted only to the changes of LAI values and soil reflectances which are required for determination of maximum LAI and background reflectance determination for the LAI estimation approaches as explained in the latter sections.

3. Methodology

3.1. Atmospheric correction

It was anticipated that atmospheric correction would be a crucial step for any quantitative analysis of satellite observations. Two procedures of atmospheric corrections and TOA compared are (1) apparent reflectance model, (2) absolute correction using four dark object subtraction (DOS) methods, and (3) absolute correction based on radiative transfer model. The main aim of the atmospheric correction is to remove atmospheric effects on remotely sensed DN in order to get ground surface reflectance. The first procedure is explained in Section 2.3.2, as retrieval method for TOA reflectance estimate. Apparent reflectance model corrects only the effects caused by the solar radiance and sun zenith angle (Eq. (2)), ignoring the effects caused by atmospheric scattering and absorption. Subsequently, procedures which further account for atmospheric scattering and also atmospheric absorptions are presented.

3.1.1. DOS method

Several techniques have been developed for image-based atmospheric correction by assuming cloud-free images and uniform path radiance (Chavez, 1996; Song et al., 2001). Although more detailed atmospheric corrections are required for generic studies using in situ measurements of atmospheric parameters, image-based methods usually provide sufficient accuracy. It is often difficult to quantify atmospheric data during the satellite overpass time to apply pixel based variable path radiance corrections. Dark object subtraction (DOS) is the simplest yet most widely used image-based atmospheric correction approach. The first step of DOS method for atmospheric correction is to convert remotely sensed DN values to at-satellite radiance based on band specific gain coefficients which comes along the image header (Eq. (1)). Then the next step is to convert apparent at-satellite radiance to surface reflectance which involves the correction of effects caused by solar and viewing angles, and the atmosphere. DOS approach corrects the path radiance caused by atmospheric molecules and aerosols. Song et al. (2001) studied the accuracy of four different DOS based methods with respect to classification and change detection. They found that the best overall accuracies were achieved by using the simpler DOS method, rather than the more complex atmospheric corrections that combine both atmospheric models and the dark object concept. The same nomenclatures as Song et al. (2001) have been adopted in this study following the improvements of DOS approaches as: DOS1, DOS2, DOS3, and DOS4. The basic assumption for all four DOS approaches is that within the image some pixels are in complete shadow and their radiances received at the satellite are only due to atmospheric scattering (path radiance).

In this study, techniques originally developed by Chavez (1996) were adapted with minor modifications. All four DOS approaches (DOS1, DOS2, DOS3 and DOS4; Table 1; Song et al., 2001) calculate surface reflectance from Eq. (3) using different assumptions for atmospheric transmittance along the path from the ground surface to the sensor \((T_v)\), atmospheric transmittance along the path from the sun to the ground surface \((T_s)\), and downwelling spectral irradiance at the surface due to scattered solar flux in the atmosphere \((E_{DOWN})\). If the sky irradiance is assumed isotropic and only atmospheric scattering and absorption are considered, the general equation for describing atmospheric interactions and retrieving a Lambertian surface reflectance can be expressed as:

\[
\rho_s = \frac{\pi (L_{s|s} - L_{p|s})}{T_v ((E_{SUN} / d^2) \cos(\theta_s) T_s + E_{DOWN})}\]

(3)

where \(\rho_s\) absolute surface reflectance and \(L_p\) is estimated path radiance. The path radiance is estimated using Eq. (4) assuming 1% surface reflectance for dark object. This assumption explains the fact that very few targets on the Earth’s surface are absolute black, so an assumed 1% minimum reflectance is better used than 0%. For this study, the dark object radiance for all bands was retrieved from the deepest and clearest water found in the image scene. The path radiance is estimated as:

\[
L_{p|s} = L_{s|s} - \frac{0.01 [(ESUN|s|/d^2) \cos(\theta_s) T_s + E{DOWN}|s|]}{\pi} T_v
\]

(4)

DOS1 assumes no atmospheric transmittance loss, the transmittance along the sun-target and target-sensor paths is assumed to be 100% \((T_v = 1)\), and the downwelling spectral diffuse irradiance is ignored \((E_{DOWN} = 0)\). In DOS2, \(\cos(\theta_v)\) and \(\cos(\theta_s)\) are used as a substitute for \(T_v\) and \(T_s\) respectively, whereas EDOWN is the same as DOS1. Unlike, Chavez (1996) and other studies who mainly used DOS approaches for nadir viewing Landsat TM images, \(T_v\) is approximated in DOS2 by \(\cos(\theta_s)\) as a surrogate for target-sensor transmittance computation.

For DOS3, Song et al. (2001) recommended a more realistic interpretation of path transmittance assuming Rayleigh scattering atmosphere for \(T_v\) and \(T_s\) (Table 1). Optical thickness for such an atmosphere is defined as:

\[
\tau_r = 0.008569 \lambda^{-4}(1 + 0.0111 \lambda^{-2} + 0.00013 \lambda^{-4})
\]

(5)

where \(\lambda\) is wavelength in \(\mu\text{m}\). For DOS3, EDOWN for Rayleigh atmosphere was estimated by 65 atmospheric radiative transfer code as zero aerosol optical depth at 550 nm (Vermote et al., 1997). DOS4 is the same as Song et al. (2001) DOS4 model, i.e., adding atmospheric aerosols for path transmittance computation assuming isotropic sky radiance. For path transmittance \(T_v\) and \(T_s\) in Table 1, \(\tau\) is solved iteratively as defined by Song et al. (2001) as:

\[
\tau = -\cos(\theta_s) \ln \left(1 - \frac{4\pi L_{s|s} - 0.01(ESUN|s|/d^2) \cos(\theta_s) T_s + E{DOWN}|s|}{ESUN|s|/d^2 \cos(\theta_s)} \right) / \pi
\]

(6)
3.1.2. Radiative transfer model

Radiative transfer code developed by Vermote et al. (1997) called 6S (second simulation of the satellite signal in the solar spectrum) was used. 6S first produces three atmospheric correction coefficients called a, b, and c that convert the measured at sensor radiance in each band to atmospherically corrected reflectance (acr). Atmospherically corrected reflectance is then obtained using the following formulas:

\[ acr = \frac{y}{1 + cy} \] (7)

where

\[ y = a \cdot sat - b \] (7.1)

The middle latitude summer standard atmospheric model available in 6S code was utilised. This was due to the fact that the atmospheric characteristics data were not available for the time of satellite overpass and studies have shown that the effect of the atmospheric correction from standard atmospheric model is comparable to that obtained from an in situ model (Moran et al., 1992). The continental standard aerosol model was used to define aerosol type. Aerosol optical depth (AOD) at 550 nm was interpolated from AOD (440) and Angstrom exponent \( \alpha \) (440–675 nm) by:

\[ AOD(550) = \frac{AOD(440)}{[550/440]^{0.0067(440–675)}} \] (8)

The AERONET AOD at 550 nm was first integrated with spatiotemporal constraints that the SPOT scene is located to two nearest available AERONET sites and the measured AOD is computed from the closest time \( t \) satellite overpass. The two AERONET sites were: CARTEL, Sherbrook, Canada (45°22′44″ N, 71°55′51″ W, and 300 m elevation); and Egbert, Canada (44°13′33″ N, 79°45′00″ W, and 264 m elevation). No bi-directional reflectance distribution function (BRDF) correction was applied.

