A method for improving hotspot directional signatures in BRDF models used for MODIS

Ziti Jiao\textsuperscript{1,2,*}, Crystal B. Schaaf\textsuperscript{3,4}, Yadong Dong\textsuperscript{1,2}, Miguel Román\textsuperscript{5}, Michael J. Hill\textsuperscript{6}, Jing M. Chen\textsuperscript{7}, Zhuosen Wang\textsuperscript{3,4,5}, Hu Zhang\textsuperscript{1,2}, Edward Saenz\textsuperscript{3}, Rajesh Poudyal\textsuperscript{8}, Charles Gatebe\textsuperscript{9}, Francois-Marie Bréon\textsuperscript{10}, Xiaowen Li\textsuperscript{1,2}, Alan Strahler\textsuperscript{4}

\textsuperscript{1}State Key Laboratory of Remote Sensing Science, Research Center for Remote Sensing and GIS, and School of Geography, Beijing Normal University. Beijing 100875, China; \textsuperscript{2}Beijing Key Laboratory of Environmental Remote Sensing and Digital City, Beijing Normal University, Beijing 100875, China; \textsuperscript{3}Department of Environmental, Earth and Ocean Sciences, University of Massachusetts, Boston, MA, USA \textsuperscript{4}Center for Remote Sensing, Department of Earth and Environment, Boston University, Boston, MA, USA \textsuperscript{5}Terrestrial Information Systems Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA \textsuperscript{6}Department of Earth System Science and Policy, University of North Dakota, Clifford hall, 4149 University Avenue, Grand Forks, ND, 58202 \textsuperscript{7}Department of Geography and Program in Planning, University of Toronto, 100 St. George St., Room 5047, Toronto, Ontario, Canada M5S 3G3 \textsuperscript{8}Science Systems & Applications, Inc., NASA Goddard Space Flight Center, Lanham, MD, USA \textsuperscript{9}Universities Space Research Association, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA \textsuperscript{10}Laboratoire des Sciences du Climat et de l’Environnement, CEA/DSM/LSCE, 91191 Gif sur Yvette, France

*email: jiaozt@bnu.edu.cn
Abstract:
The semi-empirical, kernel-driven, linear RossThick-LiSparseReciprocal (RTLSR) Bidirectional Reflectance Distribution Function (BRDF) model is used to generate the routine MODIS BRDF/Albedo product due to its global applicability and the underlying physics. A challenge of this model in regard to surface reflectance anisotropy effects comes from its underestimation of the directional reflectance signatures near the Sun illumination direction; also known as the hotspot effect. In this study, a method has been developed for improving the ability of the RTLSR model to simulate the magnitude and width of the hotspot effect. The method corrects the volumetric scattering component of the RTLSR model using an exponential approximation of a physical hotspot kernel, which recreates the hotspot magnitude and width using two free parameters ($C_1$ and $C_2$, respectively). The approach allows one to reconstruct, with reasonable accuracy, the hotspot effect by adjusting or using the prior values of these two hotspot variables. Our results demonstrate that: (1) significant improvements can be made to this method by using the inverted hotspot parameters; (2) the reciprocal nature allow this method to be more adaptive for simulating the hotspot height and width with high accuracy, especially in cases where hotspot signatures are available; and (3) while the new approach is consistent with the heritage RTLSR model inversion used to estimate intrinsic narrowband and broadband albedos, it presents some differences for vegetation clumping index (CI) retrievals. With the hotspot-related model parameters determined a priori, this method offers improved performance for various ecological remote sensing applications; including the estimation of canopy structure parameters.
Keywords: BRDF, CAR, MODIS, POLDER, multiangle remote sensing, hotspot signature, hotspot kernel, linear RTLSR model, airborne measurements
1. Introduction

Semi-empirical kernel-driven linear Bidirectional Reflectance Distribution Function (BRDF) models have been widely used to determine the properties of complex heterogeneous environments from multi-angle remote sensing. These models have been used to produce the routine BRDF/Albedo products from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Lucht et al., 2000; Schaaf et al., 2002; Schaaf et al., 2011), the Polarization and Directionality of the Earth’s Reflectances (POLDER) (Bicheron and Leroy, 2000; Bacour and Bréon, 2005), the Meteosat Second Generation (MSG) (van Leeuwen and Roujean, 2002; Geiger et al., 2005), and the Visible/Infrared Imager/Radiometer Suite (VIIRS) on board the platforms of the Suomi National Polar-orbiting Partnership (NPP) (Justice et al., 2013). They have been also used to retrieve canopy structure parameters (e.g., Chopping et al., 2008; Wang et al., 2011; Hill et al., 2011; He et al., 2012), to examine the improved accuracy of land cover classification (de Colstoun et al., 2006; Jiao, Woodcock & Schaaf et al, 2011; Jiao & Li, 2012), to accumulate and apply prior knowledge of BRDF archetypal shapes (Li, Gao, Wang & Strahler, 2001; Jiao, Hill, Schaaf et al., 2014; Jiao, Zhang & Dong et al., 2015), to couple surface reflectance with atmospheric scattering for improving atmospheric correction algorithms (Hu et al., 1999; Wang et al., 2010; Román et al., 2010; Litvinov et al., 2011), to correct for the effect of remotely-sensed anisotropic reflectance (e.g., Leroy and Roujean, 1994; Li et al., 1996), and for monitoring ecosystem disturbance and vegetation dynamics (e.g., Friedl et al., 2002, 2010; Zhang et al., 2003).

Despite their wide use, recent studies have recognized that the directional signatures near
the Sun illumination direction (aka the hotspot effect) are often underestimated by the
semi-empirical BRDF models (e.g., Chen and Cihlar, 1997; Maignan et al., 2004; Román et al.,
2011) such as RTLSR (Wanner et al., 1995; Lucht et al., 2000). In particular, the volumetric
scattering (aka Ross) component of RTLSR, originally derived from a horizontally homogeneous
plant canopy (Ross, 1981), does not include all possible correlations between the illumination
and observation geometries (Kuusk, 1991; Jupp and Strahler, 1991; Qin & Geol, 1995). Although
the geometric-optical (Li-Strahler) component derived from geometric optical (GO) models
characterizes a significant hotspot effect (Li and Strahler, 1992), the RTLSR model that linearly
combines these two (Ross and Li-Strahler) components has difficulties to simulate the both the
magnitude and signature of the hotspot effect. This model deficiency, while not significantly
impacting an albedo retrieval (Huang, Jiao & Dong et al., 2013) (based on the integral of the
entire view-illumination hemisphere), may still constrain the application of BRDF models in
retrieving canopy structure parameters (e.g., clumping index) that need the hotspot amplitude as
primary input (e.g., He et al., 2012; Zhu et al., 2012).

Various efforts have been made to improve the hotspot effect for such models. Chen and
Cihlar (1997) enhanced the hotspot effect of the kernel-driven Roujean BRDF model by
multiplying the model with an exponential approximation of a physically-based hotspot function.
He et al. (2012) and Zhu et al. (2012) suggested that the Chen and Cihlar (1997) model might
still slightly underestimate the reflectance magnitude when it is used to extrapolate the hotspot.
Recent efforts have focused on correcting the RossThick kernel with a hotspot factor (Maignan
& Bréon et al., 2004; hereafter referred to as the Maignan method) based on the geometrical
principles of the intersection of viewed and sunlit leaf areas (Jupp and Strahler, 1991). With a view to improving retrieval of clumping index (CI) from the MODIS BRDF product, He et al. (2012) also developed a correction for the MODIS hotspot amplitude by adding the difference between POLDER and MODIS hotspot BRFs, which has also been used to correct the hotspot magnitude of MISR BRFs for CI retrievals (Pisek, Ryu, Sprintsin & He et al., 2013). On the other hand, Zhu et al. (2012) corrected the hotspot amplitude by multiplying both the geometric optical scattering and volumetric scattering items by the exponential approximation of a modified hotspot function; also based on the MODIS RTLSR model.

In this study, we propose a method for improving the hotspot effect of the linear RTLSR BRDF model. The method revises the RossThick kernel using the corrected exponential approximation of the hotspot function (Chen and Cihlar, 1997, thereafter named RTCLSR model here). The principle of the formation of the hotspot is based on a canopy gap size distribution function, but is approximated by using an exponential function with two free parameters (C₁ and C₂) characterizing the height and width of the hotspot effect. The hotspot kernel within-crown and between-crown has very similar shapes, and thus can be directly applied to the scenario where a canopy cover is provided with a uniform leaf orientation of horizontally homogeneous plant canopies, as was used in the assumption in deriving this kernel from canopy radiative transfer theory (Ross, 1981). Such a correction to the RossThick kernel mainly accounts for the correlation between two gap probabilities from sun and view in the hotspot direction, which was not properly considered when deriving the original RossThick kernel.

