# Improving Clumping and LAI Algorithms Based on Multiangle Airborne Imagery and Ground Measurements

Anita Simic, Jing M. Chen, James R. Freemantle, John R. Miller, and Jan Pisek

***Abstract*—Measurements at more than one angle capture the directional anisotropy of solar radiance reflected from vegetated surfaces. According to our recent research, we propose that the best two view angles for vegetation structural mapping are the following: 1) the hotspot, where the Sun and view directions coincide, and 2) the darkspot, where the sensor sees the maximum amount of vegetation structural shadows. The Normalized Difference between Hotspot and Darkspot (NDHD), an angular index generated from Compact Airborne Spectrographic Imager (CASI) data, is found to be highly correlated with the field-measured foliage clumping index. The foliage clumping index characterizes the nonrandomness in the spatial distribution pattern of leaves. It is of comparable importance as the leaf area index (LAI) for quantifying radiation interception and distribution in plant canopies, and it also affects estimated LAI mapping using remote sensing data. As the clumping index can vary considerably within a cover type, it is highly desirable to map its spatial distribution for various ecological applications. We have generated clumping index maps based on the previous algorithms and empirical relationships between field-measured $\Omega$ and CASI-derived NDHD. Through intensive validation using field data, we demonstrate that the combination of the hotspot and darkspot reflectances has the strongest response to changes in vegetation structure. Two crown structural characteristics, namely, crown height and within-crown density, are major factors that impact the NDHD and clumping index difference between the mature and young (regrowth) coniferous forests. The study area is located near Sudbury in the northern Ontario, Canada.**

***Index Terms*—Clumping index, Compact Airborne Spectrographic Imager (CASI), darkspot, hotspot, leaf area index (LAI), multiangle, Normalized Difference between Hotspot and Darkspot (NDHD).**

## I. Introduction

CHARACTERIZED by various levels of structural organization, vegetation presents a real challenge for the remote sensing community. Leaves in plant canopies, particularly in forests and shrubs, are generally highly clumped. Leaves are more overlapped vertically than the random case because of canopy structures such as crowns, whirls, branches, and shoots. This clumping decreases the proportion of sunlit leaves and increases shaded leaves at all Sun angles, thus affecting the interaction of radiation with vegetation, plant growth, and carbon cycle.

The foliage clumping index ($\Omega$) characterizes the spatial distribution pattern of leaves. It is particularly useful for estimating radiation interception and distribution in plant canopies, and it is of comparable importance as leaf area index (LAI) for carbon/water cycle modeling [1], [2]. The clumping index is a function of the architectural properties of trees such as stem density and crown size [3], [4]. It serves as a correction factor to the effective LAI ($L_e$) to obtain the true LAI [3]. Optical instruments, such as LAI-2000, measure canopy gap fractions from the penetration of light at various angles and then convert these measurements into LAI under the assumption of a random spatial distribution of leaves, resulting in $L_e$ rather than LAI. Although the total absorbed radiation by the canopy is accurate when $L_e$ is used, the distribution of absorbed radiation in the canopy is distorted. If $L_e$ is assumed to be the true LAI, the amount of shaded leaf area is underestimated [2]. The proportion of shaded and sunlit leaves varies greatly with the clumping index. When introduced into an ecological model, the foliage clumping index caused the estimation of daily canopy photosynthesis to differ by about 20% for a black spruce site [1]. Kurcharik *et al.* [3] found that accounting for the clumping in the aspen stand results in a scaled canopy assimilation that is 39% larger than in the case of random distribution. As the clumping index can vary considerably within a cover type, it is highly desirable to map its spatial distribution for various ecological applications [2].

The need to derive the shaded fraction of leaves triggered the use of multiangle remote sensing. Based on the relationship of field data of the bidirectional reflectance distribution function (BRDF), Deering *et al.* [5] found that forest canopy structure is related to its BRDF. Rapid developments in remote sensing technologies over the last two decades inspired scientists to probe into the relationships between biophysical and structural parameters of a vegetation canopy and multiangular remote sensing data. The acquisition of measurements at more than one angle captures the directional anisotropy of solar radiance reflected from vegetation surfaces. With the advent of the POLarization and Directionality of Earth Reflectances (POLDER) instrument, the observations of the hotspot have become more available. Data from the POLDER sensor onboard the ADEOS-1 platform have the ability to measure the same ground

Manuscript received December 31, 2008; revised June 3, 2009. First published November 10, 2009; current version published March 24, 2010. This work was supported by the Canadian Space Agency.

A. Simic, J. M. Chen, and J. Pisek are with the Department of Geography, University of Toronto, Toronto, ON M5S 3G3, Canada (e-mail: simica@geog.utoronto.ca).

J. R. Freemantle and J. R. Miller are with the Department of Earth and Space Science and Engineering, York University, Toronto, ON M3J 1P3, Canada.

Digital Object Identifier 10.1109/TGRS.2009.2033383

surface at up to 14 view angles during a single overpass [6]. Due to the high frequency of revisitation, the area of interest is well covered after several days in the angular domain, increasing the chances of sampling the hotspot [7]. Both spaceborne and airborne PROBA sensors have been effectively used to show the advantage of using angular measurements for the retrieval of vegetation structural parameters [2], [4], [7]. The Multiangle Imaging Spectro-Radiometer (MISR) instrument onboard the Terra platform is capable of sampling the angular anisotropy using nine different view angles and four spectral bands. A number of studies have shown the potential of MISR measurements in retrieving information about canopy structure [8]–[13]. With a spatial resolution of 275 m, MISR is able to resolve the intrinsic structural variability of the surface, while large enough to allow for a nine-day global coverage capability [8]. The concept of combining multiangular and hyperspectral remote sensing has been successfully utilized in an existing satellite observation system named Compact High-Resolution Imaging Spectrometer (CHRIS). This technology is very promising for retrieving canopy structure. [14]–[19]. Sandmeier and Deering [20] used airborne hyperspectral multiangular data from the Advanced Solid-state Array Spectroradiometer to relate hyperspectral BRDF to structural characteristics of vegetation canopies.

BRDF modeling studies suggest a strong relationship between vegetation structural characteristics and the hotspot phenomenon [21], [22]. Chen *et al.* [23] developed the Normalized Difference between Hotspot and Darkspot (NDHD), an angular index to characterize the anisotropic behavior of the vegetated surface. The hotspot is a geometric–optical effect denoting the backscatter peak in directional reflectance hemisphere where illumination and view directions coincide, resulting in the absence of visible shadows. The darkspot reflectance, on the other hand, contains the maximum visible shadows observed in the forward scattering direction, where the reflectance is at a minimum. The combination of the hotspot and darkspot views has the strongest signals about the vegetation structure. This index and a similar index were successfully related to ground-based measurements [1], [23] and to modeled–derived clumping index [1], [14]. The relationship was found to be significant for the red and near-infrared (NIR) wavelengths when the canopy closure exceeds 25% [23].

This paper is part of a larger project related to the refinement of a multiangular and hyperspectral measurement concept developed by Simic and Chen [14]. The objective of the overall study is to explore the capability of the refined concept, which proposes simultaneous retrieval of vegetation structural and biochemical parameters, including LAI, clumping index, and leaf chlorophyll content. Estimation of LAI depends on clumping index [24]; both clumping index and LAI are important in the estimation of chlorophyll content. This paper concentrates on the retrieval of clumping index and LAI from the Compact Airborne Spectrographic Imager (CASI) multiangle data. The CASI instrument is a push-broom airborne imager that has the capacity to acquire images in visible and near-visible spectral regions (400–950 nm) [25]. The analysis presented in this paper utilizes the CASI data acquired at two off-nadir angles in addition to the nadir view. According to Chen *et al.* [23], the two best view directions for assessing the canopy structure are the hotspot and darkspot.

In this paper, we attempt to perform the following.

1) Validate and fine-tune the clumping index and LAI algorithms based on previous algorithms using a net set of ground-based and airborne measurements.
2) Demonstrate that NDHD based on the multispectral measurements at two off-nadir angles, the hotspot and darkspot, provides the optimal information for retrieving the clumping index and LAI.
3) Explore the relationships between NDHD and clumping index using a geometrical radiative transfer model 5-Scale in light of this new set of measurements.

## II. Modeling

Numerous radiative transfer models, such as GORT [26] and 4-Scale [24], enable the retrieval of scene characteristics from multiangle remote sensing data. Lacaze and Roujean [4] used the G-function and Hot SpoT model to simulate the BRDF of boreal forest and to describe the canopy geometry based upon the hotspot signature. The Four-Scale Linear Model for Anisotropic Reflectance (FLAIR), a linear kernel-like model based on 4-Scale, has been developed by White *et al.* [27] and used in [7].

In this paper, we explored the performance of the 5-Scale model [28] in retrieval of NDHD and clumping index. It is a geometric–optical radiative transfer model with emphasis on the structural composition of forest canopies at different scales, including tree groups, tree crown shapes, branches, shoots, and leaf cells. 5-Scale is a combination of 4-Scale by Chen and Leblanc [24] and a leaf-level spectral model. The LIBERTY model [29] or PROSPECT [30] is currently used as the leaf-level model. The model parameters are separated into three groups [31] as follows: 1) site parameters: domain size, LAI, tree density, solar zenith angle, viewing angle, and relative azimuth angle; 2) tree architecture parameters: crown radius and height, apex angle, needle-to-shoot ratio, foliage clumping index, and tree foliage typical size; and 3) foliage reflectance and transmittance spectra and background reflectance spectra, or band specific reflectances and transmittance for multispectral simulations. The model calculates the reflectances of four scene components: sunlit and shaded foliage and sunlit and shaded backgrounds. Each reflectivity is multiplied with the viewing proportion of each component to obtain the total reflectivity as the sum of the four components. In several studies, the 5-Scale model is demonstrated as an advanced and reliable model to be used for modeling BRDF [10], [23]. As part of the BRDF, the hotspot is modeled with two kernels. One kernel uses gap-size distribution within the crowns for the canopy hotspot, and the other kernel uses the gap-size distribution between the tree crowns for the background hotspot [24], [31].

In order to improve the performance of the 5-Scale model with respect to multiangular simulations, additional functions are introduced to the model in this paper. The original version of the model (before the modification) assumes isotropic background reflectance distributions, resulting in underestimation of reflectance near the hotspot. We have incorporated

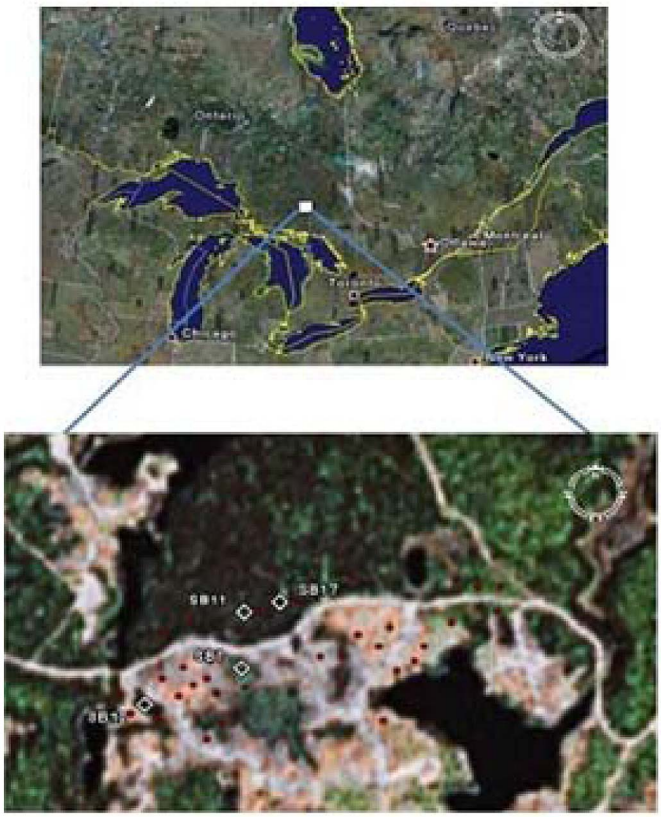


Fig. 1. Study area near Sudbury, Ontario, with black spruce study sites (SB7, SB11, SB12, and SB17). Note that, in summer 2008, 30 sites (marked with dots) were chosen across the area to validate the clumping index and LAI.

the nonlinear temporal angular model (NTAM) [32] into 5-Scale to characterize the anisotropic effect on the background reflectance measured in the field at one angle. In this way, the background reflectance becomes bidirectional, depending on the Sun–sensor geometry for each observation. The NTAM model exhibits dependence of the reflectance on satellite viewing, Sun–satellite angle, relative azimuth angle of the Sun and sensor, and solar zenith angle. For open boreal forests, this modification is particularly important. The contributions of geometric and volume scattering are directly related to the amount of vegetation weighted through polynomial relationships for different land cover types. A complete description of the model can be found in [32].