3.2. Land use map

The 6S and geometrically corrected SPOT image was used to generate land use map of the study area. The classification method was based on a supervised algorithm using a conventional maximum likelihood procedure. The training sites (polygons) were selected visually from the scene and the data was classified into four general classes initially aimed to extract vegetated areas. Apart from vegetation areas, the land use and land cover classes include: urban areas (towns, buildings and roads), water, and wet areas (shallow water and wet land). The selected pixels of training sites (an average per class of 45,000 pixels) were evaluated for class separability using Transformed Divergence (TD) (Jensen, 1996). The scale of the TD distance of the four classes was above 1999. Fig. 2 shows the final land use map of the study area. For purposes of LAI retrieval, classes such as urban, water and wet areas were masked in the SPOT image prior to inversion as discussed in the following sections. All the analysis both from SPOT and MODIS datasets are restricted to the extent of the classification image shown in Fig. 2.

3.3. Retrieval of LAI

3.3.1. Methods to retrieve fractional vegetation cover from vegetation indices

Fractional vegetation cover \( f_c \) is the percentage of vegetation cover projected vertically onto the ground or in the context of a remotely sensed image, the \( f_c \) values represent the percentage of vegetative cover present in each pixel or in a unit area (Gonsamo et al., in press). The most frequently applied SVI for agricultural and ecological applications is the normalized difference vegetation index (NDVI, Eq. (9)), which has been shown to be widely usable for vegetation analysis (Rouse et al., 1974). Several studies have successfully estimated gap fraction of \( f_c \) from satellite remote sensing using the NDVI (Richardson and Wiegand, 1977; Baret et al., 1995; Cutman and Ignatov, 1998; Xiao and Moody, 2005; Jiang et al., 2006; Sprintsin et al., 2007).

\[ \text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}} \] (9)

Jiang et al. (2006) demonstrated scale invariant \( f_c \) retrieval approach from difference vegetation index (DVI, Eq. (10)) scaled using bare soil and dense vegetation DVI values. The study showed that the derivative vegetation index inferred from the linear spectral mixing of red and NIR reflectances which in turn are the scaled difference of DVI computed from bare soil and dense vegetation is more suitable and robust approach to compute \( f_c \) over heterogeneous surfaces (Jiang et al., 2006).

\[ \text{DVI} = \text{NIR} - \text{R} \] (10)

Because of diversity both in vegetation cover and soil optical properties in natural environment, soil background have great effect on the SVIs. Hute et al. (1988) introduced soil-adjusted vegetation index (SAVI), which requires a parameter, \( L \). The determination of \( L \) needs some known information, such as vegetation cover. For different types of surface materials, one fitted value of \( L \) is not suitable in natural environment. In view of this, Qi et al. (1994) put forward modified soil-adjusted vegetation index (MSAVI, Eq. (11)). MSAVI replaces a constant, \( L \), with \( L \) being the value that can be adjusted with the variation of vegetation cover. Given a possible range of \( L \) in SAVI from zero to infinity, the expression can be derived with mathematical induction method as follows:

\[ \text{MSAVI} = \frac{2\text{NIR} + 1 - \sqrt{[2\text{NIR} + 1]^2 - [8(\text{NIR} - \text{R})]^2}}{2} \] (11)

Earlier modification of SAVI by Baret et al. (1989) proposed Transformed SAVI (TSAVI) as a measure of the angle between the soil line and the line which joins the vegetation point and the intercept of the soil line. The vegetation point is defined as the vegetation reflectances in the red and NIR bands. This study has shown that TSAVI is exponentially related with LAI based on experimental data.
MSAVI unlike TSAVI or SAVI, can be calculated directly without a scene dependent correction factor yet minimizing soil background variations.

The approaches for retrieval of $f_c$ from NDVI, DVI and MSAVI are presented in Table 2. The $f_c$ retrieval is based on the assumption that the SVI value of a given pixel is the linear combination of SVI values of green vegetation and bare soil, weighted by their relative proportions. NDVI$_{\text{back}}$ and DVI$_{\text{back}}$ are the SVI values for the bare soil (LAI = 0); and NDVI$_{\text{∞}}$, DVI$_{\text{∞}}$, and MSAVI$_{\text{∞}}$ are the SVI values for infinite LAI (asymptotic value) of a pure vegetation pixel, respectively. As discussed above, MSAVI was deemed to minimize soil background influence on the retrieval of $f_c$ thus its bare soil value was not incorporated in the equation (Table 2).

### Table 2

<table>
<thead>
<tr>
<th>Name</th>
<th>Formulation of $f_c$</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gutman and Ignatov</td>
<td>($\text{NDVI} - \text{NDVI}<em>{\text{back}}$)/($\text{NDVI}</em>{\text{∞}} - \text{NDVI}_{\text{back}}$)</td>
<td>Dense vegetation mosaic-pixel model</td>
<td>Gutman and Ignatov (1998)</td>
</tr>
<tr>
<td>Carlson and Ripley</td>
<td>($\text{(NDVI} - \text{NDVI}<em>{\text{back}})/\text{(NDVI}</em>{\text{∞}} - \text{NDVI}_{\text{back}})$)^2</td>
<td>Based on radiative transfer model</td>
<td>Carlson and Ripley (1997)</td>
</tr>
<tr>
<td>Baret et al.</td>
<td>$(1 - (\text{NDVI} - \text{NDVI}<em>{\text{back}})/\text{(NDVI}</em>{\text{∞}} - \text{NDVI}_{\text{back}}))^{0.175}$</td>
<td>Based on semi-empirical relationship between vertical gap and NDVI</td>
<td>Baret et al. (1995)</td>
</tr>
<tr>
<td>SDVI</td>
<td>($\text{DVI} - \text{DVI}<em>{\text{back}}$)/($\text{DVI}</em>{\text{∞}} - \text{DVI}_{\text{back}}$)</td>
<td>Scanned difference vegetation index from DVI</td>
<td>Jiang et al. (2006)</td>
</tr>
<tr>
<td>MSAVI</td>
<td>MSAVI/MSAVI$_{\text{∞}}$</td>
<td>Soil background invariant $f_c$ determination from MSAVI</td>
<td>This study</td>
</tr>
</tbody>
</table>

