MODIS Leaf Area Index (LAI) And Fraction Of Photosynthetically Active Radiation Absorbed By Vegetation (FPAR) Product

(MOD15)

Algorithm Theoretical Basis Document

Version 4.0

R. B. Myneni, Y. Knyazikhin, Y. Zhang, Y. Tian, Y. Wang, A. Lotsch

Department of Geography
Boston University
Boston, MA 02215
rmyneni@crsa.bu.edu

J. L. Privette, J. T. Morisette

NASA's Goddard Space Flight Center
jeff.privette@gsfc.nasa.gov

S. W. Running, R. Nemani, J.Glassy, P.Votava

School of Forestry
University of Montana
Missoula, MT 59812
swr@ntsg.umt.edu

Cite as:

Y. Knyazikhin, J. Glassy, J. L. Privette, Y. Tian, A. Lotsch, Y. Zhang, Y. Wang, J. T. Morisette, P.Votava, R.B. Myneni, R. R. Nemani, S. W. Running, MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document, *http://eospso.gsfc.nasa.gov/atbd/modistables.html*, 1999.

This document was prepared by Y. Zhang

April 30, 1999

Abstract

This Algorithm Theoretical Basis Document (ATBD) describes the algorithm to produce global Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) absorbed by vegetation from atmospherically corrected surface reflectances. The MOD15 LAI and FPAR products are 1 km at launch products provided on a daily and 8 days basis. The algorithm consists of a main procedure that exploits the spectral information content of MODIS surface reflectances at up to 7 spectral bands. Should this main algorithm fail, a back-up algorithm is triggered to estimate LAI and FPAR using vegetation indices. The algorithm requires a land cover classification that is compatible with the radiative transfer model used in their derivation. Such a classification based on vegetation structure was proposed and it is expected to be derived from the MODIS Land Cover Product. Therefore the algorithm has interfaces with the MODIS surface reflectance product (MOD09) and the MODIS Land Cover Product (MOD12).

Derivation of various empirical relationships and trends from the LAI and FPAR fields is the most likely approach which a potential user of the LAI and FPAR products will utilize in his investigation. It sets a demand on the retrieval techniques; that is, the LAI and FPAR fields must posses the same statistical properties as if they were derived from ground based measurements. Therefore we perform our retrievals by comparing observed radiances with modeled radiances for a suite of canopy structure and soil patterns that covers a range of expected natural conditions. The set of the canopy/soil patterns for which the magnitude of the residuals in the comparisons does not exceed uncertainties in observed radiances is then used to evaluate the distribution of LAI and FPAR values and to specify the most probable value of desired parameters. The key mathematics behind this technique is the measure theory which is used to establish relationships between the surface reflectances, uncertainties in their retrieval and canopy/soil patterns. This theory underlies a precise mathematical definition of the probability distribution function and thus allows us to meet the above formulated demand.

In order to better describe natural variability of vegetation canopies a three-dimensional formulation of the LAI/FPAR inverse problem underlies the algorithm. By accounting features specific to the problem of radiative transfer in plant canopies, powerful techniques in reactor theory (the Green's function and adjoint formulation of the problem) were utilized to parameterize the radiative field in terms of reflectance properties of the vacuum bounded canopy and ground, as well as to split the three-dimensional radiative transfer problem into two independent sub-problems, each of which is expressed in terms of three basic components of the energy conservation law: canopy transmittance, reflectance, and absorptance. These components are elements of the look-up table (LUT), and the algorithm interacts only with the elements of the LUT. This provides the independence of the retrieval algorithm to a particular canopy radiation model. It is

precisely derived that the dependence of canopy transmittance, reflectance, and absorptance on wavelength is described by a simple function which depends on the unique positive eigenvalue of the transport equation. This eigenvalue relates optical properties of individual leaves to canopy structure. This facilitates comparison of spectral values of the canopy reflectances with spectral properties of individual leaves, which is a rather stable characteristic of a green leaf. This allows us to fully take advantage of the spectral information content of the MODIS instrument.

The algorithm has been prototyped by Land Surface Reflectances (LASUR) and Landsat data. The prototyping results demonstrated its ability to produce global LAI and FPAR fields. The algorithm and the LUT use directly the information on the leaf canopy spectral properties and structural attributes, in stead of NDVI, to retrieve LAI and FPAR. The effects of biome misclassification between clearly distinct biomes on the algorithm can be evaluated through the Retrieval Index (RI), mean LAI and the histogram of the retrieved LAI distribution. The dependence of the algorithm on spatial resolution is illustrated using coarse and fine resolution data and LUTs.

Validation of the LAI and FPAR products is an important part of algorithm development and is in progress under the EOS validation plan. As the global validation of land remote sensing products is complicated by multiple factors, various validation techniques will be used to develop uncertainty information on EOS land products. Detailed validation methods have been proposed for global scale validation.

The important ancillary data set for this algorithm are the radiative transfer model compatible structural land cover classification, which divides the global vegetation land cover into six biomes. Efforts have been made to improve this land cover classification map from many data sources. Another important ancillary data set is the LUT of the algorithm, which is derived from both theory and data set of vegetation, soil optical properties. Detailed description of the LUT's structure and its elements is also given in this document.

The MODIS LAI and FPAR Level 3 algorithms were developed jointly by personnel at Boston University and the University of Montana SCF and NASA GSFC. The Boston University team developed the radiative-transfer (R-T) derivative science core logic and the R-T driven lookup tables comprising the core science, the direct-retrieval lookup tables, and prototype software for exercising the core logic. The University of Montana SCF team is responsible for developing, testing, and maintaining the EOSDIS Core System (ECS) production version of the software. QA tasks are shared between the two institutions, with the MOD15A1 QA activities conducted at Boston University, and the MOD15A2 QA activities run at University of Montana SCF. Validation is done by both BU and UMT SCFs in collaboration with Drs. Privette and Morisette at GSFC.

Table of Contents

MODIS FPAR AND LAI PRODUCTS ALGORITHM TECHNICAL BASIS DOCUMENT Version 4.1

1. INTRODUCTION

1.1 Identification

Table 1-1: **Product List**

MODIS Product No. 15 (MOD15)			
Parameter Number	Parameter Name	Spatial Resolution	Temporal Resolution
2680	Leaf Area Index (LAI)	1 km	Daily, 8 day
5367	Fraction of Photosynthetically Active Radiation absorbed by vegetation (FPAR)	1 km	Daily, 8 day

1.2 Overview

This document details the algorithm for producing global terrestrial leaf area index (LAI) product, and the related fraction of absorbed photosynthetically active radiation (FPAR) product. LAI defines an important structural property of a plant canopy, the number of equivalent layers of leaves vegetation displays relative to a unit ground area. FPAR measures the proportion of available radiation in the specific photosynthetically active wavelengths of the spectrum 0.4 - 0.7 μm that a canopy absorbs. It is non-linearly related to the LAI.

Both LAI and FPAR will be Level 4 MODIS products derived directly from MODIS Reflectances (MR) and ancillary data on surface characteristics such as Land cover type, background etc. These products will be produced globally at a time frequency defined by the MODIS Reflectances (MR) global compositing period, 8 days. The spatial resolution will be constrained by the MODIS reflectance dataset, and may be as fine as 250m, or standardized to 1km. The 8 day product will be produced by compositing using maximum FPAR.

2. ALGORITHM DESCRIPTION

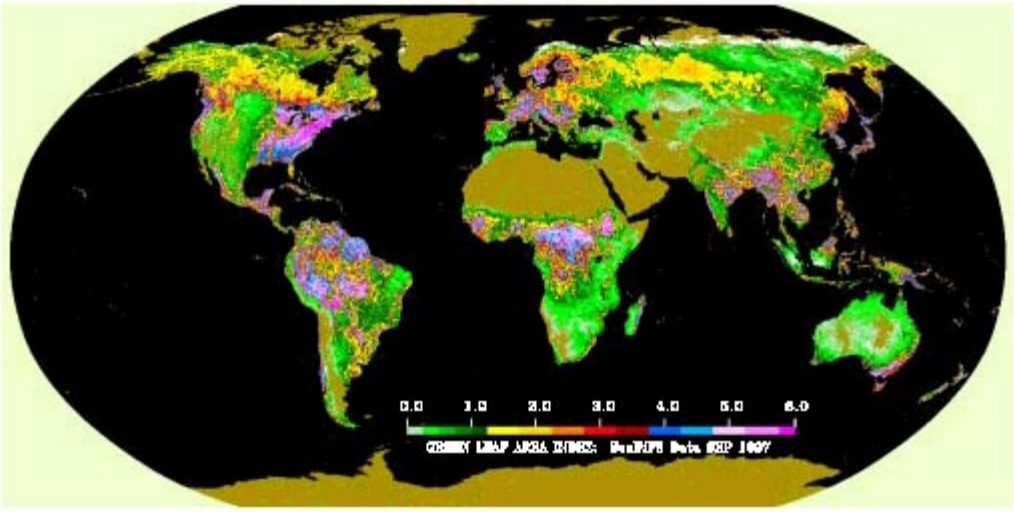
2.1 Introduction

Large-scale ecosystem modeling is used to simulate a range of ecological responses to changes in climate and chemical composition of the atmosphere, including changes in the distribution of terrestrial plant communities across the globe in response to climate changes. Leaf area index (LAI) is a state parameter in all models describing the exchange of fluxes of energy, mass (e.g., water and CO_2), and momentum between the surface and the planetary boundary layer. Analyses of global carbon budget indicate a large terrestrial middle- to high-latitude sink, without which the accumulation of carbon in the atmosphere would be higher than the present rate. The problem of accurately evaluating the exchange of carbon between the atmosphere and the terrestrial vegetation therefore requires special attention. In this context the fraction of photosynthetically active radiation (FPAR) absorbed by global vegetation is a key state variable in most ecosystem productivity models and in global models of climate, hydrology, biogeochemestry, and ecology [*Sellers et al.*, 1997]. Therefore these variables that describe vegetation canopy structure and its energy absorption capacity are required by many of the EOS Interdisciplinary Projects [*Myneni et al.*, 1997a]. In order to quantitatively and accurately model global dynamics of these processes, differentiate short-term from long-term trends, as well as to distinguish regional from global phenomena, these two parameters must be collected often for a long period of time and should represent every region of the Earth's lands. Satellite remote sensing serves as the most effective means for collecting global data on a regularly basis. The launch of EOS-AM 1 with MODIS (moderate resolution imaging spectroradiometer) and MISR (multiangle imaging spectroradiometer) instruments onboard begins a new era in remote sensing the Earth system. In contrast to previous single-angle and single-channel instruments, MODIS and MISR together allow for rich spectral and angular sampling of the radiation field reflected by vegetation canopies. This sets new demands on the retrieval techniques for geophysical parameters in order to take full advantages of these instruments. Our objective is to derive a synergistic algorithm for the extraction of LAI and FPAR from MODIS- and MISR-measured canopy reflectance data, with the flexibility to use the same algorithm in MODIS-only and MISR-only as well. Although a prototyping of the algorithm with data was also a focus of our activity, these results are not discussed in this article. Plate 2-1 demonstrates an example of the prototype of the MODIS LAI/FPAR data product.

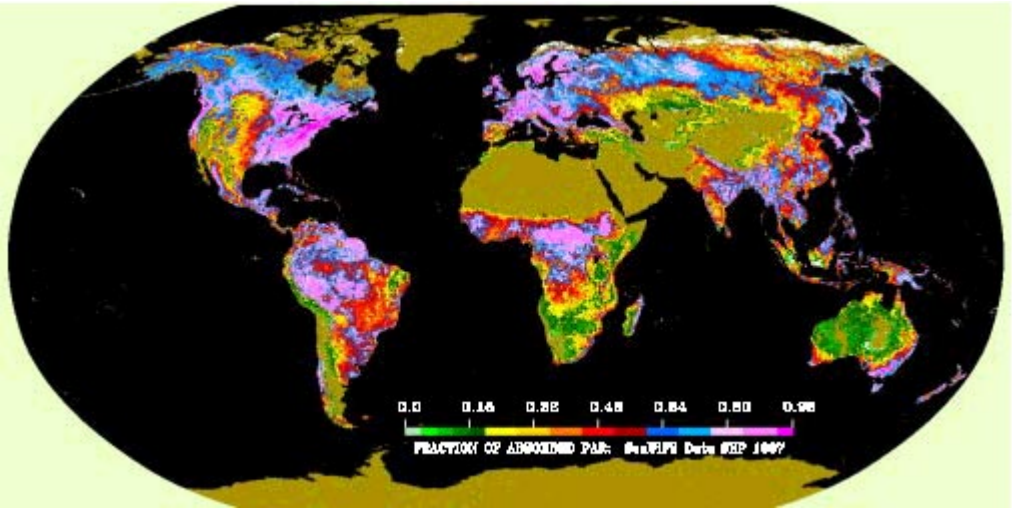
Solar radiation scattered from a vegetation canopy and measured by satellite sensors results from interaction of photons traversing through the foliage medium, bounded at the bottom by a radiatively participating surface. Therefore to estimate the canopy radiation regime, three important features must be carefully formulated. They are (1) the architecture of individual plant and the entire canopy; (2) optical properties of vegetation elements (leaves, stems) and soil; the former depends on physiological conditions (water status, pigment concentration); and (3) atmospheric conditions which determine the incident radiation field. Photon transport theory aims at deriving the solar radiation regime, both within the vegetation canopy and the radiant

exitance, using the above mentioned attributes as input data. This theory underlies numerous canopy radiation models (see, for example, reviews by *Myneni et al*. [1989] and *Ross et al*. [1992]). Usually retrieval techniques rely on a model, which provide relationships between measured data and biophysical parameters. It allows for the design of fast retrieval algorithms. However, such algorithms can retrieve only those variables that are explicitly represented in the canopy radiation models. They exclude the use of a rather wide family of three-dimensional models in which desired variables may not be in the model parameter list directly [*Ross and Marshak*, 1984; *Myneni*, 1991; *Borel et al*., 1991; *Kimes*, 1991; *Knyazikhin et al*., 1996]. They are also based on some assumptions which may not be fulfilled. For example, numerous canopy radiation models presuppose that the canopy angular reflectance measurements can be performed about the plane of the solar vertical which provides information on the hot spot effect [*Kuusk*, 1985; *Simmer and Gerstl*, 1985; *Marshak*, 1989; *Verstraete et al*., 1990; *Myneni et al*., 1991]. This suggestion may be appropriate for multiangle instruments such as MISR or POLDER (Polarization and Directionality of the Earth's Reflectance) [*Deschamps et al*., 1994]. For the single-angle and multi-channel MODIS instrument, this suggestion is not fulfilled. There is yet another problem encountered when one incorporates a particular model in the inverse mode. A rather wide family of canopy radiation models designed to account for the hot spot effect conflict with the law of energy conservation (Appendix); that is, they are not "physically based" models.

In designing the synergistic algorithm, we cast aside the idea of trying to relate a retrieval technique with a particular canopy radiation model. Our approach incorporates the following tenets: (1) a retrieval algorithm can use any field-tested canopy radiation model; that is, the retrieval algorithm is model independent; (2) the more measured information is available and the more accurate this information is, the more reliable and accurate the algorithm output would be, i.e., convergence of the algorithm; (3) the algorithm must be as simple as the one linked to a particular canopy radiation model; (4) spectral and angular information are synergistically used in the extraction of LAI and FPAR. Because three-dimensional models include all diversity of one- and two-dimensional models as special cases, property (1) of the algorithm can be achieved, if one formulates the inverse problem for three-dimensional vegetation canopies: given mean spectral, and in the case of MISR data, angular signatures of canopy-leaving radiance averaged over the three-dimensional canopy radiation field, find LAI and FPAR. It is clear that the given information is not enough to solve the inverse problem. For example, the three-dimensional canopy structure can vary considerably with LAI essentially unchanged. Therefore one needs to limit the range of variation of the variables determining the three-dimensional radiative regime in plant canopies. It can be achieved by using a vegetation cover classification parameterized in terms of variables used by photon transport theory [*Myneni et al.*, 1997]. It distinguishes six biome types, each representing a pattern of the architecture of an individual tree (leaf normal orientation, stem-trunk-branch area fractions, leaf and crown size) and the entire canopy (trunk distribution, topography), as well as patterns of spectral reflectance and transmittance of vegetation elements. The soil and/or understory type are also characteristics of the biome, which can vary continuously within given biome-dependent ranges. The distribution of leaves is described by the leaf area density distribution function which also depends on some continuous parameters. A detailed description of biome types is presented in the section 2.2.



a)



b)

Plate 2-1. (a) Global LAI and (b) FPAR in September-October 1997 derived from SeaWiFs (sea-viewing wide field-of-view sensor) data. This data set includes daily atmosphere-corrected surface reflectances at eight shortwave spectral bands. Surface reflectances at red (670 nm) and near-infrared (865 nm) at 8 km resolution were used. The algorithm was applied to daily surface reflectance data for all days from September 18 to October 12, 1997. For each pixel, LAI and FPAR values corresponding to the maximum NDVI during this period are shown in these pannels. The look-up table for biome 1 (grasses and cereal crops, Table 2-1) was used to produce global LAI and FPAR for all biome types.

Table 2-1. Canopy Structural Attributes of Global Land Covers From the Viewpoint of Radiative Transfer Modeling

	Grasses and Cereal Crops	Shrubs	Broadleaf Crops	Savannas	Broadleaf Forests	Needle Forests
Horizontal heterogeneity	no	yes	variable	yes	yes	yes
Ground cover	100%	20-60%	10-100%	20-40%	> 70%	> 70%
Vertical heterogeneity (leaf optics and LAD)	no	no	no	yes	yes	yes
Stems/trunks	no	no	green stems	yes	yes	yes
Understory	no	no	no	grasses	yes	yes
Foliage dispersion	minimal clumping	random	regular	minimal clumping	clumped	severe clumping
Crown shadowing	no	not mutual	no	no	yes mutual	yes mutual
Brightness of canopy ground	medium	bright	dark	medium	dark	dark

The canopy structure is the most important variable determining the three-dimensional radiation field in vegetation canopies. Therefore section 2.3 starts with a precise mathematical definition of this variable and how various canopy radiation models treat this variable. This allows us to specify some common properties of the present canopy radiation models. The basic physical principle underlying the proposed LAI/FPAR retrieval algorithm is the law of energy conservation. However, a rather wide family of canopy radiation models (described in the Appendix) conflict with this law. Therefore the three-dimensional transport equation which includes a nonphysical internal source is taken as the starting point for the derivation of the algorithm. In section 2.5, a technique developed in atmospheric optics is utilized to parameterize the radiative field in terms of reflectance properties of the canopy and ground, as well as to split the radiative transfer problem into two independent sub-problems, each of which is expressed in terms of three basic components of the energy conservation law: canopy transmittance, reflectance, and absorptance. These components are elements of the look-up table (LUT), and the algorithm interacts only with the elements of the LUT. This provides the required independence of the retrieval algorithm to a particular canopy radiation model. The next important step in achieving property (3) is to specify the dependence of canopy transmittance, reflectance, and absorptance on wavelength. It is precisely derived in section 2.6; this dependence is described by a simple function which depends on the unique positive eigenvalue of the transport equation. The eigenvalue relates optical properties of individual leaves to canopy structure. This result not only allows a significant reduction in the size of the LUT but also relates canopy spectral reflectance with spectral properties of individual leaves, which is a rather stable characteristic of green leaves.

In spite of the essential reduction of possible canopy representatives by introducing a vegetation cover classification, the inverse problem still allows for multiple solutions. A technique allowing the reduction of nonphysical solutions is described in section 2.7. A definition of the LUT is given in this section as well. A method to estimate the most probable LAI and FPAR, accounting for specific features of the MODIS and MISR instruments, and providing convergence of the algorithm is discussed in sections 2.8 and 2.10. The maximum positive eigenvalue and the unique positive eigenvector corresponding to this eigenvalue, detailed in section 2.6, express the law of energy conservation in a compact form. The results of this section allow us to relate the Normalized Difference Vegetation Index (NDVI) to this fundamental physical principle. Relationships between FPAR and NDVI are also used in our algorithm as a backup to the LUT approach, and so we discuss these in section 2.11.

2.2 Canopy Structural Types of Global Vegetation

Solar radiation scattered from a vegetation canopy and measured by satellite sensors results from interaction of photons traversing through the foliage medium, bounded at the bottom by a radiatively participating surface. Therefore to estimate the canopy radiation regime, three important features must be carefully formulated [*Ross*, 1981]. They are (1) the architecture of individual plants or trees and the entire canopy; (2) optical properties of vegetation elements (leaves, stems) and ground; the former depends on physiological conditions (water status, pigment concentration); and (3) atmospheric conditions which determine the incident radiation field. Photon transport theory aims at deriving the solar radiation regime, both within the vegetation canopy and radiant exitance, using the above mentioned attributes as input data. This underlies a land cover classification [*Myneni et al.*, 1997] which is compatible with the basic physical principle of transport theory, the law of energy conservation. Global land covers can be classified into six types (biomes), depending on their canopy structure (Table 2-1). The structural attributes of these land covers can be parameterized in terms of variables that transport theory admits as follows.

The heterogeneity of the plant canopy can be described by the three-dimensional leaf area distribution function u_{L}. Its values at spatial points depend on trunk distribution, topography, stem-trunk-branch area fraction, foliage dispersion, leaf and crown size, and leaf clumping [*Myneni and Asrar*, 1991; *Oker-Blom et al.*, 1991]. The three-dimensional distribution of leaves determines various models to account for shadowing effects [*Kuusk*, 1985; *Li and Strahler*, 1985; *Verstraete et al*., 1990]. The leaf area index LAI is defined as

$$\mathrm{LAI} = \frac{1}{X_{\mathrm{S}} \cdot Y_{\mathrm{S}}} \int_V u_{\mathrm{L}}(r)\, dr \quad , \tag{1}$$

where V is the domain in which a plant canopy is located; X_{S}, Y_{S} are horizontal dimensions of V. If the vegetation canopy consists of N_{c} individual trees, LAI can be expressed as

$$\mathrm{LAI} = \sum_{k=1}^{N_{\mathrm{C}}} p_k \frac{1}{S_k} \int_{V_k} u_{\mathrm{L}}(r)\, dr = \sum_{k=1}^{N_{\mathrm{C}}} p_k \cdot \mathrm{LAI}_k \ ,$$

where S_k is the foliage envelope projection (e.g., crown) of the kth plant or tree onto the ground; $p_k=S_k/(X_S \cdot Y_S)$ and LAI_k is the leaf area index of an individual plant or tree. Thus LAI is

$$\mathrm{LAI} = g \cdot \mathrm{LAI}_0 \ ,$$

where $g = \sum_{k=1}^{N_C} p_k$ is the ground cover and

$$\mathrm{LAI}_0 = \frac{1}{g} \sum_{k=1}^{N_C} p_k \cdot \mathrm{LAI}_k$$

is the mean LAI of a single plant or tree. The spatial distribution of plants or trees in the stand is a characteristic of the biome type and is assumed known. For each biome type, the leaf area density distribution function is parameterized in terms of ground cover and mean leaf area index of an individual plant or tree, each varying within given biome specific intervals $[g_{\min}, g_{\max}]$ and $[L_{\min}, L_{\max}]$, respectively. Thus the vegetation canopy is represented as a domain V consisting of identical plants or trees in order to numerically evaluate the transport equation.

To parameterize the contribution of the surface underneath the canopy (soil and/or understory) to the canopy radiation regime, an effective ground reflectance is introduced, namely,

$$\rho_{q,\mathrm{eff}}(\lambda) = \frac{1}{\pi} \frac{\int_{2\pi-} \int_{2\pi+} R_{\mathrm{b},\lambda}(\Omega',\Omega) |\mu \cdot \mu'| L_\lambda(r_\mathrm{b},\Omega',\Omega_0) d\Omega d\Omega'}{\int_{2\pi^-} q(\Omega') |\mu'| L_\lambda(r_\mathrm{b},\Omega',\Omega_0) d\Omega'} \ . \qquad (2)$$

Here L_λ is radiance at a point r_b of the canopy bottom; $R_{\mathrm{b},\lambda}$ is the bidirectional reflectance factor of the canopy bottom. The function q is a wavelength-independent configurable function used to better account for specific features of various biomes, and it satisfies the following condition:

$$\int_{2\pi^-} q(\Omega')\, d\Omega' = 1 \ . \qquad (3)$$

Note that the effective ground reflectance depends on the radiation regime in the vegetation canopy. It follows from the definition that the variation of $\rho_{q,\mathrm{eff}}$ satisfies the following inequality:

$$\min_{\Omega' \in 2\pi-} \frac{1}{\pi} \frac{\int_{2\pi+} R_{\mathrm{b},\lambda}(\Omega',\Omega) |\mu| d\Omega}{q(\Omega')} \le \rho_{q,\mathrm{eff}}(r_\mathrm{b},\lambda) \le \max_{\Omega' \in 2\pi-} \frac{1}{\pi} \frac{\int_{2\pi+} R_{\mathrm{b},\lambda}(\Omega',\Omega) |\mu| d\Omega}{q(\Omega')} \ ; \qquad (4)$$

that is, the range of variations depends on the integrated bidirectional factor of the ground surface only. The bidirectional reflectance factor of the ground surface $R_{\mathrm{b},\lambda}$ and the effective ground reflectance are assumed to be horizontally homogeneous; that is, they do not depend on the spatial point r_b. The pattern of the effective ground reflectances (ρ_1, ρ_2, …, ρ_{11}), $\rho_i=\rho_{q,\mathrm{eff}}(\lambda_i)$, at the MODIS spectral bands, is taken as a parameter characterizing hemispherically integrated reflectance of the canopy ground (soil and/or understory) and can vary continuously within the

interval defined by equation (4). The lower and upper bounds of equation (4) depend on biome type. The set of various patterns of effective ground reflectances is a static table of the algorithm, i.e., element of the look-up table. The present version of the look-up table contains 25 patterns of effective ground reflectances evaluated from the soil reflectance model of *Jacquemoud et al*. [1992], using model inputs presented by *Baret et al*. [1993]. Figure 2-1 demonstrates spectral ground reflectances $\rho_{q,\mathrm{eff}}$ for biome 1 (grasses and cereal crops).

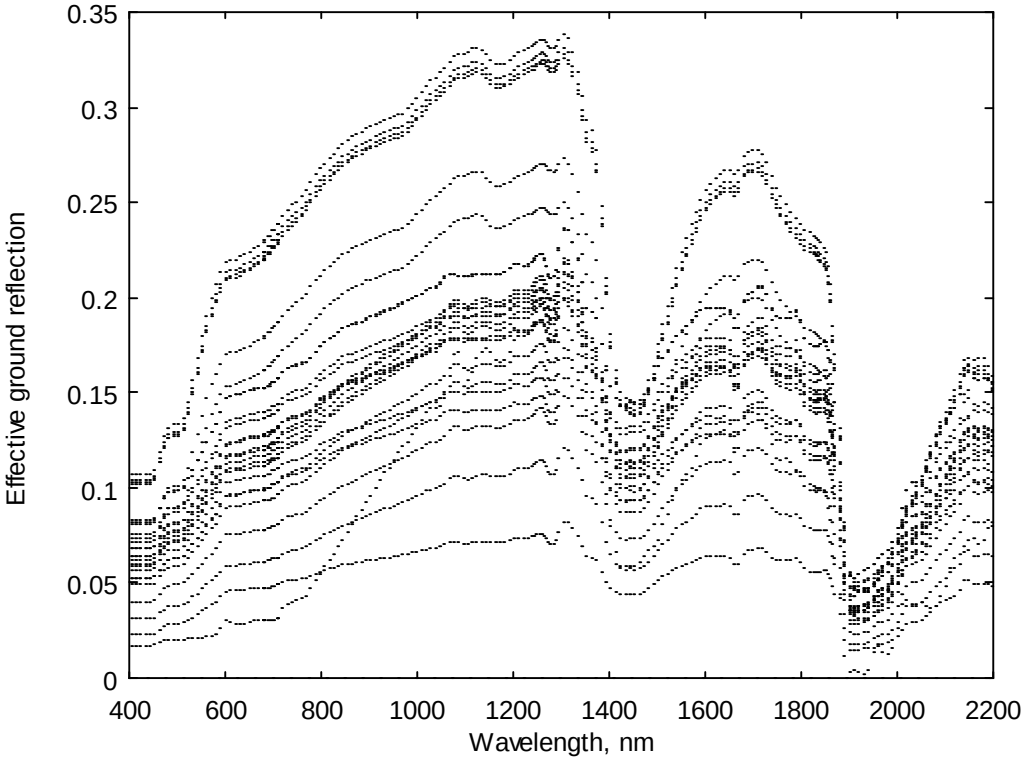


Figure 2-1. Spectral effective ground reflectance for 25 different soils. It includes three soil types described as mixtures of clay, sand, and peat. Each soil type is characterized by three moisture levels (wet, median, dry) and from two to three soil roughnesses (rough, median, smooth, or rough and smooth). These effective ground reflectances were evaluated from the soil reflectance model of *Jacquemoud et al*. [1992] using model inputs presented by *Baret et al*. [1993].

To account for the anisotropy of the ground surface, an effective ground anisotropy S_q is used,

$$S_q(r_b,\Omega)=\frac{1}{\rho_{q,eff}(\lambda)}\frac{1}{\pi}\cdot\frac{\int_{2\pi^-}R_{b,\lambda}(\Omega',\Omega)|\mu'|L_\lambda(r_b,\Omega')d\Omega'}{\int_{2\pi^-}q(\Omega')|\mu'|L_\lambda(r_b,\Omega')d\Omega'},\quad r_b\in\delta V_b,\Omega\bullet n_b<0\ ,\qquad(5)$$

where n_b is the outward normal at point r_b. The effective ground anisotropy S_q depends on the canopy structure as well as the incoming radiation field. We note the following property:

$$\int_{2\pi^+}S_q(r_b,\Omega)\,|\mu|\,d\Omega=1\ ,$$

that is, the integral depends neither on spatial nor on spectral variables. For each biome type, the effective ground anisotropy is assumed wavelength independent. The six cover types presented in Table 2-1 can now be expressed in terms of the above introduced variables.

2.2.1 Biome 1: Grasses and Cereal Crops

Canopies exhibit vertical and lateral homogeneity, vegetation ground cover of about 1.0 (g_{min}=g_{max}=1), plant height generally about a meter or less, erect leaf inclination, no woody material, minimal leaf clumping, and soils of intermediate brightness. The one-dimensional radiative transfer model is invoked in this situation. Leaf clumping is implemented by modifying the projection areas with a clumping factor generally less than 1. The soil reflection is assumed Lambertian; that is, $R_{b,\lambda}$=$R_{lam,\lambda}$. We also set q=1. The effective soil reflection and anisotropy then have the simplified form

$$\rho_{q,eff}(\lambda)=R_{lam,\lambda}\ ,\qquad S_q(r_b,\Omega)=1/\pi\ .\qquad(6)$$

2.2.2 Biome 2: Shrubs

Canopies exhibit lateral heterogeneity, low (g_{min}=0.2) to intermediate (g_{min}=0.6) vegetation ground cover, small leaves, woody material, and bright backgrounds. The full three-dimensional (3-D) model is invoked. Hot spot, i.e., enhanced brightness about the retrosolar direction due to absence of shadows [*Privette et al*., 1994], is modeled by shadows cast on the ground (no mutual shadowing because ground cover is low). This land cover is typical of semiarid regions with extreme hot (brush) or cold (tundra/taiga) temperature regimes and poor soils. For this biome we represent the bidirectional soil reflectance factor $R_{b,\lambda}$ as

$$R_{b,\lambda}(\Omega',\Omega)=R_{1,\lambda}(\Omega')\cdot R_{2,\lambda}(\Omega,\Omega_0)\ ,\qquad(7)$$

where Ω_0 is the direction of the direct solar radiance. We set

$$q(\Omega')=R_{1,\lambda}(\Omega')\big/\rho^*_{1,\lambda}\ .\qquad(8)$$

The effective soil reflection and soil anisotropy then have the form

$$\rho_{q,\mathrm{eff}}(\lambda) = \rho^{*}_{1,\lambda} \cdot \rho^{*}_{2,\lambda}(\Omega_0) \ , \qquad S_q(\Omega) = \frac{1}{\pi} \cdot \frac{R_{2,\lambda}(\Omega, \Omega_0)}{\rho^{*}_{2,\lambda}(\Omega_0)} \ , \tag{9}$$

where

$$\rho^{*}_{1,\lambda} = \frac{1}{\pi} \int_{2\pi^{-}} R_{1,\lambda}(\Omega')|\mu'|d\Omega' , \qquad \rho^{*}_{2,\lambda}(\Omega_0) = \frac{1}{\pi} \int_{2\pi^{+}} R_{2,\lambda}(\Omega, \Omega_0)|\mu|d\Omega \ .$$

The functions q and S_q are assumed wavelength independent and serve as parameter of this biome. This biome is characterized by intermediate vegetation ground cover. The use of the above model for the bidirectional soil reflectance factor means that only the incoming direct beam of solar radiation which reaches the soil can influence the anisotropy of the radiation field in the plant canopy.

2.2.3 Biome 3: Broadleaf Crops

Canopies exhibit lateral heterogeneity, large variations in vegetation ground cover from crop planting to maturity ($g_{\min}$=0.1, $g_{\max}$=1.0), regular leaf spatial dispersion, photosynthetically active, i.e., green, stems, and dark soil backgrounds. The regular dispersion of leaves (i.e., the positive binomial model) leads to a clumping factor that is generally greater than unity. The green stems are modeled as erect reflecting protrusions with zero transmittance. The three-dimensional radiative transfer model is invoked in this situation. The soil reflection is assumed Lambertian, i.e., $R_{\mathrm{b},\lambda}=R_{\mathrm{lam},\lambda}$. $_{,\lambda}$. The function q=1. The effective soil reflection and anisotropy are expressed by equation (6).

2.2.4 Biome 4: Savanna

Canopies with two distinct vertical layers, understory of grass, low ground cover of overstory trees ($g_{\min}$=0.2, $g_{\max}$=0.4), canopy optics, and structure are therefore vertically heterogeneous. The full 3-D method is required. The interaction coefficients have a strong vertical dependency. Savannas in the tropical and subtropical regions are characterized as mixtures of warm grasses and broadleaf trees. In the cooler regimes of the higher latitudes, they are described as mixtures of cool grass and needle trees. The effective soil reflection and soil anisotropy then are simulated by equation (9).

2.2.5 Biome 5: Broadleaf Forests

Vertical and lateral heterogeneity, high ground cover ($g_{\min}$=0.8, $g_{\max}$=1.0), green understory, mutual shadowing of crowns, foliage clumping, trunks, and branches are included, so the canopy structure and optical properties differ spatially. Mutual shadowing of crowns is handled by modifying the hot spot formulation. Therefore stand density and crown size define this gap parameter. The branches are randomly oriented, but tree trunks are modeled as erect structures. Both trunk and branch reflectance are specified from measurements. For this biome the three-dimensional transport equation is utilized to evaluate the effective soil reflection and

anisotropy as a function of LAI and Sun position. These are intermediate calculations and are used to precompute parameters stored in the LUT.

2.2.6 Biome 6: Needle Forests

These are canopies with needles, needle clumping on shoots, severe shoot clumping in whorls, dark vertical trunks, sparse green understory, and crown mutual shadowing. This is the most complex case, invoking the full 3-D method with all its options. A typical shoot is modeled to handle needle clumping on the shoots. The shoots are then assumed to be clumped in the crown space. Mutual shadowing by crowns is handled by modifying the hot spot formulation. The branches are randomly oriented but the dark tree trunks are modeled as erect structures. Both trunk and branch reflectance are specified from measurements. The effective soil reflection and anisotropy are evaluated the same way as for biome 5.

2.3 Radiative Transfer Problem for Vegetation Media

The domain V in which a vegetation canopy is located, is a parallelepiped of horizontal dimensions X_S, Y_S, and biome-dependent height Z_S. The top δV_t, bottom δV_b, and lateral δV_l surfaces of the parallelepiped form the canopy boundary $\delta V=\delta V_t+\delta V_b+\delta V_l$. The structure of the vegetation canopy is defined by an indicator function $\chi(r)$ whose value is 1, if there is a phytoelement at the spatial point r, and zero otherwise. Here the position vector r denotes the Cartesian triplet (x,y,z) with $(0<x<X_S)$, $(0<y<Y_S)$, and $(0<z<Z_S)$, with its origin $O=(0,0,0)$ at the top of the canopy. The indicator function is treated as a random variable. Its distribution function, in the general case, depends on both macroscale (e.g., random dimension of the trees and their spatial distribution) and microscale (e.g., structural organization of an individual tree) properties of the vegetation canopy and includes all three of its components, absolutely continuous, discrete, and singular [*Knyazikhin et al.*, 1998c]. In order to approximate this function, a fine spatial mesh is introduced by dividing the domain V into N_ε nonoverlapping fine cells, e_i, $i=1,2, \ldots, N_\varepsilon$, of size $\Delta x=\Delta y=\Delta z$. Each realization $\chi(r)$ of the canopy structure is replaced by its mean over the fine cell e_i as

$$u_L(r)=\frac{1}{m(e_i)}\int_{e_i}\chi(r)\,m(dr), \quad r\in e_i \quad . \tag{10}$$

Here m is a measure suitable to perform the integration of equation (10). The function u_L is the leaf area density distribution function. In the general case, (10) is the Lebesgue integral and it may not coincide with an integral in the "true sense." This integration technique provides the convergence process $u_L\rightarrow\chi/m(V)$ when $\varepsilon\rightarrow 0$ [*Knyazikhin et al.*, 1998c], and so equation (10) can be taken as an approximation of the structure of the vegetation canopy. The accuracy of this approximation depends on size ε of the fine cell e_i. To our knowledge, all existing canopy radiation models are based on the approximation of (10) by a piece-wise continuous function, e.g., describing both the spatial distribution of various geometrical objects like cones, ellipsoids, etc., and the variation of leaf area within a geometrical figure [*Ross and Nilson*, 1968; *Nilson*, 1977; *Ross* 1981; *Norman and Wells*, 1983; *Li et al.*, 1995]. Therefore we proceed with the suggestion that u_L is the random value whose distribution function is described by a piece-wise

continuous function. For each realization, the radiation field in such a medium can be expressed as

$$\Omega \bullet \nabla L_\lambda(r,\Omega) + G(r,\Omega)u_{\mathrm{L}}(r)L_\lambda(r,\Omega) = \frac{u_{\mathrm{L}}(r)}{\pi}\int_{4\pi}\Gamma_\lambda(r,\Omega' \to \Omega)L_\lambda(r,\Omega')d\Omega' \quad . \qquad (11)$$

Here $\Omega\bullet\nabla$ is the derivative at r along the direction Ω; L_λ is the monochromatic radiance at point r and in the direction Ω,

$$G(r,\Omega) = \frac{1}{2\pi}\int_{2\pi^+} g_{\mathrm{L}}(r,\Omega_{\mathrm{L}})\left|\Omega \bullet \Omega_{\mathrm{L}}\right| d\Omega_{\mathrm{L}} \quad ,$$

is the mean projection of leaf normals at r onto a plane perpendicular to the direction Ω; g_{L} is the probability density of leaf normal distribution over the upper hemisphere $2\pi^+$;

$$\frac{1}{\pi}\Gamma_\lambda(r,\Omega' \to \Omega) = \frac{1}{2\pi}\int_{2\pi^+} g_{\mathrm{L}}(r,\Omega_{\mathrm{L}})\left|\Omega' \bullet \Omega_{\mathrm{L}}\right|\gamma_{\mathrm{L},\lambda}(r,\Omega_{\mathrm{L}},\Omega' \to \Omega)d\Omega_{\mathrm{L}} \quad ,$$

is the area-scattering phase function [*Ross*, 1981], and $\gamma_{\mathrm{L},\lambda}$ is the leaf-scattering phase function. Unit vectors are expressed in spherical coordinates with respect to ($-Z$) axis. It follows from the above definitions that the solution of the transport equation is also a random variable. For each biome type, the angular distribution of radiance leaving the top surface of the vegetation canopy is defined to be the mean value, $<L_\lambda>_{\mathrm{bio}}$, of L_λ over different realizations of the given biome type. The following definitions of biome-specific reflectances are used in this paper.

The hemispherical-directional reflectance factor (HDRF) for nonisotropic incident radiation is the ratio of the mean radiance leaving the top of the plant canopy, $<L_\lambda(r_{\mathrm{t}},\Omega)>_{\mathrm{bio}}$, $\Omega\bullet\mathrm{n}_{\mathrm{t}}>0$, to radiance reflected from an ideal Lambertian target into the same beam geometry and illuminated under identical atmospheric conditions [*Diner et al.*, 1998a]; that is,

$$r_\lambda(\Omega,\Omega_0) = \frac{<L_\lambda(r_{\mathrm{t}},\Omega)>_{\mathrm{bio}}}{\frac{1}{\pi}\int_{2\pi-} L_\lambda(r_{\mathrm{t}},\Omega',\Omega_0)\left|\Omega' \bullet \mathrm{n}_{\mathrm{t}}\right| d\Omega'}, \quad \Omega \bullet \mathrm{n}_{\mathrm{t}} > 0 \quad .$$

Here n_{t} is the outward normal at points $r_{\mathrm{t}} \in \delta V_{\mathrm{t}}$; $<\cdot>_{\mathrm{bio}}$ denotes the averaging over the ensemble of biome realizations; and Ω_0 is the direction of the monodirectional solar radiation incident on the top of the canopy boundary.

The bihemispherical reflectance (BHR) for nonisotropic incident radiation is the ratio of the mean radiant exitance to the incident radiant [*Diner et al.*, 1998a], i.e.,

$$A_\lambda^{\mathrm{hem}}(\Omega_0) = \frac{\int_{2\pi+} <L_\lambda(r_{\mathrm{t}},\Omega)>_{\mathrm{bio}} \left|\Omega \bullet \mathrm{n}_{\mathrm{t}}\right| d\Omega}{\int_{2\pi-} L_\lambda(r_{\mathrm{t}},\Omega',\Omega_0)\left|\Omega' \bullet \mathrm{n}_{\mathrm{t}}\right| d\Omega'} \quad .$$

In order to quantify a proportion between direct and diffuse component of incoming radiation, the ratio $f_{\mathrm{dir}}(\Omega_0)$ of direct radiant incident on the top of the plant canopy to the total

incident irradiance is used. If f_{dir}=1, HDRF and BHR become the bidirectional reflectance factor (BRF), and the directional hemispherical reflectance (DHR). Here $r_\lambda(\Omega,\Omega_0)$ and $A_\lambda^{hem}(\Omega_0)$ denote, depending on the situation (f_{dir}=1 or $f_{dir}\neq$1), HDRF and BHR or BRF and DHR.