To validate this model, the study uses a variety of hotspot data to determine appropriate
hotspot-related parameters including POLDER, MODIS, airborne multiangle Cloud Absorption Radiometer (CAR) (King et al., 1986, Gatebe et al., 2003, Román et al., 2011), and two field-measured data sets (Irons et al., 1992; Deering et al., 1999). We also explore the sensitivity of two hotspot parameters \(C_1\) and \(C_2\) to hotspot-fits. Finally, we examined the influences of this new approach on the retrieval of intrinsic albedos and clumping index.

2. Hotspot data

2.1. POLDER-3 BRDF database

The spaceborne POLDER instrument can sample the land surface for viewing angles up to 60°-70° and for the full azimuth range, at a coarse spatial resolution of approximately 6 km. Comprehensive BRDF sampling with large spatial coverage enables this instrument to collect observations for the development of BRDF modeling, particularly capturing distinct hotspot signatures. The POLDER-3 sensor onboard the Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar (PARASOL) satellite acquired multiangular measurements for more than 6 years, which were used to create four BRDF databases (Bréon et al., 2007). The approach uses POLDER data from relatively homogeneous pixels to represent the dominant continental ecosystems. The data are geocoded, calibrated, atmospherically-corrected and cloudy-screened for acquiring the land surface bidirectional reflectance factors (BRFs) for each orbit. This study uses a monthly database containing 14,649 BRFs in 16 IGBP land cover classes. According to phase angle \(\xi \leq 5^\circ\), we extract \(~14,188\) BRFs. The spatial distribution of view and Sun geometries for a typical POLDER data set, i.e., IGBP_01_20060609brdf_ndvi08.0824_1671, is exemplified in Figure 1 (b). Since the ‘Snow and
Ice class tends to have maximum reflectance in forward scattering direction that current surface BRF models cannot characterize, we exclude this class from the database. The BRDF signatures are based on inversions of approximately 20 orbits worth of spectral BRF data, for each of the six POLDER bands. Up to 16 different multiangular measurements for a given POLDER pixel are included in each orbit. BRFs measured at two typical bands (red, centered 670 nm and NIR, centered 865 nm) are mainly selected to assess the hotspot effect reconstructed from the proposed model. Notably, POLDER-3 spectral measurements are not simultaneous, meaning that each channel is acquired with a slightly different viewing geometry. As such, to analyze the variability of hotspot signatures, this requires a careful assessment of measurement differences, including normalization of the POLDER viewing geometry (Bréon et al., 2005).

2.2. MODIS hotspot data

We extracted 2,275 MODIS hotspot data sets from one 10° × 10° tile (h20v11) of MODIS surface reflectance products (MOD09 and MYD09) at a spatial resolution of approximately 500-m, for 7 solar reflective bands, and using high-quality BRDF sampling distributions, represented by a phase angle range \( \zeta \leq 5° \). There are only 80 high-quality hotspot data sets within a phase angle range of \( \zeta \leq 1.5° \); indicating that MODIS seldom acquires exact hotspot observations. Each data set consists of at least 10 observations with a proper directional sampling in approximate principal plane (PP), and includes fewer hotspot measurements within the defined area (Figure 1, a). The corresponding MODIS tile h20v11 captures a wide range of grass-shrub-savanna vegetation types in the Southern Hemisphere. The 329th Day of Year (DOY) in 2010 represents a maturity season (Zhang et al., 2003). Solar zenith angles (SZAs) mainly
ranged from 20° to 40° during the study period. The full-model inversion quality of the MODIS BRDF/Albedo product (MCD43A2) accounts for approximately 70% of total pixels. The spatial distribution of MODIS observations and solar geometry from an extreme data set in this tile is presented in Figure 1 (c), where solid points are for MODIS observations, and circles are for solar geometries. The concentric circles are at 20° intervals.

Since the Terra and Aqua MODIS accumulate multiangular observations through multiple overlapping image swaths, the solar geometry corresponding to each observation will be different during a typical (16-day) retrieval period (e.g., approximately 20° for SZA in the h20v11 tile). This means that MODIS does not capture enough real-time observations in the principal plane (PP) under identical solar illumination conditions. Here, we adopt the assumption used with the MODIS BRDF/Albedo product, whereby the surface BRDF shape doesn’t change abruptly for the range of SZAs capturing during a 16-day retrieval period. With this assumption, we can accumulate some multiangular observations in the proximity of the hotspot (i.e., $\xi \leq 5°$) to constrain the RTCLSR model for acquiring the optimal free parameters describing the hotspot magnitude ($C_1$) and width ($C_2$). Figure 1 (d) presents an extreme example that uses the RTCLSR model to reconstruct BRDF shapes as a function of phase angle for IGBP savanna in the NIR band. Red solid points represent observations extracted by using this method in the approximate PP. The black curve is the reconstructed BRDF shape using the RTCLSR model with $C_1=0.4$, and $C_2=4.5°$ as the optimal values for the 2,275 MODIS data sets in this tile. The SZA was set to 30.58°, and the observation in the closest proximity to the hotspot direction is 31° (red point).
The BRDF shape reconstructed using the Maignan method (green curve) is provided here for visual comparison.

Figure 1 near here

2.3. Finer resolution measurements

We used airborne datasets collected by NASA’s Cloud Absorption Radiometer (CAR) (Gatebe et al., 2003), which captures hotspot signatures over 2 field sites. To acquire multiangular measurements, the CAR instrument is flown using a clockwise circular pattern above the surface repeatedly, and observes the reflected solar radiation at a fine angular resolution (i.e., 0.5° intervals through its 190° aperture at a rate of 100 scans per minute). This sampling scheme results in a BRDF retrieval that is based on 76,400 and 114,600 BRFs measurements per channel per complete orbit, which corresponds to a representative sample of the landscape-level (~5 km) reflected surface (Gatebe et al., 2003). At an altitude of 600m above the targeted surface area and 1° instantaneous field of view (IFOV), the pixel resolution is about 10 m at nadir and about 270 m at an 80° viewing angle (Tsay et al., 1998). The geolocation accuracy of CAR measurements is within an error margin of 0.3% (~2.0-3.5 meters as derived from the high resolution scene across the entire scan track). This accuracy holds well particularly with off-nadir looking observations (Gatebe et al., 2007). These CAR measurements used in this study are averaged at an angular resolution of 1° in the viewing hemisphere and are taken in the red (0.682 µm) and NIR (0.870 µm) bands.

CAR data source was mainly from the early Smoke Clouds and Radiation-Brazil (SCAR-B) field campaign on August, 1995. Two kinds of data sets were collected from the well-defined
surfaces of cerrado and dense forest (Tsay et al., 1998), both measured in Brazil under nearly clear-sky conditions (http://car.gsfc.nasa.gov/data/index.php?mis_id=5&n=SCAR-B&l=h). The cerrado comprises a landscape-scale mosaic of four main vegetation types ranging from campo limpo grassland, through campo sujo and campo ralo with small (< 2m) sparse-to-medium density woody plants overlaying grassland, to cerrado sensu stricto with 20 – 30 t/ha of woody biomass. The dense forest data includes two flight data (i.e., CAR Flight #1689 and CAR Flight #1693). These two forest data captured distinct hotspot signatures in the red (CAR Flight #1689) and the NIR (CAR Flight #1693) bands, and are used in this study. The area of dense forest was covered by tall trees with a large canopy where the ground surface is invisible and had a relatively homogeneous surface. Details about these airborne CAR measurements are referred to Tsay et al. (1998).

To compare these models being explored at a field scale, we also analyzed two high-quality multiangular field data sets reported in previous studies (e.g. Li et al., 2001; Strugnell et al., 2001; Huang & Jiao 2012; Jiao et al., 2014). These include soil multiangular measurements (Iron et al., 1992) acquired on a bare field located on a level alluvial plane within the United States Department of Agriculture Beltsville Agricultural Research Center, Beltsville, MD. The data were taken from full view angles and several solar illumination directions. A calibrated Barnes Model 12-1000 Modular Multiband Radiometer (MMR) with a 15° IFOV was used. Forest multiangular measurements (Deering et al., 1999) were also acquired with PARABOLA instrument with a 15° IFOV at a black spruce site that was mainly made up of old black spruce (picea mariana) with scattered emergent tamarack. The tree height was less than 10 m and the
total stem density was 8040 live stems/ha, with a basal area of 40 m²/ha. Canopy closure averaged about 55%, and the leaf area index (LAI) measured by an LAI-2000 in spring 1994 was 3.7 (Chen et al., 1997).