NDVI$_{\text{back}}$ and DVI$_{\text{back}}$ are the SVI values for the bare soil (LAI = 0); and NDVI$_{\text{∞}}$, DVI$_{\text{∞}}$, and MSAVI$_{\text{∞}}$ are the SVI values for infinite LAI (asymptotic value) of a pure vegetation pixel, respectively.

Foliage element clumping index (CI) measures the degree of aggregation of individual leaves into shoots, tree crowns and vegetation patches (Gonsamo and Pellikka, 2009). The retrieval of aggregation of canopy elements at various scales from remotely sensed data is practically not feasible. Clumping in the forest stand is typically related to stand density, which is expressed as the number of trees per unit area and is sometimes taken as a coefficient (Daniel et al., 1979). Nevertheless, measuring density by remote sensing is not easy and requires more rigorous analysis as even high resolution sensors have difficulty in detecting single trees in the stand and gaps within tree crown (Bunting and Lucas, 2006). There have been some studies mostly concentrated on the estimation of clumping index with multi-angular remotely sensed data (Lacaze et al., 2002; Chen et al., 2003, 2005) and detailed analysis using optical field instruments (Leblanc et al., 2005; Walter et al., 2003; Gonsamo and Pellikka, 2009). However multi-angular remote sensing data is not always easily available. The relationships and the meaning of CI estimates from various studies have been always ambiguous. In this study, CI was computed from SPOT HRG derived NDVI-based gap fraction in the same manner as finite-length logarithmic averaging technique (Lang and Xiang, 1986) as:

$$CI = \frac{\ln[1 - (\text{NDVI} - \text{NDVI}_{\text{back}})/\text{(NDVI}_{\text{∞}} - \text{NDVI}_{\text{back}})]}{\ln[1 - (\text{NDVI} - \text{NDVI}_{\text{back}})/\text{(NDVI}_{\text{∞}} - \text{NDVI}_{\text{back}})]}$$

where $[1 - (\text{NDVI} - \text{NDVI}_{\text{back}})/\text{(NDVI}_{\text{∞}} - \text{NDVI}_{\text{back}})]$ is the canopy mean gap fraction averaged over predefined analysis grid size, and $\ln[1 - (\text{NDVI} - \text{NDVI}_{\text{back}})/\text{(NDVI}_{\text{∞}} - \text{NDVI}_{\text{back}})]$ is the mean of logarithmic gap fractions of all SPOT HRG 10 m pixels within predefined analysis grid. This is one of the three commonly applied CI computation methods from optical data, the remaining two being one based on gap size distribution from ground measurements (Leblanc et al., 2005) and the other based on multi-angular remote sensing data (Lacaze et al., 2002; Chen et al., 2003, 2005). The finite-length averaging approach based on large scale gap fraction is the first attempt to map CI at larger scale using single view angle remote sensing data. This approach assumes that vegetation elements (tree crowns in this case) are locally randomly distributed, i.e., Poisson law is valid at the size of selected segment (SPOT HRG pixel size). The gap size here in Eq. (14) is defined as the dynamic range of gap fraction obtained from NDVI. The analysis grid size is 1 km by 1 km, equivalent of 1 MODIS pixel and containing 10,000 SPOT HRG pixels. Therefore, the CI is computed using 10,000 SPOT pixels by overlaying analysis grid of MODIS pixel size. As the logarithm of zero is undefined, for ‘empty pixels’ where vegetation is full field of view, to allow the calculation of the logarithm of gap fraction, a new value of gap > 0 is computed, using a local Poisson model by
assigning the maximum LAI to arbitrary maximum value of 12 in Eq. (12) (Gonsamo et al., in press).

The physical meaning of the CI derived from SPOT serves as the large-scale clumping related to tree distribution in the forest. NDVI has a capability for partial self-cancelation of bi-directional, atmospheric, soil, shadow and other interfering effects (Gutman and Ignatov, 1998). Therefore, the only difference is from the difference of tree distribution as other effects are either partially cancelled or assumed to be not optically varying within 1 MODIS pixel size. The resulted CI from all vegetated grids ranges from 0.4 to 0.999 excluding 0 from open fields. In real vegetation stands, CI typically ranges between 0.50 and 0.97 (Chen et al., 2003) and in good agreement with our result. The mean CI from this study was 0.7 which is in very close agreement with Leblanc et al. (2005) who found the mean of 0.72 from closed deciduous forest over Canada. The calculated CI is not expected to be intrinsically related with the magnitude of NDVI. However, the CI is computed from the pixels which contain only forest and agricultural fields, therefore the higher NDVI pixels are from forest sites which are expected to be uniform and the CI is close to unity.

There is also compelling potential that a better understanding of the relationship between clumping and density would be a convenient approach for extinction coefficient determination since stand density determination by remote sensing is likely to be easier than direct determination of the leaf angle distribution and the angle with which light transverses the canopy (required for calculating the extinction coefficient). In this regard, CI was related to the extinction coefficient. CI was also further related with scale dependency of NDVI and other SVIs-based LAI retrieval from remotely sensed image. Furthermore, the CI is used to correct the computed leaf area index by dividing LAI with CI and the resulting LAI is compared with MODIS LAI product retrieved from 3D Radiative transfer model. LAI reported in this study is not corrected unless stated ‘clumping corrected’ otherwise, as the method is experimental and yet to be validated with more analysis.