In spite of the diversity of canopy reflectance models, they can be classified with respect to how the averaging over the ensemble of canopy realizations is performed. In terms of equation (11), this is equivalent to how the averaging of $u_L(r)L_\lambda(r,\Omega)$ is performed. In the turbid medium models, the vegetation canopy is treated as a gas with nondimensional planar scattering centers [*Ross*, 1981]. Such models presuppose that

$$\left\langle u_L(r)L_\lambda(r,\Omega)\right\rangle_{bio} = \left\langle u_L(r)\right\rangle_{bio}\left\langle L_\lambda(r,\Omega)\right\rangle_{bio} \quad . \tag{12}$$

As a result, equation (10) is reduced to the classical transport equation [*Ross*, 1981] whose solution is the mean radiance $< L_\lambda(r,\Omega)>_{bio}$. This technique allows the design of conservative radiation transfer models, i.e., models in which the law of energy conservation holds true for any elementary volume. Such an approach cannot account for the hot spot phenomena because it ignores shadowing effects. This motivated the development of a family of radiative transfer models based on the following fact: the two events that a point inside a leaf canopy can be viewed from two points r_1 and r_2 are not independent [*Kuusk*, 1985]. The mean of $u_L(r)L_\lambda(r,\Omega)$ is presented as

$$\left\langle u_L(r)L_\lambda(r,\Omega)\right\rangle_{bio} = p(r,\Omega,\Omega')\cdot\left\langle u_L(r)\right\rangle_{bio}\left\langle L_\lambda(r,\Omega)\right\rangle_{bio} \quad ,$$

where p is the bidirectional gap probability [*Kuusk*, 1985; *Li and Strahler*, 1985; *Verstraete et al*., 1990; *Oker-Blom et al*., 1991]. Such models account accurately for once scattered radiance, taking $Gp<u_L>$ as the extinction coefficient. For evaluation of the multiply scattered radiance, assumption (12) is usually used [*Marshak*, 1989; *Myneni et al*., 1995b]. These types of canopy-radiation models can well simulate BRFs. However, they are not conservative (Appendix 1). The problem of obtaining a correct closed equation for the mean monochromatic radiance was formulated and solved by *Vainikko* [1973], where the equations for the mean radiance were derived through spatial averaging of the stochastic transport equation (11) in a model of broken clouds. This approach was studied in detail by *Titov* [1990]. *Anisimov and Menzulin* [1981] utilized similar ideas to describe the radiation regime in plant canopies. The stochastic models incorporate the best features of the above mentioned approaches. The aim of this paper is to derive some general properties of radiation transfer which do not depend on a particular model and which can be taken as the basis of our LAI/FPAR retrieval algorithm. Equation (11) express the law of energy conservation in the most general form. Therefore our aim can be achieved, if this equation is taken as a starting point for deriving the desired properties. In order to include canopy reflectance models with hot spot effect into consideration, a transport equation of the form

$$\Omega\bullet\nabla L_\lambda(r,\Omega) + G(r,\Omega)u_L(r)L_\lambda(r,\Omega) = \frac{u_L(r)}{\pi}\int_{4\pi}\Gamma_\lambda(r,\Omega'\to\Omega)L_\lambda(r,\Omega')d\Omega' + F_\lambda(r,\Omega) \tag{13}$$

will also be considered in this paper. Here F_λ is a function which accounts for the hot spot effect.

Equation (13) alone does not provide a full description of random realizations of the radiative field. It is necessary to specify the incident radiance at the canopy boundary δV i.e., specification of the boundary conditions. Because the canopy is adjacent to the atmosphere, and neighboring canopies, and the soil or understory, all which have different reflection properties, the following boundary conditions will be used to describe the incoming radiation [*Ross et al.*, 1992]:

$$L_\lambda(r_t,\Omega) = L^{top}_{d,\lambda}(r_t,\Omega,\Omega_0) + L^{top}_{m,\lambda}(r_t)\delta(\Omega-\Omega_0), \quad r_t \in \delta V_t, \Omega \bullet n_t < 0\ , \tag{14}$$

$$L_\lambda(r_l,\Omega) = \frac{1}{\pi}\int_{\Omega'\bullet n_l>0} R_{l,\lambda}(\Omega',\Omega)L_\lambda(r_l,\Omega')\left|\Omega'\bullet n_l\right|d\Omega' + L^{lat}_{d,\lambda}(r_l,\Omega) + L^{lat}_{m,\lambda}(r_l)\delta(\Omega-\Omega_0), \quad r_l \in \delta V_l, \Omega \bullet n_l < 0\ , \tag{15}$$

$$L_\lambda(r_b,\Omega) = \frac{1}{\pi}\int_{\Omega'\bullet n_b>0} R_{b,\lambda}(\Omega',\Omega)L_\lambda(r_b,\Omega')\left|\Omega'\bullet n_b\right| d\Omega', \quad r_b \in \delta V_b, \Omega \bullet n_b < 0\ , \tag{16}$$

where $L^{top}_{d,\lambda}$ and $L^{top}_{m,\lambda}$ are the diffuse and monodirectional components of solar radiation incident on the top surface of the canopy boundary δV_t; $\Omega_0 \sim (\mu_0,\phi_0)$ is the direction of the monodirectional solar component; δ is the Dirac delta function; $L^{lat}_{m,\lambda}$ is the intensity of the monodirectional solar radiation arriving at a point $r_l \in \delta V_l$ along Ω_0 without experiencing an interaction with the neighboring canopies; $L^{lat}_{d,\lambda}(r_l)$ is the diffuse radiation penetrating through the lateral surface δV_l; $R_{l,\lambda}$ and $R_{b,\lambda}$ (in sr^{-1}) are the bidirectional reflectance factors of the lateral and the bottom surfaces, respectively; and n_t, n_l, and n_b are the outward normals at points $r_t \in \delta V_t$, $r_l \in \delta V_l$ and $r_b \in \delta V_b$, respectively. A solution of the boundary value problem, expressed by equations (13)-(16), describes a random realization of the radiation field in a vegetation canopy.

2.4 Assumptions in Radiation Transfer Process

Theoretically, the sets D_A and D_r can be generated offline by solving the transport equation at four MISR spectral bands for various combination of Sun-sensor geometry and all canopy realizations from the set P. However, one can realize it only if the sets D_A and D_r can be reprocessed with minimum effort. The time required to precompute these sets is a direct function of the number of spectral channels used, combinations of Sun-sensor geometry, and elements in the set P. For example, the generation of the set D_r using this direct method takes approximately 192 computer hours of medium performance IBM RS/6000 RISC workstation [*Running et al.*, 1996]. The size of D_r containing BRFs for two spectral bands and for all six biomes is about 63 megabites. The inclusion of more spectral bands and view directions leads to significant demands on the core memory required to execute this algorithm. It makes this approach impractical in the case of MISR instrument. The aim of this section and section 2.5 is to formulate some assumptions allowing for a significant reduction in the size of D_A and D_r.

2.4.1 Conservativity

A radiative transfer model is defined to be conservative if the law of energy conservation holds true for any elementary volume [*Bass et al*., 1986]. Within a conservative model, radiation absorbed, transmitted, and reflected by the canopy is always equal to radiation incident on the canopy. A rather wide family of canopy radiation models [*Kuusk*, 1985; *Marshak*, 1989; *Pinty et al*., 1989; *Li and Strahler*, 1992; *Myneni et al*., 1995; *Pinty and Verstraete*, 1998] which account for the hot spot are equivalent to the solution of the above boundary value problem in which the function F_λ has the following form [*Knyazikhin et al*., 1998a]:

$$F_\lambda(r,\Omega) = [\sigma(r,\Omega) - \sigma_H(r,\Omega,\Omega_0)]L_{H,\lambda}(r,\Omega).$$

Here $L_{H,\lambda}$ is the upwardly directed once-scattered radiance produced by the hot spot, and σ_H is a model-dependent total interaction cross section, introduced in canopy radiation models to account for the hot spot effect and to evaluate $L_{H,\lambda}$. The total interaction cross section σ is used to evaluate the attenuation of both direct solar radiance and multiply scattered radiance. Because F_λ can take on negative values, it has no physical meaning in terms of energy conservation. These types of canopy radiation models are mainly used to fit simulated BRFs to measured BRFs. However, the ability of a model to simulate canopy reflection is not a sufficient requisite for the solution of the inverse problem. Canopy radiation models must also satisfy the law of energy conservation and provide the correct proportions of canopy absorptance, transmittance, and reflectance. Because the retrieval algorithm is based on energy conservation, the following "minimum" requirement, which the canopy radiation models must satisfy in order to be useful for inverse problems, is formulated:

$$\int_V dr \int_{4\pi} d\Omega\, F_\lambda(r,\Omega) = 0,$$

for any λ. This equation does not allow a nonphysical source $F_\lambda(r,\Omega)$ to influence the canopy radiative energy balance. Currently, we use a model for σ_H proposed by *Myneni et al.* [1995]. A nonconservative canopy radiation model must be corrected, as described in section 2.7.

2.4.2 Leaf Area Index

The leaf area index LAI is defined as

$$\mathrm{LAI} = \frac{1}{X_S Y_S} \int_V u_L(r)\, dr \quad .$$

If the vegetation canopy consists of N_c individual trees, LAI can be expressed as

$$\mathrm{LAI} = \sum_{k=1}^{N_c} g_k \frac{1}{S_k} \int_{V_k} u_L(r)\, dr = \sum_{k=1}^{N_c} g_k \mathrm{LAI}_k \ ,$$

where S_k is the crown projection of the kth tree onto the ground; $g_k=S_k/(X_S Y_S)$ and LAI_k is the leaf area index of an individual tree. Thus LAI is $\mathrm{LAI} = g\mathrm{LAI}_0$, where $g = \sum_{k=1}^{N_c} g_k$ is the ground cover, and

$$\mathrm{LAI}_0 = \frac{1}{g}\sum_{k=1}^{N_c} g_k \cdot \mathrm{LAI}_k$$

is the mean LAI of a single tree. The spatial distribution of trees in the stand is a characteristic of the biome type and is assumed to be random. For each biome type, the leaf area density distribution function is parameterized in terms of the ground cover and mean leaf area index of an individual tree, each varying within given biome-specific intervals $[g_{\min}, g_{\max}]$ and $[L_{\min}, L_{\max}]$, respectively. Thus the vegetation canopy is represented as a domain V consisting of identical trees in order to numerically evaluate the transport equation.

2.4.3 Anisotropy of Incoming Diffuse Radiation

A model of clear-sky radiance proposed by *Pokrowski* [1929] is used to approximate the ratio between the angular distribution of incoming diffuse radiation and its flux:

$$\frac{L_{\mathrm{d},\lambda}^{\mathrm{top}}(r_\mathrm{t},\Omega)}{\int\limits_{2\pi-} L_{\mathrm{d},\lambda}^{\mathrm{top}}(r_\mathrm{t},\Omega)|\mu|d\Omega} = \left[1-\exp\left(\frac{-0.32}{|\mu|}\right)\right]\frac{1+\Omega\bullet\Omega_0}{1-\Omega\bullet\Omega_0}, \quad \Omega\sim(\mu,\phi), \quad \mu<0\,.$$

We assume that this ratio does not depend on wavelength. The diffuse radiation $L_{\mathrm{d},\lambda}^{\mathrm{top}}$ does not depend on the top boundary space point $r_\mathrm{t}\in\delta V_\mathrm{t}$. This allows the parameterization of the incoming radiation field in terms of f_{dir} and the total (diffuse and direct) incident flux.

2.4.4 Boundary Conditions for Lateral Surface

The radiation penetrating through the lateral sides of the canopy depends on the neighboring environment. Its influence on the radiation field within the canopy is especially pronounced near the lateral canopy boundary. Therefore inaccuracies in the lateral boundary conditions may cause distortions in the simulated radiation field within the domain V. These distortions, however, decrease with distance from this boundary toward the center of the domain. The size of the "distorted area" depends on the adjoining vegetation, atmospheric conditions, and model resolution [*Kranigk*, 1996]. In particular, it has been shown that these lateral effects can be neglected when the radiation regime is analyzed in a rather extended canopy, as is the case considering the rather large MISR pixel (~1.1 km). Therefore we idealize our canopy as a horizontally infinite region. We will use a "vacuum" boundary condition for the lateral surface to numerically evaluate a solution for the case of a horizontally infinite domain,

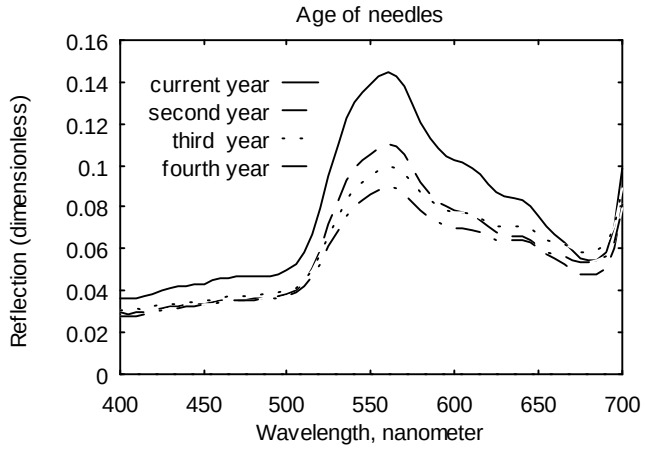
$$L_\lambda(r_\mathrm{l},\Omega) = 0, \qquad r_\mathrm{l}\in\delta V_\mathrm{l}, \qquad \Omega\bullet\mathrm{n}_\mathrm{l} < 0.$$

2.4.5 Optical Properties of Foliage

The leaf-scattering phase function $\gamma_{L,\lambda}$ is assumed to be bi-Lambertian [*Ross and Nilson*, 1968]; that is, a fraction of the energy intercepted by the foliage element is reflected or transmitted in a cosine distribution about the leaf normal,

$$\gamma_{L,\lambda}(r,\Omega_L,\Omega' \to \Omega) = \begin{cases} \pi^{-1} r_{D,\lambda}(r) \left|\Omega \bullet \Omega_L\right|, & (\Omega \bullet \Omega_L)(\Omega' \bullet \Omega_L) < 0, \\ \pi^{-1} t_{D,\lambda}(r) \left|\Omega \bullet \Omega_L\right|, & (\Omega \bullet \Omega_L)(\Omega' \bullet \Omega_L) > 0. \end{cases}$$

Here $r_{D,\lambda}$ and $t_{D,\lambda}$ are the spectral reflectance and transmittance, respectively, of the leaf element. Figure 2-2 shows an example of the sensitivity of the reflection coefficient $r_{D,\lambda}$ for the 1-year shoots (*Picea abies* (L) karst) on its location in space. In spite of this spatial variation the shapes of spectral reflectance and transmittance are rather stable. For example, compared with the mean, the deviation is, on average, about 12-15%, which does not exceed the accuracy of the canopy radiation model [*Knyazikhin et al.*, 1997]. Therefore the spatial variation of foliage optical properties can be neglected. Thus the algorithm can be parameterized in terms of spectral leaf albedo, $\omega(\lambda)=r_{D,\lambda}+t_{D,\lambda}$. For each biome the mean spectral leaf albedo is stored in the CART file. The ratio $r_{D,\lambda}/\omega(\lambda)$ is also assumed to be independent of wavelength, in any given biome type. We note that the validity of the assumptions section 2.4.3-2.4.5 was verified by comparing simulation results with field measurements [*Knyazikhin et al.*, 1997].



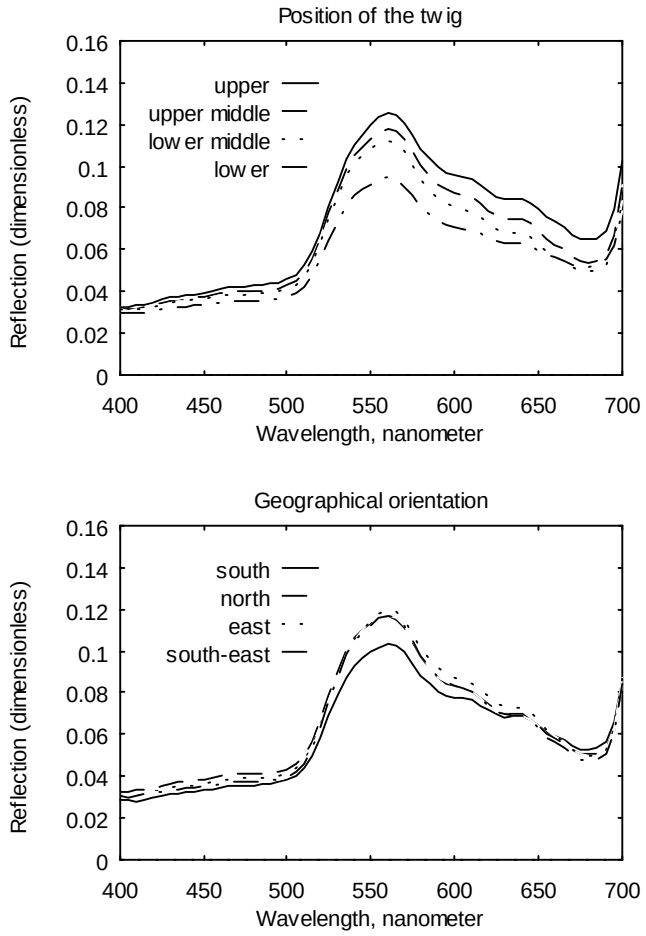


Figure 2-2. Spectral reflectance of 1-year-old spruce shoots. Three characteristics of the 1-year shoots were chosen to examine the spatial variations of foliage spectral properties, age of needles on the 1-year shoot; position within the tree crown (top, two middle, and bottom) and geographical orientation with respect to the tree stem (south, north, east and west).

2.4.6 Ground Reflectance and Anisotropy

To parameterize the contribution of the surface underneath the canopy (soil or/and understory) to the canopy radiation regime, an effective ground reflectance is introduced, namely,

$$\rho_{q,\mathrm{eff}}(r_\mathrm{b},\lambda)=\frac{\int_{2\pi-}\int_{2\pi+}R_{\mathrm{b},\lambda}(\Omega',\Omega)|\mu\mu'|L_\lambda(r_\mathrm{b},\Omega')d\Omega d\Omega'}{\pi\int_{2\pi^-}q(\Omega')|\mu'|L_\lambda(r_\mathrm{b},\Omega')d\Omega'},\quad r_\mathrm{b}\in\delta V_\mathrm{b},\quad \Omega\bullet\mathrm{n}_\mathrm{b}<0\,.$$

Here L_λ is the solution of the boundary value problem for the transport equation. The function q is a configurable function used to better account for features of biomes [*Knyazikhin et al*., 1998a], and it satisfies the following condition:

$$\int_{2\pi^-} q(\Omega')\, d\Omega' = 1 .$$

The effective ground reflectance depends on the canopy structure and the incident radiation field. It follows from the definition that the variation of $\rho_{q,\mathrm{eff}}$ satisfies the following inequality:

$$\min_{\Omega' \in 2\pi-} \frac{\int_{2\pi+} R_{\mathrm{b},\lambda}(\Omega', \Omega)|\mu| d\Omega}{\pi q(\Omega')} \le \rho_{q,\mathrm{eff}}(r_\mathrm{b}, \lambda) \le \max_{\Omega' \in 2\pi-} \frac{\int_{2\pi+} R_{\mathrm{b},\lambda}(\Omega', \Omega)|\mu| d\Omega}{\pi q(\Omega')};$$

that is, the range of variation depends on the integrated bidirectional reflectance factor of the ground surface only. For each biome type, the bidirectional reflectance factor of the ground surface $R_{\mathrm{b},\lambda}$ and the effective ground reflectance are assumed to be horizontally homogeneous; that is, they do not depend on the spatial point r_b. Effective ground reflectances at the MISR spectral bands are elements of the canopy realization $p \in P$. Various patterns of the spectral ground reflectance evaluated from the soil reflectance model of *Jacquemoud et al*. [1992] are included in the present version of the CART file.

To account for the anisotropy of the ground surface, we introduce an effective ground anisotropy S_q,

$$S_q(r_\mathrm{b}, \Omega) = \frac{1}{\rho_{q,\mathrm{eff}}(\lambda)} \frac{\int_{2\pi-} R_{\mathrm{b},\lambda}(\Omega', \Omega)|\mu'| L_\lambda(r_\mathrm{b}, \Omega') d\Omega'}{\pi \int_{2\pi-} q(\Omega')|\mu'| L_\lambda(r_\mathrm{b}, \Omega') d\Omega'}, \quad r_\mathrm{b} \in \delta V_\mathrm{b}, \quad \Omega \bullet \mathrm{n}_\mathrm{b} < 0 .$$

The effective ground anisotropy S_q depends on the canopy structure as well as the incoming radiation field. We note the following property:

$$\int_{2\pi^+} S_q(r_\mathrm{b}, \Omega)|\mu| d\Omega = 1;$$

that is, the above integral depends neither on spatial nor on spectral variables. For each biome type, the effective ground anisotropy is assumed to be wavelength independent. The ground anisotropy is used to precompute some solutions of the transport equation and thus is not stored in the CART file.

2.5 Mathematical Basis of the Algorithm

The aim of this section is to parameterize the contribution of soil/understory reflectances to the exitant radiation field. We closely follow ideas used in atmospheric physics [*Kondratyev*, 1969; *Liou*, 1980]. It follows from the linearity of equation (13) that its solution can be represented as the sum

$$L_\lambda(r,\Omega) = L_{\mathrm{bs},\lambda}(r,\Omega) + L_{\mathrm{rest},\lambda}(r,\Omega) \ . \qquad (17)$$

Here $L_{bs,\lambda}$ is the solution of the "black-soil problem" which satisfies equation (13) with boundary conditions expressed by equations (14), (15), and

$L_{bs,\lambda}(r_b,\Omega) = 0, \quad r_b \in \delta V_b, \quad \Omega \bullet n_b < 0$.

The function $L_{rest,\lambda}$ also satisfies equation (13) with F_λ=0 and boundary conditions expressed as

$L_{rest,\lambda}(r_t,\Omega) = 0, \quad r_t \in \delta V_t, \quad \Omega \bullet n_t < 0$,

$$L_{rest,\lambda}(r_l,\Omega) = \frac{1}{\pi} \int_{\Omega' \bullet n_l > 0} R_{l,\lambda}(\Omega',\Omega) L_{rest,\lambda}(r_l,\Omega') |\Omega' \bullet n_l| d\Omega', \quad r_l \in \delta V_l, \; \Omega \bullet n_l < 0 \quad , \tag{18}$$

$$L_{rest,\lambda}(r_b,\Omega) = \frac{1}{\pi} \int_{\Omega' \bullet n_b > 0} R_{b,\lambda}(\Omega',\Omega) L_\lambda(r_b,\Omega') |\Omega' \bullet n_b| d\Omega', \qquad r_b \in \delta V_b, \Omega \bullet n_b < 0 . \tag{19}$$

Note that $L_{rest,\lambda}$ depends on the solution of the "complete transport problem." The boundary condition (19) can be rewritten as

$$L_{rest,\lambda}(r_b,\Omega) = \rho_{q,eff}(\lambda) S_q(r_b,\Omega) T_{q,\lambda} \quad , \tag{20}$$

where $\rho_{q,eff}$, and S_q are defined by (2) and (5), respectively, and

$$T_{q,\lambda}(r_b) = \int_{2\pi^-} q(\Omega') L_\lambda(r_b,\Omega') |\mu'| d\Omega' \quad . \tag{21}$$

The function q is defined by (3). The coefficient $\rho_{q,eff}$ is assumed to be independent of the point r_b. It is taken as the parameter describing the reflectance of the surface underneath the canopy and can vary continuously within a biome-dependent interval (section 2.3). The biome-dependent function S_q is assumed to be wavelength independent and known (section 2.3). We replace $T_{q,\lambda}$ in (20) by its mean value over the ground surface. This implies that the variable $T_{q,\lambda}$ is independent on the space point r_b (this is automatically fulfilled if a one-dimensional radiative transfer model is used to evaluate the radiative field in plant canopies). Taking into account equation (20), we then can rewrite the solution of the transport problem, equation (17), as

$$L_\lambda(r,\Omega) = L_{bs,\lambda}(r,\Omega) + \rho_{q,eff}(\lambda) \cdot T_{q,\lambda} L_{q,\lambda}(r,\Omega) \quad , \tag{22}$$

where $L_{q,\lambda}(r,\Omega)$ satisfies equation (13) with F_λ=0, boundary condition expressed by equation (18), and

$$L_{q,\lambda}(r_t,\Omega) = 0, \qquad r_t \in \delta V_t, \quad \Omega \bullet n_t < 0 \quad , \tag{23}$$

$$L_{q,\lambda}(r_b,\Omega) = S_q(r_b,\Omega), \quad r_b \in \delta V_b, \quad \Omega \bullet n_b < 0 \quad . \tag{24}$$

Thus $L_{q,\lambda}(r,\Omega)$ describes the radiation regime in a plant canopy generated by anisotropic and heterogeneous sources $S(r_b,\Omega)$ located at the canopy bottom. We term the problem of finding $L_{q,\lambda}(r,\Omega)$ an "S problem." Substituting (22) in (21), we get

$$T_{q,\lambda}(r_b) = T^q_{bs,\lambda}(r_b) + \rho_{q,eff}(\lambda) T_{q,\lambda} \mathbf{r}_{q,\lambda}(r_b) \quad , \tag{25}$$

where

$$T^{q}_{\mathrm{bs},\lambda}(r_{\mathrm{b}}) = \int_{2\pi^-} q(\Omega')L_{\mathrm{bs},\lambda}(r_{\mathrm{b}},\Omega')|\mu'|d\Omega' \quad ,$$

$$\mathbf{r}_{q,\lambda}(r_{\mathrm{b}}) = \int_{2\pi^-} q(\Omega')L_{q,\lambda}(r_{\mathrm{b}},\Omega')|\mu'|d\Omega' \quad .$$

We then average equation (25) over the ground surface. This allows us to express $T_{q,\lambda}$ via $T^{q}_{\mathrm{bs},\lambda}$, $\mathbf{r}_{q,\lambda}$, and $\rho_{q,\mathrm{eff}}$. Substituting the averaged $T_{q,\lambda}$ into equation (22), we get

$$L_{\lambda}(r,\Omega) \approx L_{\mathrm{bs},\lambda}(r,\Omega) + \frac{\rho_{q,\mathrm{eff}}(\lambda)}{1-\rho_{q,\mathrm{eff}}(\lambda)\mathbf{r}_{q,\lambda}} T^{q}_{\mathrm{bs},\lambda} L_{q,\lambda}(r,\Omega). \tag{26}$$

Here $T^{q}_{\mathrm{bs},\lambda}$ and $\mathbf{r}_{q,\lambda}$ are averages over the canopy bottom. Note that we can replace the approximate equality in equation (26) by an exact equality if a one-dimensional canopy radiation model is used to evaluate the radiative regime. It follows from equation (26) that the BHR, $A^{\mathrm{hem}}_{\lambda}$, HDRF, r_{λ}, and the fraction of radiation absorbed by the vegetation, $\mathbf{a}^{\mathrm{hem}}_{\lambda}$, at wavelength λ can be expressed as

$$A^{\mathrm{hem}}_{\lambda}(\Omega_0) \approx \mathbf{r}^{\mathrm{hem}}_{\mathrm{bs},\lambda}(\Omega_0) + \mathbf{t}_{q,\lambda} \frac{\rho_{q,\mathrm{eff}}(\lambda)}{1-\rho_{q,\mathrm{eff}}(\lambda)\mathbf{r}_{q,\lambda}} \mathbf{t}^{\mathrm{hem},q}_{\mathrm{bs},\lambda}(\Omega_0) \quad , \tag{27}$$

$$r_{\lambda}(\Omega,\Omega_0) \approx r_{\mathrm{bs},\lambda}(\Omega,\Omega_0) + \tau_{q,\lambda}(\Omega) \frac{\pi \cdot \rho_{q,\mathrm{eff}}(\lambda)}{1-\rho_{q,\mathrm{eff}}(\lambda)\mathbf{r}_{q,\lambda}} \mathbf{t}^{\mathrm{hem},q}_{\mathrm{bs},\lambda}(\Omega_0) \quad , \tag{28}$$

$$\mathbf{a}^{\mathrm{hem}}_{\lambda}(\Omega_0) \approx \mathbf{a}^{\mathrm{hem}}_{\mathrm{bs},\lambda}(\Omega_0) + \mathbf{a}_{q,\lambda} \frac{\rho_{q,\mathrm{eff}}(\lambda)}{1-\rho_{q,\mathrm{eff}}(\lambda)\mathbf{r}_{q,\lambda}} \mathbf{t}^{\mathrm{hem},q}_{\mathrm{bs},\lambda}(\Omega_0), \tag{29}$$

where $\mathbf{r}^{\mathrm{hem}}_{\mathrm{bs},\lambda}(\Omega_0)$, $\mathbf{a}^{\mathrm{hem}}_{\mathrm{bs},\lambda}$, and $r_{\mathrm{bs},\lambda}$ are the BHR, HDRF, and the fraction of radiation absorbed by the vegetation, respectively, when the canopy ground reflectance is zero. Here

$$\mathbf{t}^{\mathrm{hem},q}_{\mathrm{bs},\lambda}(\Omega_0) = \frac{T^{q}_{\mathrm{bs},\lambda}}{\int_{2\pi^-} |\mu'| L_{\lambda}(r_{\mathrm{t}},\Omega')d\Omega'}$$

is the weighted canopy transmittance,

$$\mathbf{t}_{q,\lambda} = \int_{2\pi^+} |\mu'| L_{q,\lambda}(r_{\mathrm{t}},\Omega')d\Omega'$$

is the transmittance resulting from the anisotropic source S_q located at the canopy bottom, and

$$\tau_{q,\lambda}(\Omega) = L_{q,\lambda}(r_{\mathrm{t}},\Omega)$$

is the radiance generated by S_q which leaves the top of the plant canopy, and $\mathbf{a}_{q,\lambda}$ is the radiance generated by S_q and absorbed by the vegetation. The radiation reflected, transmitted, and absorbed by the vegetation must be related via the energy conservation law,

$$\mathbf{r}_{\mathrm{bs},\lambda}^{\mathrm{hem}} + k_{q,\lambda}(\Omega_0)\cdot\mathbf{t}_{\mathrm{bs},\lambda}^{\mathrm{hem},q} + \mathbf{a}_{\mathrm{bs},\lambda}^{\mathrm{hem}} = 1, \quad k_{q,\lambda}(\Omega_0) = \frac{\mathbf{t}_{\mathrm{bs},\lambda}^{\mathrm{hem},q\equiv 1}(\Omega_0)}{\mathbf{t}_{\mathrm{bs},\lambda}^{\mathrm{hem},q}(\Omega_0)} \quad , \tag{30}$$

$$\mathbf{r}_{q,\lambda} + \mathbf{t}_{q,\lambda} + \mathbf{a}_{q,\lambda} = 1 \quad . \tag{31}$$

Note that all the variables in equations (27) and (28) are mean values averaged over the top surface of the canopy.

It follows from equation (27) that

$$A_{\lambda}^{\mathrm{hem}}(\Omega_0) - \mathbf{r}_{\mathrm{bs},\lambda}^{\mathrm{hem}}(\Omega_0) \approx \mathbf{t}_{q,\lambda}\,\frac{\rho_{q,\mathrm{eff}}(\lambda)}{1-\rho_{q,\mathrm{eff}}(\lambda)\mathbf{r}_{q,\lambda}}\,\mathbf{t}_{\mathrm{bs},\lambda}^{\mathrm{hem},q}(\Omega_0)\,. \tag{32}$$

Thus the contribution of the ground to the canopy-leaving radiance is proportional to the square of canopy transmittance and that the factor of proportionality depends on $\rho_{q,\mathrm{eff}}$. If the right-hand side is sufficiently small, we can neglect this contribution by assigning a value of zero to the effective soil reflectance.

Thus we have parameterized the solution of the transport problem in terms of $\rho_{q,\mathrm{eff}}$ and solutions of the "black-soil problem" and "*S* problem." The solution of the "black-soil problem" depends on Sun-view geometry, canopy architecture, and spectral properties of the leaves. The "*S* problem" depends on spectral properties of the leaves and canopy structure only. At this stage, these properties allow a significant reduction in the size of the LUT because there is no need to store the dependence of the exiting radiation field on ground reflection properties. Since the solution of the "black-soil problem" and "*S* problem" determine the size of the LUT, we focus on the solution of these problems, using equation (26) as the basis of the algorithm. The next step is to specify the wavelength dependence of the basic algorithm equation.

2.6 Conservative Models

Let us consider equation (11) with boundary conditions expressed by equations (14)-(16). This boundary value problem can be reduced to the solution of the "black-soil problem" and "*S* problem." In the LAI/FPAR retrieval algorithm the boundary conditions (15) for the lateral surface of domain *V* are replaced by vacuum condition, i.e., $L_{\lambda}(r_{\mathrm{l}},\Omega)=0$ if $r_{\mathrm{l}}\in\delta V_{\mathrm{l}}$ and $\Omega\bullet n_{\mathrm{l}}<0$ [*Diner et al.*, 1998b; *Knyazikhin* et al., 1998b]. The boundary condition of the "*S* problem" expressed by equations (18), (23), and (24) are wavelength independent in this case. The incoming radiation (14) can be parameterized in terms of two scalar values: $f_{\mathrm{dir},\lambda}$ and total flux $F_{0,\lambda}$ of incoming radiation. It allows representing the "black-soil problem" as a sum of two radiation fields. The first is generated by the monodirectional component of solar radiation incident on the top surface of the canopy boundary and, the second, by the diffuse component. Dividing the transport equations and boundary conditions which define these problems by $f_{\mathrm{dir},\lambda}\cdot F_{0,\lambda}$ and $(1\text{-}f_{\mathrm{dir},\lambda})F_{0,\lambda}$, one can reduce them to transport problems with wavelength-independent boundary conditions. Thus the spectral variation of the radiative field in vegetation canopies can be described, when the spectral variation of the solution of the transport equation

with wavelength-independent boundary conditions is known. Therefore we consider the following boundary value problem for the transport equation

$$\Omega \bullet \nabla \varphi_\lambda(r,\Omega) + \sigma(r,\Omega)\varphi_\lambda(r,\Omega) = \int_{4\pi} \sigma_{s,\lambda}(r,\Omega' \to \Omega)\varphi_\lambda(r,\Omega')d\Omega' \quad , \tag{33}$$

$$\varphi_\lambda(r,\Omega) = B(r,\Omega), \qquad r \in \delta V, \, n_r \bullet \Omega < 0 \, . \tag{34}$$

Here B is a wavelength-independent function defined on the canopy boundary δV, and n_r is the outward normal at the point $r \in \delta V$. Differentiating equations (33) and (34) with respect to wavelength λ, we get

$$\Omega \bullet \nabla u_\lambda(r,\Omega) + \sigma(r,\Omega)u_\lambda(r,\Omega) = \frac{d}{d\lambda} \int_{4\pi} \sigma_{s,\lambda}(r,\Omega' \to \Omega)\varphi_\lambda(r,\Omega')d\Omega' \quad , \tag{35}$$

$$u_\lambda(r,\Omega) = 0, \qquad r \in \delta V, \, n_r \bullet \Omega < 0 \quad , \tag{36}$$

where

$$u_\lambda(r,\Omega) = \frac{d\,\varphi(r,\Omega)}{d\,\lambda} \quad .$$

The following results from eigenvector theory are required to derive a relationship between spectral leaf albedo and canopy absorptance, transmittance, and reflectance.

An eigenvalue of the transport equation is a number γ such that there exists a function φ which satisfies

$$\gamma\left[\Omega \bullet \nabla \varphi(r,\Omega) + \sigma(r,\Omega)\varphi(r,\Omega)\right] = \int_{4\pi} \sigma_{s,\lambda}(r,\Omega' \to \Omega)\varphi(r,\Omega')d\Omega' \quad , \tag{37}$$

with boundary conditions

$$\varphi(r,\Omega)=0, \qquad r \in \delta V = \delta V_t + \delta V_b + \delta V_l, \quad n_r \bullet \Omega < 0 \quad .$$

The function $\varphi(r,\Omega)$ is termed an eigenvector corresponding to the given eigenvalue γ.

The set of eigenvalues γ_k, k=0,1,2, … and eigenvectors $\varphi_k(r,\Omega)$, k=0,1,2, … of the transport equation is a discrete set [*Vladimirov*, 1963]. The eigenvectors are mutually orthogonal; that is,

$$\int_V \int_{4\pi} \sigma(r,\Omega)\varphi_k(r,\Omega)\varphi_l(r,\Omega)d\Omega\, dr = \delta_{k,l} \tag{38}$$

where $\delta_{k,l}$ is the Kroneker symbol. The transport equation has a unique positive eigenvalue which corresponds to a unique positive (normalized in the sense of equation (38)) eigenvector [*Germogenova*, 1986]. This eigenvalue is greater than the absolute magnitudes of the remaining eigenvalues. This means that only one eigenvector, say φ_0, takes on positive values for any $r \in V$ and Ω. This positive couplet of eigenvector and eigenvalue plays an important role in transport theory, for instance, in neutron transport theory. This positive eigenvalue alone determines if the fissile assembly will function as a reactor, or as an explosive, or will melt. Its value successfully relates the reactor geometry to the absorption capacity of the active zone. Because the reactor is

controlled by changing the absorption capacity of the active zone (by inserting or removing absorbents), this value is critical to its functioning. The similarity to the problem at hand is that we need to relate canopy architecture ("similar" to reactor geometry) with leaf optical properties ("similar" to the absorption capacity of the active zone). The expansion of the solution of the transport equation in eigenvectors has mainly a theoretical value because the problem of finding these vectors is much more complicated than finding the solution of the transport equation. However, this approach can be useful if we want to estimate some integrals of the solution. Therefore we apply this technique to derive a relationship between spectral leaf albedo and canopy absorptance, transmittance, and reflectance.

Equation (35) with boundary conditions (36) is a linear homogeneous differential equation with respect to λ in a functional space [*Krein*, 1972]. Its solution φ can be expanded in eigenvectors,

$$\varphi_\lambda(r,\Omega) = a_0(\lambda)\varphi_0(\lambda,r,\Omega) + \sum_{k=1}^{\infty} a_k(\lambda)\varphi_k(\lambda,r,\Omega) \quad , \tag{39}$$

where coefficients a_k do not depend on spatial or angular variables. Here we separate the positive eigenvector φ_0 into the first summand. As described above, only this summand, $a_0\varphi_0$, takes on positive values for any $r \in V$ and Ω. Substituting (39) into equation (35), we get

$$\sum_{k=0}^{\infty}\left[\Omega \bullet \nabla u_k(\lambda,r,\Omega) + \sigma(r,\Omega)u_k(\lambda,r,\Omega)\right] = \sum_{k=0}^{\infty}\frac{d}{d\lambda}\int_{4\pi}\sigma_{s,\lambda}(r,\Omega' \to \Omega)a_k\varphi_k(\lambda,r,\Omega')d\Omega' , \tag{40}$$

where $u_k = d(a_k\varphi_k)/d\lambda$. Substituting (37) into (40), further results in

$$\sum_{k=0}^{\infty}\left[\Omega \bullet \nabla + \sigma(r,\Omega)\right]\left\{\left[1-\gamma_k(\lambda)\right]u_k(\lambda,r,\Omega) - a_k(\lambda)\varphi_k(\lambda,r,\Omega)\frac{d\gamma_k(\lambda)}{d\lambda}\right\} = 0 \quad .$$

Here $\gamma_k(\lambda)$ is the eigenvalue corresponding to the eigenvector φ_k. It follows from this equation, as well as from the orthogonality of eigenvectors, that

$$\frac{d\left[a_k(\lambda)\varphi_k(\lambda,r,\Omega)\right]}{d\lambda} = \frac{\frac{d\gamma_k(\lambda)}{d\lambda}}{1-\gamma_k(\lambda)}\left[a_k(\lambda)\varphi_k(\lambda,r,\Omega)\right].$$

Solving this ordinary differential equation results in

$$a_k(\lambda)\varphi_k(\lambda,r,\Omega) = \frac{1-\gamma_k(\lambda_0)}{1-\gamma_k(\lambda)}\left[a_k(\lambda_0)\varphi_k(\lambda_0,r,\Omega)\right]. \tag{41}$$

Thus if we know the kth summand of the expansion in equation (39) at a wavelength λ_0, we can easily find this summand for any other wavelength.

We introduce e, the monochromatic radiation at wavelength λ intercepted by the vegetation canopy,

$$\mathbf{e}(\lambda) = \int_V dr \int_{4\pi} d\Omega\sigma(r,\Omega)\varphi_\lambda(r,\Omega) , \tag{42}$$

and e_0 as

$$\mathbf{e}_0(\lambda) = \int_V \int_{4\pi} \sigma(r,\Omega)\varphi_\lambda(r,\Omega)\cdot\varphi_0(\lambda,r,\Omega)drd\Omega\,. \tag{43}$$

Given e, we can evaluate the fraction *a* of radiation absorbed by the vegetation at the wavelength λ as

$$a(\lambda) = [1-\omega(\lambda)]e(\lambda), \tag{44}$$

where

$$\omega(\lambda) = \frac{1}{\pi}\frac{\int_{4\pi}\Gamma_\lambda(r,\Omega'\to\Omega)\,d\Omega}{G(r,\Omega')} \tag{45}$$

is the leaf albedo. Below an estimation of e_0 will be performed. This value is close to e. We skip a precise mathematical proof of this fact here. An intuitive explanation is as follows: Putting (39) in (42) and integrating the series results in only the positive term containing $a_0\varphi_0$. As a result, $e(\lambda)/e(\lambda_0)\approx e_0(\lambda)/e_0(\lambda_0)$. Let us derive the dependence of e on wavelength. Substituting equation (39) into equation (43) and taking into account equation (41) as well as the orthogonality of eigenvectors, equation (38), we obtain

$$\mathbf{e}_0(\lambda) = \frac{1-\gamma_0(\lambda_0)}{1-\gamma_0(\lambda)}\mathbf{e}_0(\lambda_0),$$

where γ_0 is the positive eigenvalue corresponding to the positive eigenvector φ_0. Taking into account equation (44), we can also derive the following estimation for a:

$$\mathbf{a}(\lambda) = \frac{1-\gamma_0(\lambda_0)}{1-\gamma_0(\lambda)}\cdot\frac{1-\omega(\lambda)}{1-\omega(\lambda_0)}\mathbf{a}(\lambda_0)\,. \tag{46}$$

Thus given canopy absorptance at wavelength λ_0, we can evaluate this variable at any other wavelength. Figure 2-3 shows spectral variation of the fraction of energy absorbed by the vegetation canopy a for uniform and planophile leaves. Equation (46) can also be used to specify the accuracy of a canopy radiation model to simulate the radiative field in the canopy. On can see (Figure 2-3, right) that our radiation model is erroneous in the case of planophile leaves when LAI>5 and the leaf albedo ω>0.5. At a given wavelength, a is a function of canopy structure and Sun position in the case of "black-soil problem," and a function of canopy structure only in the case of the "*S* problem." We store a at a fixed wavelength λ_0 in the LUT.

A somewhat more complicated technique is realized to derive an approximation for canopy transmittance,

$$\mathbf{t}\left(\lambda,\frac{r_{D,\lambda}}{\omega(\lambda)}\right) = \frac{1-\gamma_0(\lambda_0)}{1-\gamma_0(\lambda)}\mathbf{t}\left(\lambda_0,\frac{r_{D,\lambda}}{\omega(\lambda)}\right), \tag{47}$$

where $r_{D,\lambda}$ is the spectral reflectance of the leaf element. The ratio $r_{D,\lambda}/\omega(\lambda)$ is assumed to be constant with respect to wavelength for each biome. Thus given the canopy transmittance at wavelength λ_0, we can evaluate this variable for wavelength λ. Figure 2-4 shows spectral

variation of canopy transmittance for uniform leaves evaluated with our canopy radiation model and with equation (47). At a fixed wavelength, t is a function of canopy structure and Sun position in the case of the "black-soil problem," and a function of canopy structure in the case of the "*S* problem." We store t at a fixed wavelength λ_0 in the LUT.