## III. Methodology

### A. Site Description and Field Campaigns

The study area is located near Sudbury in the northern part of Ontario $(47°09'47.7'', 81°42'23.4'')$; it is a flat area at 350 m above sea level (Fig. 1). Two field campaigns were performed during the summers of 2007 and 2008. During the intensive field campaign in 2007, more than ten sites were selected for detailed collection of ground truth data to be used in the validation of 5-Scale. Due to the positional inconsistency between overlapping swaths, only four sites were visible in all nadir and off-nadir directions. As the goal was to demonstrate the performance of 5-Scale for open and closed canopies, two distinctive black spruce sites, SB7b and SB17, were used in the demonstration. SB7b had a more open canopy than SB17. The trees were smaller, gaps between them were larger, and the understory was more developed.

The field campaign in 2007 was completed within two days of the CASI overpass. The main purpose of the field work was to collect various parameters to be used as inputs to 5-Scale. In each forest stand, a $30 \times 30$ m$^2$ plot was determined where the measurements were performed along established transects. The effective LAI $(L_e)$ data were measured using the LAI-2000 plant canopy analyzer (Li-Cor, Inc., Lincoln, NE). LAI-2000 was operated near dusk under diffuse radiation conditions to reduce the effect of multiple scattering on the measurements. The protocol used in the measurements followed the standard instruction of the instruments [33]. Clumping index was measured using Tracing Radiation and Architecture of Canopies (TRAC) [34]. Total clumping index $(\Omega)$ is composed of two components: 1) the effect of foliage clumping at scales larger than the shoot $(\Omega_E)$ and 2) the needle-to-shoot area ratio quantifying the effect of foliage clumping within a shoot $(\gamma_E)$. The former is obtained by TRAC based on a gap-size distribution theory through measuring sunfleck widths along transects beneath the canopy [35]. The latter is a lab-measured variable based on shoot samples [35].

The ASD FieldSpectrometer (ASD, Inc., Boulder, CO) was used to measure the understory spectra from most common species. The measurements were made in the nadir direction around the solar noon under clear-sky conditions. A specially designed integrating sphere, compatible with the ASD spectroradiometer, was used to measure both needle and broad leaf spectra in a field laboratory within 10 h after the leaves were collected. The integrating sphere collects reflected or transmitted light from samples, illuminated by a collimated tungsten light source, over a full hemisphere where the energy is evenly spread over the entire interior surface of the sphere. A detector exposed to an opening in the sphere collects the radiance on the interior surface that is proportional to the average hemispherical reflectance/transmittance. A fiber optic cable is used to connect the integrating sphere and the spectroradiometer to measure the collected radiance. The leaves were placed in the Ziploc bags and refrigerated near 0 °C prior to the measurements of reflectance and transmittance. Structural parameters of trees (density, height, and diameter at breast height) were also measured. The needle reflectance and transmittance measurement techniques of previous studies [36], [37] were followed in this paper.

The second field campaign in the summer of 2008 included intensive measurements of the clumping index and effective LAI $(L_e)$ over a subarea of the study site, including three different vegetation cover types (Fig. 1). The mature coniferous forests were populated with old black spruce and jack pine. The regrowth forest contained relatively evenly distributed young jack pine trees (∼1 m high) with some open-land patches and shrubs. The understory layer was highly developed in the black spruce forests, while more ground was exposed in the old and young jack pine sites. The regrowth forest has been a controlled area for several years. Forest herbicide application promoted the survival and growth of jack pine trees valuable

for lumber. Only a few aspen and black spruce trees survived the herbicide application, giving rise to relatively pure young jack pine stands. The deciduous sites generally included tall mature aspen trees; several sites enclosed small young aspen trees under large canopy gaps. LAI-2000 and TRAC were used for the measurements of $L_e$ and clumping index, respectively.

### B. Remote Sensing Data

During the field campaign, several sets of CASI data were acquired on June 28, 2007 by the Earth Observations Laboratory at York University. The data were collected in the hyperspectral mode; the sensor was flown in the range of 105–129 knots and at an altitude of 1524 m above the ground level with a 55-ms integration time. The nadir view ($VZA = 0^\circ$), hotspot ($VZA = -40^\circ$), and darkspot ($VZA = 40^\circ$) were collected along the solar principal plane with a solar zenith angle of 40°. The CASI data were resampled to a 3-m spatial resolution; each image had 72 bands in the spectral range from 400 to 1000 nm at a spectral resolution of 7.5 nm. The data were processed by York University in the following manner: 1) Radiometric calibration was performed using latest radiance scale factors for the CASI imager; 2) the data were atmospherically corrected to at-ground-modeled reflectance using Modtran 4 found in the PCI atmospheric correction package; 3) geometrical corrections were completed using ITRES' geometric correction software and roll, pitch, and location information from an onboard inertia navigation system; 4) the nadir images were further corrected by locating in the image the positions of ground control points measured in the field; and 5) the off-nadir images were further reregistered to nadir image (from step 4) and to each other. Direct measurements of atmospheric parameters were not available during the CASI overflights; however, aerosol optical depth was estimated from nearby measurements of visibility. Methodological measurements from Environment Canada gave the conditions as clear with a visibility of 25 km. Other atmospheric parameters were allowed to default to MODTRAN midsummer midlatitude values.

### C. Overall Approach

*1) Relationship Between the NDHD and Clumping Index—NDHD and Ω Map Generation:* Chen *et al.* [2] proposed the NDHD as an angular index to characterize the anisotropic behavior of foliage components as follows:

$$NDHD = \frac{\rho_h - \rho_d}{\rho_h + \rho_d} \tag{1}$$

where $\rho_h$ and $\rho_d$ are the hotspot and darkspot reflectances, respectively. We generated the NDHD map using the CASI data acquired along the principal plane. Fig. 2 shows the main steps we used to perform validation of the clumping index map based on the existing correlation between Ω and NDHD and to create the clumping index map based on the empirical relationship between field-measured Ω and CASI-derived NDHD.

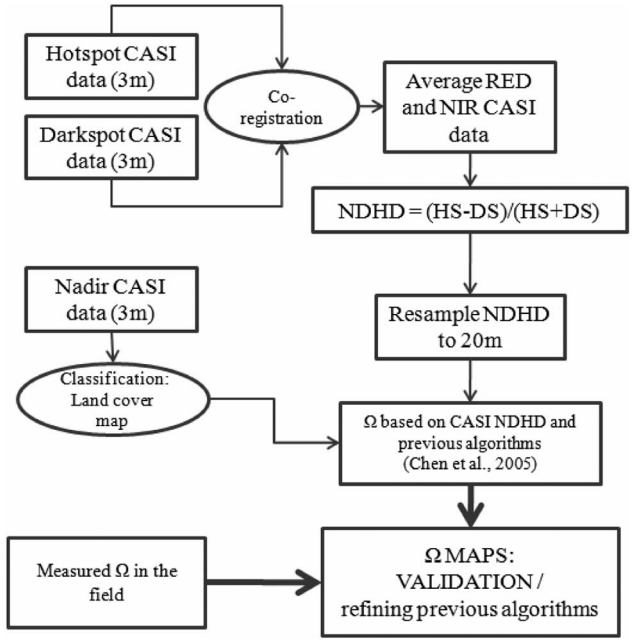


Fig. 2. Steps undertaken to develop the clumping index maps using the previous algorithms [23] and applying the empirical approach using the field-measured Ω values and NDHD.

The CASI-derived NDHD maps were generated by aggregating the red (31–38 bands) and NIR (48–65) bands needed to calculate the index. The aggregation was based on the multispectral bands of the Landsat Thematic Mapper. The NDHD maps were then resampled to 20-m resolution in order to ensure consistency of the spatial resolution of NDHD and field-measured clumping index that was measured and averaged over 20–30-m transects. In order to characterize the architecture of a canopy, the transmitted direct solar beam needs to be captured over a long transect [38].

We produced clumping index maps based on the algorithms of correlation between the NDHD and Ω for the red and NIR spectra. First, we used algorithms generated by Chen *et al.* [23] using appropriate coefficients for each cover type and created the map based on the model approach. We also generated a clumping index map based on the empirical relationship between the field-measured clumping index and CASI-derived NDHD to refine previous algorithms (Fig. 2). The algorithms were derived for three cover types (mature coniferous, deciduous, and regrowth forests) separately and were applied to 20-m NDHD maps.

In addition to the explained steps in Fig. 2, we examined the relationships between model-derived and field-measured clumping indexes. Fine-tuning of some input parameters at the nadir view was done in order to bring the 5-Scale simulations as close as possible to the CASI nadir values for selected sites. In particular, the crown radius and tree height were slightly adjusted (5%–15%) since the original values of crown radius were based on a visual approximation. After the calibration to the nadir view data, off-nadir simulations were compared with the off-nadir CASI data for all sites and all views.

The clumping index (Ω) generated from 5-Scale was derived using the corresponding sets of input parameters for each site.

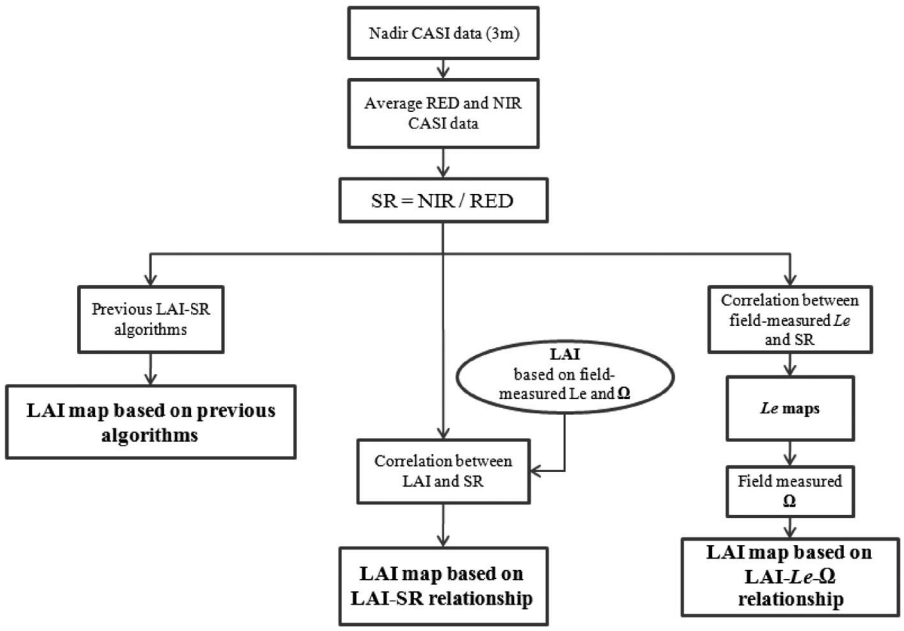


Fig. 3. Steps undertaken to develop the LAI maps using the previous algorithms [39], the correlation between SR and LAI, and the correlation between SR and $L_e$ incorporating the clumping index.

The gap fraction $P(\theta)$ at given zenith angle $\theta$, calculated in 5-Scale, was employed in Miller's theorem [39] to calculate the effective LAI $(L_e)$ first

$$L_e = -2 \int_0^{\pi/2} \ln [P(\theta)] \cos\theta \sin\theta d\theta. \quad (2)$$

The clumping index was then derived using the following equation:

$$\Omega = L_e / LAI \quad (3)$$

where $LAI$ represents an input for a given simulation. A single clumping value was produced for a canopy regardless of the spectral domain.

*2) Classification:* Maximum likelihood, a supervised classification algorithm, embedded in the Environment for Visualizing Images [40] was used to create a cover-type map from the nadir view. The nadir image was first transformed using the minimum noise fraction (MNF) transform, in order to remove noise in the data. The forward MNF with rotation was used, and eigenvalues greater than one were retained. The MNF transform segregates and equalizes the noise in the data and reduces the data dimensionality for subsequent processing. The process is based on an estimated noise covariance matrix and rescales the noise in the data. It results in transformed data in which the noise has a unit variance and no band-to-band correlation. A forward transform, performed to remove noise from data, determines which bands contain coherent images by examining the images and eigenvalues [40]. Filtering $3 \times 3$ median was applied as a postclassification process.

The regions of interest (ROIs) were created based on five cover types (mature coniferous, deciduous, and regrowth forests, and water and rocky/road areas) [Fig. 4(b)]. The covariance matrix was computed using all the pixels from the ROIs for each class. The ROIs were based on the relatively homogenous areas of the five classes using ground-based sampling. Each class consisted of several polygons and more than 300 pixels in order to ensure the homogeneity and proper sample size as required by the software (deciduous forest: 347 points; coniferous forest: 599 points; regrowth forest: 545 points; rock/road: 355 points; water: 775 points). The classification was validated in summer 2008. The results of the classification were validated based on *in situ* sampling using a traditional error matrix analysis.

*3) LAI Maps:* The steps used to produce the LAI map were based on two approaches (Fig. 3): 1) the previous simple ratio–LAI (SR–LAI) algorithms developed by Chen *et al.* [41] to explore the applicability of the algorithms and 2) the CASI sensor-specific (empirical) relationship between SR and LAI or $L_e$.