3. Model and method

3.1. RTLSR model

The semi-empirical, kernel-driven, linear BRDF model is a linear combination of three basic scattering components: isotropic scattering, volumetric scattering, and geometric-optical (GO) scattering. This model adopted a general form (Roujean et al., 1992; Lucht et al., 2000):

\[
R(\theta_v, \theta_i, \Delta \phi, \lambda) = f_{\text{iso}}(\lambda) + f_{\text{vol}}(\lambda)K_{\text{vol}}(\theta_v, \theta_i, \Delta \phi) + f_{\text{geo}}(\lambda)K_{\text{geo}}(\theta_v, \theta_i, \Delta \phi)
\]

(1)

Where \(f_{\text{iso}}(\lambda), f_{\text{vol}}(\lambda)\) and \(f_{\text{geo}}(\lambda)\) are the spectrally dependent model parameters. \(K_{\text{vol}}(\theta_v, \theta_i, \Delta \phi)\) and \(K_{\text{geo}}(\theta_v, \theta_i, \Delta \phi)\) are kernel functions of view zenith \(\theta_v\), illumination zenith \(\theta_i\), and relative azimuth \(\Delta \phi\) and provide shapes for volumetric scattering and geometric-optical scattering BRDFs; \(f_{\text{iso}}(\lambda)\) is a spectral constant for isotropic scattering; \(f_{\text{vol}}(\lambda)\) and \(f_{\text{geo}}(\lambda)\) are spectral constants, i.e., model anisotropic parameters that weight the two BRDFs; \(R(\theta_v, \theta_i, \Delta \phi, \lambda)\) is BRDF in waveband \(\lambda\).

\(K_{\text{geo}}\) and \(K_{\text{vol}}\) have been derived from physical approximation of the radiative transfer at the surface. \(K_{\text{geo}}\) is derived from the GO model (Li and Strahler, 1992; Roujean et al., 1992) and characterizes dome-shaped BRDF curves. It is a function that describes the shadowing and surface scattering from the canopy. The operational MODIS BRDF/Albedo algorithm adopted the LiSparseReciprocal kernel (\(K_{\text{LSR}}\)) that was derived from a sparsely-vegetated canopy surface.

\[
K_{\text{LSR}} = O(\theta_v, \theta_i, \Delta \phi) - \sec \theta_i \sec \theta_v + \frac{1}{2} \left( 1 + \cos \xi \right) \sec \theta_v \sec \theta_i
\]

(2)

Notably, where \(O(\theta_v, \theta_i, \Delta \phi)\) is overlap function of view and illumination shadows on the ground,
and hence the hotspot effect (i.e., the probability of seeing the sunlit background from the same
gap) is included in this kernel. Further expressions and several intermediate variables are detailed
to several papers (e.g., Wanner et al., 1995; Lucht et al., 2000).

Notably, two ratios \((h/b)\) and \((b/r)\) for describing vegetation structure are used to parameterize
overlap function in \(K_{\text{LSR}}\). Here, \(h\) is the mean height where a crown center is located, \(b\) is the
mean vertical half axis of the modeled ellipsoid, and \(r\) is the mean horizontal radius. These two
ratios are related to the hotspot effects in \(K_{\text{LSR}}\) on vegetation canopy scale (Li and Strahler, 1992).
The operational MODIS RTLSR algorithm adopts \(h/b=2\) and \(b/r=1\). In Figure 2 (top), red dashed
curve is for \(h/b=2\) and \(b/r=1.2\), and magenta dashed curve is for \(h/b=2.5\) and \(b/r=1\). Figure 2
demonstrates that, although these two ratios were in theory related to the general hotspot effect in
\(K_{\text{LSR}}\), they do not seem to be very sensitive to the changes of hotspot effect at a reasonable range.

\(K_{\text{vol}}\) is a similar function that describes the volumetric scattering component from canopy,
based on an assumption of a single-scattering approximation of the radiative transfer (RT) theory
by Ross (1981). The operational RTLSR model adopted RossThick kernel \((K_{\text{RT}})\) derived from a
homogeneously layered canopy with a large leaf area index (LAI) (Wanner et al., 1995)

\[
K_{\text{RT}} = \frac{\pi/2 - \xi}{\cos \theta_s + \cos \theta_i} \cos \xi + \sin \xi - \frac{\pi}{4}
\]

This kernel characterizes bowl-shaped BRDF curves, but does not consider the correlation
between illumination and observation processes for the observed hotspot. In theory, the
probability of observing a sunlit component can be taken as the product of two probabilities: one
for the view line from the observer and the other for illumination beam from Sun. When the view line and an illumination beam coincide, the observer either sees the sunlit foliage or the sunlit background, while the shadows of foliage and background are hidden from the view. As a result, the correlation tends to be 1:1, as the illumination and view directions get closer.

### 3.2. RTLSR hotspot correction

To consider the correlation between the view and illumination processes, Maignan et al. (2004) corrected the $K_{RT}$ with a hotspot factor:

$$K_{RTM} = \frac{\left(\frac{\pi}{2} - \xi\right) \cos \xi + \sin \xi}{\cos \theta_v + \cos \theta_i} \times \left(1 + \left(1 + \frac{\xi}{\xi_0}\right)^{-1}\right) - \frac{\pi}{4}$$

$$1 + \left(1 + \frac{\xi}{\xi_0}\right)^{-1}$$ is a simplified hotspot factor derived by Bréon et al. (2002) from the theory on the calculation of an overlay function of the intersection of viewed and sunlit leaf areas (Jupp & Strahler, 1991). This hotspot factor was used to correct $K_{RT}$, thereafter named $K_{RTM}$. Here $\xi$ is phase angle, $\xi_0$ is a characteristic angle in relation to the ratio of scattering element size and the canopy vertical density, which follows the range $\xi_0 = [1^\circ, 2^\circ]$. A $\xi_0 = 1.5^\circ$ has been suggested as a typical value representing a wide range of landscape conditions (Bréon et al., 2002; Maignan et al., 2004).

Note that the $K_{RTM}$ value near the hotspot direction will double rapidly (as much as $K_{RT}$) as the phase angle $\xi$ approaches 0°. To consider hotspot variability, we specified a $\xi_0 = 3.0^\circ$ for $K_{RTM}$ for a comparison with the standard $\xi_0 = 1.5^\circ$ (Figure 2). The results indicate that an increase in $\xi_0$ can increase hotspot width, but diverge from $K_{RT}$ beyond the hotspot region compared
with $K_{RTC}$. Therefore, we do not further examine the $\xi_0$ influence on hotspot-fits in this study, rather than adopt $\xi_0 = 1.5^\circ$ as the most appropriate value. Here, the $K_{RTM}$ has a simplified feature whereby the variation in hotspot height and width is determined by phase angle $\xi$.

For volumetric scattering kernel ($K_{vol}$), it is frequently required to meet the empirical requirement of $K_{vol} = 0$ when both the viewing and illumination geometries point to nadir. This requirement aims to ensure that the isotropic parameter is physically reflectance for the model retrieval (Roujean et al., 1992). Conversely, using the Maignan method, the nadir-view and nadir-sun reflectance are specified as $R(0,0,0,\lambda) = f_{iso}(\lambda) + f_{vol}(\lambda) \times \pi/4$.

Here, we introduce the modified exponential function of the hotspot kernel function to $K_{RT}$, hereby termed the RossThickChen kernel ($K_{RTC}$), which includes two free parameters characterizing hotspot variations (Chen and Cihlar, 1997):

$$K_{RTC} = \frac{\left(\frac{\pi}{2} - \xi \right) \cos \xi + \sin \xi}{\cos \theta_i + \cos \theta_s} \times \left(1 + C_1 e^{-\frac{\xi}{C_2}}\right) - \frac{\pi}{2}$$

(5)

Where $1 + C_1 e^{-\frac{\xi}{C_2}}$ is the modified hotspot function. The two free parameters, $C_1$ and $C_2$, allow a large dynamic range of hotspot variation, and thus facilitate the analysis of the variation of hotspot height and width in fitting hotspot BRFs. For this retrieval, we also make an empirical adjustment to retain the isotropic BRDF model parameter, $f_{iso}$, as the corresponding nadir-view and nadir-sun BRF.