The canopy reflectance r is related to the absorptance and transmittance via the energy conservation principle

$$r(\lambda) = 1 - t(\lambda) - a(\lambda) \quad . \tag{48}$$

Thus given canopy transmittance and absorptance at a fixed wavelength, we can obtain the canopy reflectance for any wavelength. Figure 2-5 demonstrates an example of equation (48).

The unique positive eigenvalue γ_0, corresponding to the unique positive eigenvector, can be estimated as [*Knyazikhin and Marshak,* 1991]

$$\gamma_0(\lambda) = \omega(\lambda)[1 - \exp(-K)] \quad , \tag{49}$$

where K is a coefficient which may depend on canopy structure (i.e., biome type, LAI, ground cover, etc.) and Sun position but not on wavelength or soil type. Its specification depends on the parameter (absorptance or transmittance) and type of transport problem ("black-soil problem" or "*S* problem"). The coefficient K, however, does not depend on the transport problem and sun position, when it refers to canopy absorptance. Figure 2-6 shows the coefficient K for the "*S* problem" and canopy absorptance as a function of LAI. This coefficient is an element of the LUT. Note that the eigenvalue γ_0 depends on values of spectral leaf albedo (45) which, in turn, depends on wavelength. It allows us to parameterize canopy absorptance, transmittance, and reflectance in terms of canopy structure, Sun position and leaf albedo.

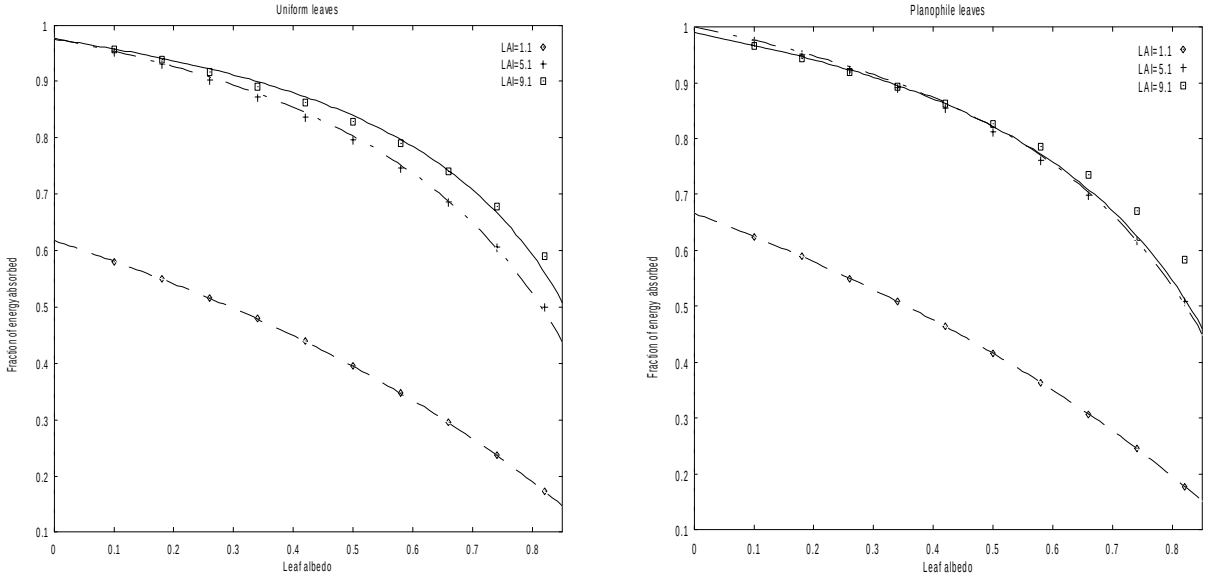


Figure 2-3. Spectral variation of fraction of absorbed radiation by vegetation for uniform (left) and planophile (right) leaves evaluated with canopy radiation model (points) and from equation (46).

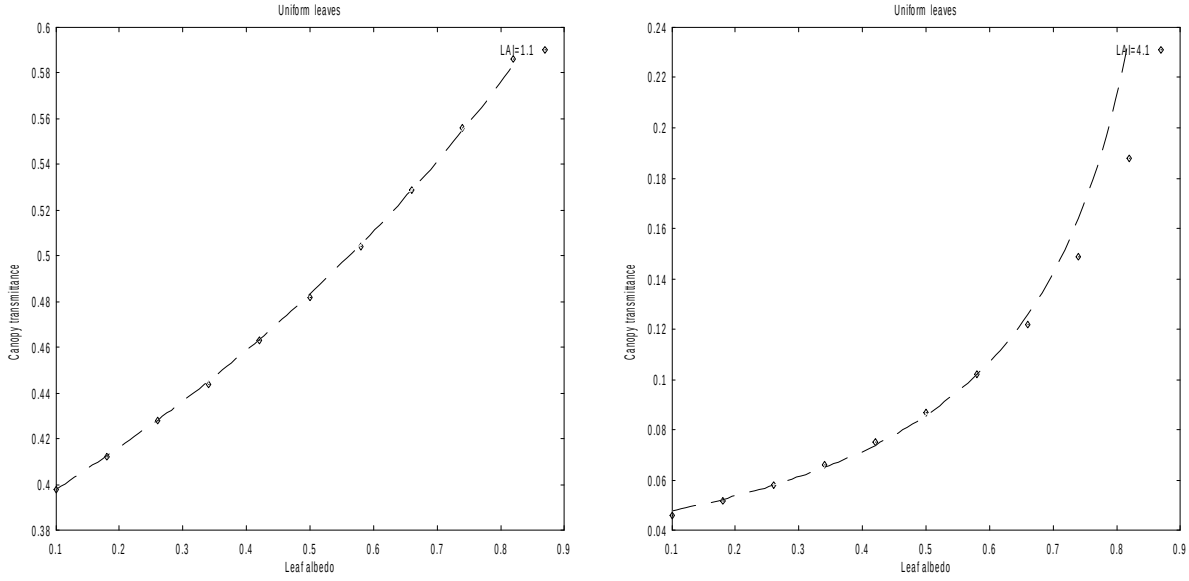


Figure 2-4. Spectral variation canopy transmittance for uniform leaves evaluated with canopy radiation model (points) and from equation (47) for LAI=1.1 (left) and 4.1 (right).

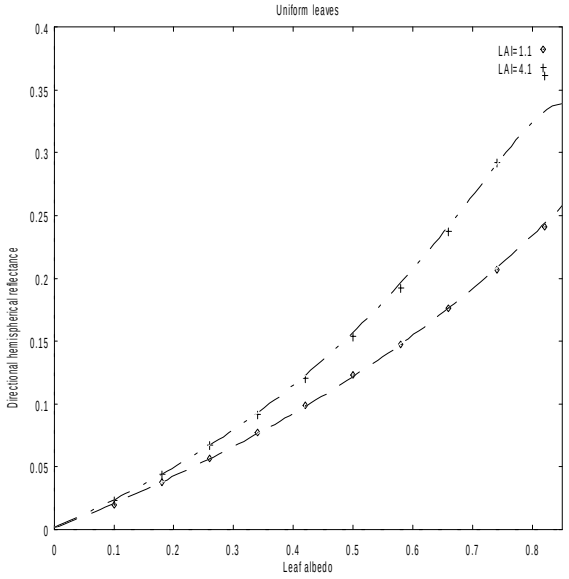


Figure 2-5. Spectral variation of the DHR for uniform leaves evaluated with canopy radiation model (points) and from equation (48) for LAI=1.1 and 4.1.

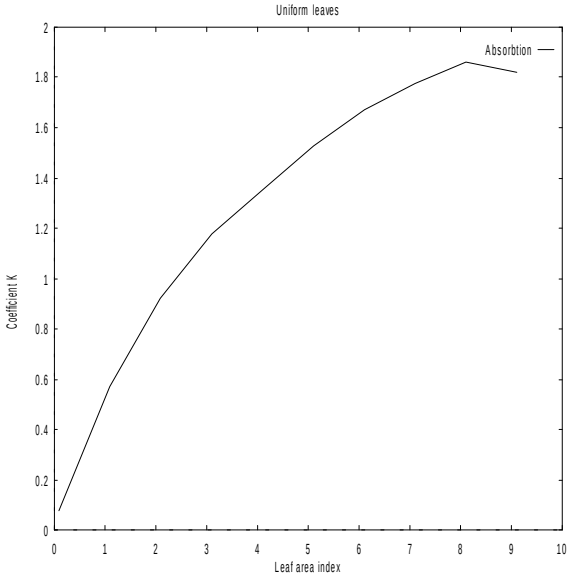


Figure 2-6. Coefficient *K* as a function of LAI for canopy absorptance.

2.7 Constraints on Look-Up Table Entries

In spite of the diversity of canopy reflectance models, their direct use in an inversion algorithm is ineffective. In the case of forests, for example, the interaction of photons with the rough and rather thin surface of tree crowns and with the ground in between the crowns are the most important factors causing the observed variation in the directional reflectance distribution. These phenomena are rarely captured by many canopy reflectance models. As a result, these models are only slightly sensitive to the within-canopy radiation regime. This assertion is based on the fact that a rather wide family of canopy radiation models are solutions to (13), including models with a nonphysical internal source F_λ (Appendix). Within such a model the sum of radiation absorbed, transmitted, and reflected by the canopy are not equal to the radiation incident on the canopy. The function F_λ is chosen such that the model simulates the reflected radiation field correctly, i.e., these models account for photon interactions within a rather small domain of the vegetation canopy. On the other hand, it is the within-canopy radiation regime that is very sensitive to the canopy structure and therefore to LAI. The within-canopy radiation regime also determines the amount of solar energy absorbed by the vegetation. Ignoring this phenomenology in canopy radiation models leads to a large number of nonphysical solutions when one inverts a canopy reflectance model. Therefore (27) and (28) must be transformed before they can be used in a retrieval algorithm.

Let us introduce the required weights

$$w_{\mathrm{bs},\lambda}(\Omega,\Omega_0)=\frac{\pi^{-1}r_{\mathrm{bs},\lambda}(\Omega,\Omega_0)}{\mathbf{r}^{\mathrm{hem}}_{\mathrm{bs},\lambda}(\Omega_0)},\qquad \int_{2\pi^+} w_{\mathrm{bs},\lambda}(\Omega,\Omega_0)|\mu|\,d\Omega=1\ , \tag{50}$$

$$w^q_\lambda(\Omega)=\frac{\tau_{q,\lambda}(\Omega)}{\mathbf{t}_{q,\lambda}},\qquad \int_{2\pi^+} w^q_\lambda(\Omega)|\mu|\,d\Omega=1\ . \tag{51}$$

With this notation, (28) can be rewritten as

$$r_\lambda(\Omega,\Omega_0)\approx\pi w_{\mathrm{bs},\lambda}\cdot\mathbf{r}^{\mathrm{hem}}_{\mathrm{bs},\lambda}(\Omega_0)+\pi w^q_\lambda\mathbf{t}_{q,\lambda}\frac{\rho_{q,\mathrm{eff}}(\lambda)}{1-\rho_{q,\mathrm{eff}}(\lambda)\mathbf{r}_{q,\lambda}}\mathbf{t}^{\mathrm{hem},q}_{\mathrm{bs},\lambda}(\Omega_0) \tag{52}$$

and from (30) and (31), the canopy reflectances $\mathbf{r}^{\mathrm{hem}}_{\mathrm{bs},\lambda}$ and $\mathbf{r}_{q,\lambda}$ can be written as

$$\mathbf{r}^{\mathrm{hem}}_{\mathrm{bs},\lambda}=1-\mathbf{t}^{\mathrm{hem},q\equiv1}_{\mathrm{bs},\lambda}-\mathbf{a}^{\mathrm{hem}}_{\mathrm{bs},\lambda}\ , \tag{53}$$

$$\mathbf{r}_{q,\lambda}=1-\mathbf{t}_{q,\lambda}-\mathbf{a}_{q,\lambda}\ . \tag{54}$$

Thus (52) is sensitive to both factors determining the directional reflectance distribution of plant canopies (the weight $w_{\mathrm{bs},\lambda}$) and to the within-canopy radiation regime $[\mathbf{t}^{\mathrm{hem},q\equiv1}_{\mathrm{bs},\lambda},\mathbf{a}^{\mathrm{hem}}_{\mathrm{bs},\lambda},\mathbf{t}_{q,\lambda},\mathbf{a}_{q,\lambda}]$. Equations (52)-(54) also allow the formulation of a test for the "eligibility" of a canopy radiation model to generate the LUT. First the weights $w_{\mathrm{bs},\lambda}$ are

evaluated as a function of Sun-view geometry, wavelength, and LAI by using a field-tested canopy reflectance model. Then with the same model, $\mathbf{r}_{\mathrm{bs},\lambda}^{\mathrm{hem}}$ and $\mathbf{r}_{q,\lambda}$ are evaluated from (53) and (54), and inserted into (52). A canopy radiation model is "eligible" to generate the LUT file if (50) and (51) are satisfied to within a given accuracy for any Sun-view combination, wavelength, and LAI. We do not know of a canopy reflectance model which can pass the above test. That is because there is no published model thus far which satisfies the energy conservation law. Although a conservative transport equation for a vegetation canopy has not yet been formulated, solutions of this equation satisfy properties derived in a previous section. These properties can be used to correct existing canopy radiation models for the "eligibility" to generate the LUT. An algorithm to correct a canopy radiation model is presented by *Knyazikhin et al*. [1998b] which was used to generate the LUT for the MISR LAI/FPAR retrieval algorithm.

It follows from (32) and (52) that the HDRF can be represented as

$$r_\lambda(\Omega,\Omega_0) \approx \pi w_{\mathrm{bs},\lambda} \cdot \mathbf{r}_{\mathrm{bs},\lambda}^{\mathrm{hem}}(\Omega_0) + w_\lambda^q [A_\lambda^{\mathrm{hem}}(\Omega_0) - \mathbf{r}_{\mathrm{bs},\lambda}^{\mathrm{hem}}(\Omega_0)] \quad . \tag{55}$$

For each pixel the MISR instrument provides the spectral BHR and DHR. Therefore this expression is used to evaluate the HDRF and BRF in the case of MISR data, setting retrieved A_λ^{hem} in (55). Equation (28) is used to evaluate the BRF in the case of MODIS data.

Thus the BHR described by (27) and the HDRF described by (55) can be expressed in terms of optical properties of a leaf and the energy conservation law, as well as in terms of solutions of the "black-soil problem" and "*S* problem" at a reference leaf albedo value of $\omega(\lambda_0)$. This facilitates comparison of spectral values of the BHR or HDRF with spectral properties of individual leaves, which is a rather stable characteristic of a green leaf. It also can be interpreted as "inclusion of additional information" into the algorithm, thus allowing a significant reduction in the number of retrieved solutions. Canopy transmittances and absorptances, and coefficients p=1-exp(-*K*) where *K* is defined by (49) for the "black-soil problem" and "*S* problem" at a reference leaf albedo value of ω^* as well as the weights (50) and (51) are precomputed and stored in the LUT. It allows the use of the same LUT for MODIS and MISR instruments. A detailed description of such a LUT is presented by *Diner et al.* [1998b].

2.8 Description of LAI Retrieval

For each pixel the MODIS instrument can provide atmosphere-corrected BRF in one view direction and at seven bands in the solar spectrum every day [*Vermote et al*., 1995]. Thus every 8 days, one has the set of pixel reflectances corresponding to 16 different Sun positions, 15 view angles, and at 11 spectral bands. These canopy reflectances and Sun-view geometry are input for the algorithm. Note that this is the maximum amount of information which may be available. In reality, however, it may be less, e.g., because of cloud cover and performance of preprocessing algorithms. Let $r_{0,\lambda}(\Omega',\Omega'_0)$ be the BRF retrieved from MODIS data and $r_\lambda(\Omega,\Omega_0)$ and $A_\lambda^{\mathrm{hem}}(\Omega_0)$ be the BRF and BHR retrieved from MISR data. Here Ω' and Ω are the view MODIS and MISR directions, Ω'_0 and Ω_0 are the direction of direct solar radiation during times of MODIS and MISR observations, and β, λ denote the center of the MODIS and MISR spectral bands,

respectively. These retrieved reflectances are the input for the algorithm which we express in the vector-matrix form as

$$\bar{r}_0 = \begin{pmatrix} r_{0,\beta_1}(\Omega'_1,\Omega'_{0,1}) & r_{0,\beta_1}(\Omega'_2,\Omega'_{0,2}) & \cdots & r_{0,\beta_7}(\Omega'_8,\Omega'_{0,8}) \\ r_{0,\beta_2}(\Omega'_1,\Omega_{0,1}) & r_{0,\beta_2}(\Omega'_2,\Omega'_{0,2}) & \cdots & r_{0,\beta_2}(\Omega'_8,\Omega'_{0,8}) \\ \vdots & \vdots & \vdots & \vdots \\ r_{0,\beta_7}(\Omega'_1,\Omega'_{0,1}) & r_{0,\beta_7}(\Omega'_2,\Omega'_{0,2}) & \cdots & r_{0,\beta_7}(\Omega'_8,\Omega'_{0,8}) \end{pmatrix},$$

$$\bar{r}(\Omega_0) = \begin{pmatrix} r_{\lambda 1}(\Omega_1,\Omega_0) & r_{\lambda 2}(\Omega_1,\Omega_0) & r_{\lambda 3}(\Omega_1,\Omega_0) & r_{\lambda 4}(\Omega_1,\Omega_0) \\ r_{\lambda 1}(\Omega_2,\Omega_0) & r_{\lambda 2}(\Omega_2,\Omega_0) & r_{\lambda 3}(\Omega_2,\Omega_0) & r_{\lambda 4}(\Omega_2,\Omega_0) \\ \vdots & \vdots & \vdots & \vdots \\ r_{\lambda 1}(\Omega_9,\Omega_0) & r_{\lambda 2}(\Omega_9,\Omega_0) & r_{\lambda 3}(\Omega_9,\Omega_0) & r_{\lambda 4}(\Omega_9,\Omega_0) \end{pmatrix}, \tag{56}$$

$$\vec{A}^{hem}(\Omega_0) = \begin{bmatrix} A^{hem}_{\lambda 1}(\Omega_0) & A^{hem}_{\lambda 2}(\Omega_0) & A^{hem}_{\lambda 3}(\Omega_0) & A^{hem}_{\lambda 4}(\Omega_0) \end{bmatrix}. \tag{57}$$

Here β_k, k=1, 2, … ,7 and λ_m, m=1,2,3,4 are centers of the MODIS and MISR spectral bands. We will use $r_{0,\lambda}(\Omega,\Omega_0)$, $r_\lambda(\Omega,\Omega_0)$, $A^{\text{hem}}_\lambda(\Omega_0)$, $\bar{r}_0$, $\bar{r}(\Omega_0)$, and $\vec{A}^{\text{hem}}(\Omega_0)$ to denote modeled canopy reflectances (i.e., evaluated from equation (52) for MODIS and equations (55) and (27) for MISR instruments) and $\tilde{r}_{0,\lambda}(\Omega,\Omega_0)$, $\tilde{r}_\lambda(\Omega,\Omega_0)$, $\tilde{A}^{\text{hem}}_\lambda(\Omega_0)$, $\tilde{\bar{r}}_0$, $\tilde{\bar{r}}(\Omega_0)$, and $\tilde{\vec{A}}^{\text{hem}}(\Omega_0)$ to denote observations of these variables.

The modeled canopy reflectances depend on the model parameters. In our algorithm we use a vegetation land cover classification parameterized in terms of variables used in photon transport theory [*Myneni et al.*, 1997]. It distinguishes six biome types, each representing a pattern of the architecture of an individual tree (leaf normal orientation; stem-trunk-branch area fractions; leaf and crown size) and the entire canopy (trunk distribution, topography), as well as patterns of spectral reflectance and transmittance of vegetation elements. The soil and/or understory type are also characteristics of the biome which can vary continuously within given biome-dependent ranges. The distribution of leaves is described by the three-dimensional leaf area density distribution function which can also depend on some continuous parameters (section 2.6). Therefore LAI may not be in the list p of model parameters directly. However, LAI can be obtained when model parameter values in the parameter list p are known; that is, LAI is a function of p: LAI=$l(p)$. The function l is assumed known. Thus the model parameter list p contains one discrete variable (biome type) which can take on six values only, continuous

variables (the soil and/or understory type), and some continuous parameters determining the leaf area density distribution function. A detailed description of canopy parameterization is presented by *Knyazikhin et al.* [1998b]. The model parameters are said to be a canopy realization if values of model variables in the parameter list are specified. We denote by P a set of all possible canopy realizations and will use p to denote a canopy realization. The set P is the sum of six subsets, $P = \bigcup_{\text{bio}=1}^{6} P_{\text{bio}}$, each representing a biome specific set of canopy realizations. Let $D_r \subset R^{4\times 9}$ and $D_A \subset R^4$ be the space of all possible values of canopy reflectances obtained by running p over the set P; that is,

$$D_A = \left\{ \vec{A}^{\text{hem}}(\Omega_0, p) : p \in P \right\}, \qquad D_r = \left\{ \bar{r}(\Omega_0, p) : p \in P \right\}. \tag{58}$$

Here we proceed with the suggestion that the sets D_A and D_r represent all possible observations of canopy reflectances; that is, any $\tilde{\vec{A}}^{\text{hem}}(\Omega_0)$ and $\tilde{\bar{r}}(\Omega_0)$ are elements of D_A and D_r, respectively. It should be noted, however, that this suggestion may be violated in real situations.

In reality any model can simulate a process to within a certain degree of accuracy only. Also, measurements cannot be carried out ideally. It means that the models predict domains $O_A \subset D_A$ and $O_r \subset D_r$ around $\vec{A}^{\text{hem}}(\Omega_0)$ and $\bar{r}(\Omega_0)$ to which the "true values" belong. The same is valid for measured quantities; that is, we can only point out neighborhoods O_A and O_r around $\tilde{\vec{A}}^{\text{hem}}(\Omega_0)$ and $\tilde{\bar{r}}(\Omega_0)$ to which the "true values" belong. The domains O_A and O_r are uncertainties in measurements and simulations: any element from these domains can be true values. We define neighborhoods O_A and O_r about measured reflectances $\tilde{\vec{A}}_{\lambda}^{\text{hem}}(\Omega_0)$ and $\tilde{\bar{r}}(\Omega_0)$ as [*Diner et al.*, 1998a]

$$O_A = \left\{ \vec{A}^{\text{hem}}(\Omega_0) \in D_A : \ \Delta_A\left[\vec{A}^{\text{hem}}(\Omega_0), \tilde{\vec{A}}_{\lambda}^{\text{hem}}(\Omega_0) \right] \le h_A \right\}, \tag{59}$$

$$O_r = \left\{ \bar{r}(\Omega_0) \in D_r : \Delta_r[\bar{r}(\Omega_0), \tilde{\bar{r}}(\Omega_0)] \le h_r \right\},$$

where

$$\Delta_A\left[\vec{A}^{\text{hem}}(\Omega_0),\tilde{\vec{A}}^{\text{hem}}(\Omega_0)\right]=\frac{\sum_{l=1}^{4}\nu_A(l)\left[\dfrac{A_{\lambda_l}^{\text{hem}}(\Omega_0)-\tilde{A}_{\lambda_l}^{\text{hem}}(\Omega_0)}{\sigma_A(l)}\right]^2}{\sum_{l=1}^{4}\nu_A(l)}, \tag{60}$$

$$\Delta_r\left[\bar{r}(\Omega_0),\tilde{\bar{r}}(\Omega_0)\right]=\frac{\sum_{l=1}^{4}\sum_{j=1}^{9}\nu_r(l,j)\cdot\left[\dfrac{r_{\lambda_l}(\Omega_j,\Omega_0)-\tilde{r}_{\lambda_l}(\Omega_j,\Omega_0)}{\sigma_r(l,j)}\right]^2}{\sum_{i=1}^{4}\sum_{j=1}^{9}\nu_r(l,j)}.$$

Here $\nu_A(l)$=1 if the BHR (or DHR) at wavelength λ_l exists and 0 otherwise; $\nu_r(l,j)$ takes on the value 1 if the HDRF (or BRF) at wavelength λ_l and in scattering direction Ω_j exists and 0 otherwise; σ_A and σ_r are uncertainties in the BHR (or DHR) and HDRF (or BRF) retrievals and h_r and h_A some configurable threshold values [*Diner et al*, 1998a]. Thus modeled quantities are defined to belong to a neighborhood around the measured values such that a model which differs from the retrieved BHR (or DHR) and HDRF (or BRF) values by an amount equivalent to or less than the retrieval uncertainty will result in values Δ_A and Δ_r of the order of unity.

Any canopy realization $p\in P$ for which $\bar{r}(\Omega_0)\in O_r$ and $\vec{A}^{\text{hem}}(\Omega_0)\in O_A$ must be considered a candidate for a true p. Let us introduce sets of candidates for the solution as follows:

$$Q_A(L,O_A;P_{\text{bio}})=\left\{\,p\in P_{\text{bio}}:l(p)<L \text{ and } \vec{A}^{\text{hem}}(\Omega_0)\in O_A\,\right\},$$

$$Q_r(L,O_r;P_{\text{bio}})=\left\{\,p\in P_{\text{bio}}:l(p)<L \text{ and } \bar{r}(\Omega_0)\in O_r\,\right\}.$$

These sets are subsets of P_{bio} and contain such p from P_{bio} for which the leaf area index LAI=$l(p)$ is less than a given value L from the interval [$\text{LAI}_{\min}$(bio), $\text{LAI}_{\max}$(bio)] and $\bar{r}(\Omega_0)\in O_r$, $\vec{A}^{\text{hem}}(\Omega_0)\in O_A$. Here

$$\text{LAI}_{\min}(\text{bio})=\inf\left\{\,l(p):p\in P_{\text{bio}}\,\right\},$$

$$\text{LAI}_{\max}(\text{bio})=\sup\left\{\,l(p):p\in P_{\text{bio}}\,\right\}.$$

The sets $Q_A(\text{LAI}_{\max}, O_A; P_{\text{bio}})$ and $Q_r(\text{LAI}_{\max}, O_r; P_{\text{bio}})$ contain all $p\in P_{\text{bio}}$ for which a canopy radiation model generates output comparable with measured data. In terms of these notations we formulate the inverse problem as follows: given atmospherically corrected canopy reflectances $\tilde{\vec{A}}^{\text{hem}}(\Omega_0)$, $\tilde{\bar{r}}(\Omega_0)$ and their uncertainties O_A, O_r find all $p\in P_{\text{bio}}$ for which

$$\vec{A}^{\text{hem}}(\Omega_0) \in O_A, \tag{61}$$

$$\bar{r}(\Omega_0) \in O_r. \tag{62}$$

The algorithm is designed to utilize all the available information of the observations by means of a two-step process. The first step involves a comparison of the retrieved spectral hemispherically integrated reflectances with those evaluated from the model, i.e., solution of (61). Only those p which satisfy this test are subject to the second step, which is a comparison of their directional reflectances at the MISR angles to the retrieved spectral directional reflectances, i.e., the solution of (62).

In order to quantify solutions of (61) and (62) we introduce measures (distribution functions) defined on the set P_{bio} as follows: The subset P_{bio} is represented as a sum of nonintersected subsets

$$P_{\text{bio}} = \bigcup_{k=1}^{N} P_{\text{bio},k}, \quad P_{\text{bio},k} \bigcap P_{\text{bio},j} = \varnothing, \quad k \neq j. \tag{63}$$

Let $N_A(L;P_{\text{bio}})$ and $N_{r,A}(L;P_{\text{bio}})$ be numbers of subsets $P_{\text{bio},k}$ containing at least one element from the set $Q_A(L,O_A;P_{\text{bio}})$ and $Q_{A,r}(L)=Q_A(L,O_A;P_{\text{bio}})\cap Q_r(L,O_r;P_{\text{bio}})$, respectively. As measures of $Q_A(L,O_A;P_{\text{bio}})$ and $Q_{r,A}(L)$, we introduce biome-specific functions $F_{A,\text{bio}}(L)$ and $F_{r,A,\text{bio}}(L)$ as

$$F_{A,\text{bio}}(L) = \lim_{N\to\infty} \frac{N_A(L;P_{\text{bio}})}{N_A(\text{LAI}_{\max};P_{\text{bio}})}, \tag{64}$$

$$F_{r,A,\text{bio}}(L) = \lim_{N\to\infty} \frac{N_{r,A}(L;P_{\text{bio}})}{N_{r,A}(\text{LAI}_{\max};P_{\text{bio}})}. \tag{65}$$

A mathematical description of the convergence process is presented in the Appendix. Intuitively, the subset $P_{\text{bio},k}$ specifies a set of canopy realizations whose range of variation is "sufficiently small." $N_A(\text{LAI}_{\max};P_{\text{bio}})$ and $N_{r,A}(\text{LAI}_{\max};P_{\text{bio}})$ are total number of solutions of (61) and (62); $N_A(L;P_{\text{bio}})$ and $N_{r,A}(L;P_{\text{bio}})$ are the number of these solutions when $\text{LAI}=l(p)$ is less then a given value L in the interval $[\text{LAI}_{\min}, \text{LAI}_{\max}]$. The functions (64) and (65) are the LAI conditional distribution functions provided $p\in P_{\text{bio}}$, and validity of (61) and (62), respectively. Note that the functions (64) and (65) depend on L and neighborhoods O_A and O_r. The values

$$L_{A,\text{bio}} = \int_{\text{LAI}_{\min}}^{\text{LAI}_{\max}} l\, dF_{A,\text{bio}}(l), \tag{66}$$

$$L_{r,A,\text{bio}} = \int_{\text{LAI}_{\min}}^{\text{LAI}_{\max}} l\, dF_{r,A,\text{bio}}(l) \tag{67}$$

are taken as solutions of (61) and (62), and the values

$$d^2_{A,\mathrm{bio}} = \int_{\mathrm{LAI}_{\min}}^{\mathrm{LAI}_{\max}} \left(L_{A,\mathrm{bio}} - l\right)^2 \, dF_{A,\mathrm{bio}}(l)\,, \tag{68}$$

$$d^2_{r,A,\mathrm{bio}} = \int_{\mathrm{LAI}_{\min}}^{\mathrm{LAI}_{\max}} \left(L_{r,A,\mathrm{bio}} - l\right)^2 \, dF_{r,A,\mathrm{bio}}(l) \tag{69}$$

are taken as the characteristics of the solution accuracy. If (61) and (62) have no solutions (i.e., F=0), we assign a default value to (68) and (69). We propose to archive (66), (67), (68), and (69) for all six biomes for diagnostic purposes.

We note some properties of the functions $F_{A,\mathrm{bio}}$ and $F_{r,A,\mathrm{bio}}$ which help to explain the definition of the solution. It follows directly from definitions (64) and (65) that if the function $l(p)$ is constant, say $l(p)\equiv L^*$, when $p\in Q_A(\mathrm{LAI}_{\max},O_A;P_{\mathrm{bio}})$ and $p\in Q_r(\mathrm{LAI}_{\max},O_{A,r};P_{\mathrm{bio}})$, then $L_{r,\mathrm{bio}}$ and $L_{r,A,\mathrm{bio}}$ coincide with L^*. This allows the use of three-dimensional canopy radiation models for which LAI is usually not in the model parameter list. In this case, canopy realizations can vary considerably, while LAI remains unchanged. This property shows that (66) and (67) are sensitive to LAI but not to the situations generating the value of LAI. It follows from this that if the inverse problem has a unique solution for given set of measurements, then (66) and (67) coincide with this solution. If model parameters from $Q_A(\mathrm{LAI}_{\max},O_A;P_{\mathrm{bio}})$ and $Q_r(\mathrm{LAI}_{\max},O_{A,r};P_{\mathrm{bio}})$ can generate several values of LAI, (66) and (67) provide a weighted mean in accordance with the frequency of occurrence of a given value of LAI. The accuracy of a solution cannot be improved if no additional information is available. These properties provide convergence of the algorithm; that is, the more the measured information is available and the more accurate this information is, the more reliable and accurate the algorithm output will be.

Figure 2-7 ~ figure 2-10 illustrate various aspects of the function (64) and retrieval results for biome 1 (grasses and cereal crops) for 40 different neighborhoods O_A. This biome type is represented by five parameters in the algorithm [*Diner et al.*, 1998b], which include the "effective" ground reflectances $\rho_i=\rho_{q,\mathrm{eff}}(\lambda_i)$ (section 2.4) in the MISR bands λ_i, i=1, 2, 3, 4, and LAI; that is, $p=(\rho_1, \rho_2, \rho_3, \rho_4, \mathrm{LAI})$. The effective ground reflection was a linear function with respect to wavelength in this example; that is, $\rho_{q,\mathrm{eff}}(\lambda_i)=s(\lambda_1-\lambda)+\rho_1$. The slope s, effective ground reflectance in the first MISR band ρ_1 and LAI, can vary continuously within given biome-dependent intervals $[\rho_{\min}, \rho_{\max}]$, $[s_{\min}, s_{\max}]$ and $[\mathrm{LAI}_{\min}, \mathrm{LAI}_{\max}]$. Thus the set P_{bio}, bio=1, is defined in our example as

$$P_{\mathrm{bio}} = \left\{ (\rho, s, \mathrm{LAI}) : \rho_{\min} \le \rho \le \rho_{\max}, \quad s_{\min} \le s \le s_{\max}, \quad \mathrm{LAI}_{\min} \le \mathrm{LAI} \le \mathrm{LAI}_{\max} \right\}.$$

The function l takes the form $l(p)$=LAI in this case. We choose 40 elements,

$$p_k=(0.025,\ 1.184\times10^{-4},\ \mathrm{LAI}_k),\ \mathrm{LAI}_k=0.1+(k-1)\times0.25,\ k=1, 2, \ldots, 40, \tag{70}$$

from the set P_{bio}. For each soil/LAI pattern p_k, we estimated DHR for the four MISR bands, which were taken as $\tilde{\tilde{A}}(\Omega_0)$. The uncertainties in (60) were simulated as

$$\sigma_A^2(l) = \varepsilon^2 \frac{\sum_{k=1}^{4} \left[A_{\lambda_k}^{\text{hem}}(\Omega_0)\right]^2}{4}, \; l=1, 2, 3, 4, \qquad (71)$$

where ε is a variable in the calculations; that is, these uncertainties can be interpreted as the mean retrieved uncertainty. If the values $A_{\lambda}^{\text{hem}}(\Omega_0)$, $\lambda = \lambda_1, \lambda_2, \lambda_3, \lambda_4$ are approximately of the same order, the value of the merit function (60) using uncertainties of the individual $A_{\lambda_k}^{\text{hem}}(\Omega_0)$ is close to that when using (71). In the general case, the merit function using (71) describes the closeness between measured and simulated values worse than the ones using the individual uncertainties, i.e., neighborhoods O_A determined by the merit function (60) with (71) are broader than the ones accounting for individual uncertainties. One object of our study was to analyze the behavior of the parameter distribution function in situations worse than what may be realized. The parameters used in this study are as follows: the polar angle of the unit direction Ω_0 was 45^0; $\rho_{\min}$=0.025, $\rho_{\max}$=0.070, $s_{\min}=1.184\times10^{-4}$, $s_{\max}=1.896\times10^{-4}$; $\text{LAI}_{\min}$=0.1; $\text{LAI}_{\max}$=9.85; f_{dir}=1, ε=0.20, and h_A=1. The number N in (63) was 1000, which was large enough to approximate the parameter distribution functions (64) and (65) sufficiently well.

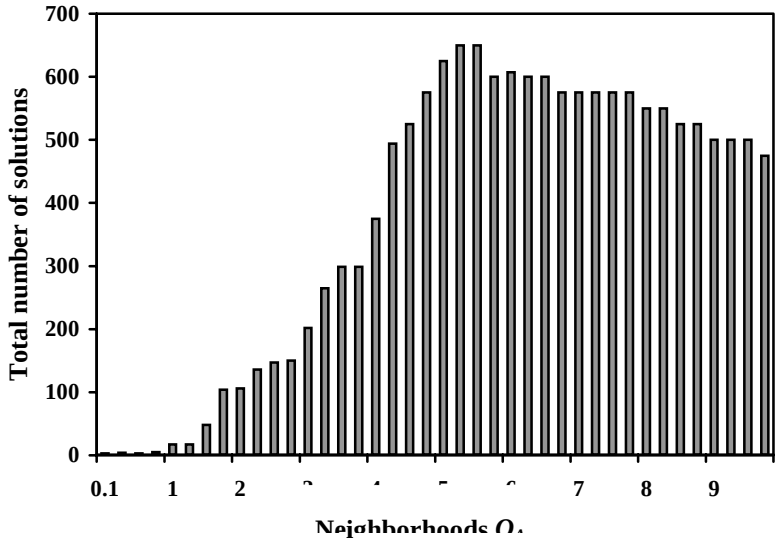


Figure 2-7. Total number of solutions for 40 patterns of neighborhoods O_A. The horizontal axis shows the values $l(p_k)$, where p_k is the soil/leaf area index (LAI) pattern for which O_A was simulated.

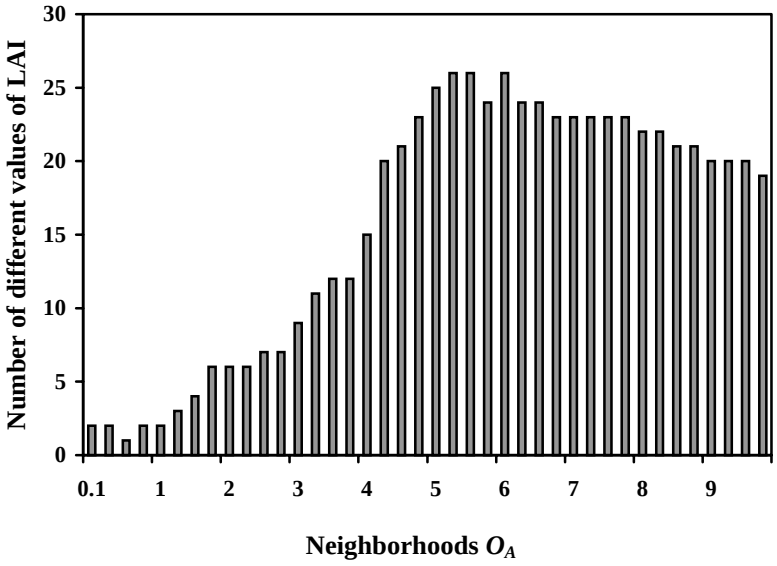


Figure 2-8. Number of different solutions for 40 patterns of neighborhoods O_A.

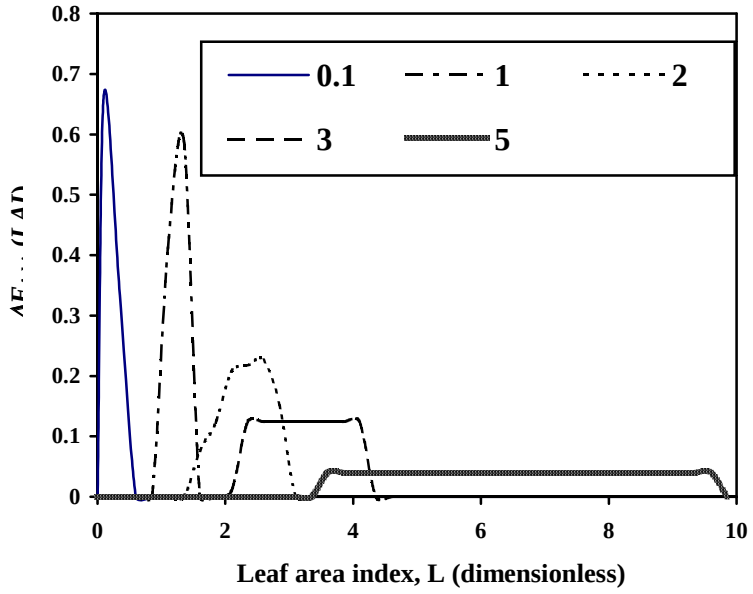


Figure 2-9. Localization of probable values of LAI. The function $\Delta F_{A,\mathrm{bio}}(L) = F_{A,\mathrm{bio}}(L+0.25) - F_{A,\mathrm{bio}}(L)$ shows the probability distribution of multiple acceptable solutions for five different patterns of O_A. Curve 0.1: $l(p_k)=0.1$; curve 1: $l(p_k)=1$; curve 2: $l(p_k)=2$; curve 3: $l(p_k)=3$; curve 5: $l(p_k)=5$.

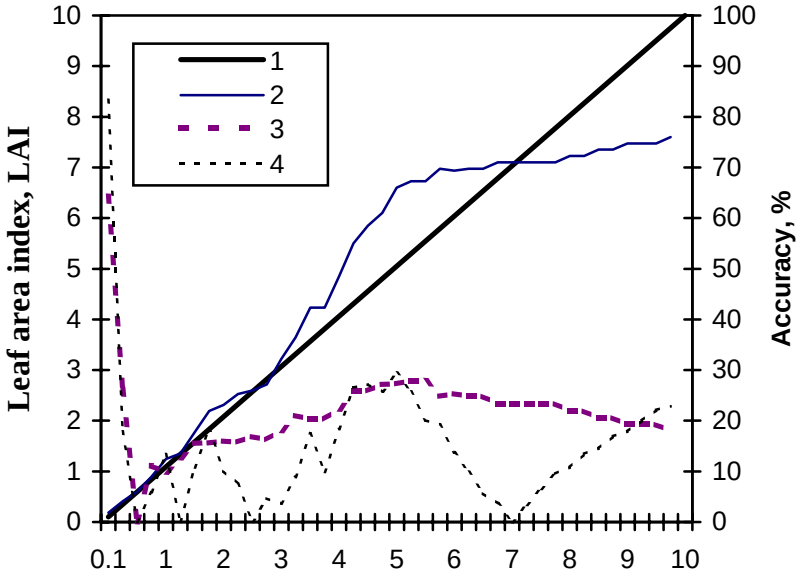


Figure 2-10. Comparison of the retrieved and exact solutions of the inverse problem for 40 patterns of neighborhoods O_A. Vertical axis on the left side: 1, exact solution; 2, retrieved solution. Vertical axis on the right side: 3, 100·|curve_1-curve_2|/curve_1; 4, 100·$d_{A,bio}$/curve_1.