The main goal was to investigate the empirical approach and to explore the effect of $\Omega$ used in the relationship between SR and $L_e$, where we incorporated $\Omega$ at two different stages. First, we developed an empirical relationship between SR and field-measured LAI, generating the LAI map $(LAI_{\text{SR-LAI}})$. In this case, the LAI values were calculated from the ground-measured $L_e$ and $\Omega$ using (3) (Fig. 3). Second, we generated the $L_e$ map based on the empirical relationship between SR and ground-measured $L_e$. Then, we incorporated the clumping index map using (3), deriving the LAI map $(LAI_{\text{SR-}L_e\text{-}\Omega})$ (Fig. 3). As discussed in the introduction, $L_e$ is obtained from canopy gap fraction, assuming that the foliage spatial distribution is random; vegetation clumping is not considered in the measurements of $L_e$. Two different approaches were compared, and map statistics were analyzed. The average reflectance values of the nadir image for red and NIR were generated according to Landsat TM. The algorithms were based on three land cover types (deciduous, mature coniferous, and regrowth forests).

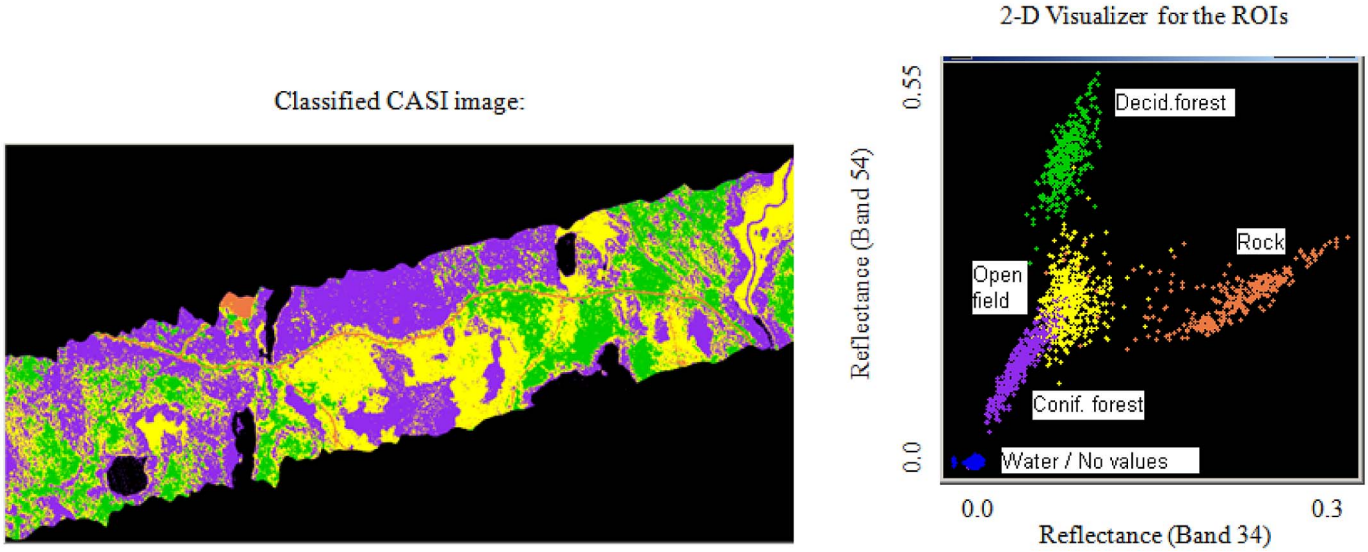


Fig. 4. (a) Classified CASI image with five classes (mature conifers, deciduous forest, regrowth forest, rocks, and water). (b) Separability graph for the classes: Snapshot of n-D visualizer (scatter plot) for band 34 (656 nm) in the red spectral region and band 54 (808 nm) in the NIR spectra region based on a subset of image data—the ROIs. Note that the water pixels are shown in blue to be distinguished from the background color.

## IV. Result and Analysis

### A. Classification

The classified image is shown in Fig. 4(a). The accuracy assessment results in a Kappa index of 85%. Two factors introduce some uncertainties in the process of validation. 1) The classification has a relatively low number of classes, which inevitably results in high accuracy. If the number of classes increases, the estimate of classification accuracy would be lower due to increased spatial heterogeneity [42]. 2) The validation sites are not randomly distributed in space, as many sites are damp and not easily accessible; the validation is often confined to the edge of forest patches.

Most classification errors are observed within the regrowth forest that contains some young aspen trees in addition to young jack pine trees. These areas are classified as deciduous forest for several sites. One reason for this could be the fact that these controlled areas had more young aspen trees during the last-year acquisition than during the current-year validation as each additional herbicide application reduces the amount of aspen trees. Fig. 4(b) suggests good separability among the classes. A slight overlap can be seen between the regrowth forest pixels and other classes. All chosen study sites and pixels in our calculations are individually validated, and they all have correct land cover classification. The land cover map is used to generate the clumping index and LAI maps using land-cover-dependent algorithms.

### B. Generation of the NDHD Maps

Based on (1) (Section III-C1), the hotspot and darkspot data sets are combined to generate the NDHD maps. As the roll and pitch effects are quite pronounced in the off-nadir images, making geometric correction difficult, we have clipped several subareas (scenes) and have performed image-to-image coregistration for each scene separately. This has further reduced the errors due to geometric distortion. The maps shown in Fig. 5 are generated for the red and NIR spectral regions using previously developed NDHD–$\Omega$ algorithms by Chen *et al.* [23].

The hotspot is affected by canopy structure and by optical properties of both foliage and background. The darkspot is also affected by optical properties of foliage and background; however, it is mostly influenced by the amount of shadow on the ground and on the shaded side of the crowns [2]. The NDHD maps in red and NIR bands have the highest values for the mature conifers. Open spaces between the crowns are common features of these forests, and the exposed understory and soil between these gaps have considerable contributions to the overall reflectance. Based on the data-set analysis, both the hotspot and darkspot values are generally lower for the black spruce than for other forests.

It is interesting to note that the area in the top-left portion of the red band image has somewhat higher hotspot values than other coniferous forest area and the higher NDHD values (Fig. 5). This location consists of mixed mature jack pine trees and black spruce trees with large gaps between tree crowns, and the understory vegetation is not as rich as in the black spruce forests. The open canopies increase the hotspot in the red spectral region and the contrast between the hotspot and darkspot, resulting in high NDHD.

In the red spectrum, the regrowth forest exhibits high values of darkspot and thus has generally low NDHD values. Poor vegetation and more soil exposure within the regrowth forest slightly increase the hotspot when compared with the mature conifers. Some young aspen trees located between the young jack pine trees likely further reduce the overall NDHD for some of these areas. Deciduous forests exhibit the lowest NDHD values. Low clumping results in less shadows and ultimately low NDHD. Background has an important effect on the overall reflectance. The bidirectional trend suggests that the developed understory most likely reduces both the hotspot and darkspot values when compared with the soil background or poorly

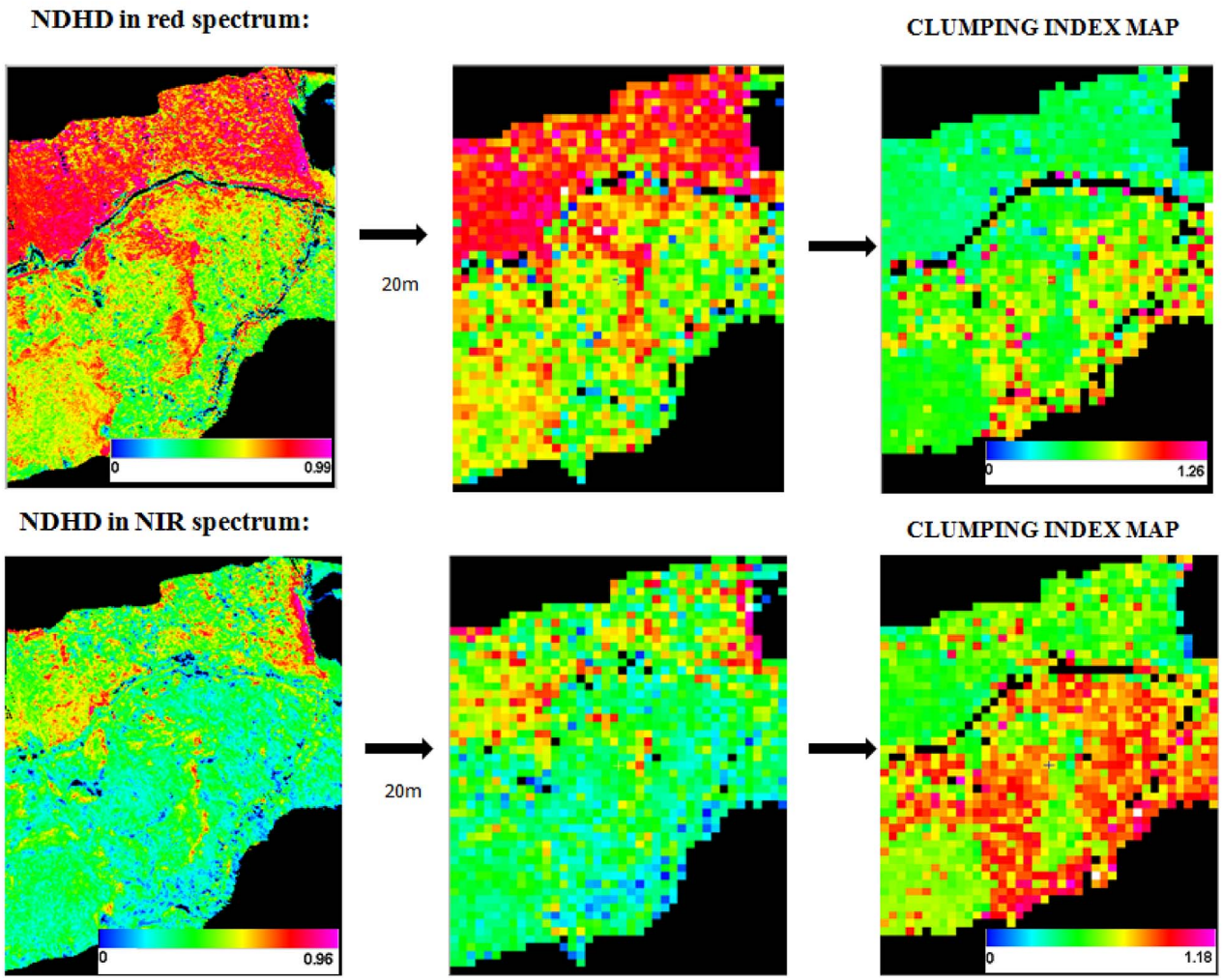


Fig. 5. NDHD maps (3- and 20-m resolutions) and clumping index map (20-m resolution) generated for Scene 2 for (a) red and (b) NIR spectral regions. Note that $\Omega > 1$ for regular foliage dispersions, equal to one for a random distribution, and less than one for clumped foliage distribution.

developed understory. However, further research is needed to determine the impact of understory vegetation at a subpixel level.

In the NIR spectral region, less contrast is observed between mature conifers and regrowth forest in the hotspot view as both understory vegetation and exposed ground have somewhat high NIR reflectance. Black spruce forest exhibits somewhat lower leaf reflectance than regrowth and deciduous forests in the NIR spectrum. Overall, the NDHD values in the red spectrum are higher and more variable than in NIR. Similar trend was found in other studies [7], [23].

### C. Generation of the Clumping Index Maps

*1) Validation of the Existing NDHD–Ω Algorithms:* The clumping index maps for the red and NIR spectra, generated by previous algorithms [23], exhibit opposite patterns from NDHD in terms of the numerical values because of the negative relationship between clumping and NDHD (Fig. 5). The clumping index of the mature coniferous forest has lower values than those of the regrowth and deciduous forests, suggesting, once again, high clumping for conifers relative to other cover types. Due to the lower clumping, the regrowth (more regular) and deciduous forests exhibit higher clumping index values.

The relationship between the clumping index generated using the previous algorithms and the field-measured clumping index for the various sites is shown in Fig. 6 and Table I. Although the root mean squared error (rmse) value is larger for the regrowth forest in the NIR spectrum, the correlations for all cover types based on NDHD in the NIR band are considerably stronger than that in the red band, suggesting that a clumping map generated in the NIR spectrum is more reliable. This agrees with the findings of Chen *et al.* [23], as the NIR band is less affected by accuracy limitations in atmospheric correction than the red band. The correlations align with the 1:1 line quite well for mature conifers and deciduous forest but not for regrowth forest. As the previous algorithms do not differentiate young and old forest types, they may overestimate the clumping index for the regrowth forest [Fig. 6(a)]. While the calculated clumping index values differ for mature conifers and regrowth forests, the field-measured clumping index values are generally similar for these two cover types.