Originally, the $K_{RTC}$ hotspot function was derived from a theoretical expression that accounts for the hotspot effect for a forest stand, based on a geometric-optical model (Chen & Leblanc, 1997). In this specific scenario, $C_1$, is linearly related to the difference between the
reflectances of the foliage and background at the wavelengths of interest (Chen & Cihlar, 1997).

In deriving this hotspot kernel, Chen and Leblanc (1997) demonstrated that this modification had very similar shapes at within-crown and between-crown scales because the gap size is scaled by the gap depth. For plane parallel layers of leaves, the average gap depth is taken as the mean distance between two layers of leaves and is related to the leaf area density. In this hotspot function ($K_{RTC}$), $C_1$ is related to the magnitude of the hotspot peak, while $C_2$ defines the half width of the hotspot in relation to the ratio of canopy height to the size of the predominant canopy structure. Notably for this study, to keep consistent with the original kernel form, the unit of $C_2$ is radians in terms of Eq. (5), but it is converted into degrees in subsequent use to describe the hotspot width.

Here, the $K_{RTC}$ kernel is empirically adjusted to meet $K_{RTC} = 0$ for the nadir-view and nadir-sun geometries, which results in a downward shift of kernel shape (Figure 2, bottom). Note that this adjustment does not affect the fitting ability of this model since the kernel shape is retained. Figure 2 aligns $K_{RTC}$ with $K_{RT}$ and $K_{RTM}$ for a convenient comparison. The results indicate that $K_{RTC}$ coincides with $K_{RT}$ in the scattering direction beyond hotspot region, and adjusts the hotspot height and width by changing $C_1$ and $C_2$ (Figure 2 bottom). A larger $C_1$ value also indicates a higher hotspot, while a larger $C_2$ value indicates a wider hotspot.

3.3. Hotspot parameter retrieval and analysis

The inversion strategy for retrieving the three parameters of this linear BRDF model is to minimize the root mean square error (RMSE) between model predicted and observed BRFs. The
full-inversion is a simple matrix inversion that is independent of land surface types and is performed pixel by pixel using all high-quality measurements available. Details are referred to the papers (Roujean et al., 1992; Lucht et al., 2000; Shuai et al., 2008).

To derive the $C_1$ and $C_2$ values, we calculated fit-RMSEs to minimize the model-observation fits using accumulated measurements in a close vicinity to the hotspot direction (i.e., $\xi \leq 5^\circ$).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left( R_{obs}^{\theta_1, \theta_2, \Delta \phi, \lambda} - R_{model}^{\theta_1, \theta_2, \Delta \phi, \lambda} \right)^2}{n-3}}$$  \hspace{1cm} (6)

Where $R_{obs}^{\theta_1, \theta_2, \Delta \phi, \lambda}$ and $R_{model}^{\theta_1, \theta_2, \Delta \phi, \lambda}$ are the model predicted and observed BRFs near the hotspot region in the viewing and solar geometries, $\theta_1$, $\theta_2$, $\Delta \phi$ at wavelength $\lambda$, as in Equation (1). We derived the optimized $C_1$ and $C_2$ values from the least RMSEs using the iteration step size of 0.1 for $C_1 = [0.3, 1.2]$ and $C_2 = [1.0^\circ, 6.0^\circ]$. This range of $C_1$ and $C_2$ can ensure realistic hotspot variations while optimizing processing time. Following this method, we retrieved the $C_1$ and $C_2$ values for the entire POLDER database; including parameterizations by IGBP cover type and phenological phase, respectively. To derive the optimal hotspot parameter values, we constrain this model by accumulating all measurements near hotspot direction ($\xi \leq 5^\circ$) for these two parameter types, including all 15 IGBP classes, excluding ‘Snow and Ice’ and 2 phenological stages: (1) maturity and (2) dormancy.

We then compared the model predicted and observed BRFs near hotspot direction by using scatterplots with regression lines. To display model-observation fits, we examined the model predicted and observed BRFs as a function of phase angle for the selected observations for
Evergreen Needleleaf Forest type (ENF) for each of POLDER’s reflective bands. Details on how to derive the shaded area (Figure 1, a) are provided in Appendix A. We also performed a comparison against finer resolution in-situ data. To analyze the sensitivity of the hotspot parameters to hotspot-fits in conjunction with the model parameters, we used simulated BRDF shapes generated from the airborne CAR data set.

Finally, we examined the intrinsic albedos (i.e., White Sky Albedo (WSA) and Black Sky Albedo (BSA)) and clumping index (CI) retrievals between models. To retrieve CI values, we made use of the algorithm by Chen et al. (2005) and assigned a simplified geometry of 45° view zenith angles in both the backward- and forward- scattering directions for the corresponding principal plane hotspot and dark spot, respectively (Zhu et al. 2012). Such a simplification is somewhat different from Chen et al. (2005), which focused on locating the optimal view geometry to capture the dark spot. Since the three models are generally consistent in fitting observations outside of the hotspot region, this simplification would not affect the inter-comparison of CI retrievals between models.

4. Results

4.1. $C_1$ and $C_2$ values for POLDER

In this section, we assess the performance of the two hotspot parameters, based on the underlying IGBP class and two vegetation phenological phases, based on a comparison of the model predicted hotspot BRFs with observed BRFs derived from the POLDER BRDF database. For this study, our analysis focuses on the RTCLSR and Maignan method.
4.1.1. Hotspot parameters in regard to surface type and phenology

Figure 3 (top left) presents the globally optimized $C_1$ and $C_2$ values (RTCLSR_GLOB) for 6 POLDER bands [490, 565, 670, 765, 865, and 1020nm]. These values indicate the variation trend of hotspot parameters for the available POLDER BRDF database. In general, the spectral variations of $C_1$ and $C_2$ present a highly negative correlation ($R^2 = 0.58$). This indicates that a higher hotspot magnitude also tends to go with a narrower shaped hotspot, and vice versa. Figure 3 (middle) presents the optimized $C_1$ and $C_2$ values derived for underlying IGBP classes in the red and NIR bands. Results indicate that the $C_1$ values in the red band are larger than in the NIR for all IGBP classes; indicating that the hotspot height is relatively prominent in red band. This is due to the higher-level chlorophyll absorption of vegetation foliage in the red band, which strengthens the sunlit-shaded component contrast, and results in a more prominent reflectance anisotropy. The $C_1$ values are significantly different between some IGBP classes; particularly in the red, and exhibits a high correlation between the red and the NIR ($R^2 = 0.62$). This reveals that a higher hotspot in the red band is frequently accompanied by a higher hotspot in the NIR band. However, the $C_2$ parameter presents a low correlation between the two bands ($R^2 = 0.21$), indicating that hotspot width is not band-dependent for these BRDF data.

Figure 3 (top right) shows the near-hotspot fit-RMSEs ($\xi < 5^\circ$) between models in 6 bands. Results indicate that the fit-RMSEs are lower in the visible than in NIR and shortwave infrared (SWIR). As expected, the RTLSR model had the largest fit-RMSEs in the vicinity of hotspot direction. We also derived globally optimized fit-RMSEs (RTCLSR_GLOB) for the entire POLDER data and the IGBP-bounded fit-RMSEs (RTCLSR_IGBP). The absolute average
difference (AAD) of the fit-RMSEs (left ordinate) between Maignan and RTCLSR_IGBP models was 0.0023 per band, while the relative average difference (RAD) of fit-RMSEs (right ordinate) between these two models was ~20-30% in the blue, red and NIR bands [490, 670, 765nm]. This reveals that the RTCLSR model with the two free parameters can provide further improved hotspot-fits. Figure 3 (bottom) presents the fit-RMSEs for 15 IGBP classes between models in the red and NIR bands. Some improvements occur for several IGBP classes using the RTCLSR. The improvements seem more pronounced for forest than for herbaceous classes in the red band, while the opposite appeared to be the case in the NIR band.