The total number, $N_A(\mathrm{LAI}_{max};P_{bio})$, of solutions of (61) for the 40 patterns of neighborhoods O_A are plotted in Figure 2-7, 2-8. The neighborhoods are sorted with respect to values $l(p_k)$ where p_k is the soil/LAI pattern (70) for which O_A was simulated. Values $l(p_k)$ are shown on the horizontal axis. Figure 2-8 presents the numbers of different values of $l(p)$ when p ran over the set $Q_A(\mathrm{LAI}_{max},O_A;P_{bio})$, i.e., numbers of different values of LAI satisfying (61). Figure 2-9 demonstrates the functions $\Delta F_{A,bio}(L)= F_{A,bio}(L+0.25)-F_{A,bio}(L)$ for five different patterns of O_A. One can discern two types of shapes for $\Delta F_{A,bio}$. The first one localizes values of LAI sufficiently well (curves 0.1, 1, 2, and 3). They correspond to neighborhoods O_A for which $l(p)\le 3$. The curve 5 shows that the set of p for which model results are nearly equivalent to the measurements is rather big. We will quantify this situation in section 2.9. Figure 2-9 contains two plots; the first one with the vertical axis on the left side demonstrates exact (curve 1) and retrieved (curve 2) values of LAI for our patterns of O_A. The meaning of the horizontal axis is the same as in Figure 2-7, 2-8. The second plot with the vertical axis on the right presents values $\delta_{A,bio}=100\cdot d_{A,bio}/L_{A,bio}$ (curve 3), and LAI_{bio} (curve 4) for 40 patterns of O_A. Here $\mathrm{LAI}_{bio}= l(p_k)$ is the value of leaf area index for which O_A was simulated and $L_{A,bio}$ is the value obtained from (66). The value $\delta_{A,bio}$ mainly varied between 11% and 28%, even in cases when the algorithm retrieves LAI accurately (compare curves 3 and 4). The range of variations in $\delta_{A,bio}$, however, is comparable to the uncertainty of O_A (recall that in our example this set

includes elements that differ from a given vector $\tilde{\tilde{A}}^{\text{hem}}(\Omega_0)$, on an average, by 20%). Therefore (68) can be taken as the characteristic of the inversion accuracy. However, this value is slightly sensitive to the two cases when the function (64) localizes LAI values (Figure 2-9, curves 0.1, 1, 2, and 3) and when such localization does not take place (curve 5). Therefore one needs an additional characteristic that distinguishes these two conditions. We must also pay attention to the case when the accuracy of the retrieved LAI exceeds the uncertainty of O_A. The neighborhood with $l(p_k)$=0.1 demonstrates such an example. In this case, N=1000 in (63) was not big enough to adequately represent the set of possible observations. There were vectors $\vec{A} \in R^4$ which are close to the simulated $\vec{A}^{\text{hem}}(\Omega_0)$, $\vec{A} \approx \vec{A}^{\text{hem}}(\Omega_0)$ and which were not elements of D_A.

2.9 Saturation Domain

Calculations presented in section 2.9 indicate that there may be "small" neighborhoods O_A and O_r in D_A and D_r which can be generated by a rather "wide" set of the canopy realizations. Curve 5 in Figure 2-9 illustrates such a condition: any p satisfying $l(p) \geq 3.5$ with equal probability can be a solution of (61). In our study, similar behavior was observed for all patterns O_A corresponding to $l(p_k) \geq 5$. The aim of this section is to quantify these situations.

Let us consider the set $S_{\text{bio}}(L^*, L)$ defined as

$$S_{\text{bio}}(L^*, L) = \left\{ p \in P_{\text{bio}} : \quad L^* \leq l(p) < L; \quad L \leq \text{LAI}_{\max} \right\}. \tag{72}$$

This set does not depend on canopy reflectances. A measured $\tilde{\tilde{A}}^{\text{hem}}(\Omega_0)$ is defined to belong to the saturation domain $D_{S,A} \subseteq D_A$, and a value, $L_A^* \in [\text{LAI}_{\min}, \text{LAI}_{\max}]$, is a saturation point if

$$\left\{ p \in P_{\text{bio}} : \quad L_A^* \leq l(p) < \text{LAI}_{\max} \text{ and } \vec{A}^{\text{hem}}(\Omega_0) \in O_A \right\} = S_{\text{bio}}(L_A^*, \text{LAI}_{\max}). \tag{73}$$

This equality shows that for given O_A, a canopy radiation model is insensitive to the canopy realizations from the set $S_{\text{bio}}(L_A^*, \text{LAI}_{\max})$. All $\tilde{\tilde{A}}^{\text{hem}}(\Omega_0)$ satisfying the condition (73) constitute the saturation domain $D_{S,A}$. Figure 2-9 demonstrates one example of an element from the saturation domain and saturation point: the neighborhood O_A corresponding to p_{21} (see (70)) belongs to the saturation domain, and any value of LAI from 3.6 to 9.85 can be a solution with equal probability. The point L=3.6 is the saturation point. Similarly, a saturation domain, $D_{S,r} \subseteq D_r$ and saturation point, $L_r^* \in [\text{LAI}_{\min}, \text{LAI}_{\max}]$, for the HDRF and BRF can be introduced.

In the algorithm, the leaf area distribution function is parameterized in terms of ground cover g and mean leaf area index L of an individual tree (section 2.4). The ranges

$g_{\min} \leq g \leq g_{\max}$, $L_{\min} \leq L \leq L_{\max}$

of their possible variation depend on the biome type and are assumed to be known [*Knyazikhin et al.*, 1998a]. Thus the function $l(p)$ has the form $l(p)=gL$, and

$$\text{LAI}_{\min} = g_{\min} L_{\min}, \ \text{LAI}_{\max} = g_{\max} L_{\max}. \tag{74}$$

We note that in the cases of biome 1 (grasses and cereal crops), vegetation is idealized as a horizontally homogeneous medium [*Knyazikhin et al*., 1998a]. For this biome, $g_{\min}=g_{\max}=1$. Analogous to (64) and (65), a solution distribution function for the saturation domain can be introduced as

$$\Phi(L^*,L)=\lim_{N\to\infty}\frac{N_S(L^*,L)}{N_S(L^*,\mathrm{LAI}_{\max})},$$

where N is defined by (63) and $N_S(L^*,L)$ is the number of subsets $P_{\mathrm{bio},k}$ containing at least one element from the set (72). Accounting for $l(p)=gL$, we get

$$\Phi(L^*,L)=\frac{\psi(L^*,L)}{\psi(L^*,\mathrm{LAI}_{\max})}, \tag{75}$$

where the function $\psi(L^*,L)$ takes on the value 0 if $L<L^*$, and

$$\psi(L^*,L)=\int\limits_{\substack{L^*\le g\cdot l<L\\ L_{\min}\le l\le L_{\max}\\ g_{\min}\le g\le g_{\max}}} dH_L(l)\,dH_G(g)=\int\limits_{\max\left\{L_{\min},\,L^*/g_{\max}\right\}}^{\min\left\{L/g_0,\,L_{\max}\right\}}\left[\int\limits_{\max\left\{L^*/l,\,g_{\min}\right\}}^{\min\left\{L/l,\,g_{\max}\right\}} dH_G(g)\right]dH_L(l)$$

if $L^*\le L<\mathrm{LAI}_{\max}$, and

$$\psi(L^*,L)\equiv\psi(L^*,\mathrm{LAI}_{\max})=\int\limits_{\max\left\{L_{\min},\,L^*/g_{\max}\right\}}^{L_{\max}}\left[\int\limits_{\max\left\{L^*/l,\,g_{\min}\right\}}^{g_{\max}} dH_G(g)\right]dH_L(l)$$

if $L\ge\mathrm{LAI}_{\max}$. Note that the function (75) is expressed in the form of the Stieltjes integral, where

$$H_G(g)=\begin{cases}0, & \text{if } g\le g_{\min};\\ g, & \text{if } g_{\min}<g\le g_{\max};\\ g_{\max}, & \text{if } g_{\max}<g.\end{cases}\quad H_L(l)=\begin{cases}0, & \text{if } l\le L_{\min};\\ l, & \text{if } L_{\min}<l\le L_{\max};\\ L_{\max}, & \text{if } L_{\max}<l.\end{cases}$$

If $L_{\min}<L_{\max}$ and $g_{\min}<g_{\max}$, then the Stieltjes integral coincides with the classical integral, and $dH_L(l)=dl$, $dH_G(g)=dg$. However, if $L_{\min}=L_{\max}$ and/or $g_{\min}=g_{\max}$, the classical integral gives a value of 0, while the Stieltjes integral provides the correct value. Thus (75) specifies the distribution of LAI in the set (72) in our algorithm. Note that if LAI and model parameters are related in another manner, function (75) may take another suitable form.

If $\tilde{\tilde{A}}^{\mathrm{hem}}(\Omega_0)$ and $\bar{\bar{r}}(\Omega_0)$ belong to the saturation domain, then L_A^* and L_r^* exist, such that

$$F_{A,\mathrm{bio}}(L)=\Phi(L_A^*,L)\ \text{ and }\ F_{r,A,\mathrm{bio}}(L)=\Phi(L_r^*,L)$$

for all L from $[\mathrm{LAI}_{\min},\mathrm{LAI}_{\max}]$. In this case, the solutions (66) and (67) and their variance coefficients (68) and (69) can be expressed as

$$L_{A,\mathrm{bio}}=s_1(L_A^*),\quad d^2_{A,\mathrm{bio}}=s_2(L_A^*)-s_1^2(L_A^*), \tag{76}$$

$$L_{r,A,\text{bio}} = s_1(L_r^*), \quad d^2_{r,A,\text{bio}} = s_2(L_r^*) - s_1^2(L_r^*) \ , \tag{77}$$

where

$$s_k(L^*) = \int_{\text{LAI}_{\min}}^{\text{LAI}_{\max}} l^k \, d\Phi(L^*, l), \ k=1,2.$$

The functions s_1 and s_2 are known and determined by canopy characteristics only and are independent of the measured quantities. The set of points $[s_1(L^*), \sqrt{s_2(L^*) - s_1^2(L^*)}]$ obtained by running L^* over $[\text{LAI}_{\min}, \text{LAI}_{\max}]$ determines a curve which is termed a saturation curve. Figure 2-11 demonstrates saturation curves for six biomes which correspond to canopy parameterization introduced by [*Myneni et al.*, 1997; *Knyazikhin et al.*, 1998a]. These relationships allow us to formulate a necessary condition for the measured reflectances belonging to the saturation domain as follows: If $\tilde{\tilde{A}}^{\text{hem}}(\Omega_0) \in D_{S,A}$ and $\tilde{\tilde{r}}(\Omega_0) \in D_{S,r}$ then the points $[L_{A,\text{bio}}, d_{A,\text{bio}}]$ and $[L_{r,A,\text{bio}}, d_{r,A,\text{bio}}]$ belong to saturation curves, or what amounts to the same thing, solutions of the equations $s_1(L_A)=L_{A,\text{bio}}$, $s_2(L_d) - s_1^2(L_d) = d^2_{A,\text{bio}}$ and $s_1(L_r)=L_{r,A,\text{bio}}$, $s_2(L_{r,d}) - s_1^2(L_{r,d}) = d^2_{r,A,\text{bio}}$ satisfy the equalities $L_A=L_d$ and $L_r=L_{r,d}$. Here the-right hand sides of these equations are evaluated during the execution of the algorithm. The left-hand sides are known functions of one variable.

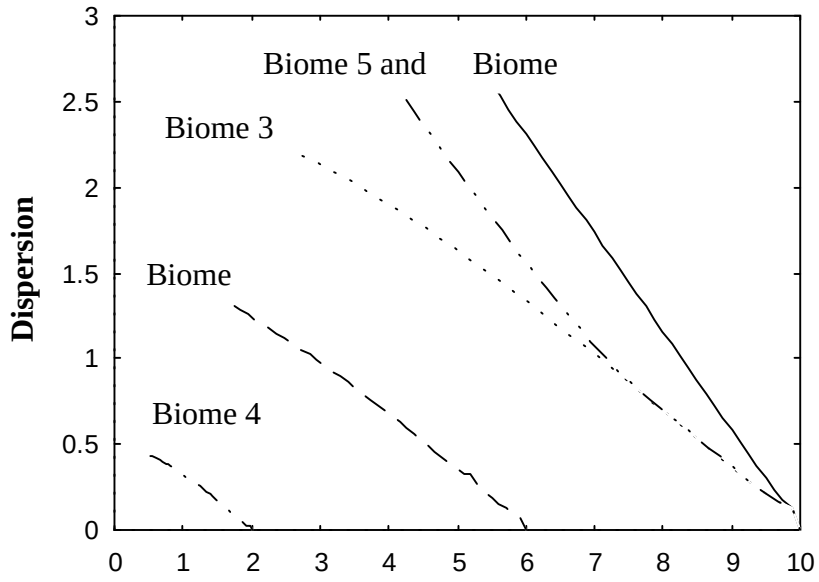


Figure 2-11. Saturation curves for six biomes. If the measured canopy reflectance belongs to the saturation domain, then the point (LAI, *d*) lies on the saturation curve. Here LAI is the retrieved value of LAI and *d* is its dispersion.

This criterion takes a simple form in the case of biome 1. It follows from (75) and $g_{\min}$= $g_{\max}$=1 that the solution distribution function for the saturation domain in these biomes is

$$\Phi(L^*, L) = \begin{cases} 0, & \text{if } L < L^*, \\ \dfrac{L - L^*}{\mathrm{LAI}_{\max} - L^*}, & \text{if } L^* \le L < \mathrm{LAI}_{\max}, \\ 1, & \text{if } \mathrm{LAI}_{\max} \le L. \end{cases}$$

Equation (76) is reduced to

$$L_{A,\mathrm{bio}}(L_A^*) = \frac{L_A^* + \mathrm{LAI}_{\max}}{2}, \qquad d_{A,\mathrm{bio}}^2(L_A^*) = \frac{1}{12}\left(\mathrm{LAI}_{\max} - L_A^*\right)^2$$

and (77) can be simplified to a similar expression, with L_A^* replaced by L_r^*. After obvious transformations, one can express the saturation criterion for biomes 1 as follows: If $\tilde{\tilde{A}}^{\mathrm{hem}}(\Omega_0) \in D_{S,A}$ and $\tilde{\bar{r}}(\Omega_0) \in D_{S,r}$, then

$$L_{A,\mathrm{bio}} + \sqrt{3} \cdot d_{A,\mathrm{bio}} = \mathrm{LAI}_{\max} \text{ and/or } L_{r,A,\mathrm{bio}} + \sqrt{3} \cdot d_{r,A,\mathrm{bio}} = \mathrm{LAI}_{\max} \tag{78}$$

where $L_{A,\mathrm{bio}}$, $d_{A,\mathrm{bio}}$, $L_{r,A,\mathrm{bio}}$, and $d_{r,A,\mathrm{bio}}$ are evaluated from (66), (67), (68), and (69). Thus after the evaluation of LAIs and their variances, condition (78) is checked. We archive -$d_{A,\mathrm{bio}}$ and -$d_{r,A,\mathrm{bio}}$ if (78) is satisfied to a given accuracy. Inclusion of the minus sign means that a solution LAI was found, but the value probably belongs to the saturation domain and any value of LAI from [2·LAI- $\mathrm{LAI}_{\max}$,$\mathrm{LAI}_{\max}$] must be considered as a true solution with equal probability.

Other biome types do not allow for the formulation of the saturation criteria in a such simple form. Therefore we store saturation curves for all biomes in the look-up table. After evaluation of LAIs from (66) and (67) and their variances from (68) and (69), conditions $m(d_{A,\mathrm{bio}},L_{A,\mathrm{bio}})$=0 and $m(d_{r,A,\mathrm{bio}},L_{r,A,\mathrm{bio}})$=0 are checked. Here

$$m(d, L) = \min_{\mathrm{LAI}_{\min} \le L^* \le \mathrm{LAI}_{\max}} \{[s_1(L^*) - L]^2 + [\sqrt{s_2(L^*) - s_1^2(L^*)} - d]^2\} \ . \tag{79}$$

We archive -$d_{A,\mathrm{bio}}$ and -$d_{r,A,\mathrm{bio}}$ if these relationships are fulfilled to a given accuracy.

2.10 Description of FPAR Retrieval

It follows from (29) and (32) that the fractional amount of incident photosynthetically active radiation (PAR) absorbed by the vegetation canopy (FPAR) can be evaluated as

$$\mathrm{FPAR}(bio, p) = \int_{400\mathrm{nm}}^{700\mathrm{nm}} \mathbf{a}_{\lambda}^{\mathrm{hem}}(\Omega_0) e(\lambda)\, d\lambda = Q_{bs}(bio, LAI, \Omega_0) + Q^q(bio, p, \Omega_0) \ , \tag{80}$$

where

$$Q_{\mathrm{bs}}(bio, LAI, \Omega_0) = \int_{400\mathrm{nm}}^{700\mathrm{nm}} \mathbf{a}_{bs,\lambda}^{\mathrm{hem}}(\Omega_0) e(\lambda)\, d\lambda, \tag{81}$$

$$Q^q(bio, p, \Omega_0) = \int_{400\text{nm}}^{700\text{nm}} \mathbf{a}_{q,\lambda}(\Omega_0) \frac{\rho_{q,\text{eff}}(\lambda)}{1-\rho_{q,\text{eff}}(\lambda)\mathbf{r}_{q,\lambda}} \mathbf{t}_{\text{bs},\lambda}^{\text{hem},q} e(\lambda)\, d\lambda \tag{82}$$

$$= \int_{400\text{nm}}^{700\text{nm}} \frac{\mathbf{a}_{q,\lambda}(\Omega_0)}{\mathbf{t}_{q,\lambda}(\Omega_0)} \left[\tilde{A}_\lambda^{\text{hem}}(\Omega_0) - \mathbf{r}_{\text{bs},\lambda}^{\text{hem}}(\Omega_0)\right] e(\lambda)\, d\lambda \quad . \tag{83}$$

The Q_{bs} term describes the absorption within the canopy for a black-soil condition, and Q^q term describes the additional absorption within the canopy due to the interaction between the ground (soil and/or understory) and the canopy. Here $p \in P_{bio}$; e is the ratio of the monochromatic flux incident at the top surface of the canopy boundary to the total downward PAR flux which can be expressed as

$$e(\lambda) = \frac{E_{0,\lambda} e_\lambda^{\text{hem}}(\Omega_0)}{\int_{400\text{nm}}^{700\text{nm}} E_{0,\lambda} e_\lambda^{\text{hem}}(\Omega_0)\, d\lambda} \quad ,$$

where $E_{0,\lambda}$ is the solar irradiance spectrum that is known for all wavelengths; e_λ^{hem} is the normalized incident irradiance defined as the ratio of the radiant incident on the surface to $E_{0,\lambda}$ [*Diner et al*., 1998a]. The mean over those $p \in P_{bio}$ which passed the test $\Delta(p) \le h$ is taken as the estimate of FPAR, i.e.,

$$\text{FPAR}_{bio} = \frac{1}{N_P} \sum_{k=1}^{N_P} \text{FPAR}(bio, p) \quad ,$$

where N_P is the number of canopy realizations $p \in P_{bio}$ passing this test. When there is no solution (i.e., F_{bio}=0), the algorithm defaults to a NDVI-FPAR regression analysis to obtain an estimate of FPAR [*Myneni et al*, 1997b].

The normalized incident irradiance and the BHR are provided by the MISR instrument at three spectral bands within the PAR region. We assume a piece-wise linear variation in these variables over regions [446 , 558 nm], [558, 672 nm], and a constant over regions [400 ,446nm], [672 ,700nm]. Substituting these piece-wise linear functions into (80) and (83), one can express FPAR as a function of e_λ^{hem} and $\tilde{A}_\lambda^{\text{hem}}$ [*Diner et al*., 1998a]. Note that the dependence of FPAR on ground reflection properties is included in $\tilde{A}_\lambda^{\text{hem}}$ which is provided by the MISR instrument;

that is, expression (80) is a function of the biome type, Sun position, ground cover, mean leaf area index of an individual plant, and retrieved BHR.

If only MODIS observations are available for a given pixel or the MODIS-only mode is executed, $e(\lambda)$ is approximated by

$$e(\lambda) = \frac{E_\lambda(5200^0 K)}{\int\limits_{400\text{nm}}^{700\text{nm}} E_\lambda(5200^0 K)} \quad ,$$

where $E_\lambda(T)$ is the Planck function [*Kondratyev*, 1969, p. 230]. In this case, the Q^q term is a function of the biome type, Sun position, ground cover, mean leaf area index of an individual plant, and pattern of the effective ground reflectance. Expression (82) is used to evaluate this term. The Q_{bs} and Q^q terms are precomputed and stored in the look-up table.

2.11 Theoretical Basis of NDVI-FPAR Relations

The measured spectral reflectance data are usually compressed into vegetation indexes. More than a dozen such indexes are reported in the literature and shown to correlate well with vegetation amount [*Tucker*, 1979], the fraction of absorbed photosynthetically active radiation [*Asrar et al.*, 1984], unstressed vegetation conductance and photosynthetic capacity [*Sellers et al.*, 1992], and seasonal atmospheric carbon dioxide variations [*Tucker et al.*, 1986]. There are some theoretical investigations to explain these empirical regularities [*Vygodskaya and Gorshkova*, 1987; *Myneni et al.*, 1995a; *Verstraete and Pinty*, 1996]. Results from the previous section allow us to relate the vegetation indexes to the fundamental physical principle, i.e., the law of energy conservation. Here we consider the normalized difference vegetation index (NDVI) whose use is included in the LAI/FPAR retrieval algorithm.

Let us consider NDVI defined as

$$\text{NDVI} = \frac{A_\alpha^{\text{hem}} - A_\beta^{\text{hem}}}{A_\alpha^{\text{hem}} + A_\beta^{\text{hem}}} \quad , \tag{84}$$

where A_λ^{hem} is the BHR or DHR, and α and β are near-IR and red spectral wavebands, respectively. These variables are a function of Sun position Ω_0, but this dependence has been suppressed in the notation of this section. For the sake of simplicity, we consider the NDVI for the "black-soil" problem and "*S* problem." It follows from equations (54), (53), and (52) that equation (84) can be rewritten as

$$\text{NDVI} = \frac{k(\alpha,\beta)\mathbf{a}(\beta) - m(\alpha,\beta)\mathbf{t}(\beta)}{2\mathbf{r}(\beta) + k(\alpha,\beta)\mathbf{a}(\beta) - m(\alpha,\beta)\mathbf{t}(\beta)} \quad . \tag{85}$$

where

$$k(\alpha,\beta)=1-\frac{1-\gamma_{0,\mathbf{a}}(\beta)}{1-\gamma_{0,\mathbf{a}}(\alpha)}\cdot\frac{1-\omega(\alpha)}{1-\omega(\beta)}\quad,\qquad m(\alpha,\beta)=\frac{1-\gamma_{0,\mathbf{t}}(\beta)}{1-\gamma_{0,\mathbf{t}}(\alpha)}-1\quad.$$

Here $\gamma_{0,\mathbf{a}}$ and $\gamma_{0,\mathbf{t}}$ are defined by equation (55) with $K=K_{\mathbf{a}}$ (for canopy absorptance) and $K=K_{\mathbf{t}}$ (for canopy transmittance), respectively. Here the ratio between the leaf spectral reflectance and the leaf albedo is assumed to be constant with respect to wavelength, and so it is excluded from the argument list of **t**. After simple transformations, one obtains

$$\mathrm{NDVI}=\mathbf{a}(\beta)\cdot\theta\left(s_{\mathbf{t},\beta},s_{\mathbf{r},\beta}\right)\ ,$$

where the function θ has the following form

$$s(x,y)=\frac{k(\alpha,\beta)-m(\alpha,\beta)\cdot x}{2y+k(\alpha,\beta)-m(\alpha,\beta)\cdot x}\quad,$$

$$s_{\mathbf{t},\beta}=\frac{\mathbf{t}(\beta)}{\mathbf{a}(\beta)}\quad,\qquad s_{\mathbf{r},\beta}=\frac{\mathbf{r}(\beta)}{\mathbf{a}(\beta)}\quad.$$

Thus NDVI is proportional to the canopy absorptance at the red band. It follows from Eqs. (53) and (85) that

$$\mathbf{a}(\lambda)=\frac{1-\gamma_{0,\mathbf{a}}(\beta)}{1-\gamma_{0,\mathbf{a}}(\lambda)}\frac{1-\omega(\lambda)}{1-\omega(\beta)}\mathbf{a}(\beta)=\frac{1-\gamma_{0,\mathbf{a}}(\beta)}{1-\gamma_{0,\mathbf{a}}(\lambda)}\frac{1-\omega(\lambda)}{1-\omega(\beta)}\frac{NDVI}{\theta(s_{\mathbf{t},\beta},s_{\mathbf{r},\beta})}\quad.$$

Let $e(\lambda)$ be the ratio of monochromatic radiant energy incident on the top surface of the canopy boundary to the total PAR flux. Integrating $e\cdot\mathbf{a}$ over the PAR region of solar spectrum, we get

$$\mathrm{FPAR}=k\cdot\mathrm{NDVI}\ ,$$

where

$$k=\frac{1}{\theta\left(s_{\mathbf{t},\beta},s_{\mathbf{r},\beta}\right)}\frac{1-\gamma_{0,\mathbf{a}}(\beta)}{1-\omega(\beta)}\left[\int_{400}^{700}\frac{1-\omega(\lambda)}{1-\gamma_{0,\mathbf{a}}(\lambda)}e(\lambda)\,d\lambda\right]\quad.$$

Thus if the canopy ground is ideally black, FPAR is proportional to NDVI. The factor of proportionality k depends on the ratios $s_{\mathbf{t},\beta}$ and $s_{\mathbf{t},\beta}$, the coefficients $K_{\mathbf{a}}$ and $K_{\mathbf{t}}$, and the leaf albedo at the red and near-IR spectral bands. A relationship between NDVI and FPAR which accounts for the soil contribution can be derived from equation (27) in a similar manner. Other types of vegetation indexes can be derived in an analogous way.

2.12 Backup Algorithm

If there are no candidate biome/canopy models which passed the comparison test for a given pixel, a back-up algorithm is triggered to estimate LAI and FPAR using Normalized Difference Vegetation Index (NDVI). This backup algorithm requires a land cover classification that is compatible with the radiative transfer model used in their derivation. Such a classification is expected to be derived from the MODIS Land Cover Product. Backup algorithm uses regression curves to estimate LAI and FPAR values, namely,

$$\mathrm{LAI} = \boldsymbol{f_1}(\mathrm{NDVI}), \quad \mathrm{FPAR} = \boldsymbol{f_2}(\mathrm{NDVI}).$$

Here $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$ are biome dependent regression curves which were derived as described below.

Definition of regression curves. Two curves $x(y)=E(X \mid Y=y)$ and $y(x)=E(Y \mid X=y)$ defined in the (x,y)-*plane* are called the regression curves of X with respect to Y and of Y with respect to X, respectively. Here, X and Y are random variables; $E(X \mid Y=y)$ is expectation of X under the condition that Y has taken the value y. The regression curve has the following interpretation: the best possible prediction of X given a realized value y of Y is $x(y)$. The regression curve $x(y)$ minimizes the expected squared error of the prediction of X on the basis of value of y, and $y(x)$ can be interpreted similarly.

Regression curves were used to build the LAI-NDVI and NDVI-FPAR relationships estimated with the MODIS LAI/FPAR algorithm applied to the SeaWiFS data as follows. The surface reflectance was treated as a biome dependent random variable. The SeaWiFS data were taken as the sets of realizations of this random variable. The MODIS LAI/FPAR algorithm is based on the estimation of the LAI distribution function under the condition that the canopy reflectances at the MODIS spectral bands have taken given values. This allows us to treat values of LAI and FPAR as random variables. The LAI/FPAR algorithm was applied to these data in order to evaluate sets of realizations of LAI and FPAR random values. The sets of NDVI realizations were obtained from the SeaWiFS data directly. Various regression curves were then derived using these sets.

Let X, Y, and Z be random variables of LAI, NDVI and FPAR values. Figure 2-12 demonstrates realizations of these random variables in the (LAI, NDVI) [left panel] and (NDVI, FPAR) [right panel] planes. We denote the regression curves of LAI with respect to NDVI and of NDVI with respect to LAI by $x(y)$ and $y(x)$, respectively. Similarly, $z(y)$ and $y(z)$ denote the best possible prediction of FPAR given a realized value y of NDVI and the best possible prediction of NDVI given a realized value z of FPAR. The regressions curves $x(y)$, $y(x)$, $z(y)$ and $y(z)$ for 6 biomes (Grasses and Cereal Crops; Shrubs, Broadleaf Crops; Savannas; Broadleaf Forests; Needle Forests) are shown in Figure 2-13. The relationships were taken as $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$, respectively. The regression curves $x(y)$, $y(x)$ and $z(y)$, $y(z)$ for biome 5 (Broadleaf Crops) are presented in Figure 2-14. One can see that the curves $y(x)$ and $z(y)$ are close to the inverse function of $x(y)$ and $y(z)$, respectively. This may be taken as evidence of the existence of LAI-NDVI and NDVI-FPAR relationships that have previously been reported in the literature. Table 2-2 shows NDVI and corresponding LAI and FPAR values used by the current version of the algorithm. This table may be updated when MODIS data are available.

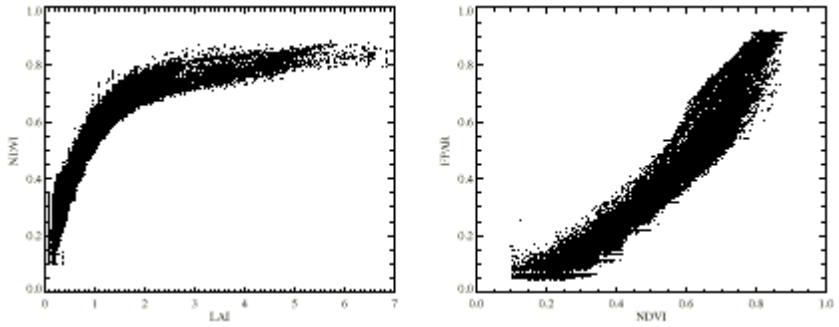


Figure 2-12. Realizations of LAI (random variable *X*), NDVI (random variable *Y*), and FPAR (random variable *Z*) in the LAI–NDVI [left panel] and NDVI–LAI [right panel] planes.

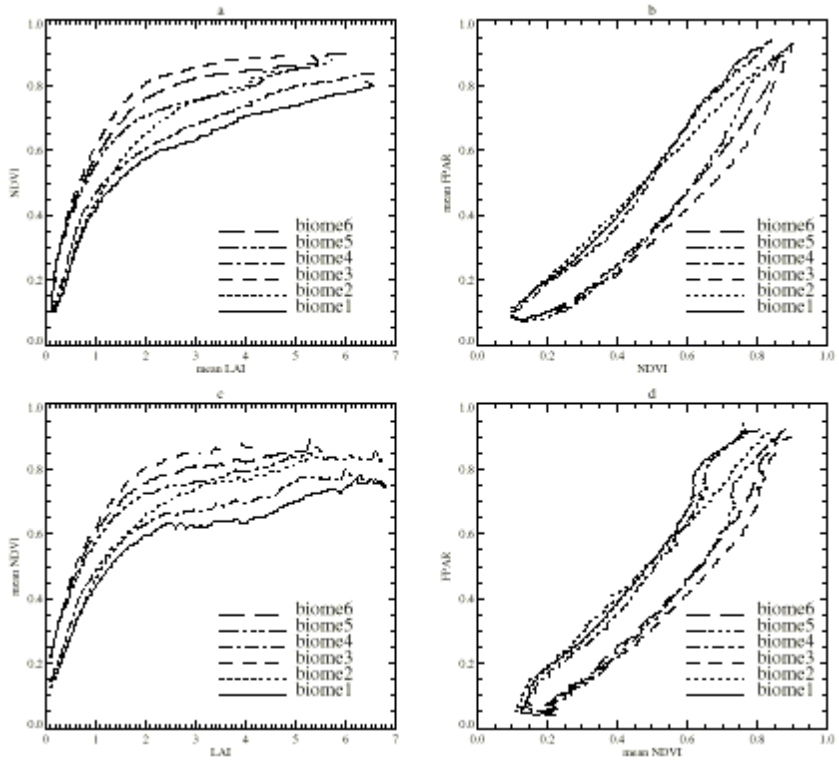


Figure 2-13. Regression curves for 6 biomes (Grasses and Cereal Crops; Shrubs; Broadleaf Crops; Savannas; Broadleaf Forests; Needle Forests) estimated with the MODIS LAI/FPAR algorithm applied to the LASUR data. Panels (a) and (b): the best possible prediction of LAI and FPAR given a realized value of NDVI. Panels (c) and (d): the best possible prediction of NDVI given realized values of LAI and FPAR.

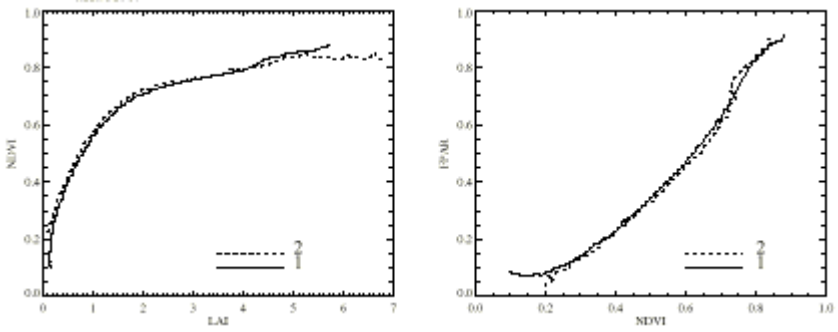


Figure 2-14. Regression curves for biome 5 (Broadleaf Forests). Left panel: the best possible prediction of LAI, $x(y)$ [curve 1], given a realized value y of NDVI, and the best possible prediction of NDVI, $y(x)$ [curve 2], given a realized value x of LAI. The regression curves $z(y)$ [curve 1] and $y(z)$ [curve 2] are shown in the right panel.

The backup algorithm consists of a straightforward direct FPAR and LAI retrieval using the pixels biome-class and the currently pixels NDVI measure as the search keys. The MOD12Q1 MODIS Landcover code is converted on the fly to one of the University of Montana biome classes, which form the primary grouping level for the LUT, within each equal width NDVI intervals are arranged. The University of Montana biome class definitions (with their MOD12Q1 approximate equivalents) are summarized in Table 2-2:

Table 2-2 Biome Class Definitions

University of Montana Biome Class Definitions		
Biome code	*Definition*	*MOD12A1 Equivalent Land Cover Type Code*
0	Water (not processed)	0
1	Grassland/cereal crops	10
2	Shrubland	9
3	Broadleaf crop	11
4	Savannah	6
5	Broadleaf forest	2
6	Needle leaf forest	1
7	Barren (not processed)	0

In the LUT itself, there are 20 equal-interval LUT records per biome class, with the intervals defined in terms of a constant NDVI interval width of 0.05. The backup LUT (see Table 2-3 for the intact LUT) is summarized in the table below:

Table 2-3 NDVI and Corresponding Values of LAI and FPAR

NDVI	Biome1		Biome2		Biome3		Biome4		Biome5		Biome6	
	LAI	FPAR	LAI	FPAR	LAI	FPAR	LAI	FPAR	LAI	FPAR	LAI	FPAR
0.025	0	0	0	0	0	0	0	0	0	0	0	0
0.075	0	0	0	0	0	0	0	0	0	0	0	0
0.125	0.3199	0.1552	0.2663	0.1389	0.2452	0.132	0.2246	0.1179	0.1516	0.07028	0.1579	0.08407
0.175	0.431	0.2028	0.3456	0.1741	0.3432	0.1774	0.3035	0.1554	0.1973	0.08922	0.2239	0.1159
0.225	0.5437	0.2457	0.4357	0.2103	0.4451	0.2192	0.4452	0.218	0.2686	0.1187	0.324	0.1618
0.275	0.6574	0.2855	0.5213	0.2453	0.5463	0.2606	0.574	0.2731	0.3732	0.1619	0.4393	0.2121
0.325	0.7827	0.3283	0.6057	0.2795	0.6621	0.3091	0.7378	0.3395	0.5034	0.2141	0.5629	0.2624
0.375	0.931	0.3758	0.6951	0.3166	0.7813	0.3574	0.878	0.393	0.6475	0.2714	0.664	0.3028
0.425	1.084	0.419	0.8028	0.3609	0.8868	0.3977	1.015	0.4425	0.7641	0.32	0.7218	0.333
0.475	1.229	0.4578	0.9313	0.4133	0.9978	0.4357	1.148	0.4839	0.9166	0.3842	0.8812	0.393
0.525	1.43	0.5045	1.102	0.4735	1.124	0.4754	1.338	0.5315	1.091	0.4402	1.086	0.4599
0.575	1.825	0.571	1.31	0.535	1.268	0.5163	1.575	0.5846	1.305	0.4922	1.381	0.5407
0.625	2.692	0.6718	1.598	0.6039	1.474	0.566	1.956	0.6437	1.683	0.568	1.899	0.6458
0.675	4.299	0.8022	1.932	0.666	1.739	0.6157	2.535	0.6991	2.636	0.702	2.575	0.7398
0.725	5.362	0.8601	2.466	0.7388	2.738	0.7197	4.483	0.8336	3.557	0.7852	3.298	0.8107
0.775	5.903	0.8785	3.426	0.822	5.349	0.8852	5.605	0.8913	4.761	0.8431	4.042	0.8566
0.825	6.606	0.9	4.638	0.8722	6.062	0.9081	5.777	0.8972	5.52	0.8697	5.303	0.8964
0.875	6.606	0.9	6.328	0.9074	6.543	0.9196	6.494	0.9169	6.091	0.8853	6.501	0.9195
0.925	6.606	0.9	6.328	0.9074	6.543	0.9196	6.494	0.9169	6.091	0.8853	6.501	0.9195
0.975	6.606	0.9	6.328	0.9074	6.543	0.9196	6.494	0.9169	6.091	0.8853	6.501	0.9195

3. Algorithm Prototyping

The objectives of this section are: by prototyping the MODIS LAI/FPAR algorithm with atmospherically corrected AVHRR (LASUR) and Landsat data to assess the effect of uncertainties in surface reflectance data on the quality of the LAI/FPAR product; to analyze global LAI/FPAR fields derived with the algorithm from LASUR data; to analyze the situations and reasons when the algorithm fails; and to understand the behavior of the algorithm as a function of spatial resolution.

3.1 Data Analysis

Before MODIS data are available, data acquired by other instruments can be used to prototype and test the functionality of the LAI-FPAR algorithm. The aim of this section is to describe and analyze the surface reflectance data used to prototype the algorithm.

3.1.1 Satellite Data

LASUR (LAnd SUrface Reflectances) is data acquired during 1989-1990 and processed at Centre d'Etudes Spatiales de la BIOsphere (CESBIO), France from the Advanced Very High Resolution Radiometer (AVHRR) onboard the NOAA 11 satellite. AVHRR is a cross-track scanning system featuring one visible (RED, 572-698 nm), one NIR (716-985 nm), one short wave infrared and two thermal infrared channels. For LASUR products, data from RED and NIR channels were used to estimate surface reflectances and vegetation index, and data from the two thermal infrared channels were used to estimate the surface temperature. LASUR data were calibrated and corrected for atmospheric effects and filtered to eliminate residual noises and perturbations. The data span from 75^{o} N to 55^{o} S in latitude, and 180^{o} W to 180^{o} E in longitude. Each image has 904 rows and 2500 columns. The spatial resolution is 1/7th degree. The temporal resolution is weekly. In this study, RED and NIR surface reflectances from July, 1989 were used to prototype the MODIS LAI/FPAR algorithm. We created a monthly layer based on maximum NDVI compositing of the four weekly layers in this month. By doing so, cloud contamination, off-nadir viewing effects, sun-angle and shadow effects, aerosol and water-vapor effects can be minimized [Holben, 1986].

A Biome Classification Map (BCM) that describes the global distribution of six canopy structural types (biomes) was used as a prototype of the MODIS Land Cover Product, which is required by the MODIS LAI/FPAR algorithm. This data set was derived from the AVHRR Pathfinder data set [Myneni et al., 1997] and it is a time-independent data set. These biome types are: grasses and cereal crops (biome 1), shrubs (biome 2), broadleaf crops (biome 3), savannas (biome 4), broadleaf forests (biome 5), and needle forests (biome 6).

We also utilized Landsat Thematic Mapper (TM) scenes of Northwest U. S. (Washington and Oregon) from June 26, 1987 at 30 meter resolution in this study to evaluate the algorithm response to high resolution data. The data are of 6 spectral bands, bands 1, 2, 3, 4, 5 and 7. This image was geometrically registered to a terrain-corrected image with an UTM (Universal Transverse Mercator) projection. The dark object subtraction method of atmospheric correction was used to correct surface reflectance for the atmospheric effect. There was also a

"sitemap", containing polygons of known ground cover, associated with this data set. This sitemap distinguishes 17 different forest densities, based on percent forest cover in a forested pixel, and 7 other types of miscellaneous landcover types. By using the Bayesian Maximum likelihood classification method, we separated three biomes in this Landsat image, namely, grasses and cereal crops, broadleaf and needle forests. Agriculture (crop, pasture) class was taken as biome 1. We attribute to broadleaf forests all the pixels in which hardwood forest makes up more than 60% of the area in one pixel. Needle forests consist of those pixels in which conifer forest makes up more than 60% of the area in one pixel. The other landcover classes not belonging to these three biomes were all defined as unknown class type. In all, grasses occupy 6.6% of the total area, and broadleaf and needle forests occupy 4.8% and 10.3% of the total area.

3.1.2 Spectral Signatures

Although all the vegetation types have relatively similar spectral properties--large absorption in RED and large reflectance in NIR, different biomes have special characteristics depending on the canopy architecture. These can be distinguished by comparing their spectral signatures. Figure 3-1(a), (b) presents histograms of canopy reflectances in RED and NIR spectral bands as a function of biome type derived from the LASUR data. In the RED band, canopy reflectances vary between 0.0 and 0.2. Broadleaf and needle forests have the strongest absorption features. On an average, they reflect only 3% and 4.5% (Table I) of the incoming

TABLE 3-1 SPECTRAL CHARACTERS OF DIFFERENT VEGETATION TYPES FOR LASUR DATA AND LANDSAT TM DATA

	LASUR DATA		
Biome Type	Mean Red	Mean NIR	Mean NDVI
Grasses and Cereal Crops	0.080	0.25	0.515
Shrubs	0.050	0.21	0.615
Broadleaf Crops	0.065	0.32	0.662
Savanna	0.050	0.23	0.645
Broadleaf Forests	0.030	0.29	0.813
Needle Forests	0.045	0.25	0.695
	LANDSAT DATA		
Biome Type	Mean Red	Mean NIR	Mean NDVI
Grasses and Cereal Crops	0.065	0.304	0.635
Broadleaf Forests	0.022	0.348	0.881
Needle Forests	0.013	0.200	0.886

** The LASUR data is July 1989 globally, which is composited from four weeks data using maximum NDVI. The Landsat TM data is from June 26, 1987, which is a scene of Northwest U.S. (Washington and Oregon).