*2) Mapping the Clumping Index Using the Empirical Approach:* In the empirical approach, the correlation between the

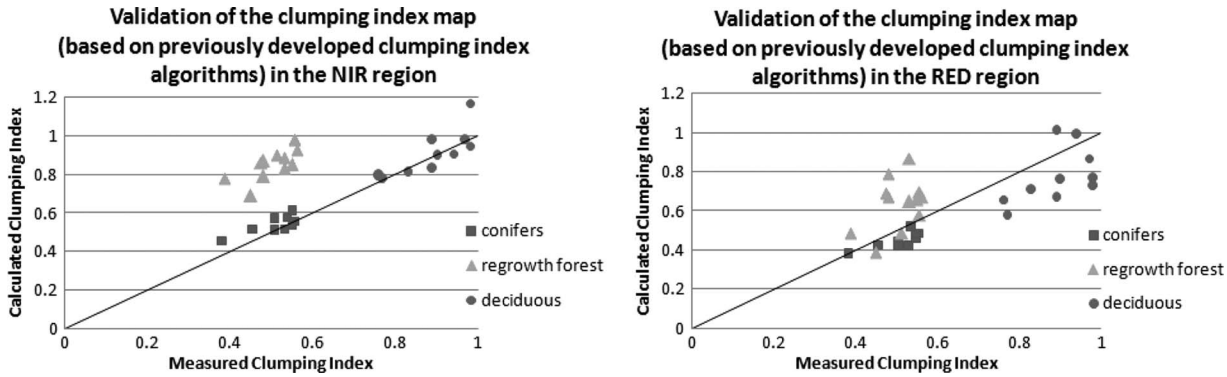


Fig. 6. Validation of the previously developed clumping index algorithms using field-measured clumping index in (a) NIR spectral region and (b) red spectral region.

TABLE I
CORRELATION FACTOR $R^2$ BETWEEN THE CLUMPING INDEX DERIVED USING PREVIOUS ALGORITHMS [23] AND FIELD-MEASURED CLUMPING INDEX FOR DIFFERENT COVER TYPES

| Cover type | $R^2$ for red spectrum | $R^2$ for NIR spectrum |
|---|---|---|
| Mature conifers | 0.593 | 0.603 |
| Deciduous forest | 0.335 | 0.657 |
| Regrowth forest | 0.186 | 0.552 |

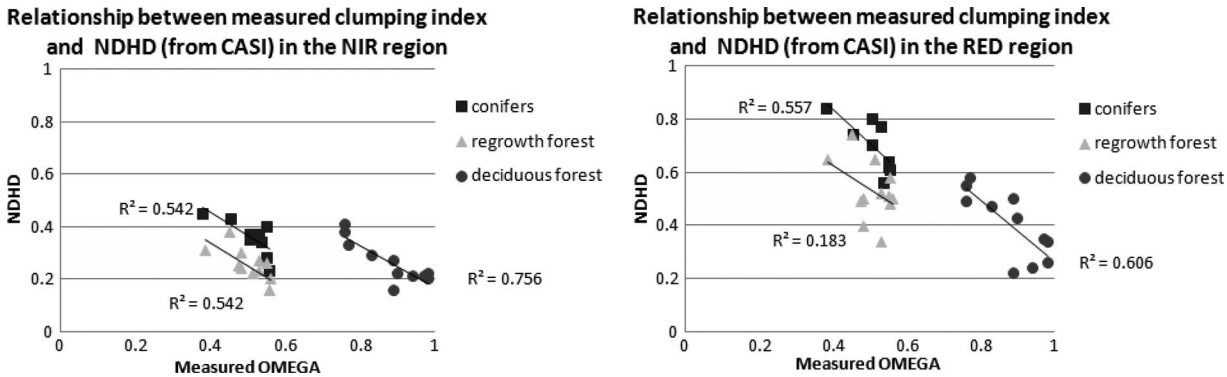


Fig. 7. Empirical approach: Relationship between CASI-generated NDHD and field-measured clumping index for (a) the red and (b) NIR spectra.

TABLE II
LINEAR COEFFICIENTS OF THE LINES OF BEST FIT AND $R^2$ FOR THE EMPIRICALLY GENERATED RELATIONSHIPS BETWEEN NDHD AND CLUMPING INDEX

| Cover type | Red spectrum | | NIR spectrum | |
|---|---|---|---|---|
| | The line of best fit | $R^2$ | The line of best fit | $R^2$ |
| Mature conifers | Y= -1.257x + 1.337 | 0.557 | Y = -0.897x + 0.814 | 0.543 |
| Deciduous forest | Y = -1.141x + 1.406 | 0.607 | Y = -0.803x + 0.97 | 0.757 |
| Regrowth forest | Y = -0.895x + 0.983 | 0.184 | Y = -0.875x + 0.691 | 0.542 |

CASI-generated NDHD and field-measured clumping index is negative for all three cover types (Fig. 7), suggesting that higher NDHD is related to lower clumping index (higher clumping). The graph shows similar trends as shown in Fig. 5. Large variability of NDHD is seen for all three types of forest.

Similar to Fig. 5, the empirical approach exhibits generally higher NDHD and higher clumping (lower clumping index) for mature conifers than regrowth and deciduous forests. The deciduous forest is clearly separated as being less clumped. While deciduous and coniferous forests have similar correlations in red and NIR spectral regions, the regrowth forest has considerably stronger correlation in the NIR than in the red band.

The correlation coefficients in the empirical approach (Table II) differ from those in [23]. These previous algorithms were developed for global applications, and no validation

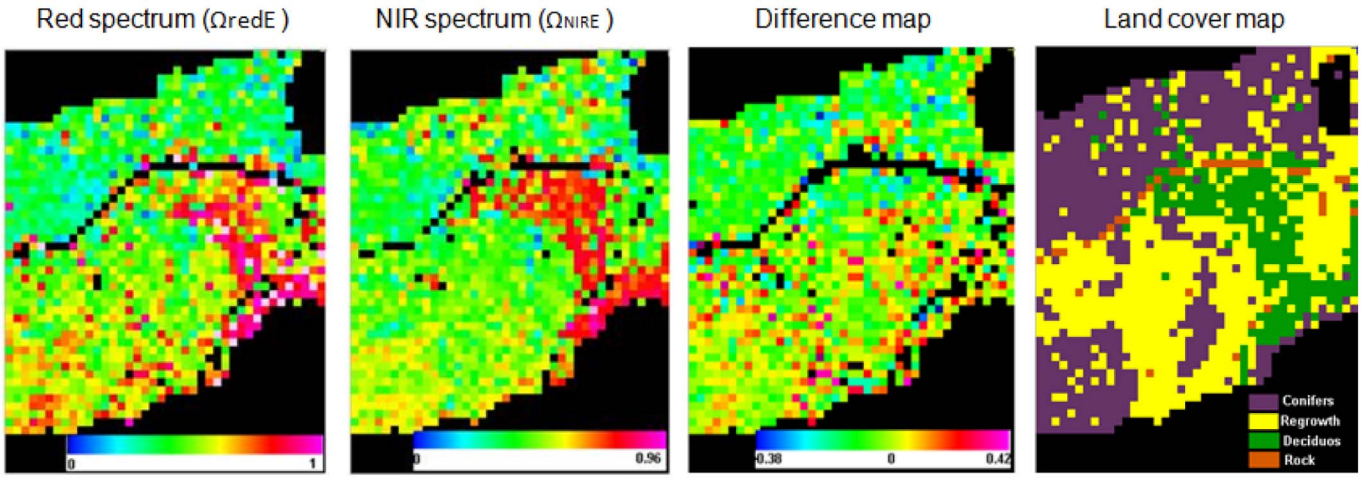


Fig. 8. Clumping index maps based on the empirical relationship between CASI-generated NDHD and field-measured clumping index for (a) red spectral region and (b) NIR spectral region, and (c) difference map between clumping index in red and NIR spectra, and (d) land cover map for Scene 2. As Scene 2 exhibits only minor coregistration errors, we use this area to present the empirical clumping index map.

using field-based measurements was made. The coefficients in this paper represent a refinement of the previous algorithms (Table II). In the study of Leblanc *et al.* [7], the coefficients are somewhat higher when they are based on 5-Scale-simulated relationships between NDHD and the total foliage clumping index.

This is the first study where extensive field-based measurements of clumping index are used to validate existing algorithms and to derive the empirically based clumping index maps. The coefficients are somewhat different for the red and NIR spectra, suggesting that the influence of leaf optical properties is not completely eliminated by NDHD. The generally strong $R^2$ between NDHD and clumping index suggests the importance of the multiangle remote sensing used in retrieving of clumping index.

Based on the empirically developed algorithms, Fig. 8 shows the clumping index maps derived from red and NIR NDHD. The red spectral region demonstrates more variable and generally higher values of clumping index for regrowth and some deciduous sites, where open spaces can be found (Fig. 8). Based on these results and those shown in Fig. 6, the clumping index map derived from NIR is deemed to be more accurate than that derived from red, and it will therefore be used for the subsequent LAI calculations.

*3) 5-Scale Site-Specific Simulations:* The canopy reflectance spectra are extracted from the CASI data and simulated by 5-Scale for two distinct black spruce sites (SB17 and SB7b) (Fig. 9). Site SB7b exhibits higher nadir values than site SB17 due to the extensive open space and greater exposure of the understory vegetation and soil background between the trees. The backward reflectances ($-40^\circ$ VZA) are higher than the nadir view reflectances, indicating the influence of the hotspot (Fig. 9). The exposed soil enhances the hotspot reflectance in the NIR spectrum [Fig. 9(a)], while the understory reduces the hotspot reflectance in the red spectrum [Fig. 9(c)]. The forward reflectances (i.e., darkspot reflectances) are generally lower than those at the nadir. Pronounced forward scattering may result from relatively sparse vegetation and strong influence of the background under the forest canopy [43]. The relative difference between the nadir and hotspot is smaller for SB7b, while the relative difference between the nadir and darkspot is greater for the same site due to more distinct clumping at SB7b. The clumping decreases the sunlit leaf component and increases the shaded leaf component in the off-nadir spectra. The general trend suggests that the hotspot, relative to the nadir, is smaller for a canopy with lower LAI and lower crown radius (SB7b) in the visible and NIR spectra, even though the tree density could be high, because of the increased probability of observing the background, which has a lower reflectivity than the foliage. This trend is consistent with the findings of Chen and Leblanc [24].

The 5-Scale model is initially calibrated against the nadir spectra. Its simulations exhibit similar patterns to the CASI data for all angles (Fig. 9). Although the trend of the simulated off-nadir spectral signatures generally coincides with the CASI measurements, most disagreement is observed between the simulated and measured hotspot reflectances $(VZA = -40^\circ)$ within the red spectrum. Even though the preprocessing of the CASI data was carefully performed in this paper, some uncertainties may still exist. A similar disagreement was observed for the forward reflectances of CHRIS data in the study of Simic and Chen [14]. The CASI spectra are somewhat lower than those simulated with 5-Scale by Chen and Leblanc [28]. Overall, the results show very good performance of the 5-Scale model. We feel confident in using this canopy reflectance model in our further research.

As we explained in Section III-C1, we also calculated the clumping index based on the site-specific 5-Scale simulations using site-specific parameters. It is found that the gap fraction decreases as the view zenith angle increases (Fig. 10). The zenith angle distribution pattern of the canopy gap fraction is variable for different sites of the same forest type, suggesting different canopy closures. The curvatures of the gap fraction against zenith angle also differ within the same forest type. This is particularly noticeable within the deciduous forest type.

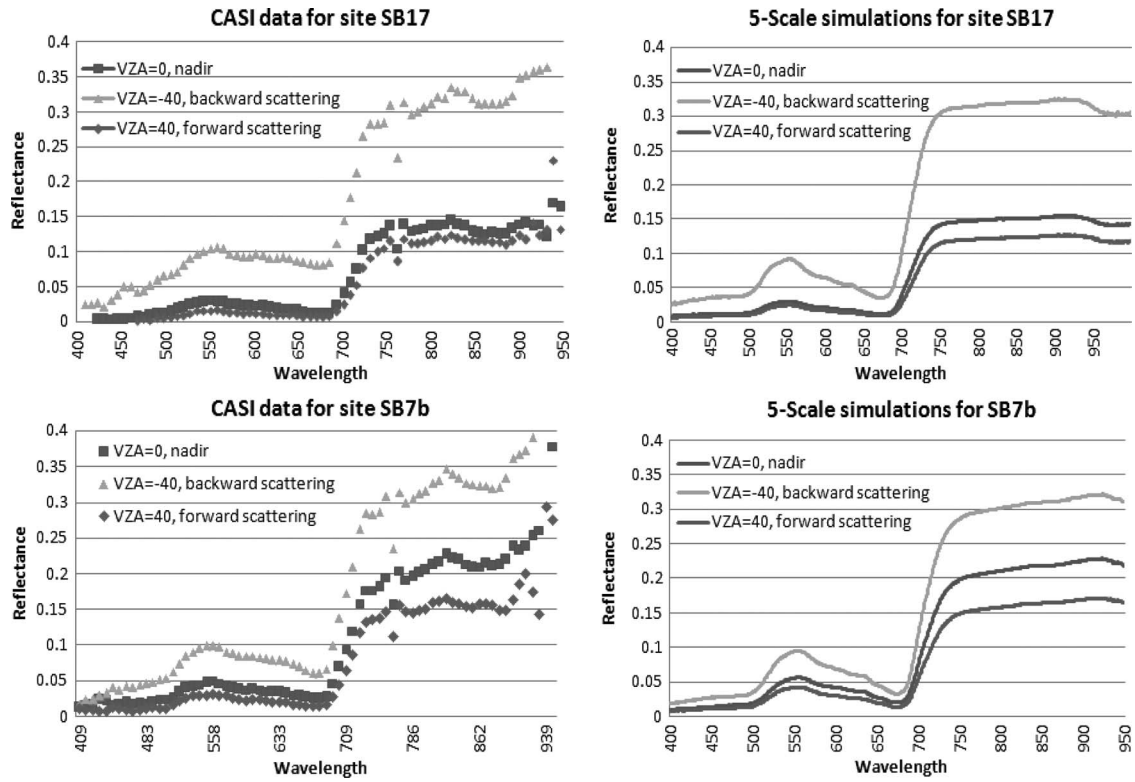


Fig. 9. CASI spectra and 5-Scale simulations for black spruce sites [(a) and (b)] SB17 and [(c) and (d)] SB7b at nadir and off-nadir angles (including the 75°, 10°, and −90° headings). Note that the agreement between the off-nadir data was based on the calibration of the nadir.