4. Results and discussion

4.1. Background and saturated SVI determination for LAI retrieval

Background reflectance can be a significant contributor to the canopy reflectance signal and care should be taken to avoid confounding effects among $f_C$-SVI (e.g., NDVI, DVI and MSAVI)-LAI especially at low vegetation coverage with varying soil optical properties. The effect of NDVI∞ and NDVIback determinations for the LAI algorithm used in this study were first analysed in a hypothetical dataset. The effect is the same for all NDVI, DVI and MSAVI algorithms except that MSAVI does not require background values (Table 2). 400 cases of random hypothetical NDVI values were generated spanning from 0 to 1, which are theoretically expected values of vegetated environments. Eq. (13) and $f_c$ cover retrieval approach shown in Table 2 were applied on a hypothetical dataset to retrieve LAI. The assessment on the NDVI is presented on Fig. 3 from hypothetical data and on Figs. 4 and 5 from the real dataset. As shown in equations listed in Table 2, the NDVI and DVI background values are divided both from the nominator and denominator therefore have minor effects. Fig. 3(B) shows that the background values have insignificant effect for the LAI estimations. However, all the NDVI values below the cut-off bounds used for the background value in the histogram of the NDVI will result in 0 LAI. For the LAI value calculated between the NDVI∞ and NDVIback cut-off bounds, LAI is systematically higher by 0.2354 for 0%, 0.2139 for 1% and 0.1445 for 5% compared with 10% cut-off bound NDVIback. All these are in acceptable range particularly for larger LAI values.

The variation of saturated NDVI value changes considerably the LAI estimates particularly for higher values whether background NDVI is the same or different (Fig. 3(A) and (C)). The lower the saturated NDVI, the higher the LAI estimates and the more measurements will be assigned to maximum LAI those which have NDVI values above the saturated value. The difference obtained for LAI values calculated using the 100% and 99% cut-off bounds as a saturated value is below 1%. However, the contrast between the cut-off bounds of 100%, and 95% or less is directly proportional to the LAI values (e.g., LAI estimated 3 and 4.5 using 100% cut-off bounds were estimated as 4 and 9.3, respectively using 95% cut-off bounds, Fig. 3(C)). In general, from a theoretical point of view, for the equations listed in Table 2, the higher cut-off bounds are used as a saturated NDVI, MSAVI and DVI values, the lower LAI estimates will be.

Subsequently, the effect of the background and saturated NDVI determination was analysed on the real dataset. As the one of the objectives of this study is to develop image-based robust algorithms, the focus was set on analysing the determination of both background and saturated values from the single date image scene. Some studies have acquired both parameters from independent field measurements which are usually not readily available and practical; others used theoretically fixed value from radiative transfer model (Carlson and Ripley, 1997); and the common practice is using the time series NDVI analysis (Gutman and Ignatov, 1998). However, leaf optical properties can vary during the year independently of LAI. Using theoretically fixed, or other values which are derived from independent dataset are very arguable as the calcu-
related values can vary due to atmospheric effects (Figs. 4 and 5), sun-view geometry and other factors which are not intrinsically related with the vegetation and soil conditions on the ground. There is also compelling reason that both parameters should be determined from the data itself, since they are specific to the region and season due to their dependency on the soil and vegetation types and vegetation chlorophyll content (Price, 1992). For instance, in deserts, background SVI can be retrieved reliably, and in evergreen vegetation saturated SVI can be retrieved reliably throughout the seasons (Gutman and Ignatov, 1998). This indicates that from measurements which are made in the peak of the growing season as this study, saturated SVI can be retrieved accurately while background values are less important as shown in Fig. 3. Therefore, studies have used the upper and lower 1–5% cut-off bounds from the NDVI histogram as NDVI_{∞} and NDVI_{back} (Sellers et al., 1996; Tang et al., 2007; Gonsamo et al., in press). The histogram of the NDVI is affected by the mixture of soil, vegetation and other land cover types which are present in most pixels resulting in ‘mixed pixels’.

Fig. 4 shows the NDVI cut-off bounds computed from 5 atmospheric correction methods and with TOA reflectance applied on the histogram of the whole image with and without masking non-vegetated pixels (see Fig. 2, for whole image scene analysis extent and mask). The atmospheric correction has considerable effect on the overall NDVI calculations while all the DOS approach resulted in very close values. The variations of the saturated values are relatively flat along the 90–99% cut-off bounds due to the fact that the forested area is in its growing peak. Due to high weight of the vegetated area (Fig. 2), the lower NDVI cut-off bounds for 1–5% are much higher than what is reported in other studies (0.02–0.05, Sellers et al., 1996; Gutman and Ignatov, 1998; Tang et al., 2007). Here again, Fig. 4(A) and (B) show that the saturated NDVI can be retrieved more accurately when the vegetation is in peak of growing season, the conclusion which is in good agreement with Gutman and Ignatov (1998). Atmospheric corrections have considerable effects on the calculation of NDVI in general and on NDVI_{∞} and NDVI_{back} in particular. MSAVI_{∞} has similar sensitivity for atmospheric correction (not presented for the sake of brevity) but SDVI normalizes out the effect and therefore has no influence on the final LAI calculation (see the next section). NDVI derived from atmospherically uncorrected reflectance values (TOA) is lower compared to corrected value (Fig. 4). This is due to the fact that scattering tends to increase the amount of red radiation received by the satellite as red is more readily scattered in the atmosphere than NIR.

The LAI calculated from varying atmospheric correction methods and NDVI_{∞} and NDVI_{back} values are presented for transect 1 and 2 in Fig. 5(A) and (B). The errors induced due to the determination of NDVI_{∞} and NDVI_{back} variations are higher than the atmospheric effect on LAI calculations. The [100%, 0%] in Fig. 5 represent the minimum and maximum NDVI values after masking out the build-up and wet areas whereas the other upper and lower 1–10% cut-off bounds were calculated from the entire analysis extent of the image. The mean ground LAI measurements were
used for the evaluation of estimated LAI. In both transects, consistently, the LAI estimated from 6S atmospheric correction method with [100%, 0%] outperformed all other options (Fig. 5) whereas all the other cut-off bounds calculations and atmospheric corrections overestimated the LAI values. The [100%, 0%] are more reasonable to use after carefully classifying the image into vegetated (including open agricultural fields) and non-vegetated classes. This highly reduces the cut-off bound dependency on the histogram weight caused by high prevalence of very low or very high density of vegetation, and the heterogeneity of land cover types, and the uncertainty caused by the atmospheric correction. Hereafter, the [100%, 0%] cut-off bounds as saturated and background values from the image masked out for non-vegetated areas were used unless stated otherwise.

4.2. The effect of atmospheric correction on LAI retrieval

The atmospheric correction is probably the most critical element for any quantitative remote sensing assessments. Remotely sensed reflectance datasets are not free of errors and the transformation of sensor DN counts into surface reflectance values will likely be associated with some extent of uncertainty. The conversion of radiance data into surface reflectance in the visible and NIR spectrum is typically most sensitive to the aerosol optical depth and type of aerosol model (Vermote and Vermeulen, 1999). The non-linear effect of different atmospheric corrections on SVIs has strong effect on the individual pixel, and on background and saturated NDVI values.