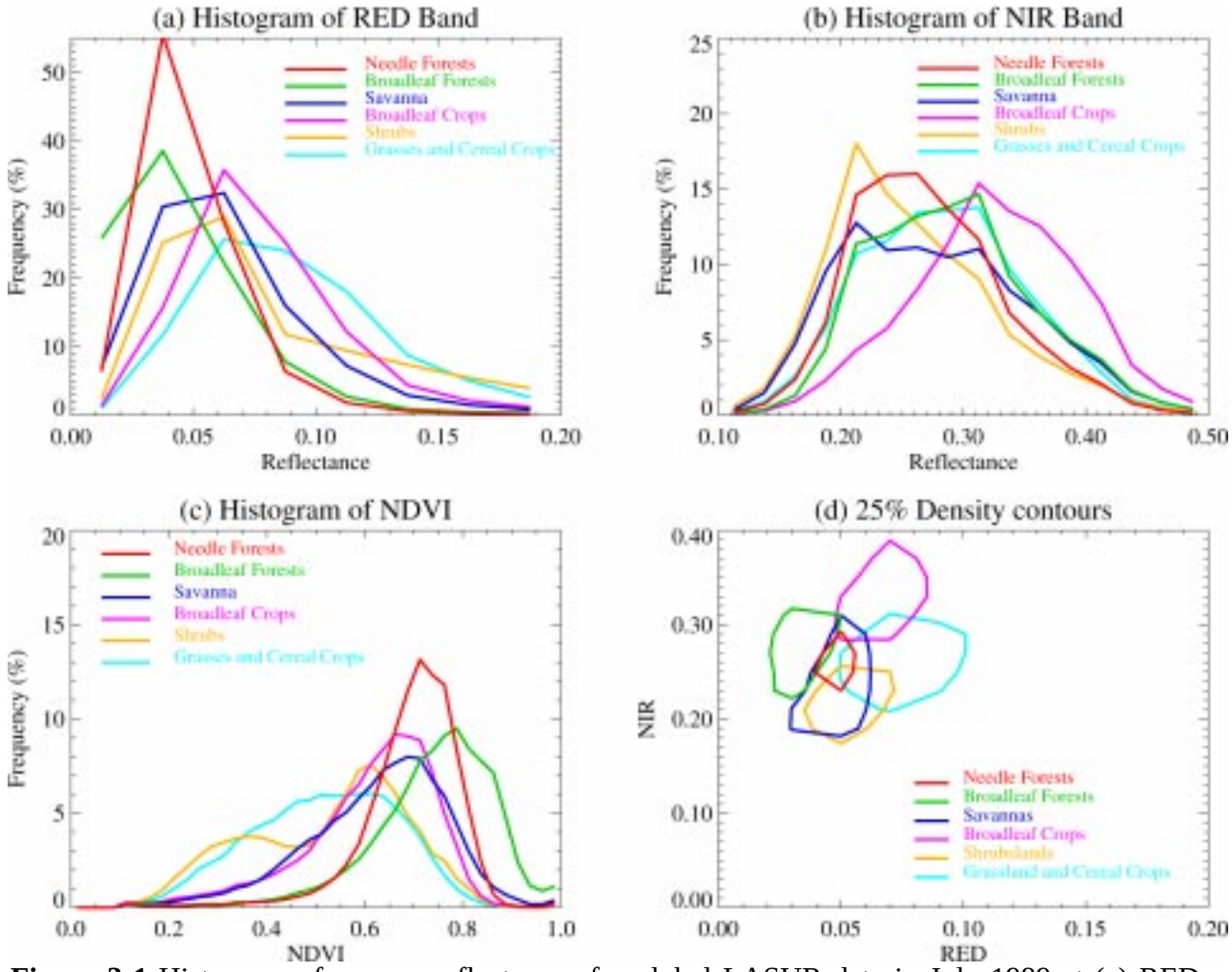


Figure 3-1 Histogram of canopy reflectances for global LASUR data in July 1989 at (a) RED band, (b) NIR band. (c) histogram of NDVI. The interval is 0.25 for above three variables. (d) 25% density contours in RED-NIR space, which shows the location of points with high density for different biomes.

radiation. Grasses and broadleaf crops are characterized as the brightest biomes; about 8% and 6.5% of the incoming radiation is reflected. In the NIR band, reflectances vary between 0.1 and 0.5. Shrubs and broadleaf crops represent the extremes. Their reflectances, on an average, are 21% and 32%, respectively. The other biomes reflect about 25% of the incoming radiation and the histograms are similar.

Vegetation indices typically capture the absorption contrast across the 0.65-0.85 μm wavelength interval through combinations of broad-band RED and NIR reflectance. The most widely used index in the processing of satellite data is the Normalized Difference Vegetation Index (NDVI) defined as $[(\rho_N - \rho_R)/(\rho_N + \rho_R)]$, where ρ_N and ρ_R are spectral reflectance at NIR and RED wavelengths, respectively. NDVI is a measure of chlorophyll abundance and

energy absorption [Myneni et al., 1995]. Figure 3-1(c) demonstrates the distribution of NDVI values derived from the LASUR data. In general, broadleaf forests have the highest NDVI values (0.813), followed by needle forest (0.695). Broadleaf crops and savannas have similar NDVI distributions, and their NDVI values are larger than those of grasses (0.515) and shrubs (0.615). Obviously, it would be difficult to distinguish broadleaf crops from savannas using NDVI only.

The data density distribution function, introduced earlier in Section II, can be used to indicate the location of data peak in the spectral space. Figure 3-1(d) shows the location of points with high density for different biomes in the RED-NIR space. Each area bounded by the contour represents an area containing the 25% density of pixels from a given biome type. We can see that the biomes tend to cluster and occupy certain well localized spaces. Broadleaf forests are located at low RED and high NIR area, while grasses at the high RED and low NIR area. Broadleaf crops and savannas occupy different locations although their NDVI distributions are comparable, not surprisingly because NDVI is a non-linear function of RED-NIR reflectances. We also note that some biomes overlap considerably. In general, the more unique the location, the better the ability to distinguish them. The influence of soil is also clear from this panel. Grasses and shrubs are biomes that are located near the soil line. Broadleaf forests are a dense vegetation type and these are located closest to the NIR axis.

Figure 3-2 presents canopy reflectance features from Landsat data. On an average, grasses, broadleaf and needle forests reflect only 6.5%, 2% and 1.3% of the incoming radiation in the RED band. This is much less than that of LASUR data. On the other hand, the NIR reflectance of grasses and broadleaf forest can be as high as 30% and 34.8%, compared with 25% and 29% for the LASUR data. Needle forests are the darkest among the three biomes, both at RED and NIR. The NDVI values for the three biomes are 0.635, 0.881, 0.886, respectively. The 25% density contours are tightly clustered occupying a small but unique location in the spectral space. At the same time, the clusters are away from the soil line, and closer to the NIR axis. The biomes are so well separated that they do not overlap even on the 75 % density contour.

Reflectances decrease in the RED band and increase in the NIR band as spatial resolution increases. The effect of soil becomes important when resolution decreases, that is, pixels located near the NIR axis (Landsat data) will move towards the soil line (LASUR data). The clustering of data also depends on resolution: the coarser the resolution, the more the biomes overlap in the two dimensional spectral space. Biomes that separate well are unlikely to be misclassified. On the other hand, biomes that overlap may be misclassified, and this thus directs the algorithm to the wrong LUT. The same input data (Red and NIR reflectances) will result in different LAI/FPAR values because the algorithm is biome dependent.

3.2 Prototyping of The Algorithm

3.2.1 Prototyping with LASUR Data

This section describes results of global LAI and FPAR fields derived with the MODIS LAI/FPAR algorithm using the LASUR-AVHRR data. The objectives were to analyze the global LAI/FPAR fields, and the situations when the algorithm failed, and to assess the influence of uncertainties in surface reflectances on the quality of the LAI/FPAR product.

The algorithm was run pixel-by-pixel using LASUR and land cover BCM data, on all pixels with NDVI greater than 0.1. The following notions are used in the discussion on algorithm performance: (1) A pixel for which the algorithm retrieves a value of LAI and FPAR is termed as a "retrieved" pixel; (2) A pixel for which the algorithm fails is termed as an "non-retrieved" pixel; (3) The ratio of the number of retrieved pixels to the total number of pixels is named the Retrieval Index (RI). All results are reported on biomes.

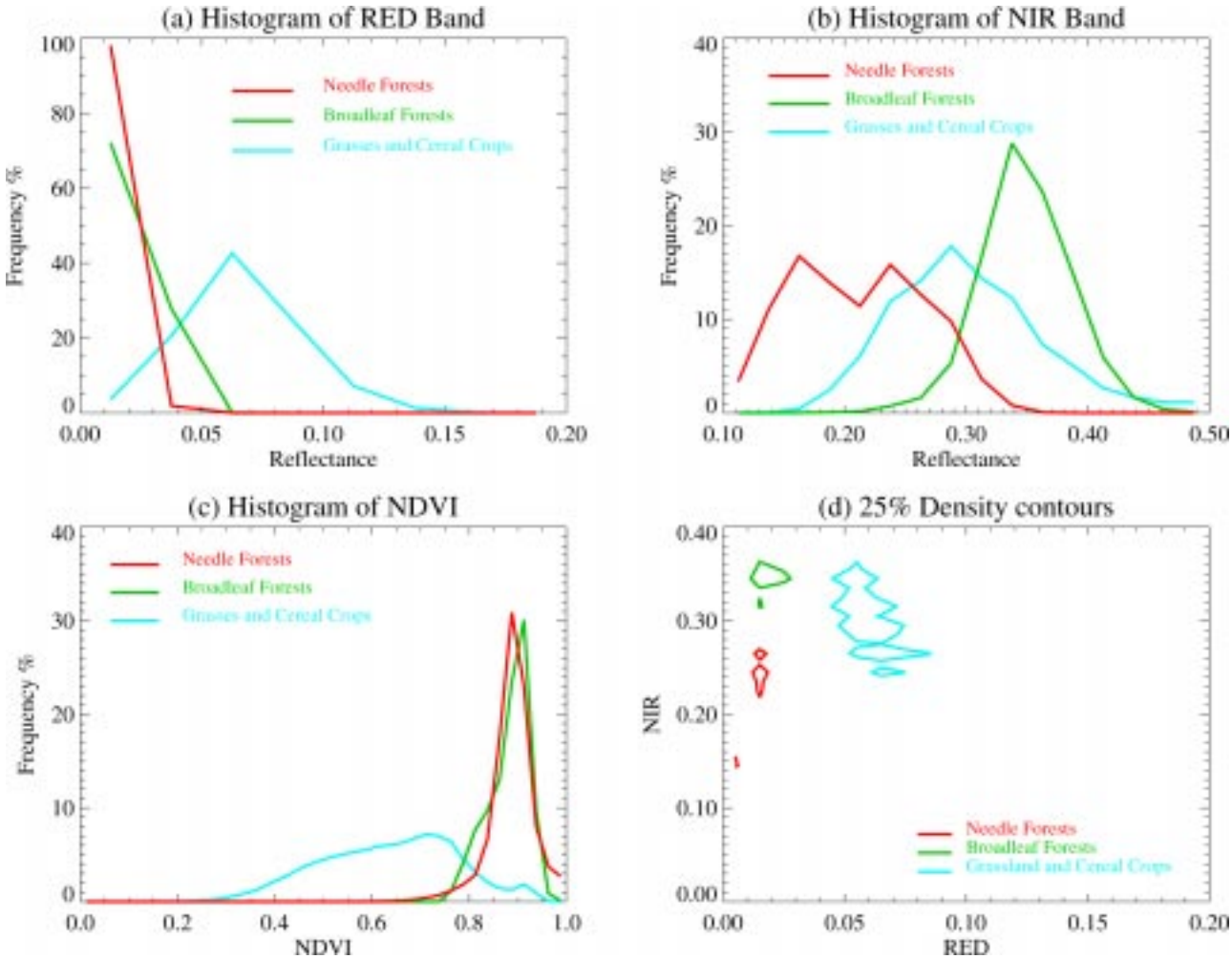


Figure 3-2 The same as Figure 3-1, but for Landsat TM data, June 1987

1) Input Data: Atmospherically corrected surface reflectances and uncertainties in measurements and simulations are inputs to the algorithm. However, LASUR reports no information on the uncertainties in reflectances. Therefore, the uncertainties were simulated as

$$\delta_{RED} = \delta_{NIR} = \varepsilon \left[d_{RED}^2 + d_{NIR}^2 \right]^{1/2} \tag{3.1}$$

Here ε is the mean uncertainty, which is assumed to be a constant in this study. Figure 3-3 demonstrates the dependence of the RI on ε. The RI increases with increase in ε, however, the quality of retrieved LAI/FPAR decreases with increase in ε. If ε is underestimated, the algorithm fails when real uncertainties in surface reflectances are greater than those determined by equation (3.1). If ε is overestimated, the algorithm can produce LAI/FPAR values for non-vegetated pixels also. Therefore, there should be a critical value of ε for which equation (3.1) optimally approximates real uncertainties. A value of 0.2 for ε was found optimal in this study for LASUR data. The RI varies with biome types at a constant ε. When ε is 0.2, the RI for the 6 biomes are 91.5%, 92.7%, 74.0%, 79.7%, 39.3%, 54.5%, respectively. The reason that broadleaf and needle forests have low RI is perhaps because of the dark soil pattern used to represent effective ground reflectance $\rho_{q,eff}(\lambda)$. Thus, if a pixel is bright, it was not considered as a pure broadleaf or needle forest pixel, and the algorithm consequently failed. Low values of RI are not necessarily indicative of poor performance by the algorithm. For the coarse resolution data such as LASUR (1/7th of a degree), the vegetation in the pixel may be a case of mixtures of different land cover classes. Therefore, biome specific spectral features may be lost. At the present time, restricting the algorithm to pure vegetation types retains the ability to discriminate biome types.

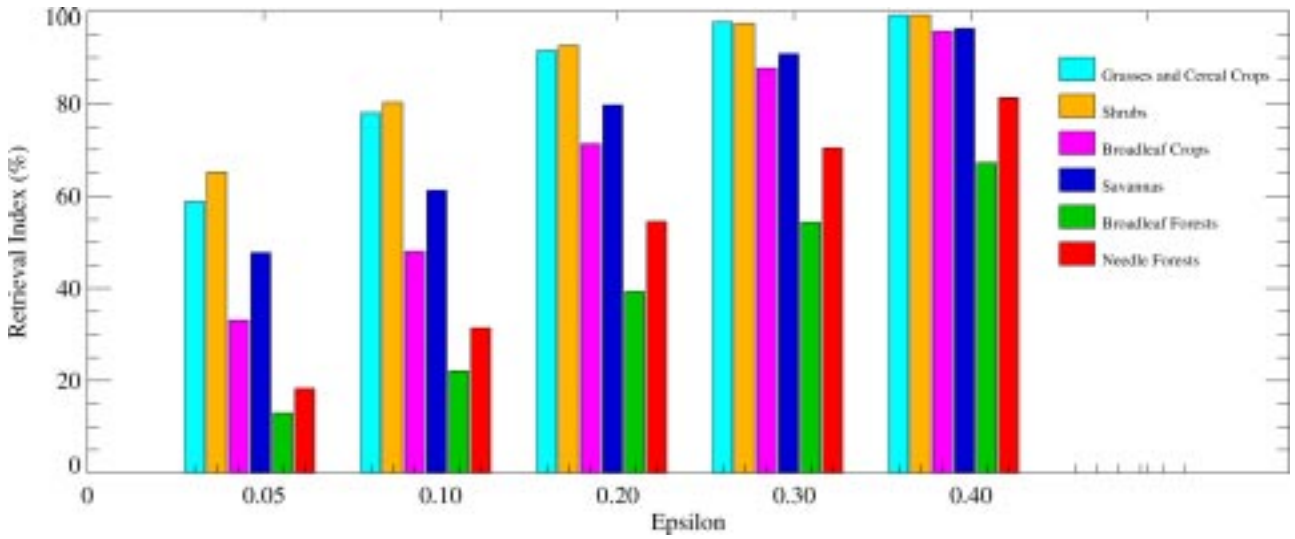


Figure 3-3 The dependence of the Retrieval Index on ε. The ε means uncertainties in measurements and simulations.

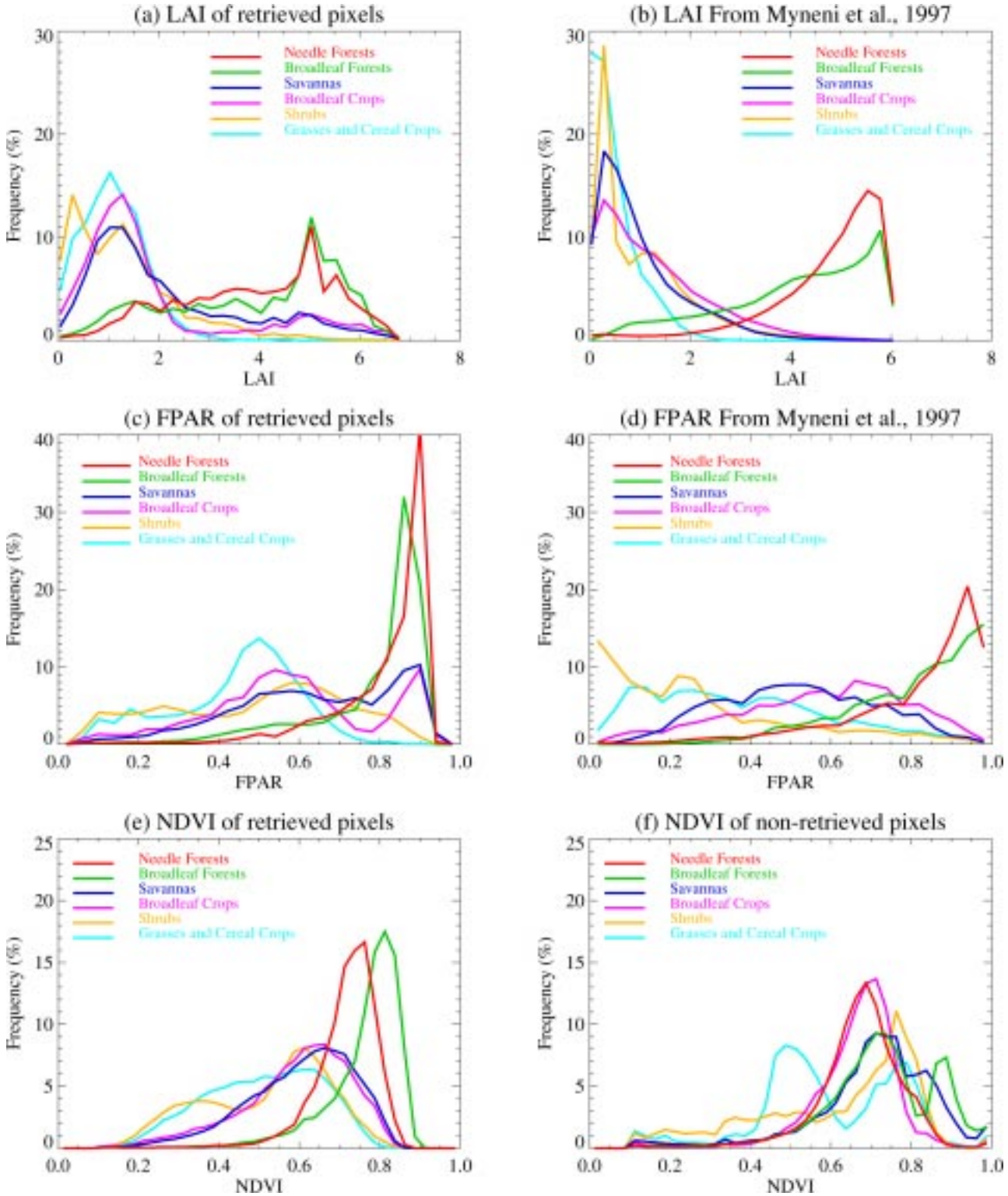


Figure 3-4 Comparison of LAI/FPAR from our algorithm (a), (c) with 10-year averaged Pathfinder [Myneni et al., 1997] (b), (d). Comparison of the histogram of NDVI from retrieved pixels (e), and from

non-retrieved pixels (f). The mean uncertainty ε used in our algorithm is 0.20 and the interval used for LAI is 0.25, for FPAR and NDVI is 0.04.

2) Histograms of LAI and FPAR: The histogram of the retrieved LAI/FPAR is a statistical characteristic that describes the distribution of these fields for various biomes. Figure 3-4(a) presents the histogram of retrieved LAI using the LASUR data. Broadleaf and needle forests have distributions distinct from the other four biomes. These biomes have relatively high LAI values, concentrated about 4.0 to 6.0. For the other biomes, the LAI values are generally less than 2. The differences among grasses, shrubs, broadleaf crops and savannas are seen in the peak and tail of the LAI histograms. The highest frequency of LAI for broadleaf crops and savannas is around 1.25, while for grasses it is at 1.0 and for shrubs at 0.75 and 1.25. For broadleaf crops and savannas, the tail contains at least 20% of the pixels whose LAI values are larger than 4.0. On the other hand for grasses and shrubs, the distribution ends at about 4.0. Thus, the mean LAI for broadleaf crops and savannas is 2.1 and 2.2, and for grasses and shrubs only 1.2 and 1.4. Shrubs have two obvious peaks, which correspond to the two peaks in the NDVI histogram shown in Figure 3-1(c), although the retrieved LAI is not based on the NDVI.

The LAI distribution from a NDVI based algorithm developed earlier by Myneni et al. [Myneni et al., 1997] is shown in Fig 3-4 (b). The data used for this NDVI based algorithm were AVHRR from July 1981 through June 1991. The average July retrievals over the 10 years period are shown in Figure 3-4(b). There are many similarities between Figure 3-4(a) and (b). Broadleaf and needle forests have much higher LAI than that of the other four biomes. The double peak in the case of shrubs is also seen in Figure 3-4(b). The similarity between the two retrievals imbues confidence in the MODIS algorithm in spite of its complexity. However, it should also be noted that LAI values of the first four biomes are larger in the case of the MODIS algorithm as compared to the NDVI based algorithm. Similarly, the histograms of FPAR of different biomes match well with the 10-year climatology (Figure 3-4(c) and (d)).

Figure 3-4(e) and (f) show the NDVI histograms from retrieved and non-retrieved pixels. The NDVI histogram of retrieved pixels is similar to the NDVI histogram of all the pixels. Therefore, the algorithm identifies most of the features in the observed data. The failures are typically of two cases. First, NDVI is too high for a particular biome. For example, the algorithm fails to retrieve when the NDVI of grasses is larger than 0.75. In the LUT, there is no information for grasses at such values of NDVI. Second, for the same NDVI value, some of the pixels are retrieved pixels, but the others not. The failure of this type is discussed later.

3) Test of Physics: There are many examples in published literature of the strong relationship between a vegetation index such as NDVI and surface parameters LAI and FPAR [Asrar et al., 1984; Tucker et al., 1986; Peterson et al., 1987; Verma et al. 1993; Myneni et al., 1994; Chen, 1996; Chen et al., 1996]. This provides us an opportunity to test the physics of the algorithm by comparing the LAI-NDVI and FPAR-NDVI relationships derived from the algorithm with those reported from field measurements. Figure 3-5(a) and (b) show the distributions of the retrieved values of LAI and FPAR with respect to the NDVI of biome 5. LAI is nonlinearly proportional to NDVI, and FPAR linearly to NDVI. This corresponds to relations reported in the literature [Myneni et al., 1997; Clevers, 1989]. Note that the NDVI in this plot is

evaluated from measured RED and NIR reflectances, while the retrieved quantities result from the algorithm which uses reflectances and not NDVI. The advantages of using the MODIS algorithm as opposed to NDVI relations are -- (1) NDVI/LAI relations are subject to changes in sun angle, background reflectance, and view angle, while the MODIS algorithm actually uses these changes as sources of information in the process of retrieval; (2) NDVI is based on two spectral bands only, while the algorithm can ingest 3, 4 or even 7 bands simultaneously to retrieve LAI and FPAR.

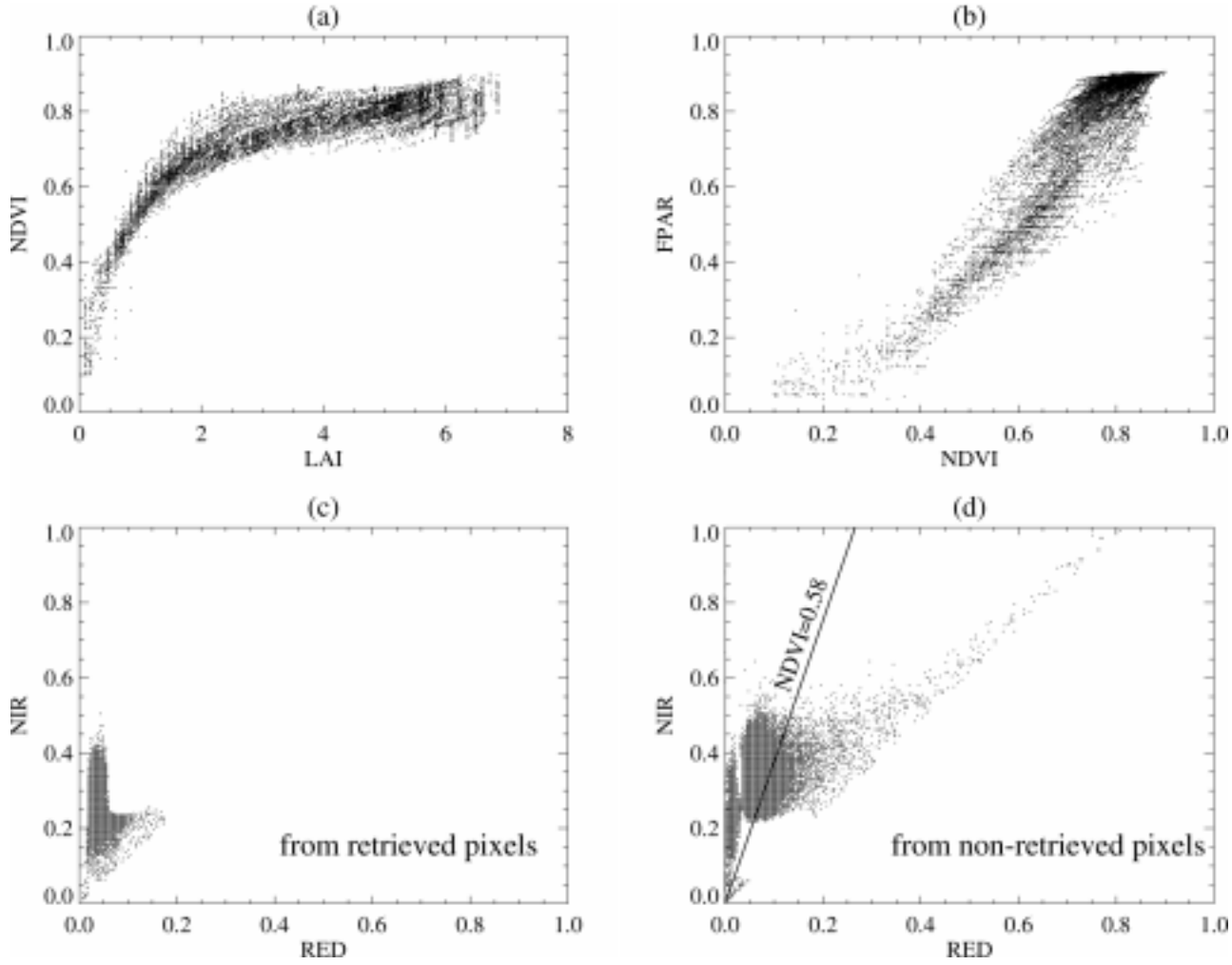


Figure 3-5 For broadleaf forests of LASUR data, the scatter plot shows (a) the LAI-NDVI relationship, (b) NDVI-FPAR relationship, (c) retrieved pixels in RED-NIR space, and (d) non-retrieved pixels in RED-NIR space.

Figure 3-5(e) and (f) show the scatter plot of the data from successful and unsuccessful pixels in the RED-NIR plane. This distribution provides insight on where and why the algorithm failed. For retrieved pixels in the RED-NIR plane, canopy reflectances range about 0.02--0.16

for the RED band, 0.1--0.42 for the NIR band. This reflectance space obviously overlaps with the 25% density contour area. From Figure 3-4(e), (f) and Figure 3-5, it appears that there are three regions where the algorithm fails: RED reflectance less than 0.03 (NDVI is very large), large RED and NIR reflectances (pixels near the soil line and NDVI is very small) and, RED and NIR are relatively large and located between these two regions. When the RED reflectance is too small, the uncertainty is large, and the probability of retrieval decreases. When a pixel is near the soil line, it is not a vegetated pixel, and the algorithm identifies such cases correctly. For the third region, consider an NDVI contour as shown in Figure 3-5(d). For the same value of NDVI, some pixels result in a retrieval while the others do not. It is clear that the algorithm uses information on the canopy spectral properties instead of NDVI, especially when there are three, four or even seven spectral bands and multi-angle data. Only when a pixel falls within the specified spectral and angular space in the LUT, can it retrieve a LAI value. Otherwise, the algorithm returns a failure, even if the NDVI is reasonable. Therefore, a correct LUT is key to algorithm performance. The non-retrieved pixels are generally biome mixtures or those mixed with non-vegetated surface. The algorithm can not identify these surface types. In general, the algorithm fails when a pixel has too small a RED reflectance and/or too large a NIR reflectance. The LUT was built to represent the core of the data, and thus is likely to fail at the margins.

4) Reliability of Retrieved LAI/FPAR: The saturation frequency and dispersion of the retrieved LAI distribution (DLAI) are two elements by which the quality of the retrieval can be assessed. The accuracy of the retrieval decreases under conditions of saturation, that is, the reflectance data contain no accurate information about the surface [Knyazikhin et al., 1998a]. Therefore, the saturation frequency and the threshold LAI value of saturation on a biome basis are important criteria for assessing the accuracy of the retrievals. For the six biomes, the overall saturation frequencies are 0.38%, 2.5%, 16%, 15%, 48.5% and 42.5%, respectively. Broadleaf and needle forests have the highest saturation frequency and, grasses and shrubs have the lowest, with broadleaf crops and savannas inbetween. Figure 3-6(c) shows the histogram of LAI under conditions of saturation for the 6 biomes. When the LAI is less than 4.0, the saturation frequency is low for all biomes. On the other hand, when LAI is larger than 4.0, the saturation frequency increases dramatically, and practically every pixel is under saturation conditions when the LAI is larger than 5.0. Therefore, with increasing LAI value, the accuracy of the retrieval decreases, as is to be expected.

Broadleaf and needle forests in general have high LAI values and therefore the saturation frequency is higher. In order to assess the reliability of the retrieved LAI/FPAR values, we examine the dispersion of the retrieved LAI distribution function in Figure 3-6(d). Here, DLAI is defined as the ratio of the standard deviation of LAI to LAI itself. In theory, the lower the DLAI value the higher the accuracy of the algorithm. In Figure 3-6(d) the standard deviation of the retrieval increases with LAI. This is not surprising, because at high LAI values, the reflectances saturate, and it is difficult to localize a single estimate, that is, the retrieved LAI distribution is under a condition of saturation. All biomes have DLAI values around 0.2 for LAIs larger than 1.0. When LAI is larger than 3.0, broadleaf and needle forests have relatively lower DLAI values than other biomes at the same LAI value. Thus, when LAI is large and saturation frequency is large, the retrieval is not necessarily poor. Obviously, DLAI can not be

less than 0.2 because the mean uncertainty in these runs was 0.2. Thus, the quality of the retrievals can not be better than the quality of the most uncertain spectral reflectance data input to the algorithm. Therefore, the availability of band specific uncertainties in atmospherically corrected surface reflectances is critical to assess the quality of the LAI/FPAR product.

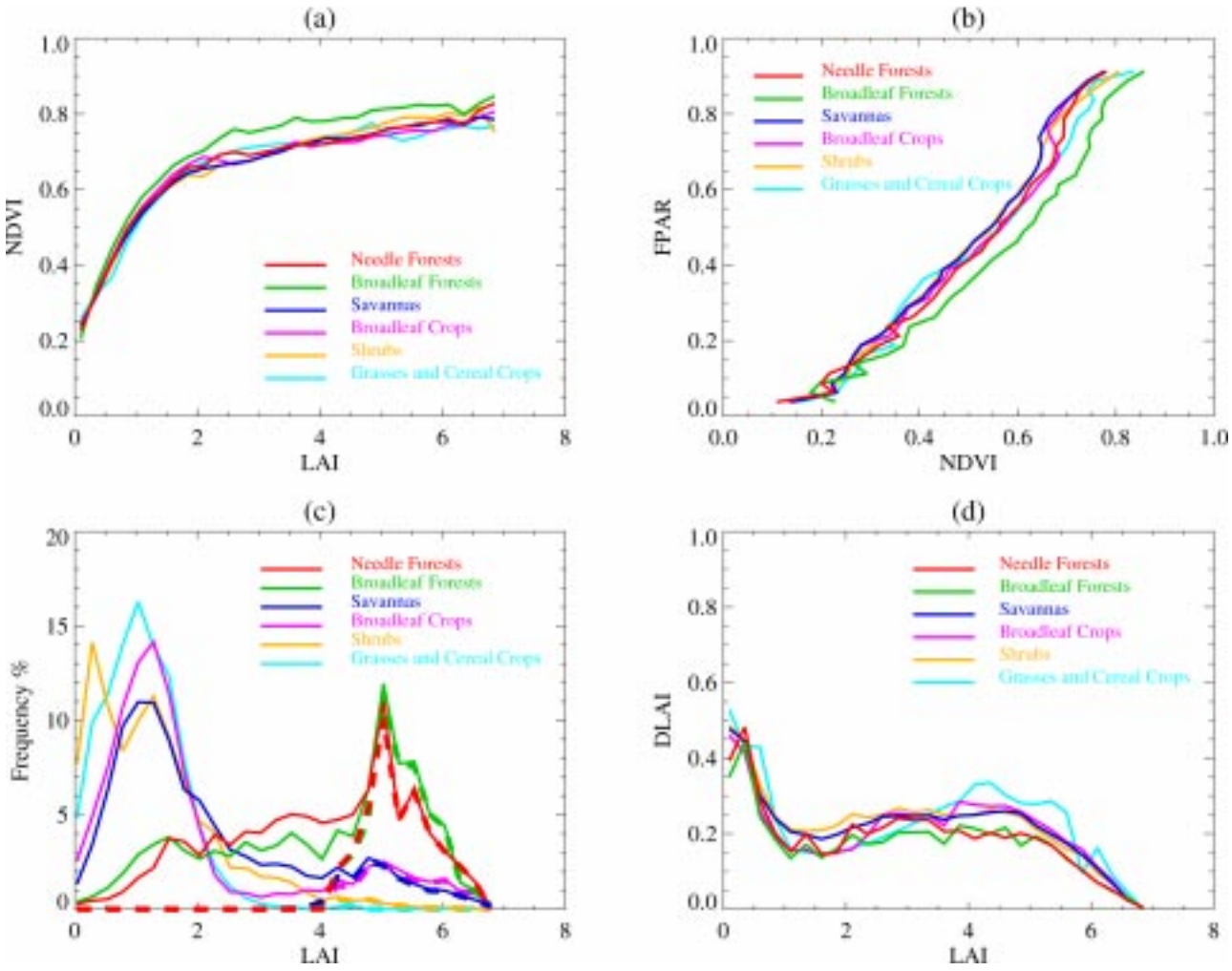
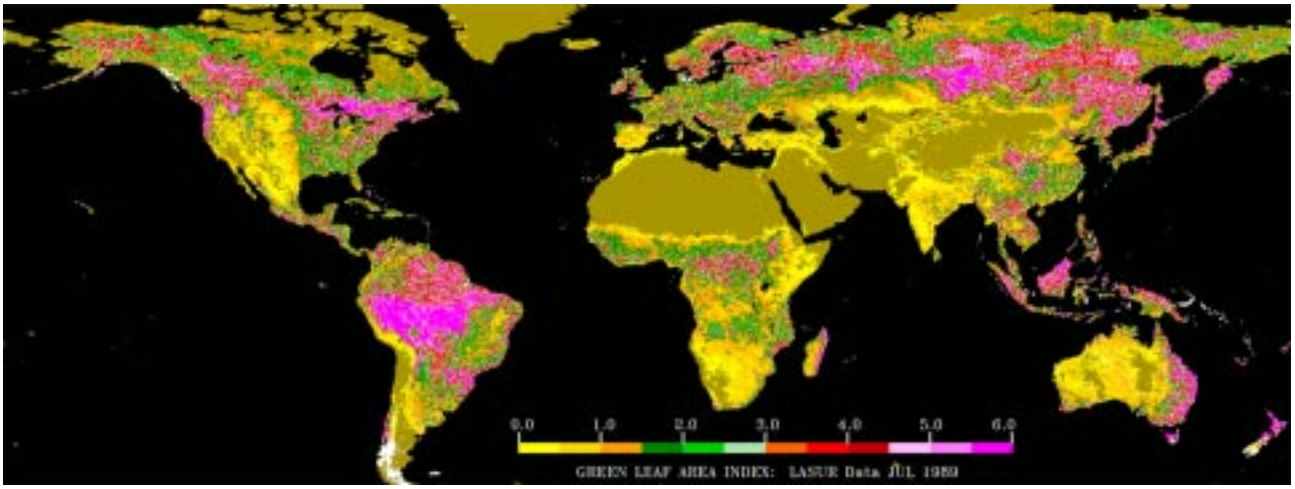
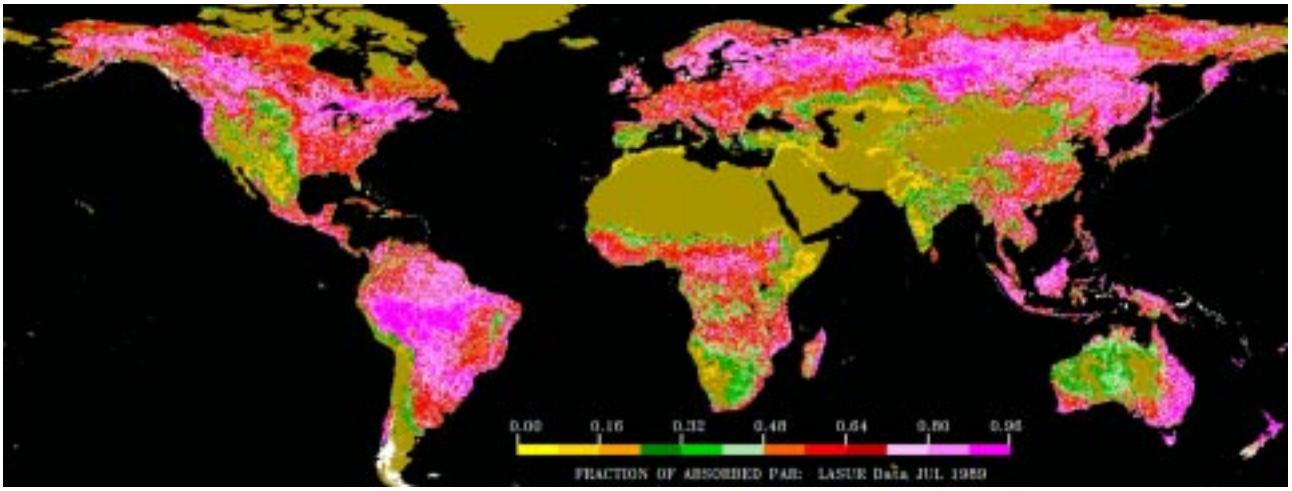


Figure 3-6 (a) The average LAI-NDVI relationship, (b) The average NDVI-FPAR relationship, (c) The histogram of LAI under conditions of saturation. The interval is 0.25. (d) The dispersion of LAI (DLAI) as a function of LAI. All these are from LASUR data, July 1989.

5) LAI and FPAR Images: The algorithm was run on the global LASUR data for the month of July 1989. For the unsuccessful pixels, the NDVI-LAI, NDVI-FPAR relations shown in Figure 3-6(a) and 6(b) were used to estimate LAI and FPAR. Figure 3-7 shows color-coded images of global LAI and FPAR. These compare well with the fields reported earlier by Myneni et al. [Myneni et al., 1997].



(a)



(b)

Figure 3-7 Global (a) LAI and (b) FPAR derived from LASUR data in July, 1989. The algorithm was run on LASUR data. For the non-retrieved pixels, the LAI-NDVI, NDVI-FPAR relations shown in Figure 6(a) and (b) were used to estimate LAI and FPAR.

6) Biome Misclassification and LAI/FPAR Retrievals: It is important to assess the impact of biome misclassification on LAI/FPAR retrievals. To understand this, we ran algorithm 6 times per pixel, each time with a different biome's LUT. This simulates the effects of biome misclassification on LAI/FPAR retrievals. The results are shown in Table II and Figure 3-8. For example, for grasses and cereal crops (Figure 3-8(a)), the histogram of grasses and cereal crops indicates that pixels of this vegetation type were retrieved using the correct LUT; the histogram of shrubs indicates that pixels of this vegetation type were misclassified as shrubs and were therefore retrieved using the LUT for the shrubs vegetation type.

When pixels are misclassified, typically, either the Retrieval Index (RI) is low and/or the retrieved LAI values are incorrect. When misclassification between clearly distinct biomes happens, the results are predictable. Consider grasses and cereal crops (biome 1) and broadleaf forests (biome 5), which are distinct in their architecture and foliage optics. When biome 1 is misclassfied as biome 5, the RI is only 27% compared to 91% without misclassification. When biome 5 is misclassified to biome 1, the retrieved LAI value decreases from 4 or 5 to 2. Thus, misclassification can be detected by RI, mean LAI and the histogram of the retrieved LAI distribution. On the other hand, when misclassification happens between spectrally and structurally similar biomes, perhaps because of coarse spatial resolution, the impact on LAI/FPAR retrievals is difficult to assess. As an example, consider shrubs (biome 2) and savannas (biome 4). The RI does not change much, and the mean LAI is also similar. Thus, the retrieved LAI/FPAR values are acceptable, although the pixels have been misclassified. Such a case also indicates that the various biome LUTs share similar entries for certain combinations of spectral reflectances.

3.2.2 Prototyping With Landsat Data

1) General Results: The MODIS LAI/FPAR algorithm was also prototyped with Landsat data for three biomes (grasses and cereal crops, broadleaf forests, needle forests) as mentioned previously. A fine resolution LUT was used to retrieve LAI and FPAR because of the finer spatial resolution of the Landsat data. The RI for the three biomes was 90.7%, 53.9%, 57.9%, respectively, and the mean LAI values were 1.87, 5.79, 4.11. Compared to LASUR data, not only did the RI increase for broadleaf and needle forests, but also the mean LAI value increased. The finer resolution pixels contain fewer mixtures which perhaps explains these results. The saturation frequencies at high LAI values for the three biomes are comparable to those reported earlier for the LASUR data, but the dispersions (DLAI) are narrower for broadleaf and needle forests indicating a better localization of the retrievals. Thus, compared to coarse resolution data, the 30 m resolution Landsat data result in qualitatively better retrievals of LAI and FPAR. This highlights the value of spatial resolution and indicates that the 1 km MODIS LAI/FPAR retrievals will be better accurate than the 8 km AVHRR retrievals currently available.