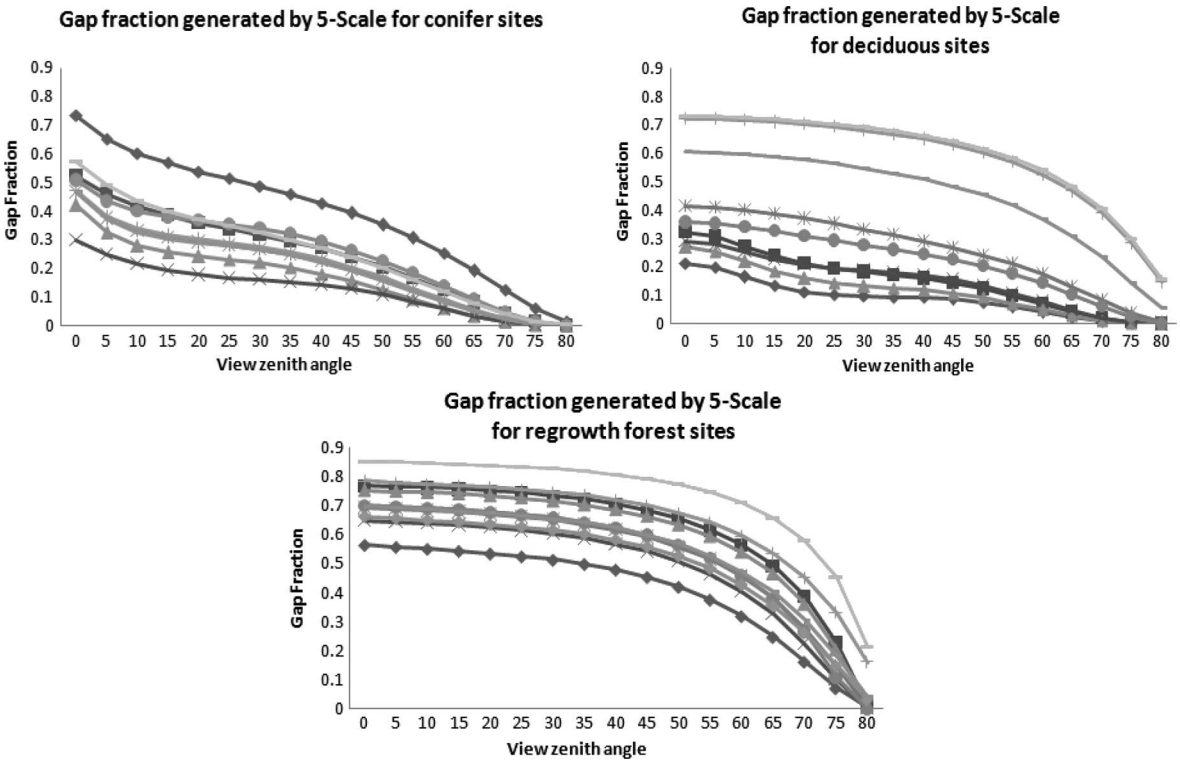


Fig. 10. Gap fractions generated by 5-Scale for three cover types: (a) Mature conifers, (b) deciduous, and (c) regrowth forest. The coefficient of variation for conifers is 53% (from 23% to 186% per view zenith angle), for deciduous sites is 81% (from 45% to 158% per view zenith angle), and for regrowth forest is 27% (from 11% to 156% per view zenith angle).

Regrowth forests have the characteristics of planophile foliage angle distributions, while mature conifers tend to show erectophile foliage angle distribution patterns. Deciduous forests have both angle distribution patterns, suggesting its variability in canopy architecture. In addition to the angular variation of foliage elements, the spatial distribution of trees and the shape

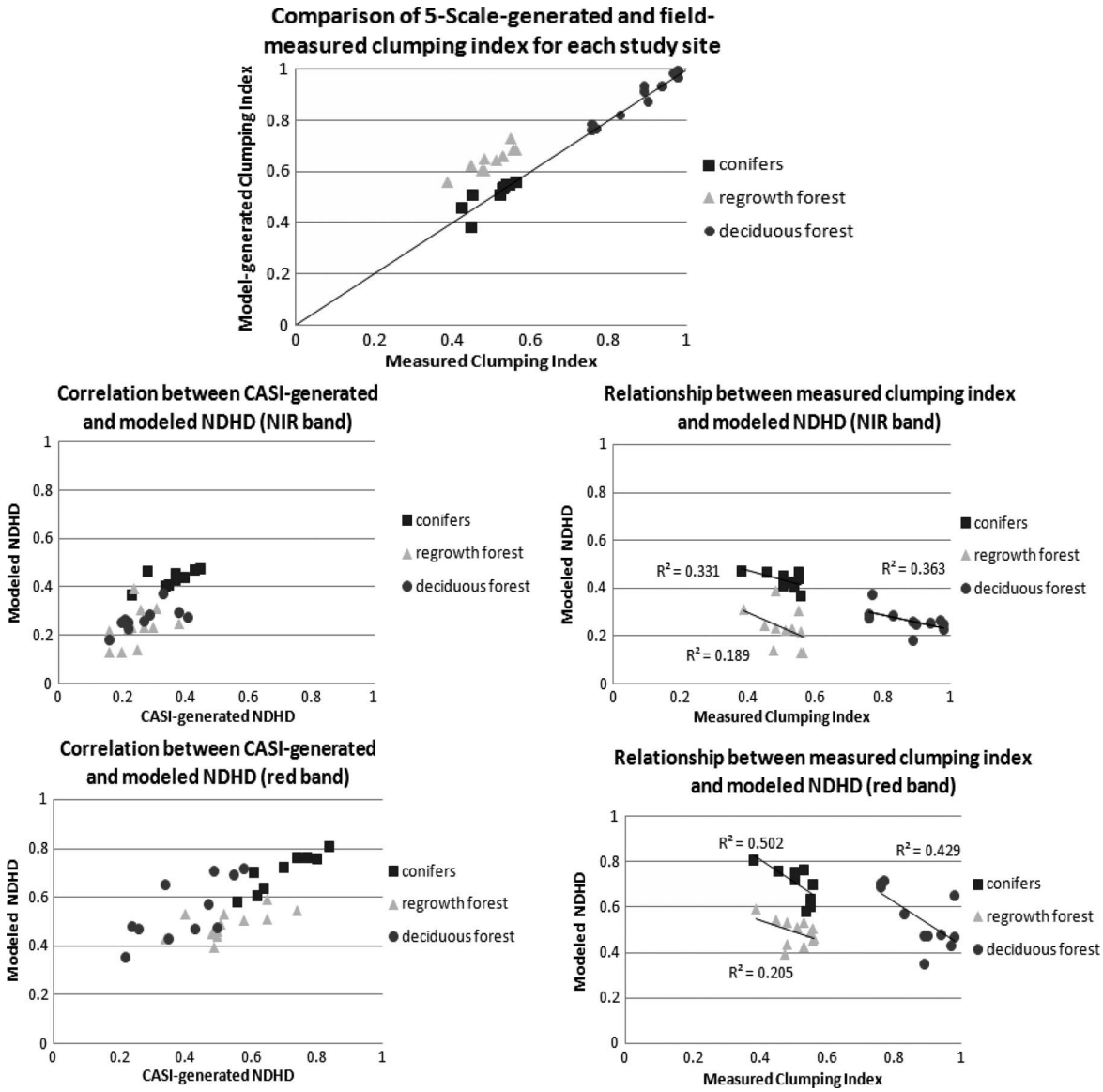


Fig. 11. Performance of 5-Scale. (a) Comparison of field-measured and 5-Scale generated clumping index using the structural parameters for near 30 sites. Comparison of NDHD generated by 5-Scale and measured by CASI for (b) NIR and (d) red spectral regions. Relationship between modeled NDHD and measured Ω for (c) NIR and (e) red spectral regions.

of crowns also influence the curvature [44], [45]. Within the conifer type, the heights of the mature black spruce and jack pine trees are generally larger than the gaps between trees, resulting in a rapid decrease of the gap fraction with view zenith angle. On the other hand, the regrowth forest has small trees where gaps are often larger than tree height, and the gap-fraction curves decrease at lower rates with view zenith angle. Transparent crowns of the young jack pine trees additionally enhance this trend. The deciduous forest, having sites with tall aspen trees and some with young aspen trees, results in the combined curve variation patterns.

The correlation between the clumping index generated by 5-Scale (a single value is produced for a canopy as explained in Section III-C1) and the field-based measurements is strong for all cover types [Fig. 11(a)]. However, the model-generated values are somewhat higher than the field-measured values for regrowth forests. Although the rmse values between the modeled and CASI-generated data within the NIR bands are smaller for the deciduous and regrowth forests, the coefficients of determination are stronger for all three cover types within the red bands [Fig. 11(b) and (d)]. 5-Scale has the capability to separate NDHD for mature conifers and regrowth forests [Fig. 11(c) and (e)]. The correlation between modeled NDHD and field-measured clumping index is somewhat lower for NIR than for red bands [Fig. 11(c) and (e)]. However, the significant correlations between modeled NDHD and measured clumping index, particularly in the red bands [Fig. 11(e)] within the small dynamic range for each cover type, give us confidence in the clumping algorithm developed using 5-Scale. The systematic difference in clumping algorithm between young and mature conifer stands suggests that forest age or tree height would be an additional parameter to be included in regional and global clumping algorithm, although these information would generally be difficult to obtain for large areas. In the study

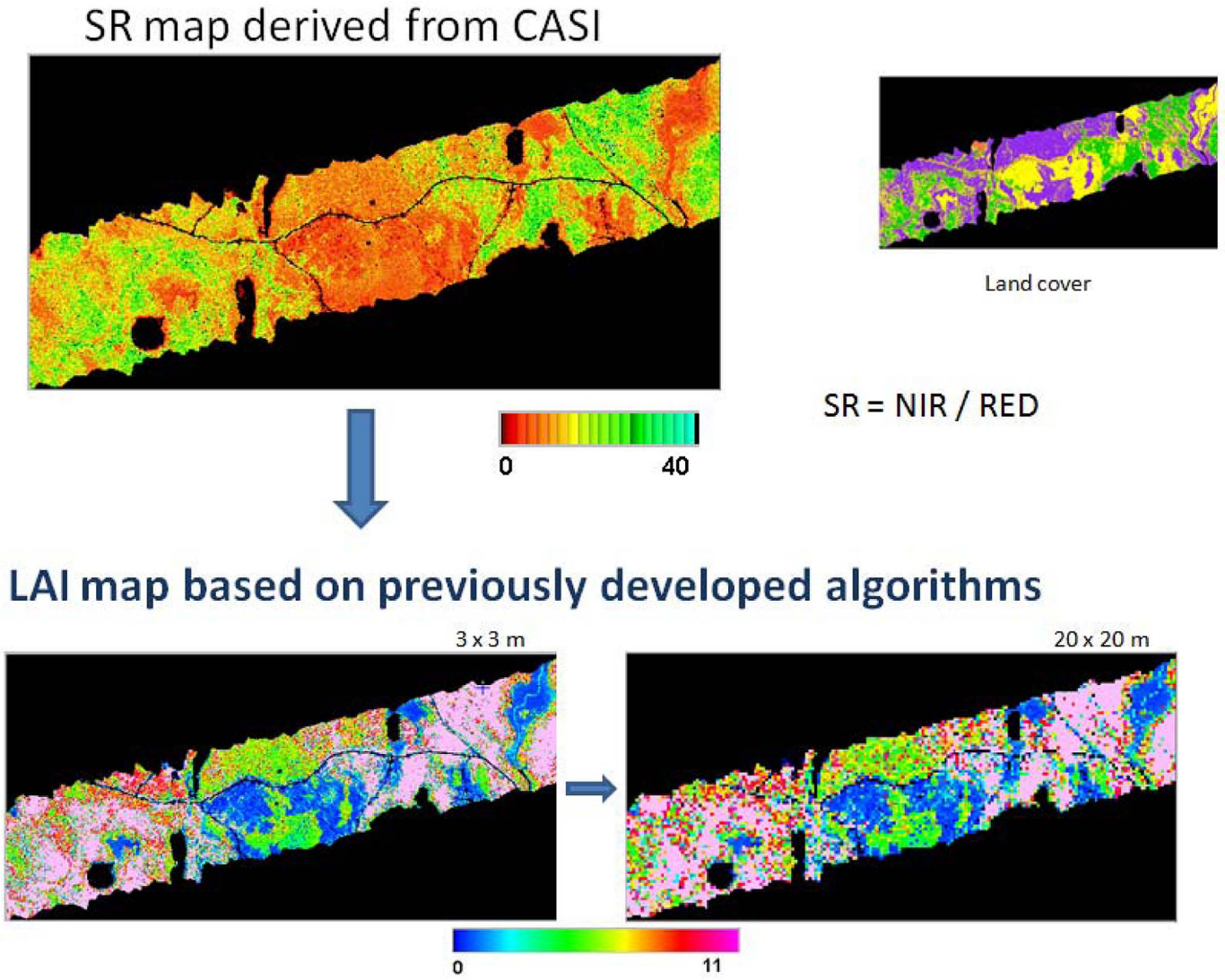


Fig. 12. (a) SR map generated from the CASI nadir data. (b) LAI maps at 3 m. (c) LAI maps at 20-m resolution based on the previous SR–LAI algorithms developed by Miller [39].

of Leblanc *et al.* [7], the FLAIR model simulations resulted in similar average NDHD values (0.5), and the values for broadleaf species did not differ considerably from those for coniferous species. In the same study, the clumping index ranges were similar for different land cover types.