The five atmospheric correction algorithms and TOA were applied to obtain the reflectance of SPOT HRG imagery over the study area. The reflectance of the visible bands (green and red) by every algorithm is reduced considerably in contrast with that of TOA (not presented here for the sake of brevity). The non-systematic differences among the four bands of the reflectance obtained from the five atmospheric correction algorithms and TOA do affect the ratio computation of SVIs such as NDVI, MSAVI and probably SDVI. All NDVI and MSAVI based methods are highly affected by atmospheric correction (Fig. 6(A)–(C) and (E)). SDVI is
not affected by the atmospheric correction (Fig. 6(D)). For all NDVI-based LAI estimation methods, the histograms of 6S corrected were very close to TOA. The maximum LAI difference averaged over all the image was occurred among MSAVI based methods (21.45%), followed by approximately the same value for NDVI-based methods (21.23%) and lastly SDVI based method (3.2%). SDVI appeared to be invariant for varying atmospheric correction methods. Nevertheless, as shown in Fig. 6(F) and the following sections, the differences in magnitude obtained among the methods are higher than the effect of atmospheric corrections. However, there was very strong correlation of estimated LAI among the three categories of methods, i.e., Gutman and Ignatov-SDVI ($R^2 = 0.94$), Gutman and Ignatov-MSAVI ($R^2 = 0.95$), and SDVI-MSAVI ($R^2 = 0.99$).

### 4.3. The effect of spatial resolution on LAI retrieval

The effect of spatial resolution on LAI retrieval was further analysed by applying varying methods of averaging of reflectance, SVI and LAI over 1 km grid on the data derived from [100%, 0%] cut-off bounds as saturated and background SVIs values from the image masked out for non-vegetated areas and corrected using 6S. Changing the scale of surface reflectance, or the SVI and retrieving the LAI using the same algorithms as the high resolution reflectance would alter the result due to the exponential relationships assumed between LAI-$P_i$, LAI-$f_c$, LAI-SVIS, and LAI-reflectance (e.g., Lang and Xiang, 1986; Table 2; Eq. (12); Fig. 3). Basis of SVIs such as NDVI are known to be sensitive to the scale effect due to non-linear formulation of the indexes. On other hand, orthogonal SVIs such as DVI are less affected and study has shown that SDVI is not affected by the scale (Jiang et al., 2006). The result of Gutman and Ignatov NDVI, SDVI and MSAVI based methods were presented, as the other NDVI-based methods have the same response as Gutman and Ignatov NDVI approach. Three commonly used averaging techniques and two methods for the determination of saturated and background SVIs values are described in Table 3. The reference method for the comparison was based on the LAI retrieved using high resolution dataset and averaged over 1 km grid. The other two averaging methods are: the SVIs are averaged over 1 km grid and the LAI algorithm is applied, and the reflectances are averaged over 1 km grid and the SVIs derived and LAI algorithm is applied.

<table>
<thead>
<tr>
<th>Method</th>
<th>Statistics</th>
<th>$1$ Average LAI</th>
<th>$2$ Average SVI</th>
<th>$3$ Average reflectance</th>
<th>$4$ Average SVI</th>
<th>$5$ Average reflectance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gutman and Ignatov</td>
<td>Mean</td>
<td>3.13</td>
<td>1.03</td>
<td>1.83</td>
<td>3.00</td>
<td>5.15</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>1.33</td>
<td>0.57</td>
<td>0.27</td>
<td>1.96</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>Fit model</td>
<td>$x$</td>
<td>$y = 0.06x + 0.11x$</td>
<td>0.984</td>
<td>0.656</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDVI</td>
<td>Mean</td>
<td>1.27</td>
<td>1.13</td>
<td>1.13</td>
<td>1.94</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>0.54</td>
<td>0.55</td>
<td>0.55</td>
<td>1.14</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>Fit model</td>
<td>$x$</td>
<td>$y = 0.91x$</td>
<td>0.978</td>
<td>0.977</td>
<td>0.977</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSAVI</td>
<td>Mean</td>
<td>1.94</td>
<td>1.60</td>
<td>1.72</td>
<td>2.11</td>
<td>2.26</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>0.87</td>
<td>0.85</td>
<td>0.87</td>
<td>1.27</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>Fit model</td>
<td>$x$</td>
<td>$y = 0.85x$</td>
<td>0.981</td>
<td>0.970</td>
<td>0.980</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

STDEV stands for standard deviation. $1$ LAI inverted from averaged reflectance values averaged over 1 km grid using the same background and saturated SVI value as $1$; $2$ LAI inverted from averaged reflectance values averaged over 1 km grid using the same background and saturated SVI value as $1$; $3$ LAI inverted from averaged reflectance values averaged over 1 km grid using the same background and saturated SVI value as $1$; $4$ LAI inverted from averaged reflectance values averaged over 1 km grid using the same background and saturated SVI value as $1$; $5$ LAI inverted from averaged reflectance values averaged over 1 km grid using the same background and saturated SVI value as $1$.

**Table 3** Comparison of estimated LAI as computed over 1 km grids using various averaging methods from SPOT image atmospherically corrected with 6S.

The assumption used for LAI inversion algorithm from the gap fraction obtained from various SVIs assumes the uniform pixel model. This is valid for 10 m pixel SPOT HRG data as it is considered high resolution and the pixels are uniform. However, NDVI of a mixed pixel as demonstrated in Jiang et al. (2006) does not vary in consistent or traceable manner with $f_c$ or LAI. This is shown also in Fig. 7 by applying varying averaging technique for NDVI-based LAI retrieval. Fig. 7(D) shows that there is a trend between the computed clumping index and the differences obtained among high resolution LAI and LAI from...
Fig. 7. Comparison of estimated LAI values using 10 m pixel size SVI values (1 Table 3) on X-axis averaged over 1 km grids with: LAI inverted from SVI average (1 Table 3); LAI inverted from reflectance average (2 Table 3); and LAI inverted from reflectance average (3 Table 3); (A) Gutman and Ignatov NDVI, (B) SDVI, (C) MSAVI, and (D) LAI values using 10 m pixel size SVI values (1 Table 3) on X-axis averaged over 1 km grids plotted against the computed clumping index. All the legends for superscripts: 1, 2, 3, 4, and 5 are given in Table 3.