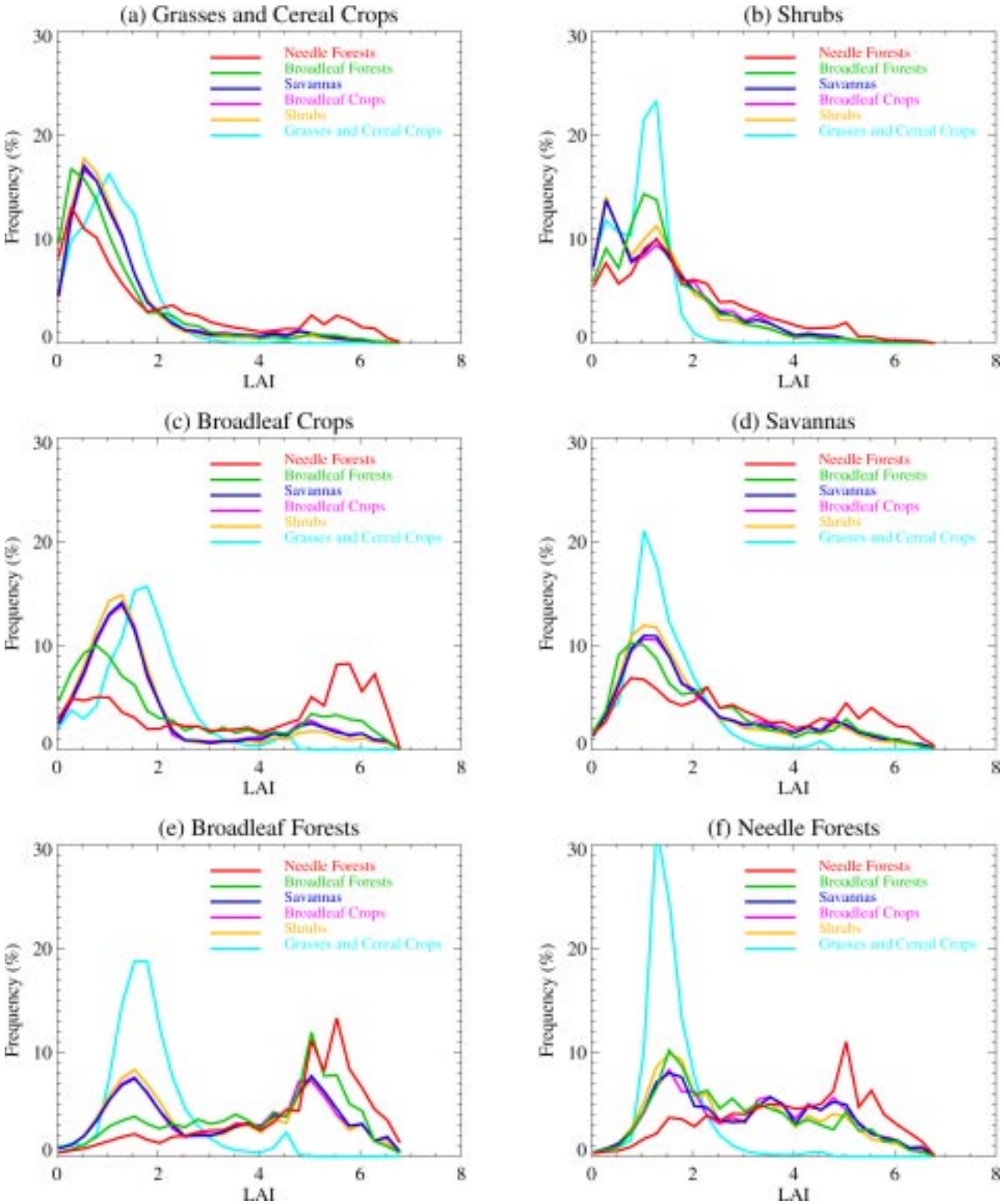


Figure 3-8 Histogram of LAI for (a) Grasses and Cereal Crops, (b) Shrubs, (c) Broadleaf Crops, (d) Savannas, (e) Broadleaf Forests, and (f) Needle Forests from LASUR data by using 6 biome' Look-Up-

Table. This simulates the effects of biome misclassification on LAI retrievals. The algorithm is run 6 times per pixel, each time with a different biome's LUT.

TABLE 3-2 EFFECTS OF BIOME MISCLASSIFICATION ON LAI RETRIEVALS FOR LASUR DATA

(A) RETRIEVAL INDEX, (b) MEAN LAI

		Misclassified Biome Type					
		Grasses and Cereal Crops	Shrubs	Broadleaf Crops	Savanna	Broadleaf Forests	Needle Forests
BCM Biome Type	Grasses and Cereal Crops	***91.53***	88.54	89.60	88.68	27.63	29.00
	Shrubs	87.67	***92.66***	91.53	91.73	47.34	46.37
	Broadleaf Crops	87.93	70.33	***74.03***	71.29	14.80	19.52
	Savanna	78.02	79.91	80.25	***79.65***	41.31	44.33
	Broadleaf Forests	55.02	63.23	61.4	61.32	***39.30***	33.59
	Needle Forests	76.75	85.74	84.92	84.78	46.38	***54.54***

(a)

		Misclassified Biome Type					
		Grasses and Cereal Crops	Shrubs	Broadleaf Crops	Savanna	Broadleaf Forests	Needle Forests
BCM Biome Type	Grasses and Cereal Crops	***1.197***	1.245	1.401	1.363	1.293	2.011
	Shrubs	1.026	***1.408***	1.542	1.514	1.505	1.987
	Broadleaf Crops	1.845	1.833	***2.097***	2.044	2.424	3.710
	Savanna	1.508	2.079	2.286	***2.250***	2.221	2.953
	Broadleaf Forests	1.921	3.299	3.439	3.451	***4.014***	4.649
	Needle Forests	1.640	2.916	3.205	3.179	2.976	***3.996***

(b)

** The algorithm is run 6 times per pixel, each time with a different biome's LUT. This simulates the effects of biome misclassification.

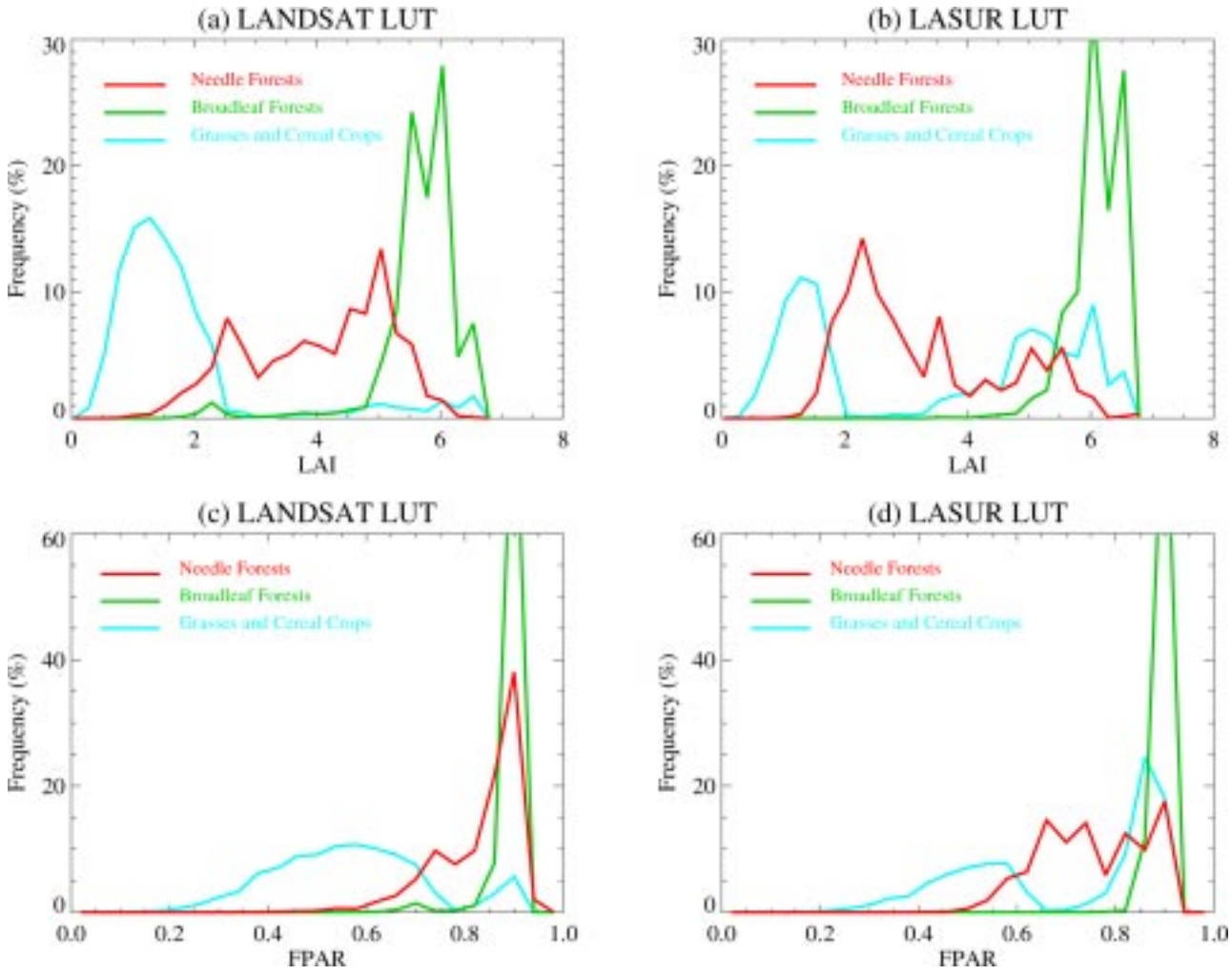


Figure 3-9 The retrievals from Landsat data as a function of spatial resolution dependent Look-Up-Table (LUT). Histogram of LAI from (a) Landsat LUT, (b) LASUR LUT. Histogram of FPAR from (c) Landsat LUT, (d) LASUR LUT.

The reason for the dependency of the LUT on spatial resolution is as follows. Depending on the resolution, leaf canopy spectral properties are very different (Figure 3-1 and Figure 3-2). At fine spatial resolution, leaves absorb more in the RED band and also reflect more in the NIR band. In the RED-NIR plane, Landsat data tend to cluster, and occupy a small region nearer to the NIR axis. The LUT should reflect these changes in vegetation canopy spectral properties with changes in resolution. To investigate this further, the algorithm was executed on Landsat data, but with LASUR LUT, that is, fine resolution data with coarse resolution LUT. Figure 3-9 shows the histogram of LAI and FPAR obtained with Landsat data and LASUR LUT, and also Landsat data with Landsat LUT. When Landsat data and LUT are used, the retrieved LAI values vary from 0.0 to 2.5 for grasses, 5.0 to 7.0 for broadleaf forest,

and 1.5 to 6.0 for needle forests (Table III). When LASUR LUT is used with Landsat data, the histogram of retrieved LAI and FPAR change greatly. The LAI of grasses/cereal crops can be as high as 4.0 to 6.0, which is unrealistic. The LAI of needle forests is concentrated between 1.5 to 4.0, which is relatively small for this biome. The RI for the three biomes also decrease to 87.5%, 39.2%, 4.7%, respectively. When the algorithm is run using LASUR data but Landsat LUT (Figure 3-10), the retrieved LAI and the mean LAI for all biomes decrease. The differences between forests (high LAI) and other biomes (low LAI) disappear. This clearly indicates the dependency between data resolution and the LUT.

TABLE 3-3 Comparison of the Results from LASUR look-up-table (lut) and LANDSAT LUT

	LASUR DATA			
	LASUR LUT		LANDSAT LUT	
Biome Type	Retrieval Index	Mean LAI	Retrieval Index	Mean LAI
Grasses and Cereal Crops	91.53	1.20	91.6	1.07
Shrubs	92.66	1.41	96.4	0.92
Broadleaf Crops	74.03	2.09	80.1	1.17
Savanna	79.65	2.25	85.4	1.61
Broadleaf Forests	39.30	4.01	41.8	2.62
Needle Forests	54.54	3.99	41.8	1.66
	LANDSAT DATA			
	LANDSAT LUT		LASUR LUT	
Biome Type	Retrieval Index	Mean LAI	Retrieval Index	Mean LAI
Grasses and Cereal Crops	90.7	1.87	87.5	3.62
Broadleaf Forests	53.9	5.79	39.2	6.21
Needle Forests	57.9	4.11	4.7	3.39

** The algorithm was executed on LASUR data with LASUR LUT and Landsat LUT, and also on Landsat data with Landsat LUT and LASUR LUT. The corresponding Retrieval Index and mean LAI indicate the dependence of the algorithm (LUT) on spatial resolution.

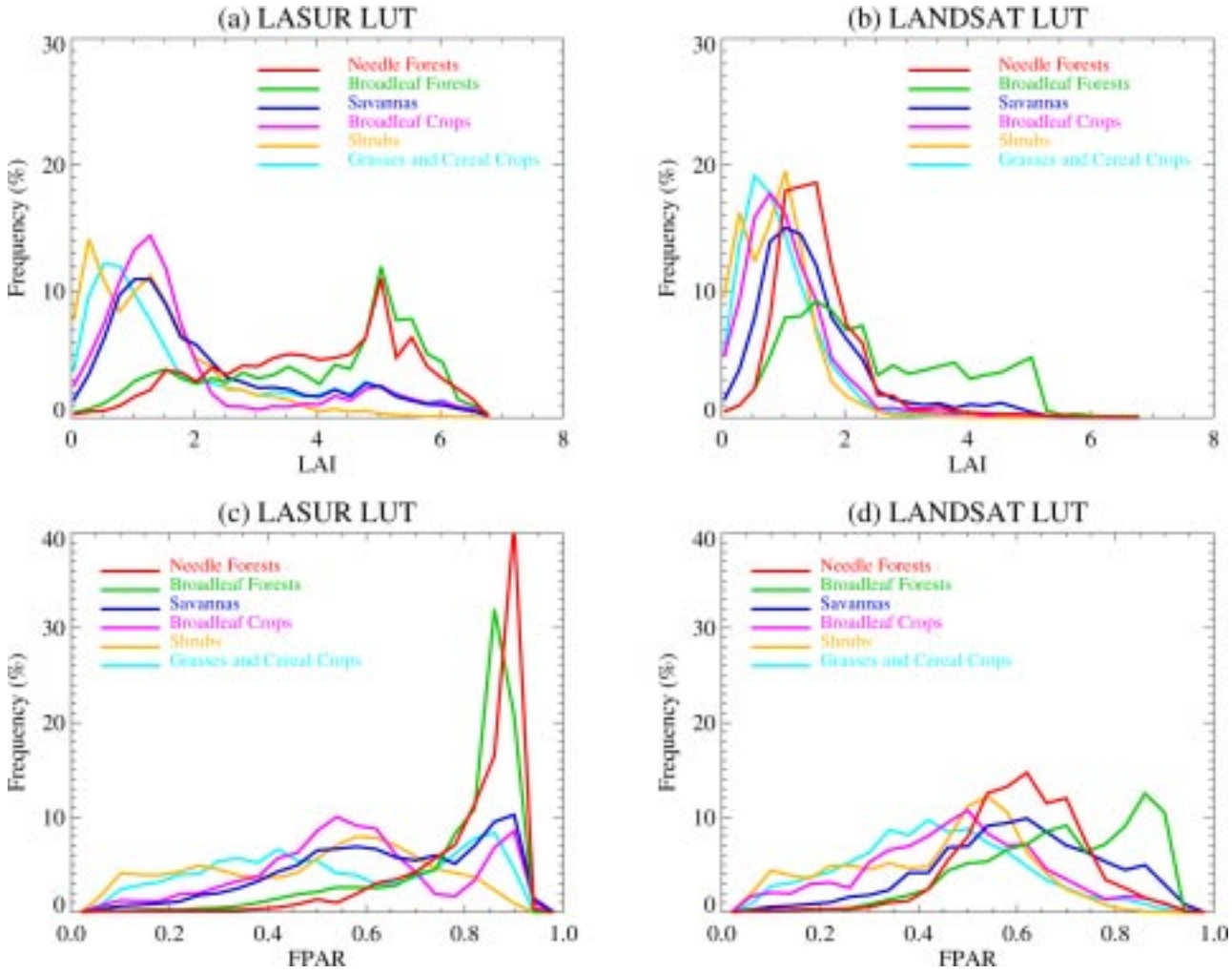


Figure 3-10 The retrievals from LASUR data as a function of spatial resolution dependent Look-Up-Table (LUT). Histogram of LAI from (a) LASUR LUT, (b) Landsat LUT. Histogram of FPAR from (c) LASUR LUT, (d) Landsat LUT.

2) Soil or Background Effects: As mentioned previously, in the design of the MODIS LAI/FPAR algorithm, the three-dimensional radiative transfer problem was split into two independent sub-problems. The first, the black soil problem, describes the radiation regime within the vegetation canopy for the case of a completely absorbing soil or background beneath the canopy. The second, the S problem, describe the radiation region of the vegetation canopy generated by anisotropic heterogeneous sources located at the canopy bottom and this characterizes the contribution of the soil or background to the canopy radiation regime. At fine resolutions, the contribution of the S problem is negligible in the case of dense vegetation such as forests. To test this assumption, we executed the algorithm using only the black soil problem on Landsat data. The RI can be as high as 50.6% (broadleaf forest) and 54.3% (needle forest) using just the black soil problem compared to 53.7% and 57.9% if the contribution from the S

problem is added. The histograms of retrieved LAI and FPAR do not change much. Thus, the fine resolution Landsat data represent pure biome types, dense vegetation in this instance, with minimal soil or background effects. For the coarse resolution LASUR data, however, the success index is only 31% and 45% for broadleaf and needle forests when only the black soil problem is used to retrieve LAI and FPAR. The soil problem is very important in this case, which can also be seen in the location of the data in the RED-NIR space, which results from a combination of the S problem and the black soil problem.

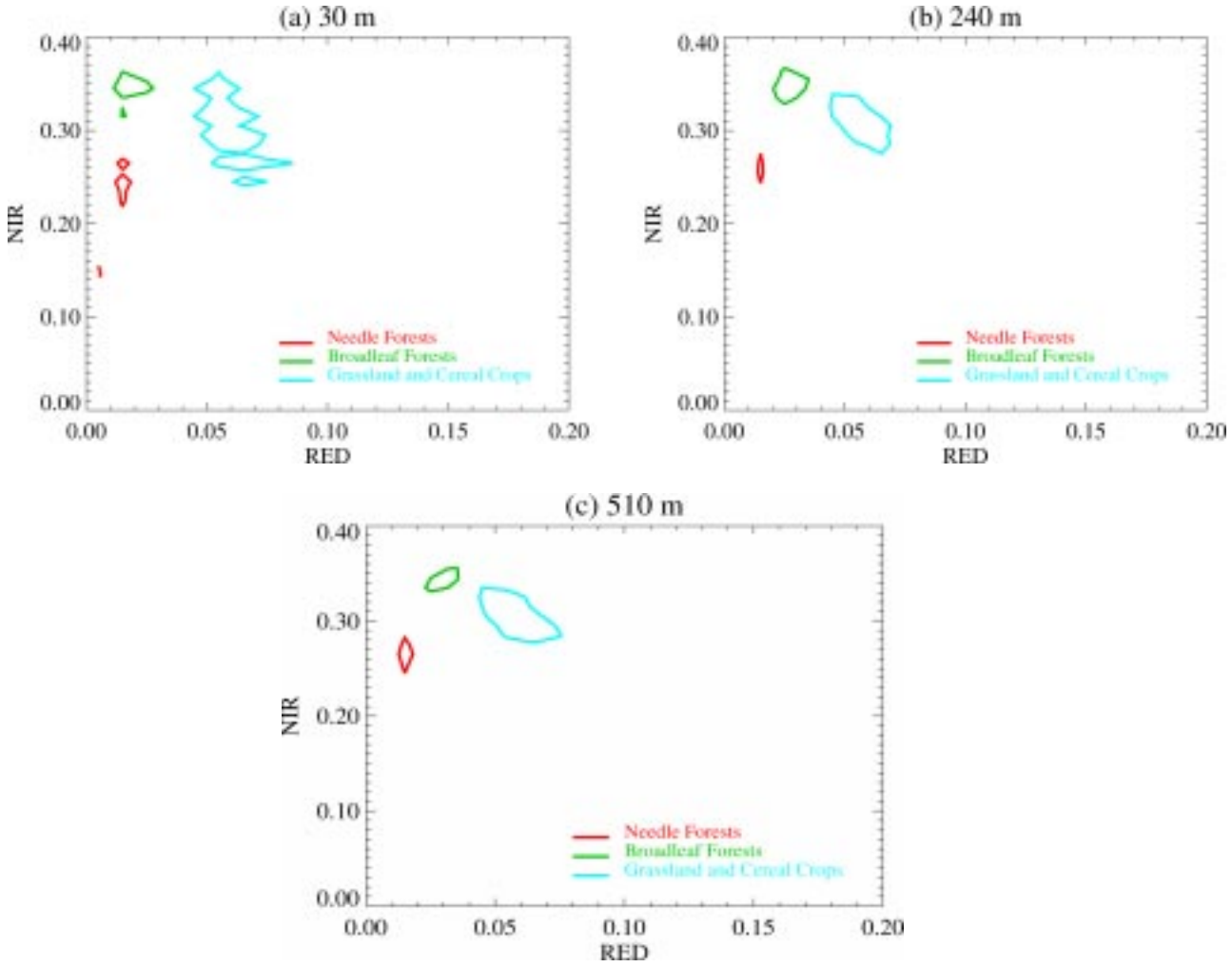


Figure 3-11 Same as Figure 3-2(d), 25% density contours of Landsat TM data on June 30, 1987 at (a) 30 m resolution, (b) 240 m resolution, and (c) 510 m resolution The 30 m resolution pixels were averaged to 240 m and 510 m resolution. A 30 m resolution landcover map was used as the baseline data set. The coarse resolution map were overlaid on the 30 m class map and the coarse grid cell was labeled based on the most frequently occurring cover type among the 30 m resolution pixels within that grid.

3) Spatial Resolution Effects: The impact of spatial resolution and aggregation of data on the MODIS LAI/FPAR algorithm was further investigated. The study focused mainly on broadleaf and needle forests. A 30 m resolution landcover map was used as the baseline data set. This map was used to label a series of coarser resolution maps of the scene produced by an aggregation procedure described below. The 30 m resolution pixels were averaged to 240 m and 510 m resolution. Each grid at the 240 m aggregation contained 64 30 m pixels. These were overlaid on the 30 m class map and the coarse grid cell was labeled based on the most frequently occurring cover type among the high resolution pixels within that grid. This procedure was performed for both resolutions. Here, we use two TM images of the same area, but from different seasons. One of these is the image mentioned previously (June 27, 1987) and the other is from September 30, 1987.

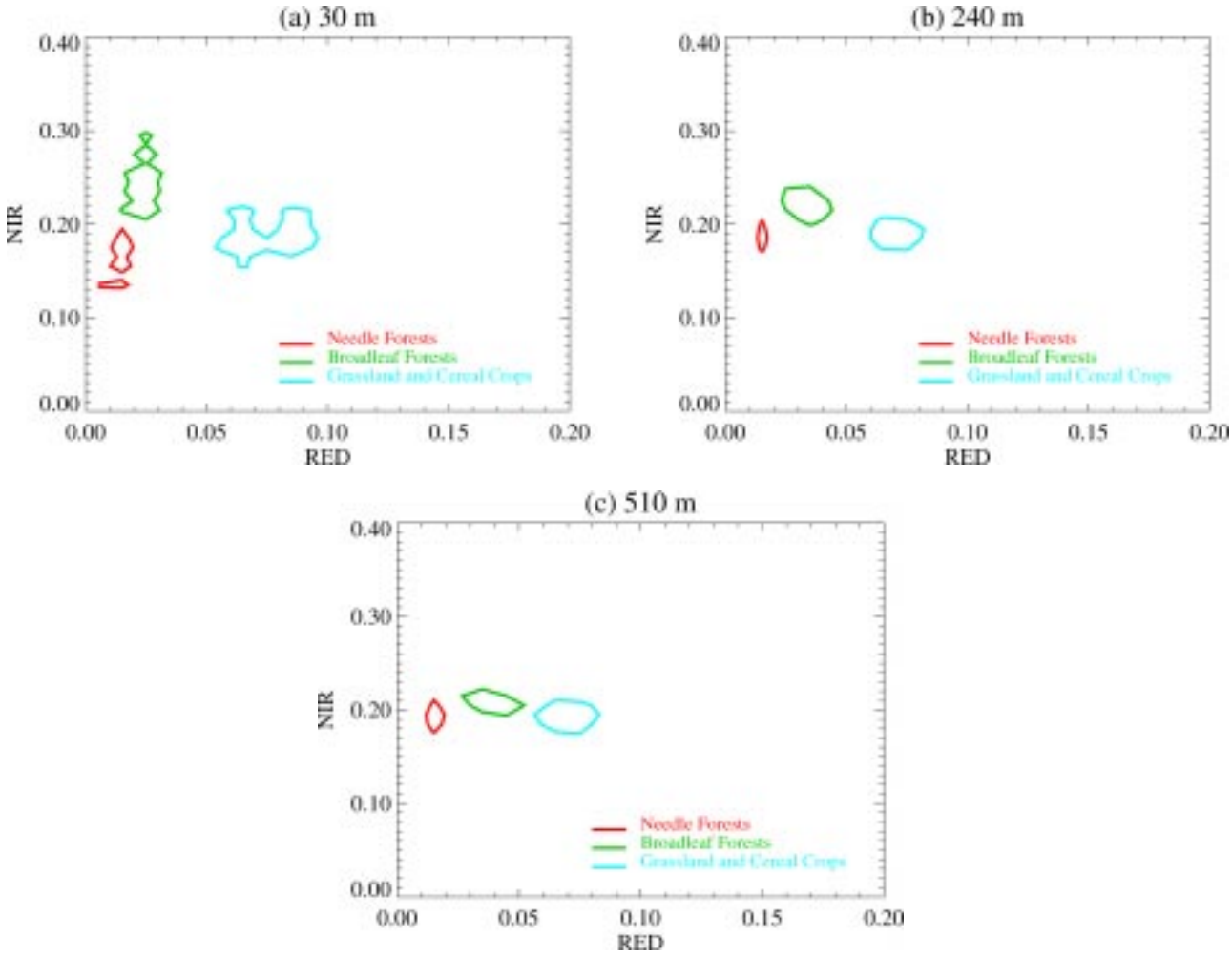


Figure 3-12 Same as Figure 3-11, data come from September 30, 1987.

From a spectral and NDVI point of view, the biomes change differently as the resolution decreases (Table IV(a) and (b)). Consider the June image. As resolution decreases, the RED reflectance of broadleaf forests increases and the NIR simultaneously decreases with the result that NDVI decreases. For needle forests, the RED as well as the NIR reflectance increases, consequently, NDVI is unchanged. In the RED-NIR space, these changes can be clearly seen (Figure 3-11). The distance between the three biome decreases, that is, the biomes become spectrally similar as resolution decreases. The September image shows similar changes (Figure 3-12), but the difference between the two images can be attributed to changing seasonality. The reason why needle forests exhibit such a large seasonal change is perhaps due to the understory [Cihlar et al., 1997]. It is important to note the distinct separation between the three biomes in both the images, suggesting that seasonal changes do not confound differences between the biomes. The reason why NIR reflectance in needleaf forests increases with decreasing resolution is their low NIR reflectance compared to the two other biomes. This is shown in Table I as well. Indeed, the spectral properties of coarse resolution data aggregated from fine resolution are also influenced by landscape characteristics. But, in general, the RED reflectance increases with the result that the biomes that tend to move towards the soil line in the RED-NIR space.

TABLE 3-4 MEAN VALUES OF RED, NIR, NDVI AND LAI AT 30 m, 240 m AND 510 m RESOLUTION FOR LANDSAT TM DATA ON (a) JUNE 30, 1987, AND (b) SEPTEMBER 30, 1987

	30 m * 30 m			
Biome Type	Mean Red	Mean NIR	Mean NDVI	Mean LAI
Grasses and Cereal Crops	0.0655	0.304	0.635	1.87
Broadleaf Forests	0.0215	0.348	0.881	5.79
Needle Forests	0.0131	0.200	0.886	4.11
	240 m * 240 m			
Biome Type	Mean Red	Mean NIR	Mean NDVI	Mean LAI
Grasses and Cereal Crops	0.0646	0.295	0.635	1.795
Broadleaf Forests	0.0275	0.346	0.853	5.385
Needle Forests	0.0141	0.232	0.889	4.488
	510 m * 510 m			
Biome Type	Mean Red	Mean NIR	Mean NDVI	Mean LAI
Grasses and Cereal Crops	0.0640	0.291	0.635	1.756
Broadleaf Forests	0.0307	0.342	0.835	5.282
Needle Forests	0.0158	0.246	0.883	4.584

(a)

	30 m * 30 m			
Biome Type	Mean Red	Mean NIR	Mean NDVI	Mean LAI
Grasses and Cereal Crops	0.0697	0.208	0.492	1.261
Broadleaf Forests	0.0310	0.234	0.763	2.438
Needle Forests	0.0138	0.154	0.842	2.365
	240 m * 240 m			
Biome Type	Mean Red	Mean NIR	Mean NDVI	Mean LAI
Grasses and Cereal Crops	0.0718	0.200	0.1472	1.043
Broadleaf Forests	0.0378	0.225	0.712	1.889
Needle Forests	0.0149	0.166	0.845	2.619
	510 m * 510 m			
Biome Type	Mean Red	Mean NIR	Mean NDVI	Mean LAI
Grasses and Cereal Crops	0.0711	0.199	0.434	1.028
Broadleaf Forests	0.0415	0.219	0.683	1.630
Needle Forests	0.0164	0.176	0.837	2.681

(b)

** The 30 m resolution pixels were averaged to 240 m and 510 m resolution. A 30 m resolution landcover map was used as the baseline data set. The coarse resolution map were overlaid on the 30 m class map and the coarse grid cell was labeled based on the most frequently occurring cover type among the 30 m resolution pixels within that grid.

The MODIS LAI/FPAR algorithm was executed with the above mentioned multiple resolution data and the 30 m LUT. The results indicate that as resolution gets coarser, the mean LAI of broadleaf forests decreases, but increases in needle forests. This is in agreement with the changes seen in the spectral data discussed previously.

4 Validation Plan

4.1 Introduction

In July, 1999, NASA will launch the Terra platform, a keystone of its Earth Observing System (EOS). Terra's five instruments include highly evolved successors to current satellite sensors (e.g., MODIS vs. AVHRR) and innovative experimental sensors (e.g., MISR). Together, the many near real-time products (e.g., leaf area index (LAI)) from Terra will provide the most comprehensive view of the Earth system to date. Moreover, the co-aligned, calibrated sensors will present the best opportunity yet for validation of remote sensing products and algorithms.

Nevertheless, global validation of land remote sensing products is complicated by multiple factors, including difficulty in measuring land surface variables over the size of a satellite pixel, inherent errors in satellite data calibration, georegistration, cloud screening and atmospheric correction, and the impracticality and expense of collecting field data over a large number of different ground/atmosphere combinations over sufficiently long time scales. These constraints have therefore limited rigorous validation efforts to several "intensive field campaigns" (e.g., FIFE, BOREAS, HAPEX) where large teams were able to collect data over reasonably large areas during discrete phenological periods. While these efforts were necessary to show that various remote sensing algorithms were working correctly, they were not sufficient to truly validate a land product or algorithm, at least not for global, year-round application.

Despite these difficulties, the EOS Project has charged its instrument teams with both the development and validation of their operational products. In 1998, the Project augmented those efforts by funding 65 competitively-selected Validation Investigators. Approximately 44 of these will rely on *in situ* measurements and comparisons with EOS products. Thus, a close dialogue is developing among field data collectors, EOS algorithm developers, and the end-user community to ensure that the validation data are collected and packaged appropriately for greatest effectiveness. Particularly notable are the SWAMP Validation Workshop in 1997 [Justice et al. 1998] and a 1998 follow-up workshop for LA1 and FPAR (canopy-absorbed radiation) validation [Privette et al. 1998].

4.2 Approach

Multiple validation techniques will be used to develop uncertainty information on EOS land products. The methods include comparisons with *in situ* data, comparisons with data from airborne and other spaceborne sensors (e.g., AVHRR, GOES), analysis of trends in products (e.g., spatial, temporal), and analysis of process model (e.g., climate model) results which are driven or constrained by EOS products. Successful validation will have been accomplished if timely and accurate product uncertainty information becomes routinely available to the product users within two years after Terra's launch.

4.3 Validation Sites

EOS products will be generated operationally for all global land areas. Validation must therefore include attention to a wide range of combined surface cover and atmospheric conditions, from tundra to deserts to tropical jungles. Clearly, the costs associated with such a program can be tremendous. In an effort to contain costs, EOS will use a variant of the Global Hierarchical Observing Strategy [GCOS/GTOS, 1997], a multi-tiered categorization of field site measurement capabilities and intensity [Suttles et al. 1996]. This categorization has an inverse number of sites in a tier relative to the measurement intensity per site. Thus, EOS will rely on few intensive field campaigns (e.g., LBA, SAFARI 2000) but on a large number of sites for which only high resolution satellite scenes ate regularly available. In this article, we focus on two strata of the hierarchy: Core Sites and Product Sites.

4.3.1 Core Sites

EOS will concentrate much of its effort around EOS Land Validation Core Sites (Figure 4-1). These 24 sites represent a consensus among the instrument teams and validation investigators and span a range of global biome types. The sites typically have a history of *in situ* and remote observations, and can expect long-term preservation. Each is nominally 100 km x 100 km in size. In most cases, a Core Site has a tower on which above-canopy instrumentation will be mounted to provide near-continuous sampling of landscape radiometric, energy flux and/or meteorological variables. Each site will also host a sunphotometer from NASA's AERONET Program for assessment of aerosol optical depth. Episodic sampling of more slowly changing land parameters (e.g., LAI) will compliment the ongoing measurements.

Significant effort has been placed on ensuring the early acquisition and open availability of relevant satellite data for Core Sites. Specifically, the ASTER and Landsat 7 teams have incorporated Core Sites into their priority scene acquisition plans, and arrangements have been made for independent archiving of ASTER, MISR, MODIS, CERES and MOPITT data. These data will be place in special archives within the EOS Data Active Archive Centers (DAACs), and be available through both traditional ordering systems and from unique Core Site WWW pages (Figure 4-2). Limited historical AVHRR and Landsat TM data will also be available, and plans are being made to include EO-1 Hyperion data as well. The Core Site data archives will thus contain the richest and most diverse colocated data sets available through EOS. These benefits are expected to facilitate both validation and early EOS science.

4.3.2 Product Sites

EOS Product Sites will provide both diversity and redundancy to the Core Sites. In contrast to the Core Sites, a given Product Site may only be used to evaluate one EOS product [Privette et al. 1998]. The Land Cover Change, Land Surface Temperature, and Snow/Sea-ice products are sufficiently unique that their validation will occur primarily at Product, rather than Core, Sites.

Despite these efforts, additional resources clearly are needed for global validation than can be supplied by EOS alone. Thus, MODLAND has applied significant effort to developing EOS-wide validation protocols and encouraging the participation of community dab collectors and

product users. In particular, significant collaboration is planned with existing measurement networks, including FLUXNET, Long Term Ecological Research (LTER) sites, AERONET, and Global Land Cover Test Sites (GLTCS). In most cases, the EOS Core Sites are members of at least one of these networks. In addition, EOS is investigating the potential usefulness of other networks, such as BSRN and SurfRad.

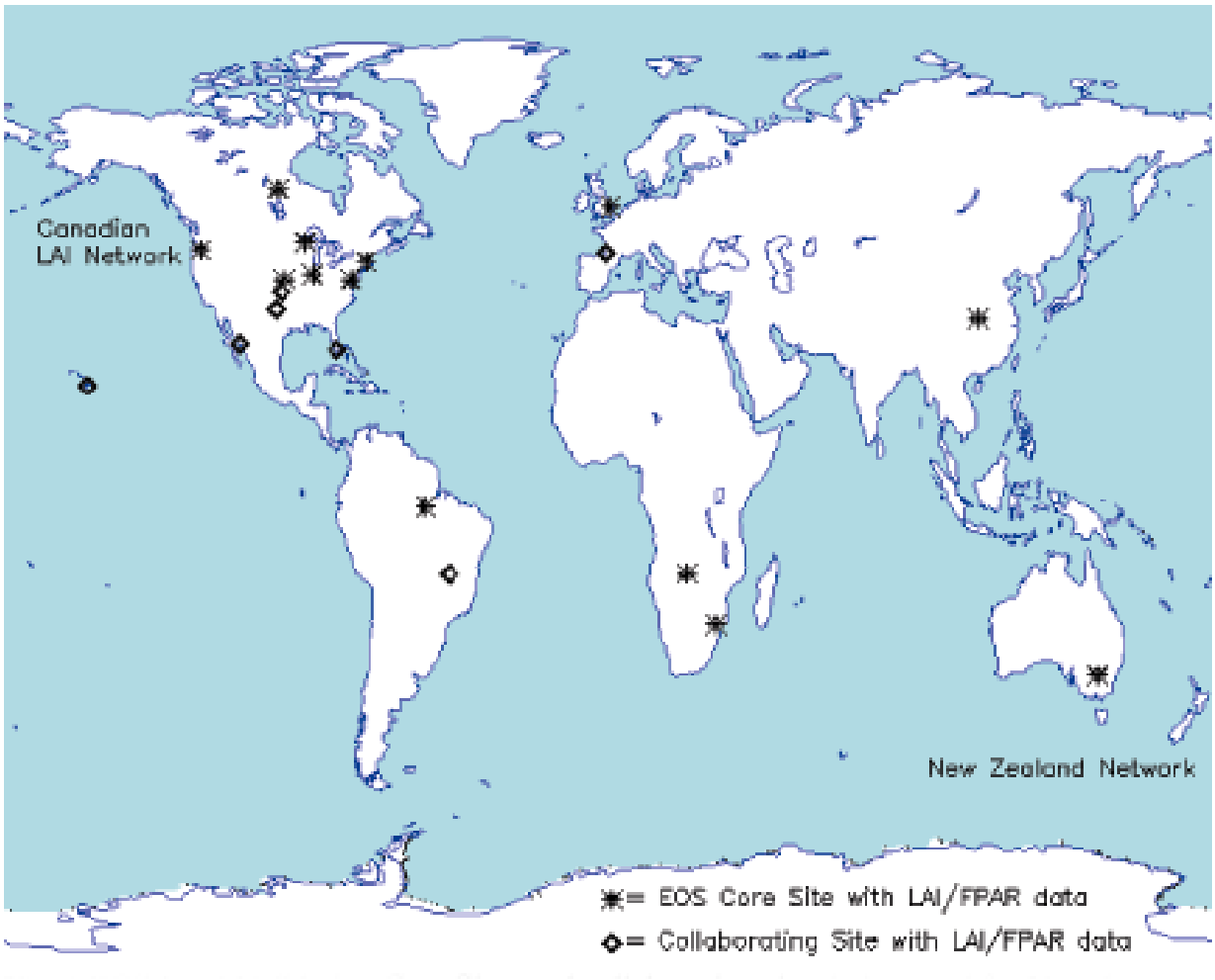


Figure 4-1 EOS Land Validation Core Sites and collaborating sites being used for Year 1 LAI/FPAR product validation.

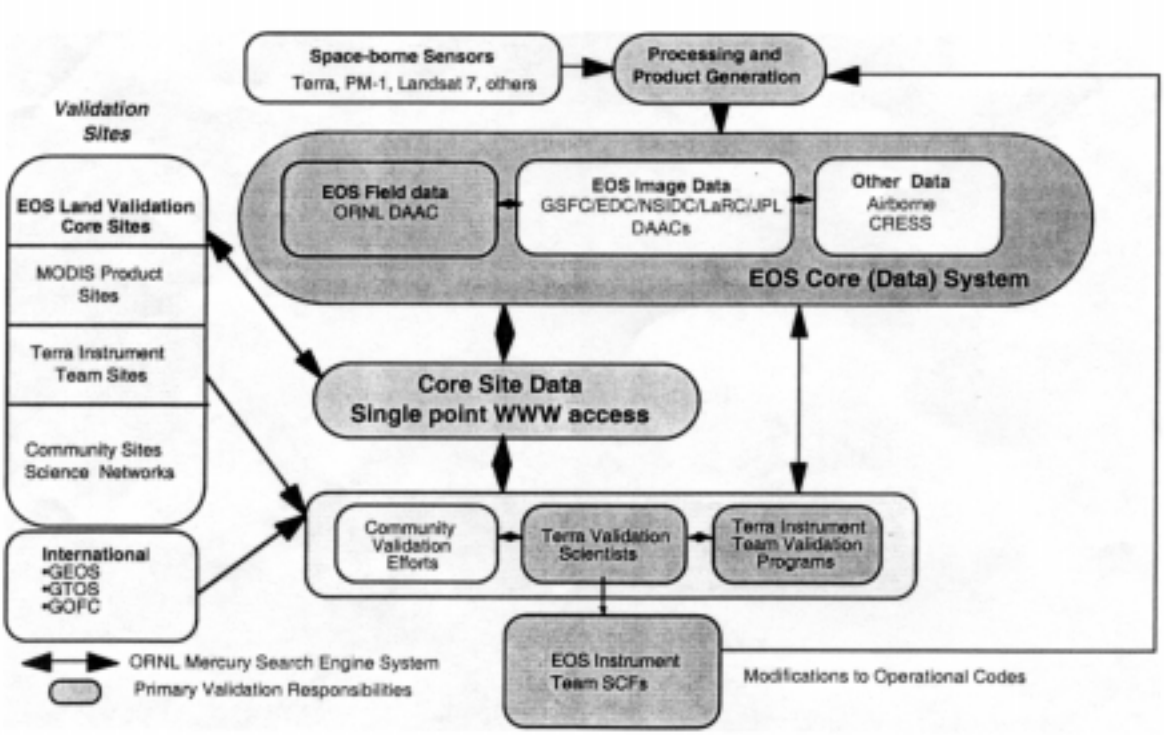


Figure 4-2 Schematic of EOS validation data resources and pathways.

Global validation requires field data from a range of sites representing a logical subset of the Earth's land covers. The EOS Land Validation Core Sites should be emphasized for this purpose (see URL in Table 4-3). These sites are foci for EOS AM and Landsat 7 land validation activities, and are high priority data acquisition and product generation targets. They are expected to facilitate both validation and early EOS science. The sites typically have a history of in-situ and remote observations, and can expect long-term preservation. Centralized WWW-based archiving of ASTER, MISR, MODIS and Landsat 7 ETM+ products in relatively easy-to-use formats are planned for these sites.

Because the Core Site network is in its infancy, LAI and FPAR validation is planned only at sites for which firm commitments to data collection have been secured. However, several "product-specific" validation sites volunteered to collect data. Those, together with the participating Core sites, are listed in Table 4-1. The list provides at least two representatives from each of the six vegetated biome types recognized by the MODLAND LAI algorithm. Participants agreed that initially at least seasonal (4 times/yr) in situ LAI/FPAR assessments were required for product validation.

4.4 Auxiliary Measurements

Although product validation can be conducted with only LAI and FPAR field data, a prioritized list of ancillary measurements needed for algorithm validation was developed. Not all

listed variables are necessary for this task, however any such measurements would be useful. The measurements are shown in Table 4-2 in order of decreasing importance.

Table 4-1 Validation Sites for LAI/FPAR Product

Name	Country	Biome	Core Site	Investigator
USDA BARC, MD	U.S.	broadleaf cropland	X	Liang
Bondville, IL	U.S.	broadleaf cropland	X	BigFoot
Gainesville, FL	U.S.	broadleaf cropland		Craig
Tapajos	Brazil	broadleaf forest	X	Asner
Hawaii	U.S.	broadleaf forest		Asner
Harvard Forest, MA	U.S.	broadleaf forest	X	BigFoot
Park Falls, WI	U.S.	broadleaf forest	X	Gower
Uardry	Australia	grassland	X	Hook
Osage, OK	U.S.	grassland		Walter-Shea
Konza, KS	U.S.	grassland	X	BigFoot
East Anglia	England	grassland	X	Barnsley
Vernon,TX	U.S.	grassland		Asner
BOREAS NSA	Canada	needleleaf forest	X	BigFoot
Cascades, OR	U.S.	needleleaf forest	X	Law
EMATREF	France	needleleaf forest		Roujean
Yaqui Valley	Mexico	shrubland		Asner
San Pedro Basin/SALSA, AZ	U.S.	shrubland	X	Qi
Skukuza South	Africa	shrubland/woodland	X	Privette
New Zealand Network	New Zealand	various		Brown
Canada Network	Canada	various		Chen
Mongu	Zambia	woodland	X	Privette
Cerrado	Brazil	woodland		Asner

Table 4-2 Land Cover Variables

Land Cover Variable or Characteristic
canopy multispectral reflectance (nadir or bidirectional)
leaf spectra (reflectance and transmittance)
background nadir spectral reflectance (soil + litter)
fraction of areal vegetation cover
vegetation crown allometry (height, width, gap)
phenology (green-up, mature, senescent stage)
vegetation composition (either by species or structural type)
wet or dry status
fraction of non-photosynthesizing vegetation (at min. photosynthetic activity stage)
meteorological data (minimum set)

4.5 Scaling

A pervasive problem for land validation is the scaling of field measurements to the more coarse resolution of satellite products. Although various schemes will be used, the newly begun BigFoot project will focus directly on EOS scaling issues. The BigFoot approach will include overlaying grids of 25 m and 1 km, extending to the 5 km x 5 km "MODIS grid," at FLUXNET tower sites. BigFoot will initially focus effortsat four EOS Core Sites. Investigators will measure and scale LAI, FPAR, net primary productivity (NPP) and landcover maps to appropriate resolutions for EOS validation. A combined program of field data collection, aircraft overflights and fine resolution satellite image acquisitions will be used. BigFoot will also attempt to characterize and parameterize the -relationship between the measured net ecosystem exchange values and the NPP product. This pathfinding activity will test various scaling methodologies and work with MODLAND to develop a WWW site outlining a recommended strategy.