To examine the phenological response of $C_1$ and $C_2$ values for surface type, we used mid-high latitude ($23.5^\circ \leq$ latitude $\leq 60^\circ$) POLDER data in terms of the timing of Jun-Jul-Aug and Dec-Jan-Feb in Northern Hemisphere to represent maturity and dormancy seasons; using opposite time periods in the Southern Hemisphere. The $C_1$ and $C_2$ values were then retrieved for each IGBP class in the red and NIR bands. Figure 4 shows that vegetation cover tends to have larger $C_1$ and $C_2$ values in maturity, but less $C_1$ and $C_2$ values in dormancy in the red band. In the NIR band, vegetation cover mainly captures less $C_1$ but larger $C_2$ in maturity and larger $C_1$ but less $C_2$ in dormancy. Interestingly, these hotspot parameter values indicate that vegetation cover mainly captures a higher and wider hotspot in maturity, but a lower and narrower hotspot in dormancy in the red band. In the NIR band, vegetation cover mainly captures a lower but wider hotspot in maturity, and a higher and narrower hotspot in dormancy. Such hotspot behaviors should result from the leaf-on and leaf-off status in combination with spectral multiple scattering.
effects within vegetation canopy in red and NIR bands. This provides direct evidence on the spectral variability of the hotspot effect with respect to vegetation phenology.

As a comparison, Figure 4 (lower left) presents the average NDVI as a function of IGBP class in maturity and dormancy seasons. NDVI values are generally bigger in maturity than in dormancy for most IGBP classes. This indicates that the selected IGBP classes for this phenological analysis are mainly in a leaf-on and leaf-off stage, respectively. Notably, the NDVI values are very close for few classes (e.g. Shrubland, Woody Savannas and Urban), likely due to the lack of green foliage across classes. Figure 4 (lower right) presents the fit-RMSEs by the RTCLSR model for maturity and dormancy seasons in the red and NIR bands. The fit-RMSEs are chiefly higher in the NIR than in the red, but do not show significant differences between these two seasons. This comparison reveals that NDVI alone doesn’t distinctly capture this type hotspot variation as was detailed by using two hotspot parameters.

**4.1.2. Hotspot BRFs**

The scatterplots in Figure 5 (top panel) show the agreement between model-predicted and observed hotspot BRFs in the red and NIR bands for the entire POLDER database. The correlations between the two are very high ($R^2 > ~0.90$) in both the red and NIR bands. In the red band, the Maignan method slightly underestimates the hotspot BRF; particularly at a range of low reflectances (i.e., bias = -0.005 for BRFs < 0.1), but overestimates hotspot BRFs at a range of high reflectances (i.e., bias = 0.011 for BRFs ≥ 0.4). In the NIR band, the Maignan method overestimates hotspot BRFs by 0.014 units at a range of BRFs ≤ 0.3, and by 0.008 at range of
These biases are slightly corrected by using the RTCLSR model. Although the RTCLSR model with two free parameters fits the hotspot BRFs a little better than the Maignan method, the differences between them in fitting observed BRFs in the close proximity of hotspot direction (i.e., $\xi \leq 1.5^\circ$) were minor.

To further examine these two models in fitting hotspot BRFs as a function of surface type, we selected four IGBP classes to reflect a forest, mixed, grass, and sparsely vegetated gradient representing different canopy physiognomies and structures. In the red band, the Maignan method tends to underestimate the hotspot BRFs by 0.014 units for the ‘Evergreen Needleleaf Forest’ (ENF) class. The underestimation is reduced to 0.011 units for the ‘Woody Savannas’ (WSa) class, and the model fits well (underestimates by 0.003) for the ‘Grasslands’ (GrL) class; but somewhat overestimates by 0.009 units for the ‘Barren or Sparsely Vegetated’ (BSV) class. The relative average difference (RAD) between these two model reaches to 12% for ENF class. These biases are generally corrected by using the RTCLSR model (Figure 5). In the NIR band, the biases in reconstructing hotspot BRFs by Maignan method are -0.007 for ENF, 0.001 for WSa, -0.017 for GrL, and 0.018 for BSV. These biases are reduced to the range from -0.001 to -0.003 by using the RTCLSR model. This demonstrates that the RTCLSR model can provide further improvement for hotspot-fits for a subset of IGBP classes.

To qualitatively compare the Maignan and RTCLSR models, Figure 6 presents model predicted and observed BRFs for the ENF class as a function of phase angle in the proximity of the principal plane in 6 bands. We only use the observations falling within the shaded area in
Figure 1 (a) to approximate the principal plane. Data sources are marked on each subplot with a SZA=44.89° recording the hotspot direction. The NDVI values can be directly derived from the data sources. The optimized $C_1$ and $C_2$ values for the ENF class in the corresponding bands are used for each data set. This figure shows that the Maignan method and the RTCLSR model generally capture the hotspot signatures very well for these POLDER data; however, Maignan method seems to overestimate the hotspot height and underestimate hotspot width in blue band, but somewhat underestimate hotspot width in other bands for this dataset. The RTCLSR model presents more flexibility for capturing the hotspot observations. This is one of the key features of the RTCLSR model; i.e., its ability to leverage the reciprocal nature of anisotropic reflectances to improve retrieval quality.

Figure 6 near here

**4.2. Hotspot parameter values for one tile of MODIS hotspot data**

We retrieve two hotspot parameter values using one tile of MODIS hotspot data (h20v11) (Table 2). In general, $C_1$ values are larger in lower reflectance, which indicates relatively prominent hotspot effects due to lack of multiple scatterings from vegetation cover, which shows a consistent variation tendency with two hotspot parameter values for POLDER data. $C_1$ and $C_2$ values do not present significant differences between three major land cover types in this tile, dominated by savannas. However, the optimized $C_1$ and $C_2$ values present some differences in the corresponding bands between MODIS and POLDER, likely because MODIS hotspot data are derived only from one tile. From Figure 1 (d), we can also see that the BRDF shapes reconstructed using the Maignan (green curve) and RTCLSR models (black curve) in the NIR
band are very close except for the hotspot peak. The optimal $C_1=0.4$ for 2,275 MODIS data sets in this tile characterizes a low hotspot height, compared with the Maignan method for this extreme example. Note that the MODIS has difficulty in acquiring sufficient hotspot signatures on a global scale. As such, analysis of potential scale inconsistencies between MODIS and POLDER was not comprehensively performed using current hotspot data.

### 4.3. Results with in-situ and airborne data

We used airborne CAR SCAR-B cerrado and forest measurements (Tsay et al, 1998), as well as two high-quality field BRDF data sets (Iron et al., 1992; Deering et al., 1999) to evaluate the RTCLSR approach. Measurements are limited in the proximity of the principal plane, and hotspot observations are selected for $\xi \leq 1.5^\circ$ according to our sampling design (Figure 1, a). The CAR cerrado dataset was collected on a forest-grass vegetation system that is known for its distinct hotspot signature in the red and NIR bands. Dense forest covered by tall trees with a close canopy had two-flight measurements. For these data sets, we only make use of the band where the hotspot signature was the most prominent.

Using CAR SCAR-B cerrado measurements, Figure 7 presents the model predicted and observed BRFs in the red and NIR bands (not shown for the RTLSR model). We provide three specific cases to identify the potential difference between these two methods. Figure 7 (a-b) present the optimized $C_1$ and $C_2$ for the RTCLSR model for this specific data set. Figure 7 (c-d) use the $C_1 = 1.0$ and $C_2 = 3^\circ$ as default values because $K_{RTC}$ with the default $C$ values captures a very close kernel shape with $K_{RTM}$ (Figure 2). Figure 7 (e-f) present two-model consistencies by adjusting two hotspot parameters of the RTCLSR model to approach to Maignan predicted
shapes for this specific data set (observations not shown). This comparison shows that the RTCLSR model with two free hotspot parameters can approach to Maignan method to a very great degree. Using CAR SCAR-B forest measurements, we compared the difference between the model-predicted and observed BRFs in the principal plane, and over the entire viewing geometries between two models (Figure 8). The difference in reconstructing the hotspot BRFs between these two models are distinctly exhibited over the entire viewing hemisphere. We further performed the statistical analysis for these two CAR datasets for BRFs near the hotspot region of phase angle $\xi \leq 5^\circ$ and within the viewing hemisphere of VZA $\leq 75^\circ$ (Figure 9). While these two approaches do not present difference for the entire CAR datasets, the significant improvement occurs in recreating hotspot signatures using the optimized hotspot parameters of the RTCLSR model for the RMSEs, biases and correlation coefficients (Figure 9).