### *D. Generating LAI Maps*

The SR map for the whole study area and the LAI maps generated (3 and 20-m resolutions) using previously developed algorithms by Chen *et al.* [41] were based on Landsat TM (Fig. 12). It has been noted that the LAI values, particularly for deciduous forest, are considerably high. Although the maximum value is set to 11, most of the deciduous pixels could reach values higher than 20 without this limitation. The difference in the sensor calibration of Landsat TM and CASI may cause such large LAI values. We therefore need to develop CASI sensor-specific LAI algorithm for LAI mapping using CASI data.

Along with the clumping index measurements, extensive $L_e$ and LAI measurements were also made in the 2008 campaign.

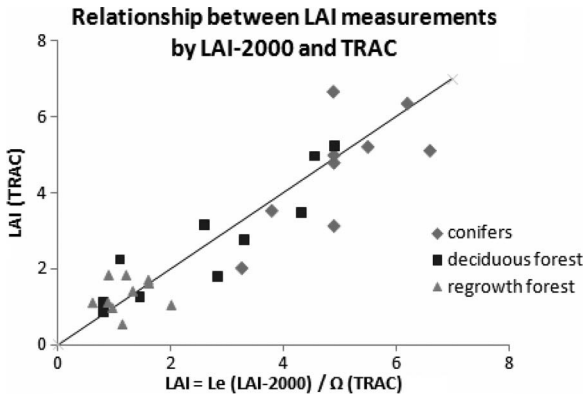


Fig. 13. Relationship between LAI calculated from LAI-2000 measurements of $L_e$ and TRAC measurements of $\Omega$, and LAI measured by TRAC directly.

We used these data to develop new LAI algorithms for CASI application. To increase the confidence in our measurements with LAI-2000, we compare them with the measurements performed by TRAC while measuring the clumping index. The

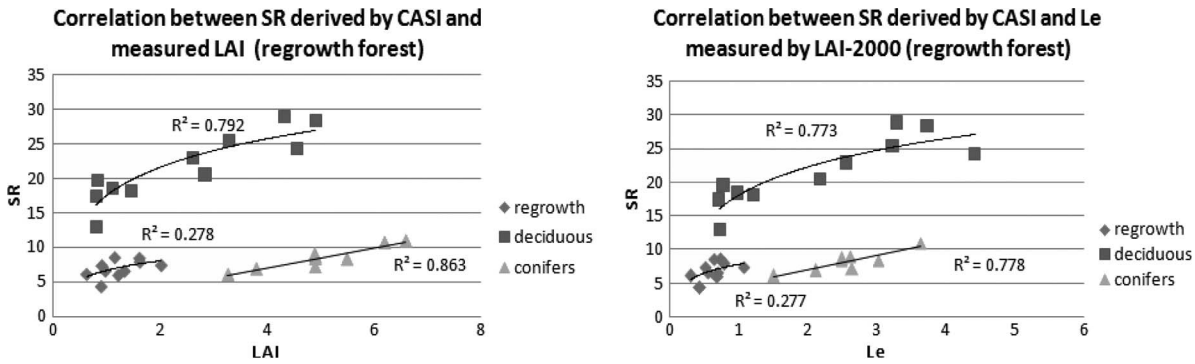


Fig. 14. Correlation between SR and (a) LAI calculated from field-measured $L_e$ and $\Omega$, and (b) field-measured $L_e$.

correlation between measurements from these two instruments is relatively strong (Fig. 13), having $R^2$ of 0.69. TRAC exhibits somewhat lower values than LAI-2000 for mature conifers. Considering the inhomogeneity of these forests and the fact that these two instruments measure different portions of a stand (TRAC for the canopy in the Sun's direction, while LAI-2000 for the hemispherical average), this level of agreement between these two instruments is satisfactory, although Chen *et al.* [46] found a better agreement for 17 tower flux sites which are more homogeneous.

We incorporated $\Omega$ into the LAI calculations in two different ways as explained in Section III-C3. In order to explore at which stage the clumping index should be included into the LAI calculations, we correlate SR with measured LAI [Fig. 14(a)] and with measured $L_e$ [Fig. 14(b)]. The LAI is derived from the measured $L_e$ and $\Omega$ using (3). Strong correlations are seen on both figures for deciduous and coniferous forests ($R^2 = 0.79$ and $R^2 = 0.77$ for deciduous forest; $R^2 = 0.86$ and $0.78$ for mature conifers). Low correlations are seen for regrowth forests in both cases ($R^2 = 0.28$) because of the small dynamic ranges. For mature conifers, the relationship is linear, and the SR–LAI relationship is stronger than the SR–$L_e$ relationship. The relationships for deciduous and regrowth forests are logarithmic, and SR–LAI relationships are also stronger. These findings are in the general agreement with that in [44]. Although they found that the correlation between SR and $L_e$ for conifers is stronger than that between SR and LAI in late spring when the background effect is relatively small, they also found that the SR–LAI relationship is stronger than the SR–$L_e$ relationship in midsummer when the background vegetation contributes significantly to the canopy-level SR.

Large $L_e$ variability is observed within the deciduous and mature coniferous forests (Figs. 14 and 15). This is a result of large-scale variations in the density of the deciduous forest (less homogenous), while the regrowth forest exhibits a smaller range variation in the cover density [34]. This is in general agreement with the findings of Chen [45], where lower $L_e$ is observed for young jack pine stands than in old conifer stands.

The LAI maps are calculated using different approaches, as explained in Section III-C3. Fig. 16(a) shows a part of the LAI image clipped for Scene 2 (developed using previous algorithms as shown in Fig. 12). As noted, it shows overestimated LAI values for the deciduous forest. Both $LAI_{\text{SR-}L_e\text{-}\Omega}$ and $LAI_{\text{SR-LAI}}$ maps exhibit similar mean values (Fig. 16(b)

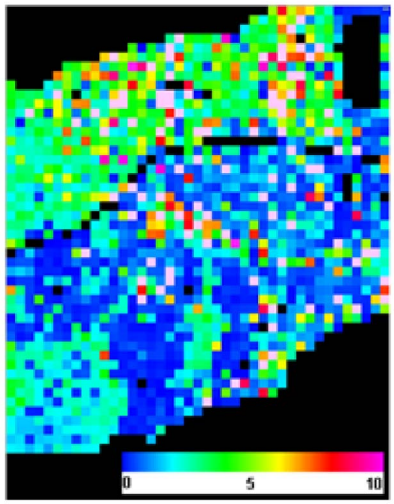


Fig. 15. $L_e$ map empirically derived based on the SR–$L_e$ relationship for Scene 2.

and (c), Table III) and improvement over the LAI map based on the previous algorithms. However, we cannot conclude that the clumping index map combined with the $L_e$ map enhance the results of the LAI map. The $LAI_{\text{SR-}L_e\text{-}\Omega}$ map shows somewhat higher values within the mature conifers (Fig. 16(c), Table III) than the $LAI_{\text{SR-LAI}}$ map (Fig. 16(b) and Table III). Although our results are not yet conclusive, we feel that separate mapping of $L_e$ and clumping may not lead to improved LAI mapping due to the confounding influence of the background. However, ground clumping measurement is the necessary step to obtain the correct LAI for LAI algorithm development, and mapping clumping is critical when LAI is used for radiation and productivity estimations.

The measured LAI values are based on gap-fraction measurements and include only the overstory foliage, while the SR responds to both overstory and understory. The understory increases the overall reflectance and affects the SR–LAI relationship. The impact is stronger in the open-canopy mature conifer stands where more light is available to the understory vegetation and where the understory is less spectrally distinct from the overstory canopy (e.g., SB 7—Fig. 9). The CASI data were collected in the summer when the understory was fully developed and when its contribution was maximum. The conifer sites were dominated by the black spruce, where the seasonal

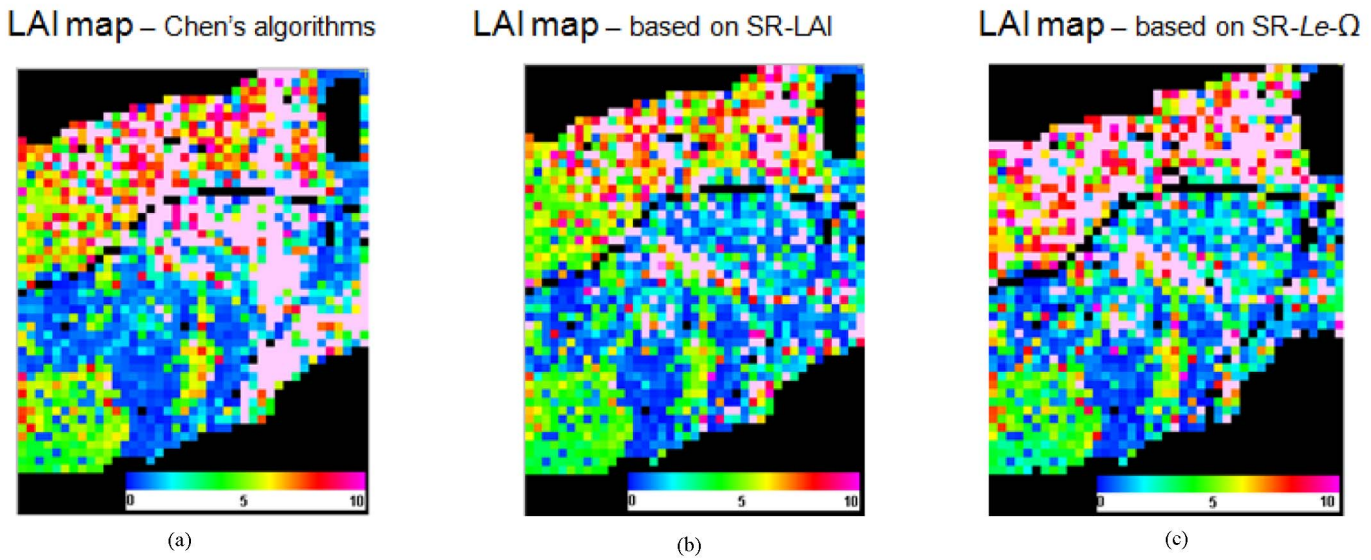


Fig. 16. LAI maps for Scene 2 generated by using (a) previous algorithms by Miller [39], (b) empirical SR–LAI algorithm $(LAI_{\text{SR-LAI}})$, and (c) empirical SR$-L_e-\Omega$ algorithm $(LAI_{\text{SR-}L_e\text{-}\Omega})$. Pixels shown in light pink have values over ten and represent generally overestimated values.

TABLE III
MAP STATISTICS: DIFFERENCES BETWEEN THE LAI MAPS GENERATED USING DIFFERENT ALGORITHMS. NOTE THAT MAXIMUM AND MINIMUM REPRESENT 5% TRIMMED VALUES OF EACH TAIL OF THE HISTOGRAM

| | MAP STATISTICS – Difference between maps | | | |
|---|---|---|---|---|
| **Maps** | **Land Cover** | **Min pixel value*** | **Max pixel value*** | **Mean pixel value** |
| $LAI_{CHEN} - LAI_{SR\text{-}LAI}$ | Deciduous | 4.49 | 9.23 | 7.57 |
| | Mature conifers | 0 | 1.27 | 0.69 |
| | Regrowth | -0.65 | 0.13 | -1.47 |
| $LAI_{SR\text{-}LAI} - LAI_{SR\text{-}Le\text{-}\Omega}$ | Deciduous | -2.27 | 0.07 | -0.59 |
| | Mature conifers | -3.58 | 0.71 | -1.13 |
| | Regrowth | -1.1 | 0.4 | -0.31 |

variability of the understory reflectance can be significant [44]. To eliminate the impact of the understory in retrieving LAI, spring images, when the understory is not developed, would be more useful for estimating the overstory LAI. Chen and Cihlar [44] found that the coefficient of determination $(R_2)$ from the linear regression between SR and LAI decreased from 0.53 in late spring to 0.49 in midsummer.

### *E. Uncertainties*

The accuracy of both the field and remote sensing data is important for the overall accuracy of the results. Minor uncertainties could be related to different Sun zenith angles during the TRAC measurements [7].