4.6 Modland QUick Airborne LookS (MQUALS)

To facilitate BigFoot's and others' scaling efforts, MODLAND developed the MODLAND Quick Airborne Looks (MQUALS). This program is based on an aircraft remote sensing package that includes three digital cameras (red, blue and near-infiared,640x480 pixels), shortwave and near-infraredalbedometers, a calibrated radiometer with four MODIS spectral bands, and an optional thermal radiometer. All data are simultaneously collected and stored on a laptop computer. A ground-based 4-band radiometer will be used to simultaneously measure irradiance. The complete package was designed to be shipped easily to small aircraftoperators near validation sites for low cost site and reflectance characterization. Initial MQUALS products should be available within seven days of data collection. A duplicate MQUALS package will operate in southern Afiicaas part of SAFARI 2000. Within the greater EOS validation

framework, MQUALS will be complimented by remote observations from sensors on NASA's high altitude ER-2 and a light aircraft package developed by the CERES team.

4.7 Data Protocols and Dissemination

Successful validation will in part depend on easy access to accurate and documented field data. The MODIS Land Team has worked extensively with the Oak Ridge. (ORNL) and EROS Data Center DAACs to develop validation data protocols and pathways. Based on distributed WWW mini-archives, the ORNL-based Mercury system will conduct daily data set cataloging and can provide single point access to all EOS validation data (Figure 4-2), The minimal effort required to interface with Mercury should allow rapid data submission and public release. This evolving data system accommodates a diverse user community in which EOS instrument teams, Validation Investigators, and independent investigators are simultaneously collecting, archiving, distributing and using validation and remote sensing data. Moreover, the ORNL DAAC has initiated a vigorous effort to mine historical field data. These data will be used to establish expected values and reasonable ranges of EOS products in some cases.

EOS land product validation is being planned as part of a long-term implementation plan. Initial validation efforts will both estimate product accuracy and prototype validation scheme components. It is hoped the components currently planned, upon successful post-launch evaluation, will be substantially extended to provide more rigorous and comprehensive product evaluation.

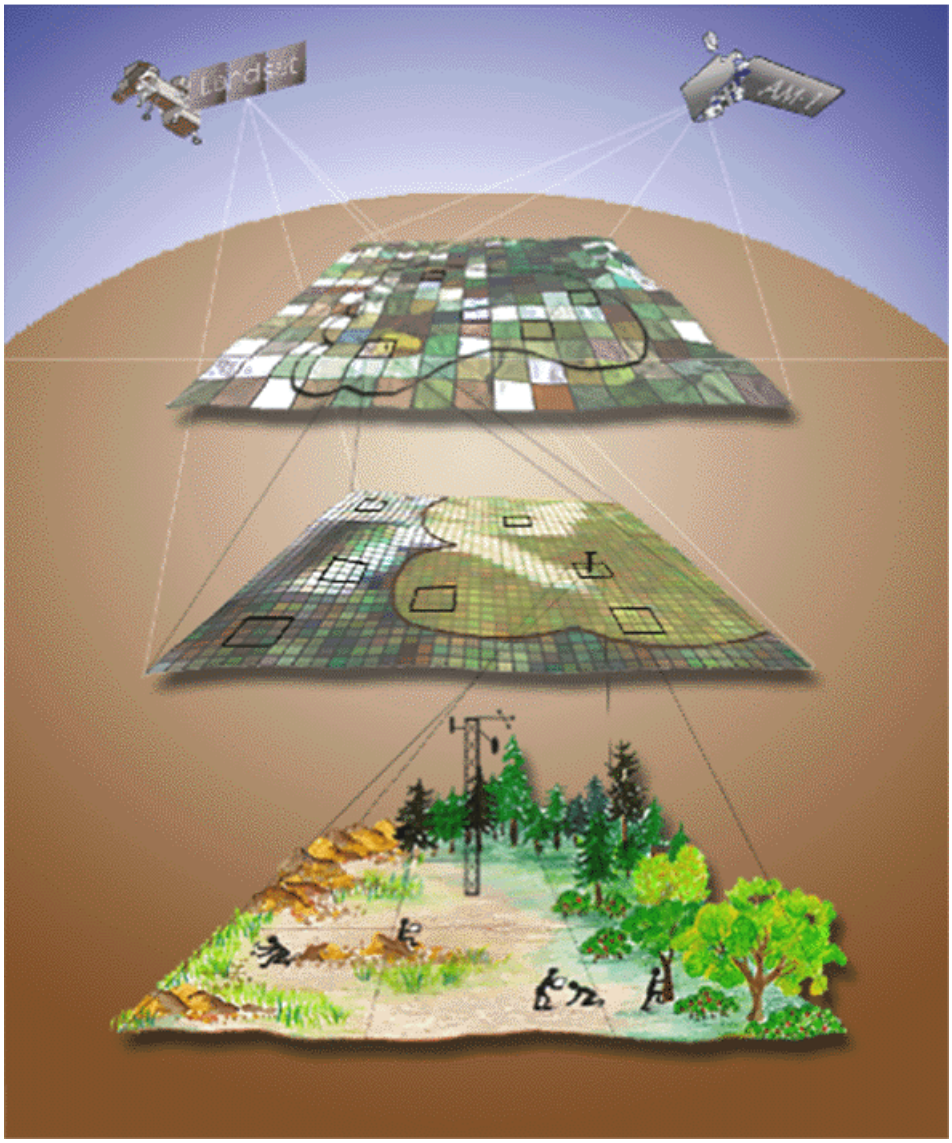


Figure 4-3 Conceptual design for scaling in the BigFoot project. See URL in Table 4-3

We anticipate that the validation procedures started by the EOS instrument teams and validation investigators will act as a catalyst for broader involvement by the research community in product evaluation. Clear protocols for data collection and WWW archives and access will give all researchers a simple mechanism for participation. With the recent increase and planned launch of new moderate resolution sensors (e.g., VEGETATION, MODIS AM/PM, GLI, NPP, NPOESS) by different space agencies and the increased availability of higher order standard products, the benefits of standard measurement protocols and validation site data sharing are considerable. The CEOS Calibration/Validation Working Group is an obvious mechanism to expand the early developments and lessons learned in EOS land validation into a truly global validation initiative.

Table 4-3. WWW Site Addresses

Site	URL
BigFoot	www.fsl.orst.edu/spacers/bigfoot/plan.html
Committee for Earth Observation Satellites	ceos.esrin.esa.it
EOS Validation Program	eospso.gsfc.nasa.gov/validation/valpage.html
FLUXNET	daacl.ESD.ORNL.gov/FLUXNET
Global Climate Observing System	www.wmo.ch/web/gcos/gcoshome.html
Global Observations of Forest Cover	www.ccrs.nrcan.gc.ca/ccrs/tekrd/internet/gofc/gofce.html
Global Terrestrial Observing System	www.fao.org/gtos/Home.htm
Long Term Ecological Research	lternet.edu
EOS Land Validation Core Sites	modarch.gsfc.nasa.gov/MODIS/LAND/VAL/core_sites.html
MODLAND Validation	modarch.gsfc.nasa.gov/MODIS/LAND/VAL
Myneni's LAI/FPAR Site	cybele.bu.edu/research/modismisr/
ORNL DAAC Validation Site	www-eosdis.ornl.gov/eos_land_val/valid.html

4.8 Proposed Validation Tests

The following five tests are proposed to be carried out in order to validate LAI/FPAR algorithm against ground measurements. All the tests are designed for validating biome specified ground measurements, therefore, biome type information of data is a necessity.

4.8.1 TEST 1 : (LAI/FPAR TEST)

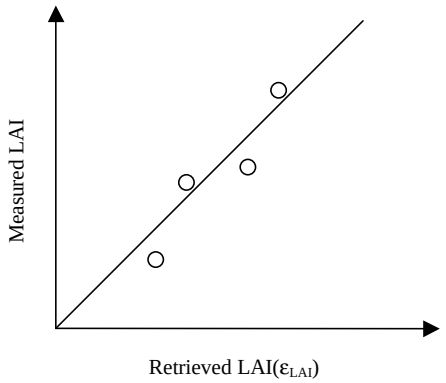
Given spectral $<HDRF_\lambda>$ or $hdrf_\lambda$ run the LAI/FPAR algorithm for different values of ε (0.05; 0.1; 0.15; 0.20; 0.25; 0.30) to get LAI and FPAR values as a function of ε. Find ε_{LAI}, ε_{FPAR}, and ε_{TOT} which minimize the expressions

$$\delta_{LAI}(\varepsilon_{LAI}) = \min\left\{\sum_{i=1}^{N_{mes}}\left[LAI_i - LAI_i(\varepsilon)\right]^2\right\},$$

$$\delta_{FPAR}(\varepsilon_{FPAR}) = \min\left\{\sum_{i=1}^{N_{mes}}\left[FPAR_i - FPAR_i(\varepsilon)\right]^2\right\},$$

$$\delta_{LAI}(\varepsilon_{LAI}) = \min\left\{\delta_{LAI}(\varepsilon) + \delta_{FPAR}(\varepsilon)\right\},$$

where LAI_i and $FPAR_i$ are measured values of LAI and FPAR; N_{mes} is number of measurements.



Plot $LAI_i(\varepsilon_{LAI})$, $LAI_i(\varepsilon_{TOT})$ and $FPAR_i(\varepsilon_{FPAR})$, $FPAR_i(\varepsilon_{TOT})$ versus measured values of LAI and FPAR.

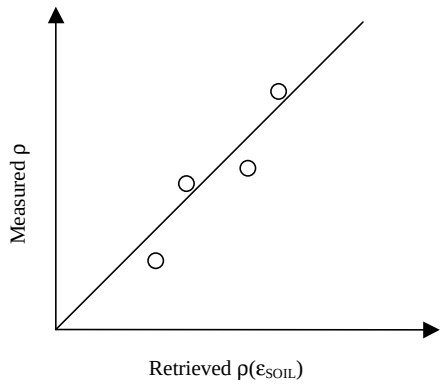
Figure 4-4 Validation test 1: LAI/FPAR test.

4.8.2 TEST 2 : (SOIL TEST)

Given spectral <$HDRF_\lambda$> or $hdrf_\lambda$ run the LAI/FPAR algorithm for different values of ε (0.05; 0.1; 0.15; 0.20; 0.25; 0.30) to get the effective ground reflectance as a function of ε.
Find ε_{SOIL} which minimizes the expression

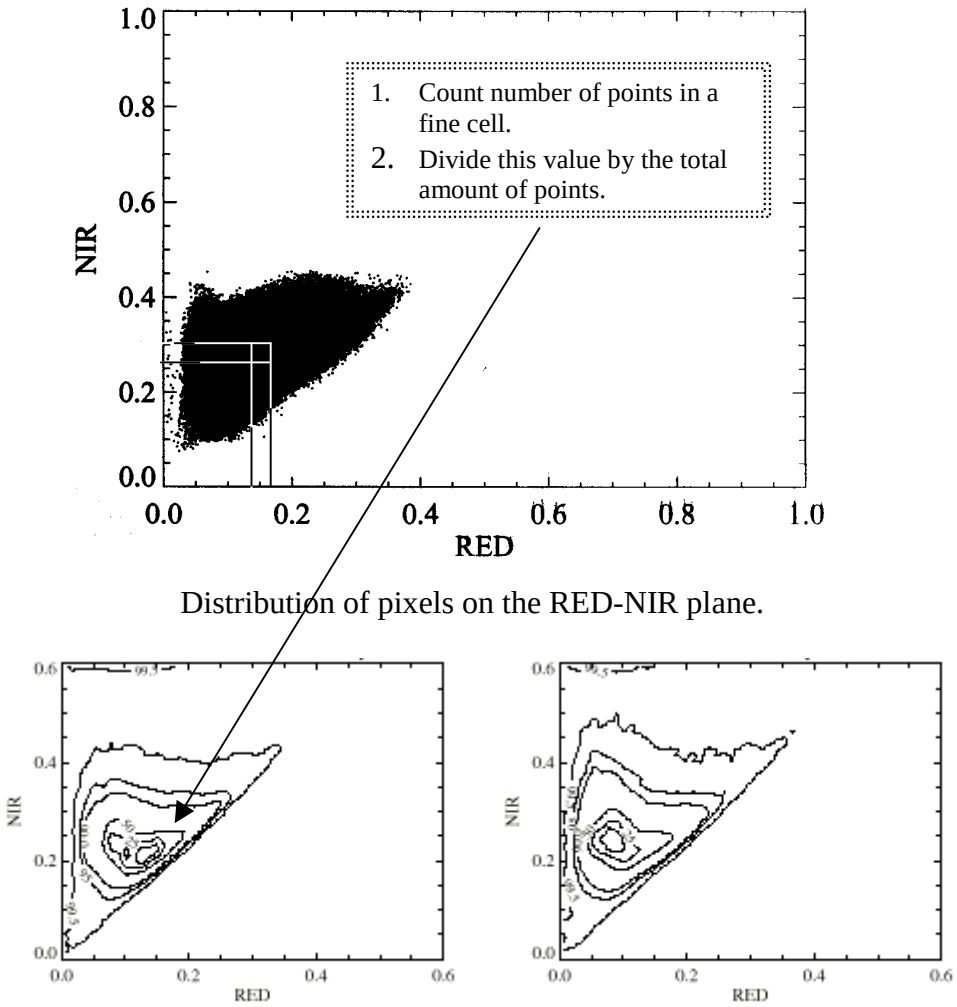
$$\delta_{SOIL}(\varepsilon_{SOIL}) = \min\sum_{i=1}^{N_{mes}}\left[\rho_i - \rho_i(\varepsilon)\right]^2,$$

where ρ_i is measured value of the effective ground reflectance; N_{mes} is number of measurements.



Plot $\rho_i(\varepsilon_{LAI})$ versus measured values of effective ground reflectance.

Figure 4-5 Validation test 2: Soil test.



Distribution of pixels on the RED-NIR plane.

The point density distribution function.

Figure 4-6 Validation test 3: Reflectance density plots.

4.8.3 TEST 3: (RED-NIR DENSITY PLOTS)

Derive the density distribution function from the distribution of pixels with respect to their reflectances at NIR and Red wavelengths using MODIS data and field measured data. Compare locations of areas containing the points of high density.

4.8.4 TEST 4 : (NDVI-LAI and FPAR-NDVI relationships)

Derive NDVI-LAI and FPAR-NDVI regression curves using field measured canopy reflectances, MODIS data and the LAI/FPAR algorithm and compare them.

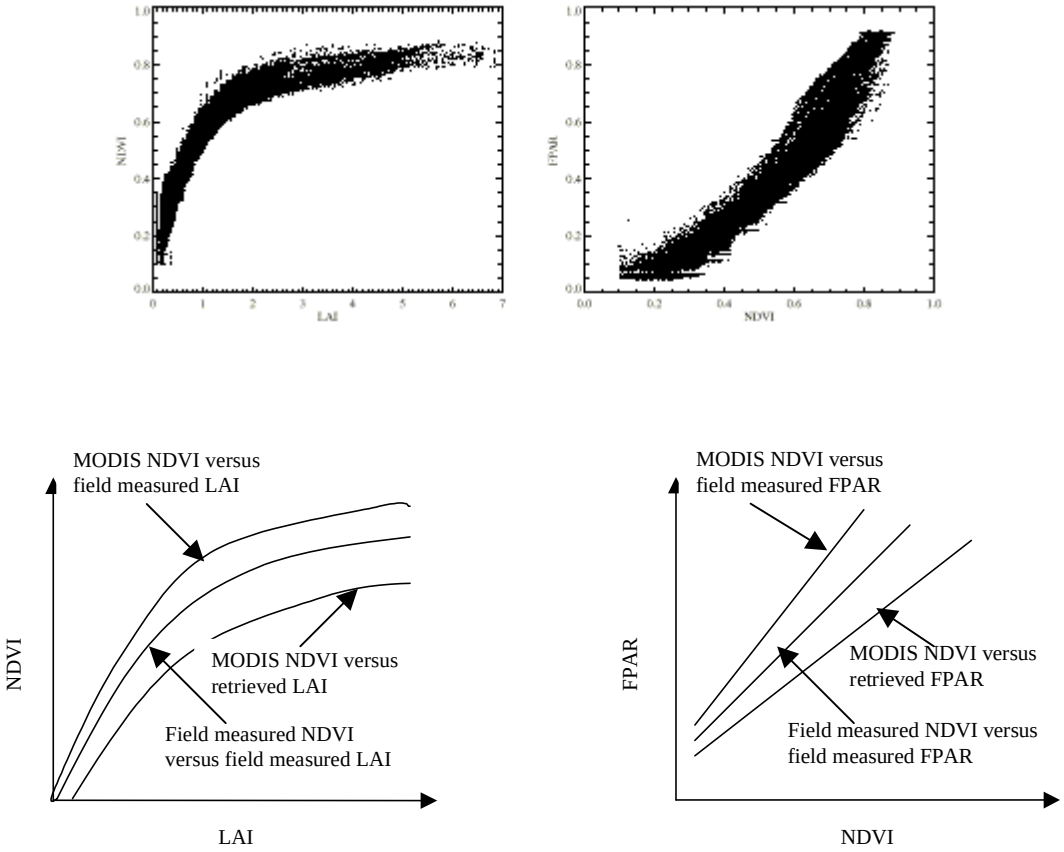
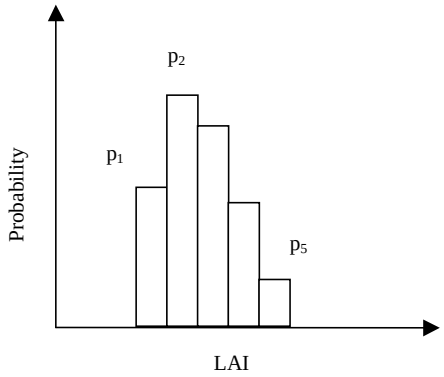


Figure 4-7 Validation test 4: NDVI-LAI and FPAR NDVI relationships.

4.8.5 TEST 5 : (INFORMATION CONVEYED ABOUT LAI BY HDRF)

Steps:

1. Find LAI distribution function using all available values of measured LAI.

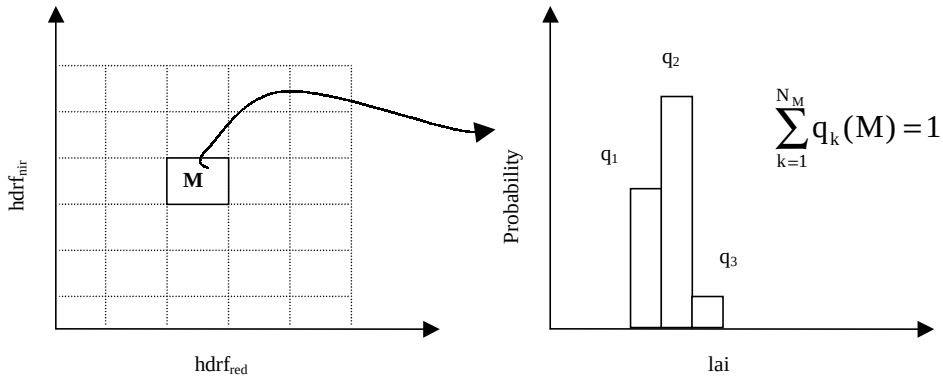


$$\sum_{k=1}^{N_{LAI}} p_k = 1$$

N_{LAI} is number of measurements

Figure 4-8 Validation test 5: LAI distribution function.

2. Evaluate the entropy: $E_0 = -\sum_{k=1}^{N_{LAI}} p_k \log p_k$.
3. Evaluate the conditional LAI probability function *F*(lai|M), where M is the set of spectral canopy reflectances.



$$\sum_{k=1}^{N_M} q_k(M) = 1$$

Figure 4-9 Validation test 5: LAI probability function.

Example of the evaluation of the conditional probability function using canopy reflectances at Red and NIR wavelength.

4. Evaluate the joint LAI-lai(M) probability function, i.e.,

$$p_{i,j}(M) = \mathrm{Prob}(LAI = LAI_i \text{ and } lai(M) = LAI_j).$$

5. Evaluate the "information conveyed about LAI by HDRF" as:

$$E(M) = E_0 + \sum_{i=1}^{N_{LAI}} \sum_{j=1}^{N_M} p_{i,j} \log \frac{p_{i,j}(M)}{q_j(M)}.$$

6. Do the same for retrieved values of LAI.
7. Compare results.

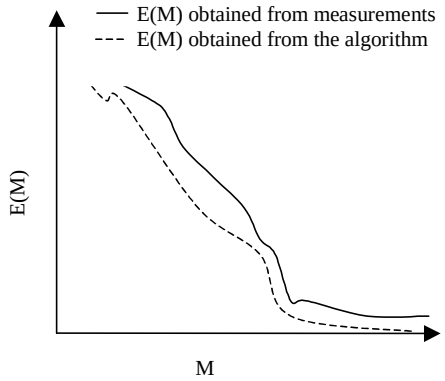


Figure 4-10 Validation test 5: Information convey curves.

5. Ancillary Data

5.1 At-Launch Land Cover Classification

As the earlier discussion indicates, an accurate land cover map is a pre-requisite for choosing the appropriate relation between surface parameters (LAI and FPAR) and the satellite derived reflectances. Global land cover maps in a 1-km resolution are currently available from University of Maryland (UMD) [Hansen et al., 1999] and the Earth Resources Observation System (EROS) Data Center (EDC) [Loveland et al., 1995].

Both UMD and EDC follow the International Geosphere Bioshere Program (IGBP) Classification logic [Loveland et al., 1997, Belward et al., 1996]. The 17 land cover classes defined by IGBP as well as the corresponding class labels and names in the UMD map are shown in table 5-1 below. Note that the UMD scheme does not include permanent wetlands, cropland mosaics and snow/ice and uses a somewhat different terminology. For a full description of the biome classes refer to section 2.

Table 5-1: IGBP land cover classification scheme.

	IGBP Classes	**UMD-IGBP Classes**	**Ground Cover**	**Canopy Height**	**Description**
1	Evergreen Needleleaf Forest (ENF)	Evergreen Needleleaf Forest	> 60\%	> 2m	woody, green year-round
2	Evergreen Broadleaf Forest (EBF)	Evergreen Broadleaf Forest	> 60\%	> 2m	woody, green year-round
3	Deciduous Needleleaf Forest (DNF)	Deciduous Needleleaf Forest	> 60\%	> 2m	Woody, shed leaves during dry season
4	Deciduous Broadleaf Forest (DBF)	Deciduous Broadleaf Forest	> 60\%	> 2m	Woody, shed leaves in annual cycle
5	Mixed Forest (MXF)	Mixed Forest	> 60\%	> 2m	woody, needleleaf/broadleaf mixture, neither component > 60\%
6	Closed Shrubland (CSH)	Closed Shrubland	> 60\%	< 2m	woody, herbaceous understory, evergreen or deciduous
7	Open Shrubland (OSH)	Open Shrubland	< 60\%	< 2m	woody, sparse herbaceous understory, evergreen or deciduous
8	Woody Savannas (WSA)	Woodland	30-60\%	> 2m	tree/shrub, herbaceous understory, evergreen or deciduous
9	Savannas (SAV)	Wooded Grassland	10-30\%	> 2m	tree/shrub, herbaceous understory, evergreen or deciduous
10	Grasslands (GRL)	Grassland	<10\%	< 2m	herbaceous
11	Permanent Wetlands (PWL)		> 60\%		water mosaic, herbaceous/woody, salt, brackish or fresh water
12	Croplands (CRL)	Croplands	> 60\%	< 2m	broadleaf crops, cereal crops
13	Urban and Built-Up (URB)	Urban and Built-Up			man-made structures, buildings
14	Cropland Mosaics (CRM)				croplands/nat. vegetation mosaic, neither component > 60\%
15	Snow and Ice (SNI)				snow/ice covered most of the year
16	Barren/Sparsely Vegetated (BSV)	Bare Ground			exposed soil, sand, rocks
17	Water Bodies (WAT)	Water			oceans, lakes, reservoirs, rivers

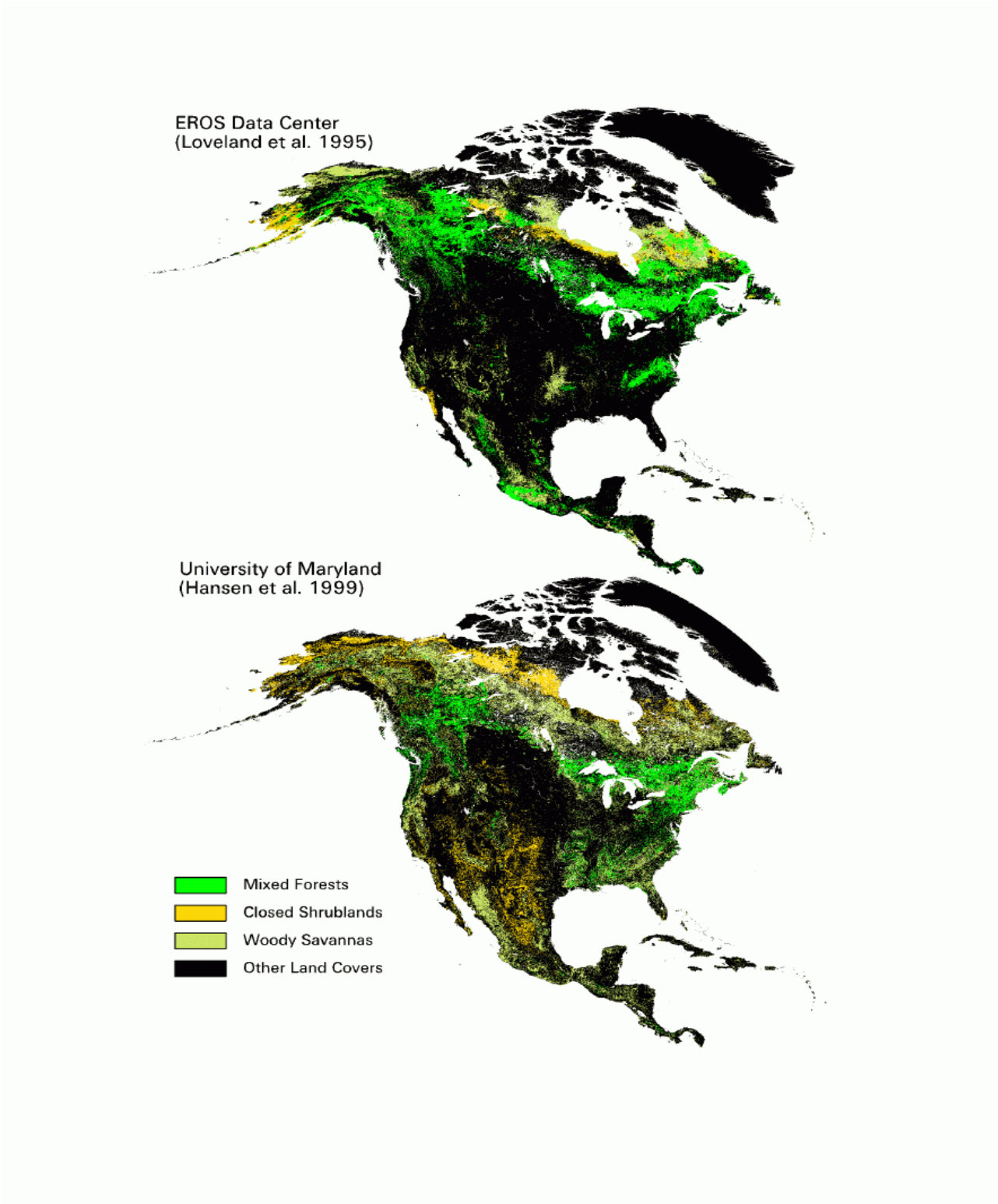


Figure 5-1: IGBP land cover classes for North America in the EDC and UMD maps that cannot be directly translated into biome classes.

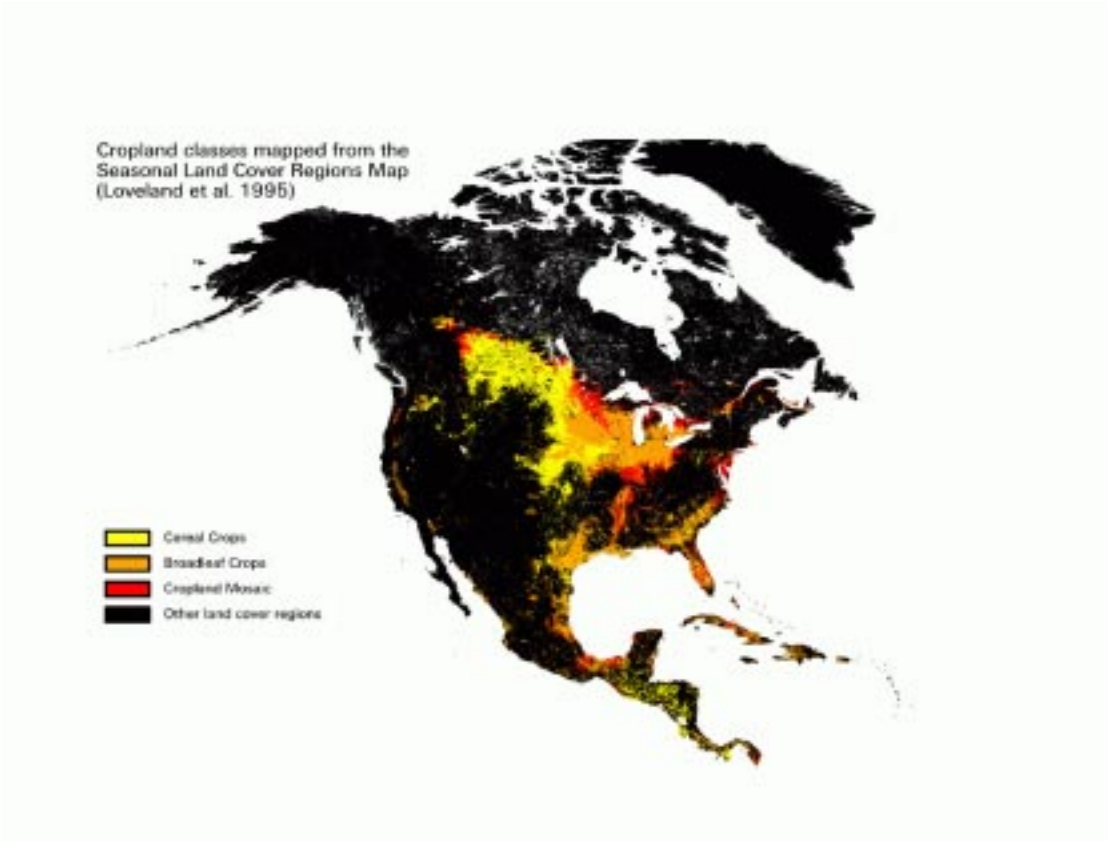


Figure 5-2: Cereal crops, broadleaf crops, and cropland mosaics mapped from the North America EDC SLCR map.

Unfortunately, the translation of IGBP classes into the classes in the biome scheme required for the MODIS/MISR LAI/FPAR algorithm is ambiguous with respect several classes. In particular, direct translation of the 17 IGBP classes into the six biome classes is not possible for the IGBP classes 5, 6, 8, 12, 14 (mixed forest, closed shrublands, woody savannas, croplands and croplands mosaic, respectively). Aggregation of the IGBP maps into the biome scheme therefore introduces distortions of the land cover proportions and may cause classification errors that are relatively severe from the viewpoint of radiative transfer modeling.

To illustrate the magnitude of distortion introduced by simple aggregation, mixed forests, closed shrublands and woody savannas were mapped from the UMD and EDC maps (Figure 5-1). The map comparison shows significant differences in both extent and spatial distribution of mixed forests, closed shrublands and woody savannas. Table 5-2 shows the areal extent of these classes in the EDC and UMD maps. Furthermore, the IGBP classification scheme does not distinguish between broadleaf crops and cereal crops. From the viewpoint of radiative transfer, however, this distinction is very important. A disaggregation of croplands and cropland mosaics (IGBP class 11 and 14) into cereal crops (biome 1) and broadleaf crops (biome 3) is not possible.

Using a simple aggregation scheme, e.g. assigning all croplands to biome 1, major areas would be spuriously assigned to one biome class (Figure 5-2 and Table 5-2).

The only digital map source at global scales providing more detailed information on the distribution of land cover is the Global Land Cover Characterization Data Base (GLCC) available from EDC [Loveland et al. 1995]. This data set is derived from 1-km Advanced Very High Resolution Radiometer (AVHRR) data spanning a 12-month period (April 1992- March 1993) and is based on a flexible data base structure and seasonal land cover regions concepts. Seasonal land cover regions (SLCR) provide a framework for presenting the temporal and spatial pattern of vegetation. The regions are composed of relatively homogeneous land cover associations, which exhibit distinctive phenology (i.e., onset, peak, and seasonal duration of greenness). A time series of 12-month NDVI composites was used in an unsupervised clustering algorithm. The resulting clusters were then labeled using extensive post-classification stratification and refinement techniques. The UMD map, on the other hand, was generated with a supervised classification approach using decision trees [Hansen et al. 1999]. Also, post-classification was an important methodological step to remove major misclassifications in the UMD map.

Table 5-2: A: Areal extent of mixed forests, closed shrublands, woody savannas and croplands in the EDC and UMD map. B: Areal Extent of broadleaf crops, cereal crops and cropland mosaics in the SLCR map (both North America).

A	**EDC**	**UMD**
Mixed Forests	*2,854,132 km^2*	*1,349,244 km^2*
Closed Shrublands	*579,582 km^2*	*1,703,853 km^2*
Woody Savannas	*1,658,740 km^2*	*3,113,640 km^2*
Croplands	*1,852,240 km^2*	*1,781,694 km^2*

B	**SLCR**
Broadleaf Crops	*1,055,278 km^2*
Cereal Crops	*1,487,628 km^2*
Cropland Mosaic	*544,887 km^2*

The GLCC project defined 205 SCLR for North America, 255 for Eurasia, 167 for South-America, 197 for Africa and 137 for Australia-Pacific. The narrow definition of the SLCR classes allows their aggregation into broader classes of other desired classification schemes, e.g., the 6-biome scheme. Look up tables (LUT) for the aggregation of SLCR classes into various existing classification schemes as well as ancillary data sheets, were provided by EDC and used as a guideline for the translation to 6 biomes performed for the at-launch MODIS/MISR LAI/FPAR land cover data set.

To generate a map in the 6-biome scheme for the MODIS/MISR LAI/FPAR algorithm, the map from UMD was used in association with the SLCR map. More specifically, for those classes

in the UMD-IGBP scheme that can be directly translated into one of the biome classes, a direct class assignment was performed. This applies to evergreen needleleaf forests, evergreen broadleaf forests, deciduous needleleaf forests, deciduous broadleaf forests, wooded grassland, open shrubland, grasslands, bare ground, and urban/built-up. For classes, that do not allow a direct translation, the respective SLCR label was retrieved and a biome class label was assigned using the IGBP-biome LUT described above. This routine was performed on a per-pixel basis.

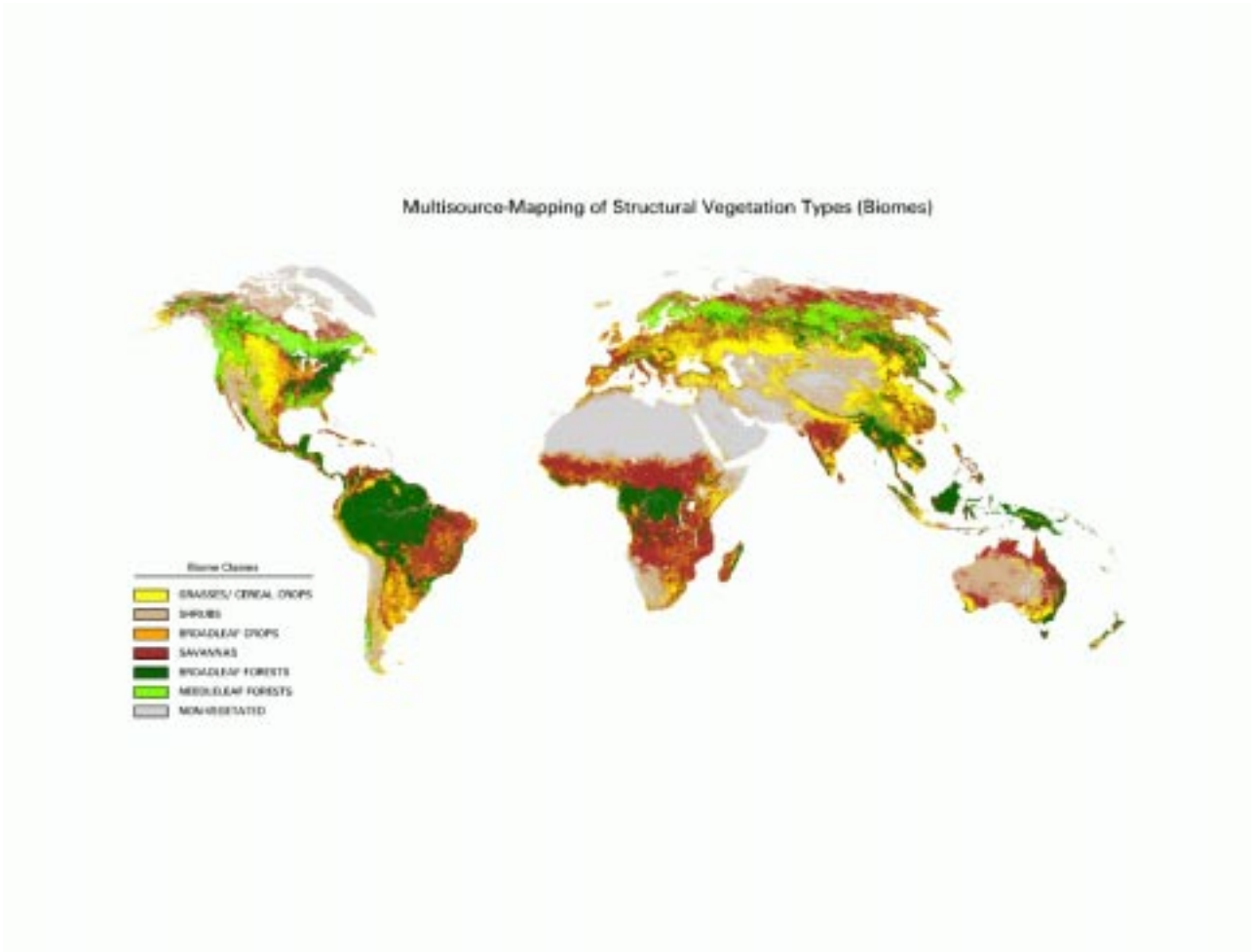


Figure 5-3: At-launch global biome map derived from the UMD land cover map and the EDC SLCR map.

The use of the SLCR maps in conjunction with the ancillary information in the global land cover characterization data base allows to resolve ambiguities in translating the UMD map into the 6-biome scheme. The SLCR labels were particularly useful for disaggregating the cropland class into broadleaf crops and cereal crops. The ancillary information about structural properties of the land cover types in the GLCC database also helped disaggregating the closed shrubland and woody savannas classes. In the case of mixed forests the disaggregation into either

needleleaf forests of broadleaf forests remained ambiguous in many cases, since many areas are in fact characterized by a mixture of both forest types. Using the LUT for each of the five continents, a pixel labeled according to the biome scheme can be related to its original class label in the UMD land cover map. This allows comparisons of various MODIS products (e.g. net primary productivity or leaf area index) that used modified versions of the UMD map in their algorithm.

5.2 Look-Up Table Structure

Summary of parameters determining the size of the LUT

Parameter	*Value*	*Description of Content*
Maxbands	7	Number of MODIS spectral bands for which the LUT was created.
Nbiome	6	Number of biome types.
Specdim	152	Number of points at which the spectral leaf albedo is given.
Maxviewpol	5	Number of sensor zenith intervals for which the LUT was created.
Maxviewazi	6	Number of sensor azimuth intervals for which the LUT was created
Maxsunpol	4	Number of solar zenith intervals for which the LUT was created
Laidim	40	Number of LAI values for which the LUT was created
Maxsat	41	Number of saturation curve points for which the LUT was created
Maxsoildim	29	Number of patterns of spectral effective ground reflection defined at the MODIS spectral bands which the LUT was created
Maxpixel	1000	Maximum pixels for run. *This parameter does not effect the LUT size*.

The size of the LUT can be reduced by setting laidim=29 (instead of 40). However it does not effect the program execution time. Therefore I recommend to take laidim=40. It will give us a possibility to influence the algorithm by varying laidimbio and laimax in Bu_anc_soil.txt.

Summary of the input files for the 00MODIS_v4F.exe

Input File	*Description of Contents*
Bu_anc_pcf_v4.txt	All runtime session (single scalar) input parameters
Bu_anc_slbar.txt	Saturation curve, varying by {sat_point=maxsat}
Bu_anc_lai.txt	Describes values of the plant LAI, number of the ground reflection patterns, number of saturation curve points as a function of biome
Bu_anc_rosoil.txt	Patterns of the effective ground reflection at MODIS spectral bands, varying by {bio=nbiome,soil_pattern=maxsoildim}
Bu_anc_constants.txt	MODIS spectral bands, sun-view geometry, ratio of direct over (direct+diffuse) radiation, accuracy factor for each spectral band
Bu_anc_ADLUT.txt	Inputs of the "S problem", varying by {bio=nbiome, lai=laidim}
Bu_anc_BSLUT.txt	Inputs of "Black soil" problem, varying by {bio=nbiome, lai=laidim, sun_azimuth=maxsunpol}
Bu_anc_WBS.txt	WBS inputs, varying by {bio=nbiome, lai=laidim, sun_zenith= maxsunpol, sensor_zenith=maxviewpol, sensor_azimuth =maxviewazi}
Bu_anc_WTA.txt	WTA inputs, varying by {biome=nbiome, lai=laidim, sensor_azimuth= maxviewpol},
Bu_anc_albedo.txt	Leaf albedo, varying by {bio=nbiome, lambda=specdim}
Bu_anc_spatial.txt	All pixel-wise inputs, varying spatially {x,y},n_x,n_y

Input file "Bu_anc_pcf_v4.txt"

Bu_anc_pcf_v4.txt		
Field	***Maximum value***	***Description of Content***
npixels	maxpixel	No. of pixels (e.g., n_rows x n_columns) for run.
nband	maxbands	Number of spectral bands used.
viewpoldim	maxviewpol	Number of sensor zenith angles used.
viewazidim	maxviewazi	Number of sensor azimuths used.
sunpoldim	maxsunpol	Number of solar azimuth angles used.
eps00	1	threshold value to execute comparison test for BRF's of ideal quality.
eps01	1	threshold value to execute comparison test if BRF's are less then ideal quality.
epssat	0.10	threshold value to execute the saturation test.
w0	0.10	Reference leaf albedo value (solution of the minimization problem.
threshold	0.50	not used in this version.