Various averaging methods shown in Fig. 7(A), (B) and (C). For example, regression analysis reveals that: (a) from NDVI-based method, 1LAI = 0.4343(2LAI/CI) ($R^2 = 1$, RMSD = 0); (b) from SDVI based method, 1LAI = 2/4 or 3LAI/CI ($R^2 = 0.98$, RMSD = 0.06); and (c) from MSAVI based method, 1LAI = 1.4251(4LAI/CI) ($R^2 = 0.99$, RMSD = 0.06). Therefore, the clumping index developed in this study can be used for correction of SVIs or LAI calculated using various averaging methods.

Since there is no information about the extinction coefficient ($k$), for Eq. (12) and (13), it is indispensable to approximate $k$ from the image itself therefore all the variables needed for the LAI calculation based on the information obtained from image is acquired analytically. $k$ from satellite observation depends on the leaf orientation and distribution, as well as stand and canopy architecture. The former two are difficult to obtain from satellite observations. The stand and canopy architecture can be approximated to stand and canopy distribution which in turn can be expressed with the clumping index. Light extinction under forest canopy decays exponentially with amount of green biomass as does fractional of absorbed photosynthetically active radiation. As shown in Fig. 7(D), the clumping index also decays exponentially with LAI (the higher value indicating the lower clumping). This information from clumping index can be related with $k$ for gap fraction correction rather than LAI, as also Baret et al. (1995) and Carlson and Ripley (1997) tried to correct $f_c$. The gap fraction in Eq. (12) is related to contact frequency:

$$P_0 = e^{-K \text{LAI}} = e^{-N}$$  \hspace{1cm} (15)

where $K$ is the mean number of contacts between a light beam and a vegetation element.

$$-\ln(P_0) = k\text{LAI} = K$$  \hspace{1cm} (16)

Carlson and Ripley (1997) described that the variation of NDVI with respect to the global LAI in partially vegetated areas would be mostly controlled by the variation in the fraction of vegetated surface area illuminated by the sun and visible to the radiometer. Verstraete and Pinty (1991) discussed the nature and extent of NDVI variations in partially vegetated lands, and argued that NDVI is more strongly controlled by changes in vegetation cover than by changes in the optical thickness of canopies. Therefore, theoretically, it is fair to assume that the NDVI-based method (Gutman and Ignatov) gives $K$ rather than LAI if there is no correction for $f_c$ or gap fraction applied as demonstrated by Baret et al. (1995) and Carlson and Ripley (1997). Fig. 7(A) confirms these assertions as the LAI from reflectance averaging methods (3 and 5 Table 3) overestimated compared to high resolution LAI for lower LAI values (1 Table 3). The overestimations and the deviation of the LAI from high resolution LAI are clearly explained by CI (Fig. 7(D)). Therefore, if the clumping index shown in this study is related with correction exponential parameter for gap fraction, the LAI from corrected gap
fraction without the knowledge of \( k \) will be:

\[
\text{LAI} = -\ln(P_0^c) \tag{17}
\]

If CI = 1, then the average, non-intercepted fraction \( P_0 \) of a beam of radiation travelling is related to \( f_c \) from NDVI without correction, and therefore the corrected and non-corrected \( P_0 \) are the same. Baret et al. (1995) showed that the exponential correction parameter varies depending on the density of vegetation, and it becomes small in areas with a low vegetation density which is in good agreement with the CI computed in this study (Fig. 7(D)). The corrected gap fraction in Eq. (17) may explain the confounding effects of crown clumping, extinction coefficient and various averaging differences.

Eq. (17) was applied for the gap fraction data obtained from NDVI by averaging reflectance over 1 km grid (Table 3). The result is compared with the high resolution LAI estimated from NDVI based on constant \( k \) (Table 3). The result shows exponential relationships between two LAI estimations with the similar trend of reflectance averaging LAI with constant \( k \) from NDVI and high resolution LAI. The exponential relationship is mainly due to the differences of NDVI and NDVI

\[\text{MSAVI} \rightarrow 2 \rightarrow 16.14 \rightarrow 12\]  
\[\text{SDVI} \rightarrow 3 \rightarrow 6 \rightarrow 40.24 \rightarrow 12.45\]  
\[\text{Gutman and Ignatov} \rightarrow 2 \rightarrow 8 \rightarrow 29.93 \rightarrow 3.21 \rightarrow 25.01\]  
\[\text{Carson and Ripley} \rightarrow 2 \rightarrow 8 \rightarrow 21 \rightarrow 17.98 \rightarrow 3.21 \rightarrow -7.32\]  
\[\text{Baret et al.} \rightarrow 2 \rightarrow 8 \rightarrow 2 \rightarrow 5 \rightarrow 28.01 \rightarrow 4.14 \rightarrow -22.81\]  
\[\text{SDVI} \rightarrow 2 \rightarrow 5 \rightarrow 36 \rightarrow 40.24 \rightarrow 12.45 \rightarrow -37.30\]  
\[\text{MSAVI} \rightarrow 2 \rightarrow 5 \rightarrow 2 \rightarrow 2 \rightarrow 16.14 \rightarrow 12 \rightarrow -4.23\]  

\[\text{CV}\% = \frac{100}{n-1} \sum_{i=1}^{n} (O_i - \mu)^2 \]  

\[\text{Bias} = \frac{100}{n-1} \sum_{i=1}^{n} (O_i - P_i)^2 \]  

\[\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2} \]  

\[\text{Mean Error} = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) \]  

\[\text{CV} = \frac{\text{RMSD}}{\text{Mean}} \times 100\% \]  

\[\text{Bias} = \frac{\text{Mean}}{\text{RMSD}} \times 100\% \]  

\[\text{Mean} = \frac{1}{n} \sum_{i=1}^{n} O_i \]  

\[\text{SDVI} \rightarrow 2 \rightarrow 5 \rightarrow 36 \rightarrow 40.24 \rightarrow 12.45 \rightarrow -37.30\]  

\[\text{MSAVI} \rightarrow 2 \rightarrow 5 \rightarrow 2 \rightarrow 16.14 \rightarrow 12 \rightarrow -4.23\]  