Finally, we compared the RTLSR, RTCLSR, and Maignan models fits to the field measurements collected from soil and vegetation surface (Figure 10). A $C_1=0.4$ and a $C_2=5.2^\circ$ are the optimal values for the soil measurements in the red, and $C_1=1.3$ ($C_2=10^\circ$) and $C_1=1.0$ ($C_2=8.0^\circ$) capture the best hotspot-fits for the black spruce (*picea mariana*) in the red and NIR bands. Results indicate a significant difference between the model-predicted hotspot BRFs for the soil and the black spruce surfaces. The RTCLSR model with a $C_1=1.0$ reconstructs the hotspot height for the black spruce as same as Maignan method; but a $C_2=8.0^\circ$ can characterize a wider hotspot for matching this measurement in the NIR band. The old back spruce captures a more prominent hotspot signature in the red than in the NIR. For the in-situ data, the fit-RMSEs
for the RTLSR, the Maignan and the RTCLSR are 0.0073, 0.0067 and 0.0058 in the red, and 0.0530, 0.0515 and 0.0506 in the NIR, respectively. Since these models are consistent in fitting these measurements except for hotspot region, the difference in overall fit-RMSEs results from their hotspot effects.

Figure 10 near here

### 4.4. Sensitivity of hotspot parameters

To examine the sensitivity of hotspot parameters to hotspot-fits, we provide the plots showing the fitting errors (color contour) as functions of $C_1$ and $C_2$ at each iteration for the entire POLDER sampled dataset in the red and the NIR bands (Figure 11, top). Results indicate that the $C_1$ values are more sensitive to hotspot-fit than $C_2$, because the fitting errors change more rapidly along $C_1$ than $C_2$. For a given $C_2$, fitting errors present certain symmetry about $C_1$ because large or small $C_1$ values generate the comparable magnitude of fitting errors. The minimum RMSE occurs at the optimal $C_1$ and $C_2$ values (the white point). Similar results were found across IGBP classes.

Since the BRDF sampling can have an effect on the sensitivity of model parameter retrievals for kernel-driven models (Lucht et al., 2000), we also examined the sensitivity of the two hotspot parameters to the model parameters. To do this, we evaluated the modeled hotspot reflectance as a function of $C_1$ (taking $C_2 = 3^\circ$ as the default value) for three sample sizes (i.e., 12, 60 and 161 measurements) in principal plane. These BRDF observations are sampled from the airborne CAR cerrado measurements, which contained 29,160 BRF samples. In each case, 6 uniformly-distributed observations in the vicinity of the hotspot region (i.e., $\xi \leq 5^\circ$) were used,
and the rest of the observations outside of this hotspot region were randomly sampled.

Figure 11 (bottom) presents the sensitivity of modeled hotspot reflectance as a function of $C_1$ in the red band ($C_2 = 3^\circ$). This figure demonstrates that the BRDF sample sizes have a certain effect on the sensitivity of the $C_1$ values to the modeled hotspot BRFs, since the hotspot BRFs as a function of $C_1$ values presents varying slopes for these three cases of BRDF sample sizes. This leads to slightly varying $C_1$ values, even when using identical hotspot observations for different total BRDF sample numbers. However, the variation range of the modeled BRFs as a function of $C_1$ can effectively cover the observed hotspot peak (i.e. $\text{HS}_{\text{BRF}} = 0.23$). Result also indicate that the modeled hotspot BRFs, using the Maignan method, are sensitive to the total BRDF sample numbers in this examination; but with a slight overestimation of the hotspot BRFs. This is consistent with result shown for the NIR band when using MODIS, POLDER and CAR.

This investigation demonstrates that the sensitivity of hotspot parameters is somewhat related to varying BRDF sample numbers, which in turn tend to have an effect on model parameter retrievals. This effect holds true even when using the same observations in the vicinity of the hotspot direction (i.e., $\xi \leq 5^\circ$). This implies that the model predicted hotspot BRFs using a set of optimized prior $C_1$ and $C_2$ values in RTCLSR model (combining $K_{\text{LSR}}$ and $K_{\text{RTC}}$) would be sensitive to the three model parameters (i.e., $f_{\text{iso}}$, $f_{\text{geo}}$, and $f_{\text{vol}}$) to a certain extent that the adjustment of hotspot BRF dynamics by three model parameters is no longer effective. In such a situation, the two free hotspot parameters in RTCLSR model provide an improved capacity to capture more accurate hotspot signatures.

Figure 10 near here
4.5. Influence on retrieval of intrinsic albedo and clumping index

To evaluate the performance of the RTCLSR method for potential applications, we evaluated the retrieval of intrinsic albedos and clumping index (CI) from available POLDER and MODIS data. The BSA values were calculated for SZAs ranging from $0^\circ$ to $75^\circ$ at the interval of $15^\circ$. In general, the intrinsic albedo retrievals between the RCLSR and RTCLSR models are in a very high correlation for both POLDER ($R^2 = \sim 1.0$) and MODIS ($R^2 > \sim 0.9$) data; with negligible biases ($< \sim 0.0003$ in red and $< \sim 0.0006$ in NIR), while the difference in CI values between models was rather significant (Table 1 and Figure 12). As compared with the RTCLSR model results, a major overestimation of CI values occurs with the RTLRSR model, but a slight underestimation of CI values occurs with the Maignan method. This is attributed to the difference in reconstructing hotspot BRFs between models. Figure 12 exhibits the scatter plots between models for ~11,632 CI values for 15 IGBP classes in the NIR (Figure 12, a-b) and the red bands (Figure 12, c-d) using POLDER data. According to the CI inversion algorithm, the CI biases between models (despite a high correlation, $R^2 = \sim 0.9$) result from biases in modeling hotspot BRFs. For most IGBP classes from POLDER data, about 20% of the relative average difference (RAD) in retrieving CIs occurs between the RTLRSR model and the RTCLSR model, while about 10% of RAD occurs between the Maignan method and the RTCLSR in the NIR (Table 1). In the red band (Figure 12, c-d), the RAD in CI retrievals between the Maignan method and the RTCLSR model is generally small, while it still remains about 20% between the RTLRSR and the RTCLSR. For the MODIS data, the RAD in CI retrievals arrive at about 6% and 12% in the red and the NIR bands between the RTCLSR and the Maignan (Figure 12, e-f).
Notably, we use more measurements in CI estimates because we can, in theory, reconstruct the hotspot BRFs using two prior IGBP-based hotspot parameter values that are derived from the high-quality observations in Section 4.1.1. In addition, we merely compare the retrieved CI values between models following the objective of this paper. To extensively validate these CI estimates with ground measurements would require the expanded use of additional measurements beyond the ones used in this study.

5. Conclusions and discussion

In this study, we proposed a new method to correct the hotspot effects for the RTLSR model known for its use within the operational MODIS BRDF/Albedo product. The method makes use of the exponential function with two free parameters ($C_1$ and $C_2$) to characterize hotspot height and width. Our results indicate that, although the Maignan method with no free hotspot parameters characterizes the hotspot effect in a relatively high accuracy (particularly for POLDER), the RTCLSR model can provide a further improvement in hotspot-fits. This is mainly attributed to the two free hotspot parameters that can be adjusted to reach their optimal values for the near-hotspot measurements available. Further analysis of the two hotspot parameters using hotspot data reveals that the hotspot signatures are somewhat related to surface type and vegetation phenology for available POLDER data. We found that the hotspot height ($C_1$) value is bigger in the red band than in the NIR band for most vegetation types, indicating a more prominent hotspot in the red band possibly due to the strong chlorophyll absorption of vegetation foliage.

The RTCLSR model is quite consistent with the other two models in the intrinsic albedos
retrievals, but is somewhat different in CI retrievals through the use of the inversion algorithm of
Chen et al. (2005). This reveals that the albedo retrievals using the RTCLSR would be consistent
with the archived albedo product using the routine RTLSR algorithm; however, the RTCLSR
model would provide more accurate hotspot signatures that may lead to an improved
understanding of vegetation biophysical parameter retrievals in relation to hotspot BRFs. A
similar analysis of the RTCLSR with the MODIS hotspot data, airborne CAR data and field
measurements reveals the broad adaptability of this new method for different spatial resolutions;
however, further investigation into the scale issue of hotspot effect using the RTCLSR model is
still a challenge mainly due to the lack of sufficient hotspot data at different spatial resolutions.
Investigation into the sensitivity of these two hotspot parameters shows that \( C_1 \) is more sensitive
to hotspot-fits than \( C_2 \). The total BRDF sampling design can also play a role on the sensitivity of
the hotspot parameters to hotspot-fits, because the total sample numbers (even having a good
distribution) can have an effect on the model parameter retrievals, in particular the sparse BRDF
sampling cases typified by the MODIS retrieval.