The CASI sensor can be adjusted by 5° at a time relative to the nadir. In order to acquire an image as close to the hotspot as possible, the timing of the acquisition was selected to be as close to 40° solar zenith angle as possible. As a result, both view zenith and solar zenith angles matched almost exactly at 40° (within 0.5°). However, across the swath of the scanned image, the view angle from the nadir varied ±17°, and this may cause error for hotspot sampling as the reflectance decreases exponentially from the hotspot [22]. The chosen areas for clumping analysis are part of the swath with an angle range of 10° from the principal solar plane. We estimate that the absolute error in hotspot measurements due to this departure from the principal solar plane is in the range of 0%–0.94% and that the relative error is in the range of 0%–12.4%, and we therefore did not make correction for this error.

Heterogeneity of some sites may also impact the field measurements and image classification. For instance, regrowth forests are dominated by jack pine trees but also have some patches of open land where these trees are absent. These patches have an impact on the NDHD values in the process of aggregation from 3- to 20-m pixels on the images.

Due to relatively low reflectance in the red spectra and high reflectance in NIR spectra, in some pixels mostly designated as deciduous pixels, some SR values are too high, resulting in inappropriate LAI. Chen and Cihlar [44] found that SR shows somewhat stronger sensitivity to LAI than normalized difference vegetation index; however, the reflectance from the understory vegetation may exert strong influence on the overall signal when SR is used. In addition, the SR map is derived from the nadir view data, while the NDHD maps are based on the off-nadir spectra. This may result in less accurate overall results when errors in the image coregistration are significant; vegetation indexes derived from off-nadir data may provide stronger correlations. The combination of the nadir images (SR and land cover) and off-nadir images (NDHD and clumping index) allows for some feature to exhibit slightly different shapes,

resulting in ambiguous pixels along the edges of vegetation patches. It is also important to note that the CASI data were acquired in 2007 and the clumping index field measurements were collected in 2008; our results may be influenced to a small extent by the possible changes in tree size, particularly in young jack pine stands.

## V. Discussion

The highest clumping (the lowest clumping index) is observed for the mature conifer forests (Figs. 5–8 and Fig. 11) as they exhibit more shadows and therefore have darker darkspots. The shaded tree crown component is perhaps the most important structural element of the clumping in the darkspot. Leblanc *et al.* [7] did not find evidence that the clumping index is directly related to the hotspot but found some evidence that it is related to the darkspot. The clumping index information seems to be more related to the darkspot, while the hotspot serves as the normalizing factor when used in the NDHD [7]. The foliage distribution within tree crowns is another structural element that has an impact on the overall clumping. Highly clumped conifer canopies have denser crowns, which cause more shadows at the darkspot [2]. The deciduous forests exhibit distinctive and lower clumping (high clumping index) than other forests due to the canopy and leaf structure. It is of importance to note that, in both cases of the modeled clumping index values, using previously developed algorithms and site-specific calculations (Figs. 6 and 11, respectively), the clumping index demonstrates minor differences between mature conifer and regrowth forests. This can be related to the differences between NDHD for the mature conifers and regrowth forests (as will be explained hereinafter). The similarity in the TRAC measurements between mature conifers and regrowth forests suggests that both forest types have the same gap sizes relative to shoot sizes. Chen [45] found that only the large gaps between the tree crowns are responsible for the nonrandom foliage spatial distribution and are critical in the element clumping index calculations measured by TRAC. Kurcharik *et al.* [3] also argue that small gaps may not be measured accurately with the TRAC instrument because of the problem with the penumbra effect; they explored a multiband vegetation imager using the gap fraction and gap-size distribution theory where small gaps can be resolved. However, the small gaps between shoots within branches may just slightly change the gap-size distribution, having a minimal effect on the clumping index [45]. Chen and Cihlar [34] reported that the accuracy of deriving the element clumping index from gap-size distribution is approximately 97%. The effect of clumping within the shoots was quantified using the needle-to-shoot area ratio values of 1.6 ($\gamma_E \sim 1.6$). In the darkspot direction, one can be expected to have the maximum contribution of specularly reflected radiation, which, for the most canopy modeling, has been neglected. It may be an additional factor confounding the results.

Two crown structural characteristics, namely, crown height and within-crown density, are major factors that impact the NDHD difference between mature conifers and regrowth forests. As the young pine trees are short ($\sim$1.0 m), their crowns are less dense than the crowns of old conifer trees, allowing for vertical transmission through the crowns. These structures exhibit less shadowing and, ultimately, higher (brighter) darkspot values. This results in reduced contrast between the hotspot and darkspot values. The young pine trees generally have a higher clumping (aggregation) of needles to shoot than black spruce, suggesting that, for the same amount of foliage, light transmittance is higher and the NDHD is lower [45]. Overall, the young pine trees transmit more light, which results in lower NDHD. On the other hand, the old conifer trees have lower within-crown vertical transmittance than the young jack pine trees. These trees have denser crowns, and their height is larger than the between-crown openings. This reduces the penetration of incident radiation in the viewing direction [1], and the sensor sees more shadows, which results in a low darkspot reflectance and, ultimately, higher NDHD.

The understory plays an important role in the NDHD variability. With respect to the mature conifers, the black spruce forest has more developed understory vegetation than old jack pine trees where more soil is exposed. As the soil exposure increases the hotspot values at a higher rate than understory vegetation, the regrowth forest generally exhibits higher hotspot than mature conifer forests. This component of the overall reflectance generally increases NDHD for regrowth forests. While soil exposure, seen more in regrowth forests, can cause high variability of the NDHD values in the red region, much less effect of the soil and understory vegetation on the overall reflectance and NDHD is seen in the NIR spectrum as both soil and understory have high reflectance in this spectral region. Furthermore, the contribution from multiply scattered radiation is less emphasized in the red band than in NIR; the red spectral region exhibits darker shadows, and directional effects are larger. This additionally increases the NDHD values in the red band (Figs. 5 and 7). Thus, less variability in the NIR region is a result of the reduced contrast between the signatures observed at hotspot and darkspot due to higher multiple scattering effects [1], [4]. Overall, it can be concluded that the background optical properties have stronger impacts on the hotspot while canopy structural properties, related to both element and within-crown clumping, affect more pronouncedly the shadow components and, thus, the darkspot. The balance between these two effects causes the difference in contrast between hotspot and darkspot, and this contrast affects changes in NDHD. From the analysis of our data and results, we conclude that canopy structural arrangements are more responsible for the NDHD differences between regrowth and mature conifer forests in this paper, resulting in lower NDHD for regrowth forests and higher NDHD for mature conifers.

We have incorporated anisotropy of the background by introducing the NTAM model to 5-Scale; this has further enhanced the performance of the model. After performing a sensitivity analysis of 5-Scale (not shown), we may conclude that the model is quite sensitive to background input spectra. Understory spectra may have a considerable impact on canopy reflectance modeling, and the ability to separate understory from canopy reflectance data would increase the accuracy [18]. Atmospheric correction is another important factor affecting the validity of the visible range. As input background and foliage spectra to 5-Scale are reliable, the shape of the simulated

canopy-level reflectance spectra should also be reliable, although the absolute values may be in error due to uncertainties in the scene component modeling. The 5-Scale model is reliable in calculating both NDHD and clumping index (Fig. 11). In addition, it is capable of distinguishing the clumping index values to such an extent that it successfully mimics the satellite data and provides more precise values than those measured by TRAC. Several studies suggest that 5-Scale simulations can help in assessing the relationship between clumping index and angular properties of directional reflectance [7], [14].

The fact that the $LAI_{\text{SR-LAI}}$ map exhibits more reasonable values than the $LAI_{\text{SR-}L_e\text{-}\Omega}$ map (Fig. 16, Table III) can be likely related to the stronger correlation between SR and LAI for conifers during the summer season (Fig. 14). As $L_e$ is a gap-fraction-dependent parameter, it is more critical in spring when there is high contrast between the overstory vegetation and open-space gaps. During summer time, SR increases due to two factors: new leaves on the top of old leaves on trees and understory vegetation. At the same time, canopy LAI increases due to the new leaves, while $L_e$, as gap dependent, stays almost unchanged [44]. This simultaneous increase of both LAI and SR, while $L_e$ stays almost the same, strengthens the SR–LAI relationship during summer. When stem density and crown size increase, the clumping decreases (element clumping index increases, approaching unity and random leaf distribution), but foliage clumping within shoots still remains and becomes more important for the calculation of LAI (nonrandom distribution) than element clumping [44]. The concept should be further explored with the data acquired in spring, when a stronger correlation between SR and $L_e$ is expected for conifers [44].

In this paper, the utilization of the hotspot and darkspot spectra is shown to be very effective in retrieving the clumping index. The strong relationship between NDHD, derived from the hotspot and darkspot measurements, and the field-measured clumping index justifies the new emphasis on multiangle measurements. The correlation between NDHD and clumping index may also serve as a surrogate for the validation of the hotspot and darkspot reflectance measurements, as no multiangle reflectance measurements of forest canopies are commonly available. The field-measured clumping index exhibits high variability for all cover types, particularly for deciduous and regrowth forests. This confirms the importance of the clumping index mapping; accepting one value for a cover type in ecological models [47] would result in a coarse estimation. Although no significant improvement of clumping index incorporated with $L_e$ is observed in this paper, it is important to emphasize that, besides LAI, clumping index mapping is essential to calculate accurate sunlit and shaded leaf components of canopy and, thus, accurate net primary productivity and the retrieval of other vegetation biophysical and biochemical information. Accounting for clumping is essential to scaling leaf photosynthetic rates to the canopy [1], [3], [23]. The element clumping index varies with the solar zenith angle to some extent [3], [45]. In this paper, the clumping index measurements were generally performed in early afternoon, within the time frame of the CASI acquisition. The solar zenith angle was 40°, and therefore, this measurement geometry provided a reasonable average clumping index values [45].

## VI. Conclusion

Through the intensive process of validation, we demonstrated that the combination of the hotspot and darkspot reflectances had the strongest signals related to the vegetation structure quantified using the foliage clumping index. High correlation between NDHD and field-measured clumping index suggested the importance of multiangle measurements in retrieving structural vegetation parameters. Through this intensive validation using multiangle airborne data and field measurements, the following conclusions are drawn.

1) The clumping index algorithm initially developed by Chen *et al.* [23] for global applications and refined in this paper has two shortcomings: 1) The difference in the effect of background (understory, litter, and soil) on the hotspot and darkspot reflectances was not considered, leading to considerable variation of hotspot and darkspot (NDHD) and clumping index for open stands where the understory is abundant, and 2) the influence of vegetation height on NDHD is not explicitly considered, leading to biased estimates of clumping for newly regenerated forests. The first shortcoming is minimized in this paper by introducing a simple BRDF model (NTAM) for the background. The second shortcoming can be addressed through using an additional parameter, tree height or stand age, in the algorithm. We therefore need to make efforts to obtain the spatial information of one of these parameters for the purpose of vegetation structural mapping.
2) Separate $L_e$ and clumping mappings do not seem to improve LAI mapping due to the confounding influence of the understory and possibly due to errors in coregistering multiangle images. However, clumping should be considered in the ground measurement of LAI for its algorithm development. A clumping index map would nevertheless be useful in conjunction with a LAI map for estimating the radiation distribution within the canopy and vegetation productivity.

## Acknowledgment

The authors would like to thank L. Gray from York University for efficient airborne campaign with the CASI instrument.

## References

[1] R. Lacaze, J. M. Chen, J.-L. Roujean, and S. G. Leblanc, "Retrieval of vegetation clumping index using hot spot signatures measured by POLDER instrument," *Remote Sens. Environ.*, vol. 79, no. 1, pp. 84–95, Jan. 2002.

[2] J. M. Chen, J. Liu, S. G. Leblanc, R. Lacaze, and J.-L. Roujean, "Multiangular optical remote sensing for assessing vegetation structure and carbon absorption," *Remote Sens. Environ.*, vol. 84, no. 4, pp. 516–525, Apr. 2003.

[3] C. J. Kurcharik, J. M. Norman, L. M. Murdock, and S. T. Gower, "Characterizing canopy nonrandomness with a multiband vegetation imager (MVI)," *J. Geophys. Res.*, vol. 102, no. D24, pp. 29 455–29 473, 1997.

[4] R. Lacaze and J.-L. Roujean, "G-function and hot spot (GHOST) reflectance model: Application to multi-scale airborne POLDER measurement," *Remote Sens. Environ.*, vol. 76, no. 1, pp. 67–80, Apr. 2001.

[5] D. W. Deering, T. F. Eck, and B. Banerjee, "Characterization of the reflectance anisotropy of three boreal forest canopies in spring–summer," *Remote Sens. Environ.*, vol. 67, no. 2, pp. 205–229, Feb. 1999.

[6] P. Y. Deschamps, F. M. Breon, M. Leroy, A. Podaire, A. Bricaud, J. C. Buriez, and G. Seze, "The POLDER mission: Instrument characteristics and scientific objectives," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 598–615, May 1994.