4.4. Validity of high resolution LAI (SPOT) estimations as compared to ground measurements

Since the entire model parameters for LAI estimations used in this study were derived from the image itself, the validity assessment of the product is crucial. The LAI estimated from SVIs corresponds to the green part of the vegetation which is slightly different from the classical definition and the LAI obtained from the ground-based hemispherical photography which is “plant area index”. Therefore, the validity assessment was focused more on the comparison rather than direct validation. The mean error calculated as the average deviations of SVIs-based LAI from ground-based LAI, and the percentage of root mean square deviation (RMSD), standard deviation and bias of all five SVIs methods are presented in Table 4. The mean error of MSAVI was the lowest and the estimates were the closest to the ground-based measurement with smallest underestimation (~4% of bias) and with coefficient of variation (CV) close to the ground-based measurement (Table 4). MSAVI also resulted in least RMSD of 16.14%. SDVI resulted in the highest absolute bias (Table 4) indicating that the method underestimated the LAI. Jiang et al. (2006) also found out that SDVI has underestimated \( f_c \) cover compared to NDVI-based methods. The underestimation of the SDVI was attributed to a non-linear increase of NIR reflectance for higher \( f_c \) due to rapidly accumulating green biomass with only gradual lateral percent cover increase (Huete et al., 1985). However, SDVI based method explained the variation of LAI on the field better than all other methods (Table 4). All NDVI-based method (Gutman and Ignatov, Carson and Ripley and Baret et al., the latter two being derived from the first) resulted in very narrow range of LAI compared to the ground-based and other SVIs-based methods. NDVI-based method without correction (Gutman and Ignatov) was the only method which resulted in positive bias.

Nevertheless, except that of the coefficient of variation, the other evaluation parameters (Table 4) may give bias for the comparison of SVIs-based LAI with ground-based measurements due to the difference of the definitions assumed and practically could be achieved. Hence, the evaluation assessment was further extended by comparing with MODIS LAI product and applying the methods developed here in the independently simulated data.

Table 4

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean error (%)</th>
<th>RMSD (%)</th>
<th>CV (%)</th>
<th>Bias (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gutman and Ignatov</td>
<td>28</td>
<td>29.93</td>
<td>3.21</td>
<td>25.01</td>
</tr>
<tr>
<td>Carson and Ripley</td>
<td>21</td>
<td>17.98</td>
<td>3.21</td>
<td>-7.32</td>
</tr>
<tr>
<td>Baret et al.</td>
<td>5</td>
<td>28.01</td>
<td>4.14</td>
<td>-22.81</td>
</tr>
<tr>
<td>SDVI</td>
<td>36</td>
<td>40.24</td>
<td>12.45</td>
<td>-37.30</td>
</tr>
<tr>
<td>MSAVI</td>
<td>2</td>
<td>16.14</td>
<td>12</td>
<td>-4.23</td>
</tr>
</tbody>
</table>

*Coefficient of variation (CV) of ground-based measurement = 16.39%. The percent mean error, root mean square deviation (RMSD), CV, and Bias are defined below. \( P \) is predicted and \( O \) is ground-based measured LAI value of plot \( i \) and \( n \) is the number of measurements and \( \mu \) is average value. Mean error \( = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) \) \% 100, RMSD \( = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2} \) \% 100, CV \( = \frac{\text{RMSD}}{\text{Mean}} \times 100\% \) and Bias \( = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) \) \% 100. Table 4 Evaluation of different methods to derive LAI (n = 54).
4.5. Comparison with MODIS LAI product

The comparison between MODIS LAI product and LAI estimated using the five formulas shown in Table 2 are given in Fig. 9. The comparison is made only for those pixels which are retrieved with main algorithm without cloud contamination. The LAI from SPOT data is derived using the high resolution and averaged over the MODIS pixels. Fig. 9 (lower row) shows the SPOT LAI which is corrected for clumping index. It is theoretically possible that the LAI of a single MODIS pixel is not retrieved accurately due to uncertainties in inputs and the algorithm. MODIS LAI product is overestimated compared to all five methods of LAI retrieval used in this study. However, the correlations are statistically significant and RMSD are reasonably low. The deviation from regression line in Fig. 9 (above) can partially be explained by the uncertainties in coregistration between MODIS and SPOT pixels. A statistical comparison of the overall averages shows that, MODIS LAI overestimated by 43% for Gutman and Ignatov, 100% for Carson and Ripley, 131% for Baret et al., 232% for SDVI, and 134% for MSAVI methods. For the LAI corrected for clumping index, MODIS LAI overestimated by 4% for Gutman and Ignatov, 46% for Carson and Ripley, 68% for Baret et al., 140% for SDVI, and 70% for MSAVI methods. The overestimation is very severe for higher LAI values (Fig. 9). During the late summer, MODIS LAI products show great and progressive overestimation (Pisek and Chen, 2007). MODIS LAI products are known to overestimate LAI, for example, up to 75% (Yang et al., 2006), 200% over Canada (Abuelgasim et al., 2006), by 2–3 LAI over BOREAS study area in Canada for broadleaved forest (Fang and Liang, 2005) and up to 323.6% across north American forest sites (Heinsch et al., 2006). MODIS LAI resulted in large standard deviation compared to any of the methods used in this study and in agreement with Pisek and Chen (2007). The spatial distribution of MODIS LAI measured as a coefficient of variation was 0.33, with the five methods without correction for clumping was 0.34–0.39 and with correction 0.23–0.28. This indicates that the range of LAI with both MODIS and the five methods are in good agreement to explain the distribution of LAI.

There was better relation with MODIS LAI and the LAI estimated using SPOT data for lower LAI values where the MODIS pixels are retrieved without saturation. The overestimation in higher LAI values and in general can be explained by prevailing amount (72%) of MODIS pixels were retrieved with saturation. The new clumping index might help better relate the LAI derived from SVIs with structural property of the forest canopy, and the geometry and the shade of the crown distribution. MODIS uses a stochastic radiative transfer model to account for 3D effects of vegetation heterogeneity (clumping and species composition) to assign a single clumping value to each cover type in LAI algorithm (Wang et al., 2004). This does not account for its spatial variation and relies on biome classification and the level of abstraction inherent in the classification procedure. Therefore, theoretically, however, there was no reason that the LAI corrected with clumping should be better related with MODIS product. However, the correction has improved the overall mean agreement with MODIS LAI but decreased slightly the correlation and increased the RMSD. The algorithm for the clumping index should be further tested particularly with high resolution multi directional remote sensing data and refined with more ground control parameters if there is any need for such kind of parameter for functional–structural plant models to represent physiological processes in 3D tree representations. There are few studies which estimated clumping index from remotely sensed multi-angular data (Lacaze et al., 2002; Chen et al., 2003, 2005; Leblanc et al., 2005) even though the meaning of the index is often arguable. The methodology presented in this study for clumping index estimation have a physical meaning to explain the architecture of the vegetated canopy, which remains to be discussed in further studies.