The original design of the kernel-driven linear Ross-Li BRDF model comprised a collection
of kernels for different scenarios of land surface types. Thus, kernel functions are derived from
different assumptions for vegetation canopy structures, and view and illumination geometries.
These assumptions may result in differences in modeling the radiation field (Wanner et al., 1995)
especially in the hotspot direction (Huang & Jiao et al., 2012). Methodologically, the exponential
approximation of this hotspot kernel function may also be used to correct the other volumetric
scattering kernel (i.e., RossThin), because this exponential hotspot function makes use of two
free parameters ($C_1$ and $C_2$) to fit hotspot signatures, rather than use surface biophysical parameters as inputs to drive a theoretical hotspot model. The exponential function form is rooted in gap probability theory and has been one of major contributions to hotspot modeling (Qin et al., 1995). The corrected volumetric kernel can be combined with various geometric optical kernels for potential applications. However, the use of multi-kernel combination models would require recalibration of the $C_1$ and $C_2$ parameters.

In the situation where hotspot signatures are not available, the $C_1$ and $C_2$ values would need to be estimated on an *a priori* basis for it to initialize an RTCLSR model inversion. This has been attempted by fitting the hotspot data of several spatial resolutions via two means. Globally optimized $C_1$ and $C_2$ values can characterize the overall accuracy of the hotspot effect for certain satellite hotspot data, while variable-related optimized $C_1$ and $C_2$ values can help explore hotspot variation as a function of some underlying variables such as surface type and vegetation phenology. We found that a prior $C_1 = 1$ and $C_2 = 3^\circ$ provide a stable initialized value in the RTCLSR model that is comparable in performance to the Maignan method.

Because the hotspot effect is very sensitive to the phase angle between the view and illumination in the retro-solar direction (Bréon et al. 2002), its analysis for remote sensing applications requires a very high geometric accuracy. This implies that it is especially difficult to capture accurate hotspot observations in field measurements due to the shadows of the goniometer and sensors that must have a small enough IFOV (e.g., $\leq 1^\circ$). Use of the field measurements with an IFOV= $15^\circ$ in this study aims to stress that the reciprocal nature allows the RTCLSR model to acquire hotspot-fits with the least RMSEs for this data set. However, with
two free hotspot parameters determined \textit{a priori} by using enough hotspot data, this new method provides an improved understanding of the hotspot effect, and thus has potentials for certain ecological applications in regard to the hotspot BRFs for complex heterogeneous environments.
6. Appendix A

The shaded area in Figure 1 (a) contains four parts. As the 1st constraint, all observations falling within the phase angle having the radius of R should be included in set A. Here, R = 1.5° is used, but R can be adjusted properly for different hotspot data.

\[ A = \{ \xi | \xi \leq R \} \]

To select the observed BRFs close to principal plane (PP) for visually comparing with modeled BRFs in PP, we define the distance, H, which is perpendicular to the principal plane, as the 2nd constraint. Obviously, a less H value can select observations in the shaded area to get closer to PP. This generates set B as follows. Here, \( H = 10^\circ \) is attempted for MODIS and POLDER data, but \( H = 1^\circ \) is attempted for airborne CAR and field data.

\[ B = \{ \phi \| \theta, \sin \varphi \leq H \} \]

As the 3rd constraint, we should consider that, in a 2-D plot that exhibits the observed BRFs in approximate PP, spurious visualizations, particularly in the proximity of hotspot direction in PP, possibly result from some observed BRFs that are not actually near hotspot region in PP, e.g., a \( H = 10^\circ \) without other constraints will allow the observed BRFs in the unshaded sector domain taking Sun as the center (Figure 1, a) to be exhibited as near-hotspot BRFs in PP. To remove these observations in 2-D plots that present observed and modeled BRFs in approximately PP (e.g., Figure 1, d), we define an intersection angle (\( \theta \)) between PP line and the line passing through Sun (Figure 1, a), and derive set C and set D as follows. Here, \( \theta = 45^\circ \) is attempted for these hotspot data used in this study.
Finally, the shaded area in Figure 1 (a) can be derived by implementing set operation for four sets above:

\[ A \cup (B \cap (C \cup D)) \]
7. Acknowledgments

This work was supported by the NSFC (41171261) and the National Basic Research Program (973 Program, 2013CB733400). The POLDER-3/PARASOL BRDFs databases are elaborated by the LSCE, and provided by the POSTEL Service Centre. Many thanks are due to Dr. Francois-Marie Bréon (LSCE, France) for his kindly providing a piece of code to reprocess the POLDER data, and for his great comments and suggestions on this paper.
### Table 1. Statistics of the clumping index values retrieved by three models

<table>
<thead>
<tr>
<th>Class</th>
<th>Number</th>
<th>RTLSR</th>
<th>Maignan</th>
<th>RTCLSR</th>
<th>RTLSR - RTCLSR</th>
<th>Maignan - RTCLSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
<td>Std.</td>
<td>Mean - RTCLSR</td>
</tr>
<tr>
<td>ENF</td>
<td>793</td>
<td>0.57</td>
<td>0.04</td>
<td>0.49</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>EBF</td>
<td>898</td>
<td>0.76</td>
<td>0.08</td>
<td>0.56</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>DNF</td>
<td>219</td>
<td>0.66</td>
<td>0.02</td>
<td>0.57</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>DBF</td>
<td>682</td>
<td>0.80</td>
<td>0.06</td>
<td>0.63</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>MiF</td>
<td>690</td>
<td>0.76</td>
<td>0.10</td>
<td>0.57</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>CSh</td>
<td>549</td>
<td>0.77</td>
<td>0.08</td>
<td>0.58</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>OSh</td>
<td>1247</td>
<td>0.86</td>
<td>0.10</td>
<td>0.68</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Wsa</td>
<td>1035</td>
<td>0.83</td>
<td>0.06</td>
<td>0.62</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Sav</td>
<td>734</td>
<td>0.87</td>
<td>0.06</td>
<td>0.67</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>GrL</td>
<td>1075</td>
<td>0.88</td>
<td>0.09</td>
<td>0.68</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Pwe</td>
<td>21</td>
<td>0.95</td>
<td>0.06</td>
<td>0.65</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>CrL</td>
<td>960</td>
<td>0.89</td>
<td>0.07</td>
<td>0.67</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Ubu</td>
<td>835</td>
<td>0.84</td>
<td>0.06</td>
<td>0.65</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>CNVM</td>
<td>727</td>
<td>0.89</td>
<td>0.06</td>
<td>0.66</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>BSV</td>
<td>1167</td>
<td>0.99</td>
<td>0.16</td>
<td>0.85</td>
<td>0.19</td>
<td>0.17</td>
</tr>
</tbody>
</table>

### Table 2. The globally optimized $C_1$ and $C_2$ values derived from one 10° × 10° tile (h20v11) of MODIS surface reflectance products (MOD09 and MYD09) in 7 reflected solar bands

<table>
<thead>
<tr>
<th>Band (nm)</th>
<th>Band1 (620-670)</th>
<th>Band2 (841-876)</th>
<th>Band3 (459-479)</th>
<th>Band4 (545-565)</th>
<th>Band5 (1230-1250)</th>
<th>Band6 (1628-1652)</th>
<th>Band7 (2105-2155)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>0.7</td>
<td>0.4</td>
<td>0.7</td>
<td>0.7</td>
<td>0.6</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>$C_2 (^\circ)$</td>
<td>5.2</td>
<td>4.5</td>
<td>5.2</td>
<td>5.2</td>
<td>3.5</td>
<td>5.2</td>
<td>5.2</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1. (a) The method to select observations near principal plane and in the proximity of hotspot direction to be showed in 2-D plots, its formulation is detailed to appendix A; (b) spatial distribution of view and sun geometries for a typical POLDER data set, i.e., IGBP_01_200609brdf_ndvi08_0824_1671, and (c) for a typical MODIS data set in h20v11 for savanna; (d) the observed and modeled BRFs using Maignan method, and RTCLSR model with $C_1=0.4$, and $C_2=4.5^\circ$ as a function of phase angle for this MODIS data in the NIR. A minus sign is assigned to the phase angle when $\theta_s \cos \varphi < \theta_s$.