[7] S. G. Leblanc, J. M. Chen, H. P. White, R. Latifovic, R. Lacaze, and J.-L. Roujean, "Canada-wide foliage clumping index mapping from multiangular POLDER measurements," *Can. J. Remote Sens.*, vol. 31, no. 5, pp. 364–376, Oct. 2005.

[8] J.-L. Widlowski, B. Pinty, N. Gobron, M. M. Verstraete, D. J. Diner, and A. B. Davis, "Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies," *Clim. Change*, vol. 67, no. 2/3, pp. 403–415, Dec. 2004.

[9] B. Pinty, J.-L. Widlowski, N. Gobron, M. M. Verstraete, and D. J. Diner, "Uniqueness of multiangular measurements—Part 1: An indicator of subpixel surface heterogeneity from MISR," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 7, pp. 1560–1573, Jul. 2002.

[10] F. Canisius and J. Chen, "Retrieving forest background reflectance in a boreal region from multi-angle imaging spectroradiometer (MISR) data," *Remote Sens. Environ.*, vol. 107, no. 1/2, pp. 312–321, Mar. 2007.

[11] Y. Knyazikhin, J. V. Martonchik, D. J. Diner, M. M. Verstraete, B. Pinty, and N. Godron, "Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data," *J. Geophys. Res.*, vol. 103, no. D24, pp. 32 239–32 256, 1998.

[12] M. J. Chopping, A. Rango, K. M. Havstad, F. R. Schiebe, J. C. Ritchie, T. J. Schmugge, A. N. French, L. Su, L. McKee, and M. R. Davis, "Canopy attributes of desert grassland and transition communities derived from multiangular airborne imagery," *Remote Sens. Environ.*, vol. 85, no. 3, pp. 339–354, May 2003.

[13] N. Gobron, B. Pinty, M. M. Verstraete, J.-L. Widlowski, and D. J. Diner, "Uniqueness of multiangular measurements—Part 2: Joint retrieval of vegetation structure and photosynthetic activity from MISR," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 7, pp. 1574–1592, Jul. 2002.

[14] A. Simic and J. M. Chen, "Refining a hyperspectral and multiangle measurement concept for vegetation structure assessment," *Can. J. Remote Sens.*, vol. 34, no. 3, pp. 174–191, 2008.

[15] M. Rautiainen, M. Lang, M. Mottus, A. Kuusk, T. Nilson, J. Kuusk, and T. Lukk, "Multi-angular reflectance properties of a hemiboreal forest: An analysis using CHRIS PROBA data," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2627–2642, May 2008.

[16] P. Lewis, M. J. Barnsley, and M. Cutter, "CHRIS-PROBA: Mission status and prospects for mapping surface biophysical parameters," in *Proc. IGARSS*, Jul. 2001.

[17] F. Vuolo, L. Dini, and G. D. Urso, "Assesment of LAI retrieval accuracy by inverting art model and a simple empirical model with multiangular and hyperspectral CHRIS/PROBA data from Sparc," in *Proc. 3rd CHRIS/Proba Workshop*, Jun. 2005.

[18] G. D. Urso, L. Dini, F. Vuolo, L. Alonso, and L. Guanter, "Retrieval of leaf area index by inverting hyperspectral multi-angular CHRIS/PROBA data from SPARC 2003," in *Proc. 2nd CHRIS/Proba Workshop*, Jul. 2004.

[19] M. Schlerf and J. Hill, "Estimation of forest biophysical characteristics through coupled atmosphere-reflectance model inversion using hyperspectral multi-directional remote sensing data—A contribution to future forest inventory strategies," in *Proc. 3rd CHRIS/Proba Workshop*, Jun. 2005.

[20] S. Sandmeier and D. W. Deering, "Structure analysis and classification of boreal forests using airborne hyperspectral BRDF data from ASAS," *Remote Sens. Environ.*, vol. 69, no. 3, pp. 281–295, Sep. 1999.

[21] W. Qin and Y. Xiang, "On the hotspot effect of leaf canopies: Modeling study and influence of leaf shape," *Remote Sens. Environ.*, vol. 50, no. 2, pp. 95–106, Nov. 1994.

[22] J. M. Chen and J. Cihlar, "A hotspot function is a simple bidirectional reflectance model for satellite applications," *J. Geophys. Res.*, vol. 102, no. D22, pp. 25 907–25 913, 1997.

[23] J. M. Chen, C. H. Menges, and S. G. Leblanc, "Global mapping of foliage clumping index using multi-angular satellite data," *Remote Sens. Environ.*, vol. 97, no. 4, pp. 447–457, Sep. 2005.

[24] J. M. Chen and S. G. Leblanc, "A four-scale bidirectional reflectance model based on canopy architecture," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 5, pp. 1316–1337, Sep. 1997.

[25] C. D. Anger, S. Mah, and S. Babey, "Technological enhancements to the compact airborne spectrographic imager (CASI)," in *Proc. 1st Int. Airborne Remote Sens. Conf. Exhib.*, 1994, pp. 205–213.

[26] X. Li, A. Stahler, and C. E. Woodcock, "A hybrid geometric optical-radiative approach for modeling albedo and directional reflectance of discontinuous canopies," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 466–480, Mar. 1995.

[27] H. P. White, J. R. Miller, and J. M. Chen, "Four-scale linear model for anisotropic reflectance (FLAIR) for plant canopies. Model description and partial validation," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 5, pp. 1072–1083, May 2001.

[28] J. M. Chen and S. G. Leblanc, "Multiple-scattering scheme useful for geometric optical modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 5, pp. 1061–1071, May 2001.

[29] T. P. Dawson, J. P. Curran, and S. E. Plummer, "LIBERTY—Modeling the effects of leaf biochemical concentration on reflectance spectra," *Remote Sens. Environ.*, vol. 65, no. 1, pp. 50–60, Jul. 1998.

[30] S. Jacquemoud, S. L. Ustin, J. Verdebout, G. Scmuck, G. Andreoli, and B. Hosgood, "Estimating leaf biochemistry using the PROSPECT leaf optical properties model," *Remote Sens. Environ.*, vol. 56, no. 3, pp. 194–202, Jun. 1996.

[31] S. G. Leblanc, P. Bicheron, J. M. Chen, M. Leroy, and J. Cihlar, "Investigation of directional reflectance in boreal forests using an improved 4-scale model and airborne POLDER data," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1396–1414, May 1999.

[32] R. Latifovic, J. Cihlar, and J. M. Chen, "A comparison of BRDF models for the normalization of satellite optical data to a standard Sun-target-sensor geometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 8, pp. 1889–1898, Aug. 2003.

[33] LI-COR Inc., LAI-2000 Plant Canopy Analyzer Operating Manual, Lincoln, NE, Apr. 1992.

[34] J. Chen and J. Cihlar, "Plant canopy gap-size analysis theory for improving optical measurements of leaf-area index," *Appl. Opt.*, vol. 34, no. 27, pp. 6211–6222, Sep. 1995.

[35] J. M. Chen, P. M. Rich, S. T. Gower, J. M. Norman, and S. Plummer, "Leaf area index of boreal forests: Theory, techniques, and measurements," *J. Geophys. Res.*, vol. 102, no. D24, pp. 29 429–29 443, 1997.

[36] I. Moorthy, J. R. Miller, and T. L. Noland, "Estimating chlorophyll concentration in conifer needles with hyperspectral data: An assessment at the needle and canopy level," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2824–2838, Jun. 2008.

[37] Y. Zhang, J. M. Chen, J. R. Miller, and T. L. Noland, "Leaf chlorophyll content retrieval from airborne hyperspectral remote imagery," *Remote Sens. Environ.*, vol. 112, no. 7, pp. 3234–3247, Jul. 2008.

[38] S. G. Sylvain, J. M. Chen, and M. Kwong, *Tracing Radiation and Architecture of Canopies TRAC Manual*. Ottawa, ON, Canada: Natural Resources Canada, 2002.

[39] J. B. Miller, "A formula for average foliage density," *Aust. J. Bot.*, vol. 15, pp. 141–144, 1967.

[40] RSI, Envi User's Guide, Boulder, CO, Sep. 2004, ENVI Ver. 4.1.

[41] J. M. Chen, G. Pavlic, L. Brown, J. Cihlar, S. G. Leblanc, H. P. White, R. J. Hall, D. R. Peddle, D. J. King, J. A. Trofymow, E. Swift, J. Van der Sanden, and P. K. E. Pellikka, "Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 165–184, Apr. 2002.

[42] D. L. Verbyla and T. O. Hammond, "Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids," *Int. J. Remote Sens.*, vol. 16, no. 3, pp. 581–587, 1995.

[43] M. Kneubuhler, B. Koetz, R. Richter, M. Schaepmann, and K. Itten, "Geometric and radiometric pre-processing of CHRIS/PROBA data over mountainous terrain," in *Proc. 3rd CHRIS/Proba Workshop*, Jun. 2005.

[44] J. Chen and J. Cihlar, "Retrieving leaf area index of boreal conifer forests using Landsat TM images," *Remote Sens. Environ.*, vol. 55, no. 2, pp. 153–162, Feb. 1996.

[45] J. M. Chen, "Optically-based methods for measuring seasonal variation of leaf area index in boreal conifer stands," *Agric. For. Meteorol.*, vol. 80, no. 2–4, pp. 135–163, Jul. 1996.

[46] J. M. Chen, A. Govind, O. Sonnentag, Y. Zhang, A. Barr, and B. Amiro, "Leaf area index measurements at Fluxnet-Canada forest sites," *Agric. For. Meteorol.*, vol. 140, no. 1–4, pp. 257–268, 2006.

[47] J. Liu, J. M. Chen, J. Cihlar, and W. Chen, "Net primary productivity distribution in the BOREAS study region from a process model driven by satellite and surface data.," *J. Geophys. Res.*, vol. 104, no. D22, pp. 27 735–27 754, 1999.

**Anita Simic** received the H.B.Sc. degree in environmental science and the M.Sc. and Ph.D. degrees in remote sensing from the University of Toronto, Toronto, ON, Canada, in 1998, 2002, and 2009, respectively.

From 2002 to 2005, she was an Intermediate Research Scientist with the Canada Centre for Remote Sensing, Ottawa, ON. She is currently a Sessional Instructor with the University of Toronto and Ryerson University, Toronto, where she has been teaching remote sensing and geographic information systems since 2005. Her doctoral research was concerned with refining the concept of combining multiangle and hyperspectral sensors for retrieval of structural and biochemical vegetation parameters. Her previous work included applications of remote sensing to hydrology and hydrogeology and carbon budget modeling.

**Jing M. Chen** received the B.Sc. degree in applied meteorology from the Nanjing Institute of Meteorology, Nanjing, China, in 1982 and the Ph.D. degree in meteorology from Reading University, Reading, U.K., in 1986.

From 1989 to 1993, he was a Postdoctoral Fellow and Research Associate with the University of British Columbia, Vancouver, BC, Canada. From 1993 to 2000, he was a Research Scientist with the Canada Centre for Remote Sensing, Ottawa, ON, Canada. He is currently a Professor with the University of Toronto, Toronto, ON, and an Adjunct Professor with York University, Toronto. He has published over 160 papers in refereed journals. His recent research interests are in the remote sensing of biophysical parameters, plant canopy radiation modeling, terrestrial water and carbon cycle modeling, and atmospheric inverse modeling for global and regional carbon budget estimation.

Dr. Chen is a Fellow of the Royal Society of Canada and a Senior Canada Research Chair. He served as an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 1996 to 2002.

**James R. Freemantle** received the B.Sc. degree in physics from the University of Waterloo, Waterloo, ON, Canada, and the M.Sc. degree in physics from York University, Toronto, ON, Canada.

Since 1990, he has been a Project Scientist with the Earth Observations Laboratory, Centre for Research in Earth and Space Technology, Department of Earth and Space Science and Engineering, York University. His research interests include atmospheric correction of airborne and satellite imagery, software development for the processing of airborne imagery from compact airborne spectrographic imagers, and real-time analysis of remote sensing data.

**John R. Miller** received the B.E. degree in physics and the M.Sc. and Ph.D. degrees in space physics from the University of Saskatchewan, Saskatoon, SK, Canada, in 1963, 1966, and 1969, respectively.

He spent two years on a postdoctoral fellowship with the Herzberg Institute, National Research Council, Ottawa, ON, Canada. Since 1972, he has been with York University, Toronto, ON, where he was a Professor of physics and earth and space science until his retirement in 2008. His remote-sensing interests include atmospheric correction and extraction of biophysical surface parameters through radiative transfer models from water-color reflectance and from canopy reflectance for forestry and agriculture applications. Over the past two decades, his primary focus has been on the application of reflectance spectroscopic techniques in remote sensing using imaging spectrometer sensors.

**Jan Pisek** received the M.Sc. degree in geoinformatics and cartography from Masaryk University, Brno, Czech Republic, in 2004 and the M.Sc. degree in physical geography from the University of Toronto, Toronto, ON, Canada, in 2005, where he is currently working toward the Ph.D. degree.

He is also currently with the Tartu Observatory, Toravere, Estonia. His work deals with the retrieval of vegetation background reflectance and clumping index and their incorporation in the global vegetation models. His research interests cover *in situ* and remote sensing vegetation structure retrieval, with multiangle data in particular.