4.6. Correction factor for SVI-LAI relationship and robustness assessment using simulated data

The NDVI, SDVI and MSAVI based methods applied on the simulated spectrum data in order to assess the performances and derive the correction parameters for f_c-LAI relationships as implemented in Baret et al. (1995) and Carlson and Ripley (1997). NDVI_{back} and DVI_{back} were calculated as average SVIs from the three soil spectrum as described earlier. NDVI_{∞}, DVI_{∞}, and MSAVI_{∞} were calculated with PROSPECT + SAIL simulated spectra with an LAI value of 12. The interactive correction was determined by applying different exponential factors on the derived f_c for the three methods. The performance of interactive determination for correction factor was controlled using the true and estimated LAI regression slope, R^2, and...
RMSD. The factors which gave the regression slope close to unity, highest $R^2$ and lowest RMSE were deemed to be the correction factor. As shown in Fig. 10, SDVI and MSAVI based methods were found to be the best fit without any correction and outperformed NDVI-based method. Therefore, there was no need to apply any correction factor for SDVI and MSAVI based methods. LAI estimated using NDVI-based methods both with and without correction showed sigmoidal relationships with true LAI with smooth curves. The sigmoidal relationship between the estimated and true LAI from NDVI-based method can be explained by that as NDVI is ratio based, it is very susceptible for scaling using background and saturated values for the exponential fit of LAI estimation. Based on the interactive determinations, the correction factor for LAI estimation using NDVI-based method was found to be 6.15, i.e., the $f_c$ derivation should be corrected as $[(\text{NDVI} - \text{NDVI}_{\text{back}})/\text{(NDVI}_{\infty} - \text{NDVI}_{\text{back}})]^{6.15}$. The correction lowered the RMSD, increased the $R^2$ and resulted in the regression slope close to unity (Fig. 10). However, the correction factor has underestimated the lower LAI and overestimated the higher LAI which were otherwise without correction (Fig. 10). The mean LAI obtained using the three methods computed over the three soil types for uncorrected models are statistically different (t-Test, $p<0.001$). The soil has small effect for the LAI (Fig. 10), however except for that of NDVI-based method, the mean LAI for the three soil types are statistically different (t-Test, $p<0.001$). The differences of mean LAI among the three soil types were 0.5–3% for NDVI, 6–22.5% for SDVI and 4–15% for MSAVI based methods. These indicate that SDVI and MSAVI based methods are vulnerable for the influence of the soil; although, the differences are in acceptable ranges.

The relationships between the gap fractions derived from the three SVIs with LAI are robust for the simulated dataset. However, the result is dependent on the choice of the input background and saturated SVI values as explained in Section 4.1. The exponential correction factors which would be applied on $f_c$ derived using the equations in Table 2 in order to best fit best LAI estimations were however found out to be less of importance. Using entirely different approaches and data sources, Choudhury et al. (1994) and Gillies and Carlson (1995) and Carlson and Ripley (1997) independently obtained identical square root relationship between $f_c$ and NDVI-based method (Table 2). Baret et al. (1995) obtained varying correction factors both for the same and among different SVIs for the simulated and real datasets. Baret et al. (1995) had studied solely gap fraction without inverting into LAI, and the others have studied $f_c$ (Choudhury et al., 1994; Gillies and Carlson, 1995; Carlson and Ripley, 1997). However, there is no uniform relationship between $f_c$ and LAI. Besides, the exponential correction factors are dependent on the choice of different combination of background and saturated SVI values. This indicates that, using any of the generic correction factors introduce biases and should be avoided. The focus rather should be on the choice of appropriate background and saturated SVI. The vegetation SVIs which give indication of the optical thickness of vegetated biomass can be exponentially related with LAI using Eq. (12) without any correction. As demonstrated in this study and earlier by Baret et al. (1995), applying the exponential correction term which is fitted in a certain season, vegetation type, and atmospheric condition and correction approach, may not be directly applicable to other areas.
5. Conclusions

On the basis of simulated datasets, ground-based and satellite measurements, the validity of an approach for the computation of LAI based on the information solely contained on image scene itself was proven to be reasonably accurate. The approach uses the classical red and NIR radiometric channels and is applied to mostly temperate hardwood forest and open fields. The methodology developed here can be extended to other types of vegetation, provided that there is accurate estimation of SFI for the upper-end value of saturated LAI. The role of accurately specifying the background and saturated SVIs, the spatial scale and atmospheric effects, the robustness test, and fitness with global LAI product was considered for effectiveness assessments. The simple approach was also demonstrated to characterize the spatial heterogeneity of canopy architecture from a single view angle radiometric data.

The varying definitions and assumptions used for LAI obtained from ground-based measurements, SPOT image retrieval and MODIS product, and the validity of using simple 1D radiative transfer model for robustness assessment make any complete validation of the approaches almost impossible. The ground-based measurements give the “plant area index” which includes non-photosynthetic parts of plants such as braches and stems. On the other hand, the LAI retrieved from SPOT imagery is related to the “green leaf area index” which is highly relevant from application and vegetation function point of view for photosynthesis, evapotranspiration and carbon balance studies. The green LAI can be accurately estimated from the image as shown in this study. The most important parameter for LAI determination was the saturated SVI value as the background SVI has very small effect for the final LAI. The saturated SVI value can be retrieved from the image particularly from large swath imagery during the growing season. SDVI was found out to be both scale and atmospheric correction invariant provided that the atmosphere all over the image scene is assumed to be constant. This is important attribute towards developing simple algorithm for LAI retrieval as the atmospheric correction can be rigorous process. MSAVI is also scale invariant. NDVI was found to be not appropriate index for large scale LAI inversion due to the sensitivity for atmospheric correction and scale. In the cases of well defined a priori information are lacking for radiative transfer inversion, or, sufficient ground measurements for robust empirical equations are not available, the method shown in this study can be utilised to retrieve LAI on operational bases particularly during the growing season.

Acknowledgements

The research was funded by the Academy of Finland through the TAITATOO-project and personal grant to Dr. Petri Pellikka (120190), and by Natural Sciences and Engineering Research Council of Canada funding to Dr. Douglas King. The SPOT image was obtained from OASIS program financed by the European Commission, DG Research. AG is financed partly by CIMO, Department of Geography of University of Helsinki, and the Finnish Graduate School of Geography.

References