Figure 2. LiSparseReciprocal kernel ($K_{LSR}$) at SZA of 15°, 30°, 45° and 60° (top) and three volumetric kernels in PP (bottom). In the top subplot, the red and magenta dashed curves around the red solid curve are from different $h/b$ and $b/r$ ratios in $K_{LSR}$. The red solid curve is for the operational $K_{LSR}$ that adopts $h/b=2$ and $b/r=1$; red dashed curve is for $h/b=2$ and $b/r=1.2$; magenta dashed curve is for $h/b=2.5$ and $b/r=1$. In the bottom subplot, three volumetric kernels are for the operational RossThick kernel ($K_{RT}$, black curve), the Maignan kernel ($K_{RTM}$ with green solid curve for $\xi_0=1.5^\circ$ and green dashed curve for $\xi_0=3.0^\circ$), and RossThickChen kernel ($K_{RTC}$) with $C_1=1$ and $C_2=3^\circ$ (red dashed curve). The upwardly-shifted $K_{RTC}$ with $C_1=1$ and $C_2=3^\circ$ (red solid curve), $C_1=1$ and $C_2=5^\circ$ (blue dashed curve), and $C_1=0.6$ and $C_2=3^\circ$ (magenta dashed curve) for a SZA=30° is aligned with the $K_{RT}$ and $K_{RTM}$ for a convenient comparison.

Figure 3. The globally optimized $C_1$ and $C_2$ values, and the fit-RMSEs derived from the entire POLDER BRDF data in 6 bands (top); the $C_1$ and $C_2$ values (middle) and the corresponding fit-RMSEs (bottom) for underlying IGBP class. The dashed lines (top right) present the relative fit-RMSEs (right ordinate) between Maignan method and RTCLSR model. 16 IGBP classes are Evergreen Needleleaf forest (ENF),
Evergreen Broadleaf Forest (EBF), Deciduous Needleleaf Forest (DBF), Deciduous Broadleaf Forest (DBF), Mixed Forest (MiF), Closed Shrublands (CSh), Open Shrublands (OSh), Woody Savannas (WSa), Savannas (Sav), Grasslands (GrL), Permanent Wetlands (PWe), Croplands (CrL), Urban and Build-up (UBu), Cropland/Natural Vegetation Mosaic (CNVM), Barren or Sparsely Vegetated (BSV). Snow and Ice (SI) is excluded from this study because of its strong forward scattering.

**Figure 4.** Two hotspot parameters (top) as a function of IGBP class in terms of the timing of Jun-Jul-Aug/Dec-Jan-Feb in the northern hemisphere to represent maturity and dormancy season, respectively (opposite in southern hemisphere) in the red and the NIR bands. The corresponding average NDVI and fit-RMSEs (bottom) are presented as a comparison.

**Figure 5.** Comparison of the model predicted and observed hotspot BRFs for the entire POLDER BRDF data (upper panel), and two IGBP classes for ENF and WSa (lower panel) in the red band.

**Figure 6.** POLDER observations (red points) and the reconstructed BRDF shapes by Maignan (green) and RTCLSR model (black) in 6 bands for the ENF as a function of phase angle in terms of the sampling design (i.e., Figure 1, a)

**Figure 7.** Comparison of Maignan with RTCLSR model using CAR/SCAR-B cerrado measurements in principle plane in the red and the NIR bands. We elaborate on the differences between these two models in terms of three cases: (a) and (b) show RTCLSR with the optimal $C_1$ and $C_2$ in case 1; (c) and (d) show RTCLSR using $C_1=1$ and $C_2=3^\circ$ as default values in case 2; (e) and (f) adjust RTCLSR model to approach to Maignan result for deriving the optimal $C_1$ and $C_2$ values in case 3.

**Figure 8.** Using CAR SCAR-B forests to examine the difference between model predicted and observed BRFs in PP (top) and the difference between RTCLSR and Maignan over the entire viewing hemisphere.
(bottom) using Flight # 1689 in the red (c) and using Flight # 1693 in the NIR (d).

**Figure 9.** Scatterplots showing the difference between modeled and observed BRFs near the hotspot region ($\xi \leq 5^\circ$, red points) and for the viewing hemisphere ($VZA \leq 75^\circ$, black points), using CAR Flight # 1689 in the red band (a, b) and CAR Flight # 1693 in the NIR band (c, d).

**Figure 10.** Comparison of field soil measurements (top) and field old black spruce measurements (bottom) with three models.

**Figure 11.** The fitting errors as functions of $C_1$ and $C_2$ for the entire POLDER data in the red (left) and in the NIR (right), and the white points on the contour plots represent the least fit-RMSEs with the optimal $C_1$ and $C_2$ values; the modeled hotspot reflectance as a function of $C_1$ (given $C_2 = 3^\circ$) using CAR data to simulate three BRDF sample sizes (12, 60 and 161 samples) in the red band (bottom).

**Figure 12.** Comparison of CI retrievals between models using POLDER data in the NIR band (top) and in the red band (middle), as well as using MODIS data (bottom) in the red and the in NIR bands. The dashed and solid lines represent the one-to-one lines and the fitted lines, respectively.
Figure 1. (a) the method to select observations near principal plane and in the proximity of hotspot
direction to be showed in 2-D plots, its formulation is detailed to appendix A; (b) spatial distribution of
view and sun geometries for a typical POLDER data set, i.e., IGBP_01_200609brdf_ndvi08. 0824_1671;
(c) for a typical MODIS data set in h20v11 for savanna; (d) the observed and modeled BRFs using
Maignan method, and RTCLSR model with $C_1=0.4$, and $C_2=4.5^\circ$ as a function of phase angle for this
MODIS data in the NIR. A minus sign is assigned to the phase angle when $\Theta_s \cos \varphi < \Theta_i$. 

Figure 2. LiSparseReciprocal kernel ($K_{LSR}$) at SZA of $15^\circ$, $30^\circ$, $45^\circ$ and $60^\circ$ (top) and three volumetric kernels in PP (bottom). Details are expounded in figure captions.
Figure 3. The globally optimized $C_1$ and $C_2$ values, and the fit-RMSEs derived from the entire POLDER BRDF data in 6 bands (top); the $C_1$ and $C_2$ values (middle) and the corresponding fit-RMSEs (bottom) for underlying IGBP class. The dashed lines (top right) present the relative fit-RMSEs (right ordinate) between Maignan method and RTCLSR model.
Figure 4. Two hotspot parameters (top) as a function of IGBP class in terms of the timing of Jun-Jul-Aug/Dec-Jan-Feb in the northern hemisphere to represent maturity and dormancy season, respectively (opposite in southern hemisphere) in the red and the NIR bands. The corresponding average NDVI and fit-RMSEs (bottom) are presented as a comparison.
Figure 5. Comparison of the model predicted and observed BRFs for the entire POLDER BRDF data and two IGBP classes, i.e., ENF and WSa in the red band.
Figure 6. POLDER observations (red points) and the reconstructed BRDF shapes by Maignan (green) and RTCLSR model (black) in 6 bands for the ENF class as a function of phase angle.
Figure 7. Comparison of Maignan with RTCLSR model using CAR/SCAR-B cerrado measurements in principle plane in the red and the NIR bands. Details were expounded in figure captions.
Figure 8. Using CAR SCAR-B forests to examine the difference between model predicted and observed BRFs in PP (top) and the difference between RTCLSR and Maignan over the entire viewing hemisphere (bottom) using Flight # 1689 in the red (c) and using Flight # 1693 in the NIR (d).
Figure 9. Scatterplots showing the difference between modeled and observed BRFs near the hotspot region (\(\xi \leq 5^\circ\), red points) and for the viewing hemisphere (VZA \(\leq 75^\circ\), black points), using CAR Flight #1689 in the red band (a, b) and CAR Flight #1693 in the NIR band (c, d)
Figure 10. Comparison of field soil measurements (top) and field old black spruce measurements (bottom) with three models.
Figure 11. The fitting errors as functions of $C_1$ and $C_2$ for the entire POLDER data in the red (left) and in the NIR (right), and the white points on the contour plots represent the least fit-RMSEs with the optimal $C_1$ and $C_2$ values; the modeled hotspot reflectance as a function of $C_1$ (given $C_2 = 3^\circ$) using CAR data to simulate three BRDF sample sizes (12, 60 and 161 samples) in the red band (bottom).
Figure 12. Comparison of CI retrievals between models using POLDER data in the NIR band (top) and in the red band (middle), as well as using MODIS data (bottom) in the red and the in NIR bands. The dashed and solid lines represent the one-to-one lines and the fitted lines, respectively.
References


53


Zhu, G., Ju, W., Chen, J. M., Gong, P., Xing, B. and Zhu, J,(2012). Foliage Clumping Index Over China’s