Input file "Bu_anc_slbar.txt"

Definition. The parametric plot $\left\langle \bar{L}(L^*), \bar{d}(L^*) \right\rangle$, $g_{\min} \cdot L_{\min} \le L^* \le g_{\max} \cdot L_{\max}$, is defined to be a saturation curve. Here

$$\bar{L}(L^*) = \int_{LAI_{\min}}^{LAI_{\max}} l \, d\Phi(L^*, l), \quad \bar{d}(L^*) = \sqrt{\int_{LAI_{\min}}^{LAI_{\max}} [l - \bar{L}(L^*)]^2 \, d\Phi(L^*, l)} \,,$$

and $g_{\min}$, $g_{\max}$, $L_{\min}$ and $L_{\max}$ are introduced in {Section 2.9}.

Bu_anc_slbar.txt		
Field	***Value Range***	***Description of Contents***
isat	1,2,…,maxsat	Number of the arameter L^*
satwlai for biome 1	5,…,10	Value of $\bar{L}(L^*)$ for biome 1
satdwlai for biome 1	0…10	Value of $\bar{d}(L^*)$ for biome 1
satwlai for biome 2	5,…,10	Value of $\bar{L}(L^*)$ for biome 2
satdwlai for biome 2	0…10	Value of $\bar{d}(L^*)$ for biome 2
satwlai for biome 3	5,…,10	Value of $\bar{L}(L^*)$ for biome 3
satdwlai for biome 3	0…10	Value of $\bar{d}(L^*)$ for biome 3
satwlai for biome 4	5,…,10	Value of $\bar{L}(L^*)$ for biome 4
satdwlai for biome 4	0…10	Value of $\bar{d}(L^*)$ for biome 4
satwlai for biome 5	5,…,10	Value of $\bar{L}(L^*)$ for biome 5
satdwlai for biome 5	0…10	Value of $\bar{d}(L^*)$ for biome 5
satwlai for biome 6	5,…,10	Value of $\bar{L}(L^*)$ for biome 6
satdwlai for biome 6	0…10	Value of $\bar{d}(L^*)$ for biome 6

Input file "Bu_ans_lai.txt"

Bu_anc_soil.txt		
Field	***Value Range***	***Description of Contents***
Biome	{1,…,nbiome}	Biome land cover class code
Soildimbio	≤ maxsoildim	Number of patterns for the effective ground reflectances {Section 2.2, Eq. (2)} at MODIS spectral bands
Laidimbio	≤ laidim	Number of LAI values LAI (Sect. 2.2).
nsat	≤ Maxsat	Number of saturation curve points
Biome1	Between 0.1 and 9.85	Lai values for biome 1
Biome2	Between 0.1 and 9.85	Lai values for biome 2
Biome3	Between 0.1 and 9.85	Lai values for biome 3
Biome4	Between 0.1 and 9.85	Lai values for biome 4
Biome5	Between 0.1 and 9.85	Lai values for biome 5
Biome6	Between 0.1 and 9.85	Lai values for biome 6

Input file "Bu_ans_rosoil.txt"

This file contains the effective ground reflection defined by {Section 2.2, Eq. (2)} at the MODIS spectral bands.

Bu_anc_rosoil.txt		
Field	***Value Range***	***Description of Contents***
Biome	{1,…,nbiome}	Biome land cover class code
Pattern	1,…, maxsoildim	Soil pattern number
Band 1	between 0 and 1	The effective ground reflection at the 1st MODIS band.
Band 2	between 0 and 1	The effective ground reflection at the 2nd MODIS band.
Band 3	between 0 and 1	The effective ground reflection at the 3rd MODIS band.
Band 4	between 0 and 1	The effective ground reflection at the 4th MODIS band.
Band 5	between 0 and 1	The effective ground reflection at the 5th MODIS band.
Band 6	between 0 and 1	The effective ground reflection at the 6th MODIS band.
Band 7	between 0 and 1	The effective ground reflection at the 7th MODIS band.

Input file "Bu_anc_constants.txt"

Bu_anc_constants.txt		
Field	***Value Range***	***Description of Contents***
lam	{645, 859, 469, 555, 1240, 1640, 2130}	MODIS spectral bands used by algorithm; lam=1,2,…,nbands
bio, wMODIS	1 {7 values between 0 and 1}	Biome #, values of leaf albedo at the MODIS spectral band.
bio, wMODIS	2 {7 values between 0 and 1}	Biome #, values of leaf albedo at the MODIS spectral band.
bio, wMODIS	3 {7 values between 0 and 1}	Biome #, values of leaf albedo at the MODIS spectral band.
bio, wMODIS	4 {7 values between 0 and 1}	Biome #, values of leaf albedo at the MODIS spectral band.
bio, wMODIS	5 {7 values between 0 and 1}	Biome #, values of leaf albedo at the MODIS spectral band
bio, wMODIS	6 {7 values between 0 and 1}	Biome #, values of leaf albedo at the MODIS spectral band
fdir0	7 values between 0 and 1	Ratio of direct over radiation for each band
epl	7 value between 0 and 1	Uncertainty factor for each band
viewpol	{8.5, 22.5, 37.5, 52.5, 67.5}	Sensor zenith angles. Number of angles modeled is maxviewpol.
viewazi	{25., 55., 85., 115., 145., 180.}	Sensor azimuths. Number of angles modeled is maxviewazi
sunpol	{22.5, 37.5, 52.5, 70.}	Solar zenith angles. Number of angles modeled is maxsunpol.

See routine "Prep_angles_v4.f" for detailed description of Sun-sensor geometry.

Input file "Bu_anc_ADLUT.txt"

Bu_anc_ADLUT.txt		
Field	***Value***	***Description of Contents***
Biome	{1,2,3,4,5,6}	Biome land cover class code
LAI values	{1,2,3…laidim}	Number of LAI values is laidim.
pAAD	0.0< pAAD<1	Solution of the minimization problem { Knyazikhin et al. 1998b, Eq. (49)} for the "S problem."
pABS	0.0< pABS<1	Solution of the minimization problem { Knyazikhin et al. 1998b, Eq. (49)} for the "Black soil problem."
AAD	0.0< AAD<1	Fraction of absorbed PAR by vegetation at the reference leaf albedo w0 for the "S problem." This is value $\mathbf{a}^{q}_{iso,\lambda_0}$ introduced on the right side of { Section 2.5, Eq. (29)}.
pTAD	0.0< pTAD<1	Solution $\mathbf{pt}_{\mathrm{iso}}$ of the minimization problem { Knyazikhin et al. 1998b, Eq. (50)} for the "S problem."
TAD	0.0< TAD<1	Canopy transmittance at the reference leaf albedo w0 for the "S problem." This is value $\mathbf{t}^{q}_{iso,\lambda_0}$ introduced on the right side of {Section 2.5, Eq. (25)}.

Input File "Bu_anc_BSLUT.txt"

Bu_anc_BSLUT.txt		
Field	***Value***	***Description of Contents***
Biome	{1,2,3,4,5,6}	Biome land cover class code
LAI values	{1,2,3…laidim}	Number of LAI values is laidim.
sunpol	{22.5, 37.5, 52.5, 70.}	Solar zenith angles. Number of angles modeled is maxsunpol.
pTBSdir	0.0< pTBSdir<1	Solution of the minimization problem { Knyazikhin et al. 1998b, Eq. (50)} for the "Black soil problem." Radiation field is generated by direct component of incoming solar radiation.
pTBSdif	0.0< pTBSdif<1	Solution of the minimization problem { Knyazikhin et al. 1998b, Eq. (50)} for the "Black soil problem." Radiation field is generated by diffuse component of incoming solar radiation.
ABSDIR	0.0< ABSDIR<1	Fraction of absorbed PAR by vegetation at the reference leaf albedo w0 for the "Black soil" problem. Radiation field is generated by direct component of incoming solar radiation.
ABSDIF	0.0< ABSDIF<1	Fraction of absorbed PAR by vegetation at the reference leaf albedo w0 for the "Black soil" problem. Radiation field is generated by diffuse component of incoming solar radiation.
TBSDIR	0.0< TBSDIR<1	Canopy transmittance at the reference leaf albedo w0 for the "Black soil" problem. Radiation field is generated by direct component of incoming solar radiation.
TBSDIF	0.0< TBSDIF<1	Canopy transmittance at the reference leaf albedo w0 for the "Black soil" problem. Radiation field is generated by diffuse component of incoming solar radiation.

The variables ABSDIR, ABSDIF, TBSDIF and TBSDIF form the values $\mathbf{a}_{bs,\lambda_0}^{hem}$ introduced by { Section 2.5, Eq. (29)} and $\mathbf{t}_{bs,\lambda_0}^{hem,q}$ introduced by { Section 2.5, Eq. (27)} at the reference leaf albedo $\omega_0=\omega(\lambda_0)$ as

$$\mathbf{a}_{bs,\lambda_0}^{hem} = \text{fdir} * \text{ABSDIR} + (1-\text{fdir}) * \text{ABSDIF},$$

$$\mathbf{t}_{bs,\lambda_0}^{hem} = \text{fdir} * \text{TBSDIR} + (1-\text{fdir}) * \text{TBSDIF}.$$

Input file "Bu_anc_WBS.txt"

This input file contains coefficient $\pi \cdot w_{bs,\lambda}$ at the MODIS spectral bands. Here $w_{bs,\lambda}$ is the weight defined by {Section2.7, Eq. (50)}.

Bu_anc_WBS.txt		
Field	***Value***	***Description of Contents***
BiomeClass	{1,2,3,4,5,6}	Biome class code
LAI values	{1,2,3…laidim}	Number of LAI values modeled is laidim.
sunpol	{22.5, 37.5, 52.5, 70.}	Solar zenith angles. Number of angles modeled is maxsunpol.
viewpol	{8.5, 22.5, 37.5, 52.5, 67.5}	Sensor zenith angles. Number of angles modeled is maxviewpol.
viewazi	{25., 55., 85., 115., 145., 180.}	Sensor azimuths. Number of angles modeled is maxviewazi
Band1	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 1st MODIS band.
Band2	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 2nd MODIS band.
Band3	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 3rd MODIS band.
Band4	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 4th MODIS band.
Band5	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 5th MODIS band.
Band6	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 6th MODIS band.
Band7	between 0 and 10	Value of $\pi \cdot w_{bs,\lambda}$ at the 7th MODIS band.

Input file "Bu_anc_WTA.txt"

This input file contains coefficient $\pi \cdot w_\lambda^q$ at the MODIS spectral bands. Here w_λ^q is the weight defined by { Section2.7, Eq. (51)}.

Bu_anc_WTA.txt		
Field	***Value***	***Description of Contents***
BiomeClass	{1,2,3,4,5,6}	Biome class code
LAI values	{1,2,3…laidim}	Number of LAI values modeled is laidim.
viewpol	{8.5, 22.5, 37.5, 52.5, 67.5}	Sensor zenith angles. Number of angles modeled is maxviewpol.
Band1	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 1st MODIS band.
Band2	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 2nd MODIS band.
Band3	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 3rd MODIS band.
Band4	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 4th MODIS band.
Band5	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 5th MODIS band.
Band6	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 6th MODIS band.
Band7	between 0 and 10	Value of $\pi \cdot w_\lambda^q$ at the 7th MODIS band.

Input file "Bu_anc_spatial.txt"

This table contains one record per pixel, mirroring the spatially defined (e.g. 2D) inputs of the MODIS production algorithm.

Bu_anc_spatial.txt		
Field	***Value***	***Description of Contents***
Biome	{1,2,3,4,5,6}	Biome, land cover class code
qA	00; 01; 02; 03	00 ideal quality of BRF; 01 less then ideal; 02 and 03 BRF is not produced
sen_azi	{0.0 <= sen_aazi < 360.0}	Sensor azimuth at this pixel
sen_pol	{0.0 <= sen_pol < 90.0}	Sensor zenith at this pixel
sol_azi	{0.0 <= sol_azi < 360.0}	Solar azimuth at this pixel
sol_pol	{0.0 <= sol_pol < 90.0}	Solar zenith at this pixel
Band1	{0 < b1 < 100}	BRF at the 1st MODIS band
Band2	{0 < b2 < 100}	BRF at the 2nd MODIS band
Band3	{0 < b3 < 100}	BRF at the 3rd MODIS band
Band4	{0 < b4 < 100}	BRF at the 4th MODIS band
Band5	{0 < b5 < 100}	BRF at the 5th MODIS band
Band6	{0 < b6 < 100}	BRF at the 6th MODIS band
Band7	{0 < b7 < 100}	BRF at the 7th MODIS band

Input file "Bu_anc_albedo.txt"

This input file contains the spectral leaf albedo varying by biome. The definition of this parameter is given by {Section2.6, Eq. (45)}. See also {Section 2.4.5, Fig. 2-2}

Bu_anc_albedo.txt		
Field	***Value***	***Description of Contents***
Biome	{1,2,3,4,5,6}	Biome land cover class code
Lambda	$400 \le \lambda \le 702$	Wavelength (nm). Number of spectral points is specdim.
Leaf albedo	$0 < \omega(\lambda) \le 1$	Leaf albedo at this wavelength

Output file "Bu_output.txt"

Bu_anc_output.txt		
Field	***Value***	***Description of Contents***
FPAR	$\{0 \le f \le 1\}$	FPAR, fraction of photosynthetically active radiation absorbed by vegetation.
QAFPAR	$\{0 \le q \le 1\}$	Variance of FPAR
LAI	{0 < lai < 10}	LAI value
QALAI	{0 < q < 1}	Variance of LAI
qA	0,1,2,3	0 - algorithm was executed: 1 - algorithm was executed. Saturation; 2 – Algorithm was not executed because BRF's were not available (cloud or "or other reasons"); or algorithm was not executed because view or sun zenith was too law; 3: algorithm fails.

6. Programming and Procedural Considerations

The MODIS land FPAR, LAI Level 3 algorithms were developed jointly by personnel at Boston University and the University of Montana SCF. The Boston University team developed the radiative-transfer (R-T) derivative science core logic and the R-T driven lookup tables comprising the core science, the direct-retrieval lookup tables, and prototype software for exercising the core logic. The University of Montana SCF team is responsible for developing, testing, and maintaining the EOSDIS Core System (ECS) production version of the software. QA tasks are shared between the two institutions, with the MOD15A1 QA activities conducted at Boston University, and the MOD15A2 QA activities run at University of Montana SCF.

The FPAR, LAI software codes are collectively organized into two software components:

- a daily executable (MOD_PR15A1, as PGE-33), used to produce a series of candidate daily products, and
- an 8-day executable (MOD_PR15A2, as PGE-34), used to produce the 8-day composite FPAR,LAI product archived at the Eros Data Center DAAC.

Each of these components defines a registered NASA EOSDIS Earth Science Data Type (EDST), providing users with a consistent product metadata organization supported by various end-user browse and data order environments. The daily product produced by MOD_PR15A1 generates the MOD15A1 ESDT, and the 8-day product generated by MOD_PR15A2 generates the archive MOD15A2 ESDT. Note that the daily product is used exclusively by the compositing process, and is not archived by the DAAC over time.

6.1 Programming Issues

6.1.1 Implementation Software Environment

The MOD15A1 and MOD15A2 FPAR, LAI software codes are written to comply with a number of standards, and to interoperate with several supporting binary libraries. The codes are written in ANSI C and are POSIX 1.x compliant. Both daily and 8-day codes link to the mandatory NASA Software Data Production Toolkit library (SDPTK v.5.2.4), the NASA HDF-EOS library v.2.4), and our SCF API -- the MODIS-University of Montana (MUM) API v. 2.2 . The table below breaks down the lines-of-code for our at-launch FPAR,LAI algorithms:

FPAR, LAI At-launch Lines-of-code (LOC) Metrics		
Code Layer	*Lines Of Code*	*Percent of Code*
MOD15A1 (PGE33)	13,336	28 %
MUM library	33,588	72 %
PGE-33 subtotal:	46,924	-
MOD15A2 (PGE34)	7275	18 %
MUM library	33,588	82 %
PGE-34 subtotal:	40,863	-
Total, All Codes	54199	-

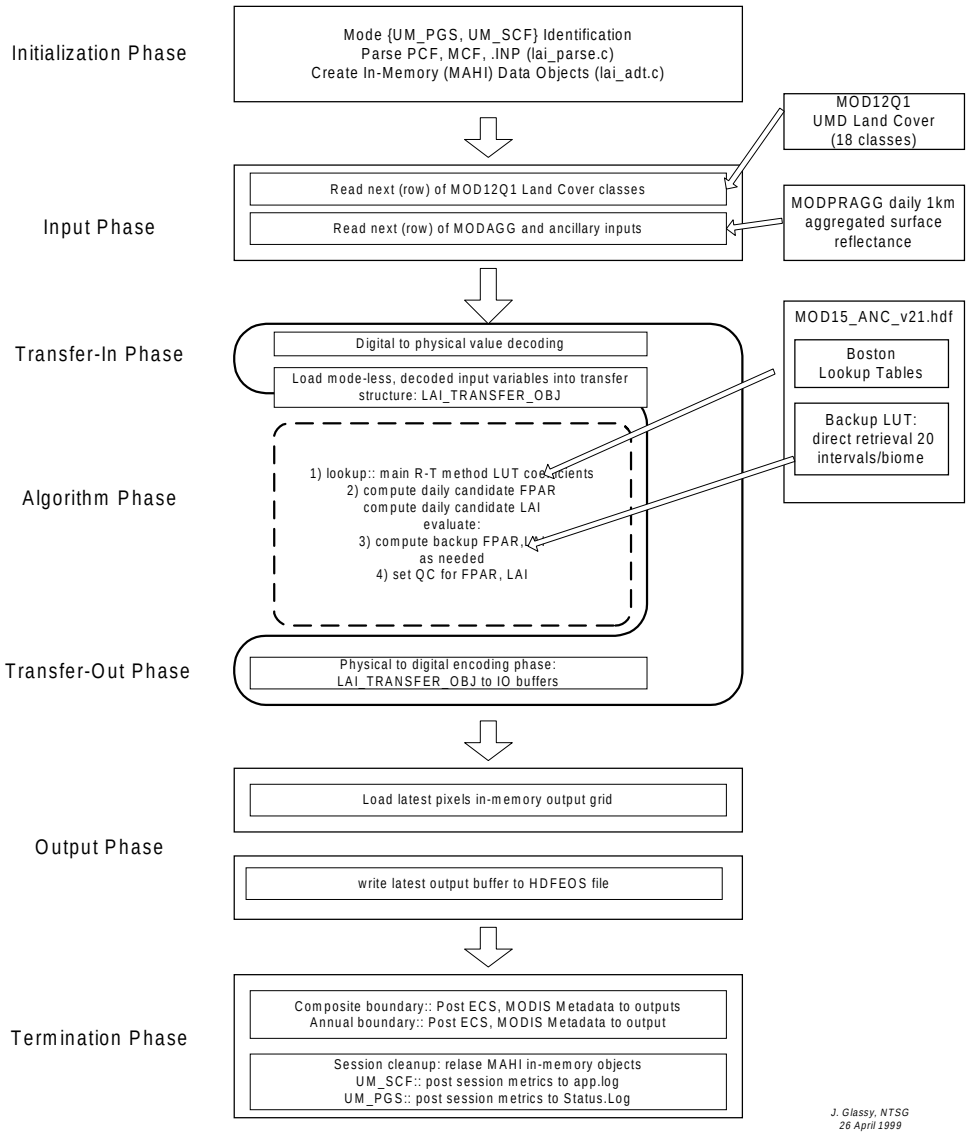
6.1.2 Software Design

In designing our at-launch algorithms, emphasis was placed on code robustness, reliability and maintainability. A significant fraction of the ECS production code developed at the Montana SCF rests on the common MUM API foundations. Larger amounts of code re-use generally promote life-cycle reliability in several ways: a) the total number of potential points-of-failure (system-wide) are reduced across all algorithms sharing a common foundation, and b) the life-cycle maintenance effort required to support the shared foundation (API) code is spent just once while each "client" benefits from the software services it provides.

Another common design thread running throughout our implementations is the emphasis on data-driven parameterization of the algorithm software. By "data driven" we refer to the externalization of key software inputs, to allow some revisions in program behavior without having to recompile and link the software. The lookup table (LUT) orientation of MOD15A1 is the best example of this type of externalization, stored in the static ancillary HDFEOS (MOD15_ANC_v21.hdf) file available to the algorithm at runtime. We also store all the defining characteristics of all ESDT product gridfields in this ancillary file, allowing minor file specification changes to be effected from outside the software.

The following diagram illustrates the high level organization (schema) used in MOD15A1; an almost identical schema is used for MOD15A2:

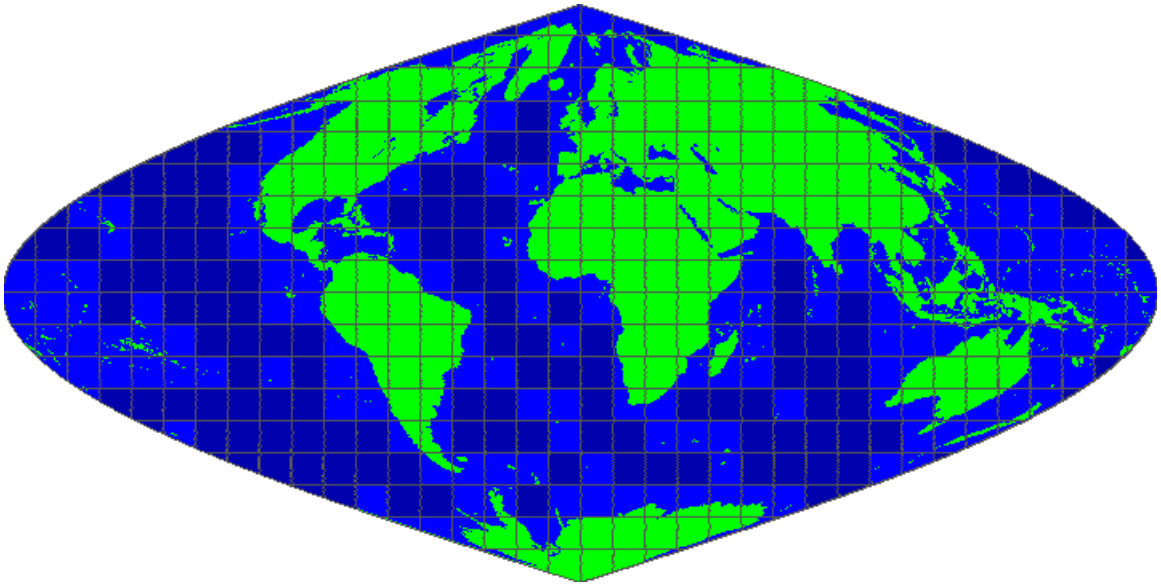
MOD15A1 v 2.1 Algorithm Component:
Daily FPAR, LAI
PGE 33 Software Schema



J. Glassy, NTSG
26 April 1999

6.1.3 Spatial Projection

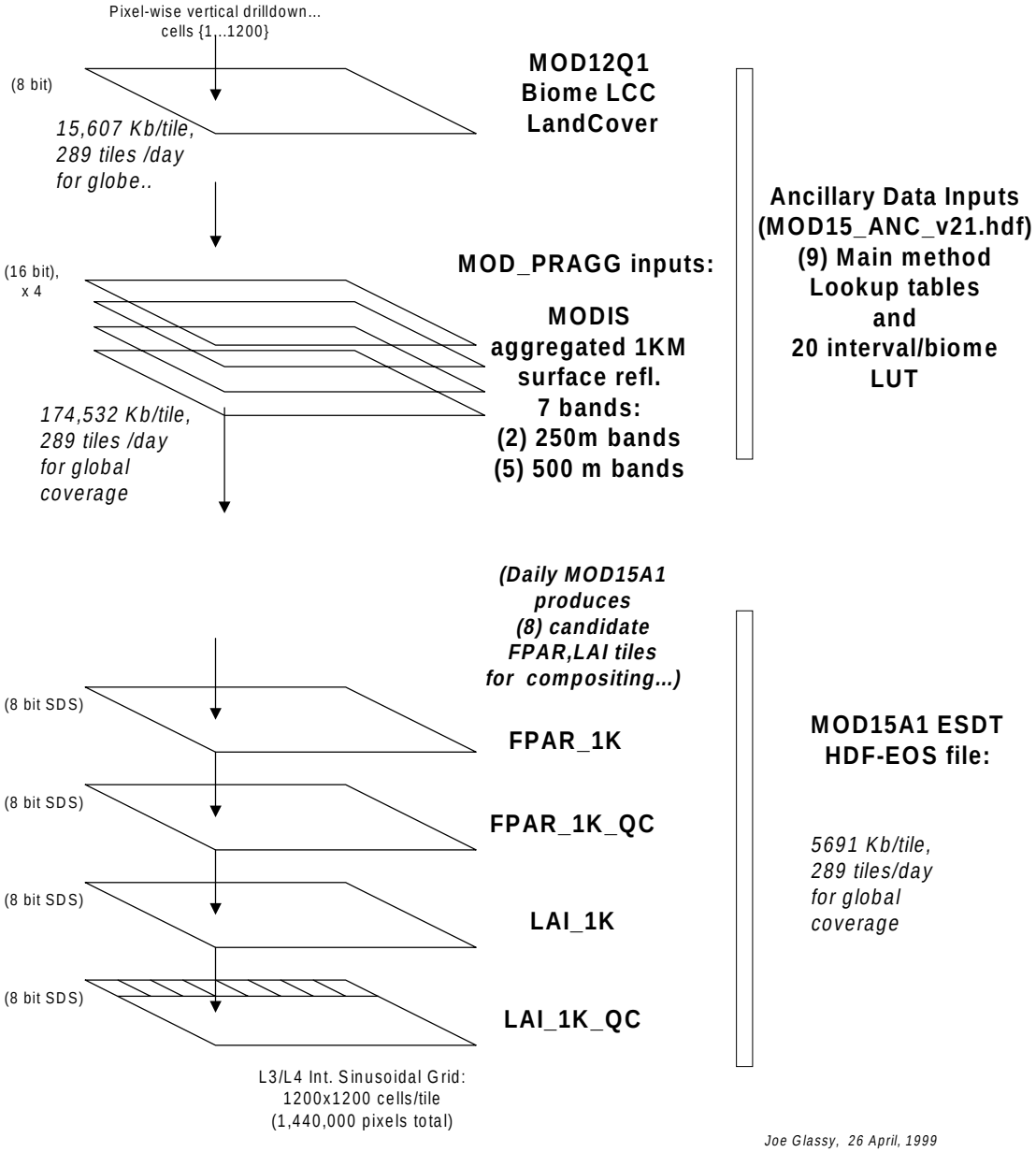
The MOD15A1 and MOD15A2 algorithms, like most MODIS Land processes, are organized to accept global coverage inputs, and produce global coverages either daily (PGE-33) or on an 8-day (PGE-34) timestep. Rather than process synoptic 1 KM spatial resolution images, the MODIS Land team has adopted a contiguous land tile scheme, based on the Integerized Sinusoidal Grid -- a map projection derived from the sinusoidal map projection (with the General Cartographic Map Projection code of GCTP_ISINUS). This projection defines a total of 648 tiles globally, at 10 degree resolution. We currently estimate that ***289 tiles of 648*** will be classified as "land" tiles, and thus represent the maximum spatial extent our global algorithms will process. Refer to http://modland.nascom.nasa.gov/developers/bndrytb2.html for additional details. The figure below graphically depicts the IS grid (assuming the standard 10 degrees processing tiles) we expect to work with at-launch. Each individual tiles in this grid includes approximately 1200 x 1200 1 km pixels:



6.1.4 Data Flows and Dependencies

The MOD15A1 (fired daily) and MOD15A2 (fired once each 8-day period) executables are positioned towards the end of the Land L4/L4 processing chain. Our algorithms are therefore quite dependent on the quality and correct functioning of all upstream processes. The following diagram illustrates the input I/O dependencies for MOD15A1 and MOD15A2:

MOD15A1 V2.1 Spatially Defined
Input/Output Data Layers
ECS Mode: 1 KM

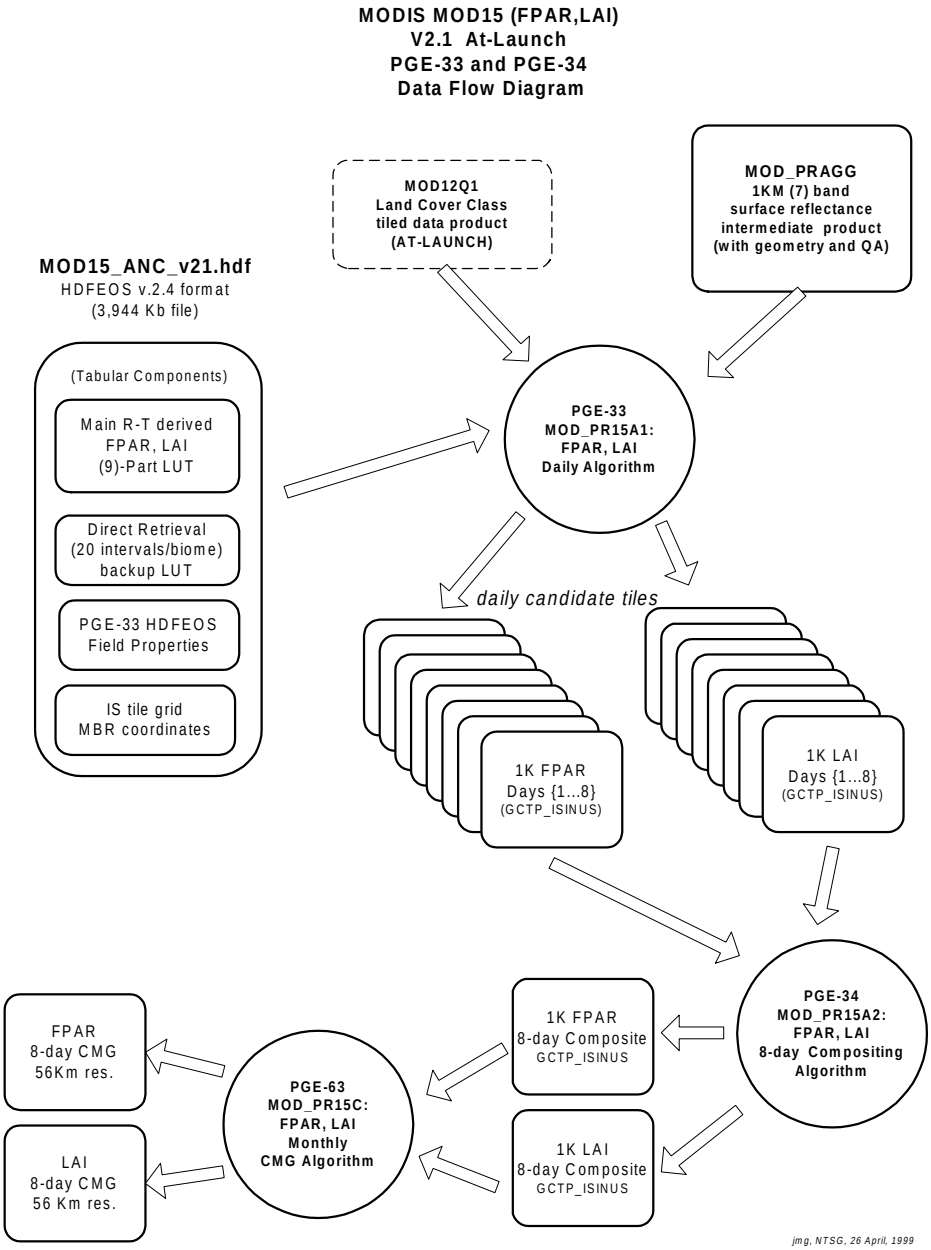


Joe Glassy, 26 April, 1999

6.2 Processing Issues

6.2.1 Processing Context

The key MODIS inputs to the MOD15A1 (PGE-33) and MOD15A2 (PGE-34) algorithms are the 1KM land cover (MOD12Q1, gridfield "Land_Cover_Type1", and the aggregated 1KM surface reflectance product (MOD_PRAGG: "Level 3 BRDF Preprocessing Database at 1km resolution"). The daily 1KM surface reflectances are derived from the MOD09 250m and 500m surface reflectance product, and are atmospherically corrected. The following figure illustrates



the high level data flows for these algorithms:

6.2.2 Performance: MFLOPS and Storage Load Estimates

The at-launch computational metrics for MOD15A1 and MOD15A2 are shown in the table below:

Computational Metrics: MFLOPS and Storage Loads		
Algorithm	MFLOPS	Storage Load (Gb/day)
MOD15A1 (PGE-33)	7.009200	1.67 Gb/day (not archived)
MOD15A2 (PGE-34)	0.000233	1.67 Gb/8-day period (archived)
Totals:	7.001133	1.67 per 8-day period (archived)

Load Methods

IS grid tiles are classified as "land" when at least 1 or more 1KM pixel in the 0.50 degree tile region is classified as "land". For the daily storage load estimate, we assume (a worst case) that all (289) "land" tiles globally are produced daily, and that each product tile is 5,691 Kb in size (((5691 Kb/1000)*289 tiles)/1000=1.64 Gb/day.

MFLOPS Methods

The methodology used to estimate the MFLOPS is specific to the SGI 64 bit workstation class (using the IRIX 6.4.x operating system), with 2 or more R10000 CPUs running at 195mhz. To obtain the MFLOPS estimate, we built and ran a special form of the algorithm on our SGI Octane (dual R10000,195mhz) workstation using the SGI IRIX "perfex" utility (e.g. for the MOD15A1.exe algorithm, we used:

perfex -a -mp -y >& MOD15A1_MFLOPS.report

Note that the MFLOPS value obtained from this utility technically refers to millions of floating point ***instructions*** per second not millions of floating point ***operations*** per second, and due to the fact that there are sometimes multiple instructions per "operation", the value reported is not exact.

6.2.3 FPAR, LAI Algorithm Daily Logic (PGE-33)

The following section briefly summarizes the high-level runtime logic implemented in MOD15A1:

- Upon instantiation of the session (once per tile), the runtime inputs are retrieved from the process control file (.pcf) file provided by the high level scheduler/loader environment.
- The PGE allocates all in-memory data structures it requires, opens the static ancillary HDFEOS file (MOD15_ANC_v21.hdf) and builds an in-memory dictionary of its contents.
- All spatial input HDFEOS files are opened (land cover and aggregated 1KM surface reflectance files) and in-memory dictionaries are built of their contents.
- The daily candidate output file is created, using field properties read at runtime from the ancillary file.
- A row-wise processing loop is established, and each subsequent row from all input files is buffered into memory.
- A pixel-wise processing loop is established, and for each pixel, input fields are decoded from their digital representation (integer) to their biophysical form (floating point) using {gain,offset} calibration factors. QA fields from the inputs are also checked here, in order to pass-through all pixels not classified as "land" or otherwise of unsuitable quality.
- The main R-T based retrieval method is performed on the pixel, using the channel-wise reflectances and pixel geometry. An FPAR, LAI estimate (and QA for each) is then calculated for the pixel.
- The biophysical estimates are encoded to their digitial (integer) representations, and are placed in the output buffer.
- When the current row is complete, it is written out to the open HDFEOS archive product file.
- When all rows are complete, final ECS metadata processing is performed, whereby the ECS metadata fields are set into their blocks, and finally written out to the product output file.
- Final session cleanup tasks are then performed, where all dynamic memory is released, files are closed.

6.2.4 FPAR, LAI Algorithm 8-day Compositing Logic (PGE-34)

The MOD15A2 (PGE-34) executable accepts a set of up to candidate tiles produced by the daily MOD15A1 process, and composites these using a simple selection rule whereby the pixel with the maximum FPAR (across the 8 days) is selected for inclusion in the output tile of identical format. Relevant QA and other tile-level metadata is taken from the day chosen to contribute its primary FPAR. The same day chosen to represent the FPAR measure also contributes the pixels LAI value.

The MOD15A2 (PGE-34) executable is launched once every 8-day period, and is provided with the current 8-day set of MOD15A1 candidate daily tile products. It performs the following actions:

- Upon instantiation of the session (once per tile), the runtime inputs are retrieved from the process control file (.pcf) file provided by the high level scheduler/loader environment.
- The PGE allocates all in-memory data structures it requires; note this executable ***does not*** require the ancillary file, MOD15_ANC_v21.hdf.
- The set of 1-8 candidate daily MOD15A1 product files are opened, and in-memory dictionaries are built of their contents.
- A row-wise processing loop is established, and each subsequent row from each days input file is buffered into memory.
- A pixel-wise processing loop is established, and for each pixel, the raw digital values are stored.
- An inner-most, temporal loop is established, looping through the (1 to 8) days of candidate data.
- From each set of 8 days pixels (e.g. using a vertical drill-down through the day dimension), the pixel with the highest FPAR value is identified. The base-0 index for this day is stored and used to retrieve the spatially coincident FPAR, LAI estimates associated with this day.
- The FPAR and LAI (and QC) values for the day identified above are then placed in the single 2D output buffer at the appropriate {line,sample} position.
- When the current row is complete, it is written out to the open HDFEOS archive product file.
- When all rows are complete, final ECS metadata processing is performed, whereby the ECS metadata fields are set into their blocks, and finally written out to the product output file.
- Final session cleanup tasks are then performed, where all dynamic memory is released, files are closed.

6.3 Quality Assurance

6.3.1 Quality Control and Diagnostics

The level of attention to Quality Assurance tasks, on the whole, is expected to directly correlate with the usefulness of the EOS "Terra" AM-1 instrument data for terrestrial products. Thus, a considerable fraction of the output data is classified as "QA data". This contribution by QA activities is matched within the algorithm logic itself: a significant portion of the logic is dedicated to verifying the quality of the internal data pipeline. In total, the QA subsystem with an algorithm represents some complex data flows, since input data from "upstream" is also supplied with both tile and pixel-level QA data and must be appropriately evaluated in the course of our PGE33 process.

The QA information in NASA EOS product files is loosely classified as "metadata", or data about data. Tile level metadata fields are further divided into Earth Observing System Distributed Information System (EOSDIS) Core System (ECS) metadata fields, and SCF supplied metadata. ECS metadata fields will be archived in separate NASA databases, and are considered "searchable" metadata, in that tools will be provided to users to allow them to search and filter tiles on the basis of the appropriate ECS metadata fields.

Within the group of ECS "searchable" metadata fields, an additional category is offered called ***product specific metadata*** (PSA). These are additional, special metadata fields unique to each Earth Science Data Type (ESDT) or product. Each type of metadata component plays an important role and are processed somewhat differently. All metadata fields are also fully described in the given ESDT's file specification. The ECS metadata fields are organized within one of (3) compound (character) attributes stored at the top logical level within HDFEOS files. These are:

- "StructMetadata.0" (HDFEOS grid specific and map projection parameters)
- "CoreMetadata.0" (Core metadata)
- "ArchiveMetadata.0" (Archive metadata)

The StructMetadata.0 field contains a number of map projection related parameters relevant to the product file. The CoreMetadata.0 block contains a set of browse-able metadata fields that will eventually serve as selection criteria for users ordering the data, and the ArchiveMetadata.0 block contains a number of system parameters, such as the limits of the tile's spatial extent (in geodetic coordinates), the granule day/night beginning and ending date and times, etc.

Tile level metadata fields within these large compound text blocks are further organized using the NASA Object Data Language(ODL)/Parameter Value Language (PVL) object hierarchy scheme. A special set of software functions --the "MET" subsection of the Software Data Production Toolkit (SDPTK) library -- are used to perform all manipulations on ECS metadata fields within this hierarchy. The user may discover all the metadata associated with a given product by referring to:

- The product's "MCF" file (ours is MOD15A1.mcf). In this structured text file all the ECS metadata fields (archived, searchable, and PSA) are enumerated along with most of their defining properties.

- The MOD15A2 archive product's file specification (MOD15A2.fs), which describes all aspects of a given ESDT product, including the number, layout of primary data planes in an HDFEOS file, as well as the full set of metadata fields (ECS and SCF).

- Refer to NASA web sites dedicated to ECS metadata information base: e.g. the Users may refer to a very complete information base on the ECS metadata concepts may be found at: http://ecsinfo.hitc.com/metadata/metadata.html

We define a number of product specific (PSA) attributes, stored in the CoreMetadata.0 block, set under the internal name ADDTIONALATTRIBUTES. Refer to the table describing these in detail, in the section below titled **Output Product**.

The PGE33 (daily MOD15A1) and PGE 34 (8-day composite MOD15A2) pixel wise QA information consists of a single 8-bit unsigned character compound bit-field, organized as follows:

MOD15A1/MOD15A2 PGE33 QA Pixel Descriptions			
Bit Field	Bits	Range	Bit-code Definitions
MODLAND QA	00-01	00-11	00= Product pixel produced at ideal quality, 01= Product pixel produced, less than ideal quality 10= Product pixel not produced due to cloud effects 11=Product pixel not produced for other reason
ALGORITHM_PATH	02-03	00-11	00= pixel could not be calculated using any method (usually due to missing or bad inputs) 01=calculate using main Boston method 10=calculated using backup method
PGE33: not used v.2.1	04-05	00-11	
FPAR,LAI Summary QA	06-07	00-11	00= Highest quality (76-100 percentile quality score) 01= Good quality (49-75 percentile quality score) 10= Poor or Questionable quality (26-50 percentile score) 11= Unacceptable quality, (0-25) recommend avoid using

We anticipate that the official NASA data ordering tool(s) will incorporate methods by which users may mask (e.g. define selection criteria) for tiles of interest using various combinations of these pixel level flags.

6.3.2 Pixel Level Qualification Criteria and Selection Logic

PGE33 subjects each input pixel to an independent qualification test based on the collective QA for the MODAGAGG data supplied. The result of the series of tests is that the input pixel is scored on a simple 4-level scale. These levels are:

Univ.Montana SCF 4-Level QA Scores		
Code	Value	Definition
IsQaBestQuality	0	Best possible, preferred quality
IsQaMinAcceptQuality	1	Minimum acceptable quality
IsQaBadQuality	2	Poor quality, not recommended for use (not proc)
IsQaWorstQuality	3	Worst possible, pixel not processed further.

All input-pixel level QA logic is implemented in the ***lai_tran_input_pix()*** function in the **lai_tran.c** module file. The logic used in the series of pixel-wise tests is summarized next, by order of the test:

1. The UMD Land Cover class code is first converted from the original UMD (18 or 14) class scheme to the University of Montana/Boston Univ. (7) biome classification. Note that UM biome code 0 is water/rock and code 7 is barren--pixels with these types are not processed further.

2. If the converted biome class lies in the range {1,2,3,4,5,6}, processing continues for the pixel; otherwise, fill values are placed in the output buffer to indicate that the pixel is not processed, and appropriate ECS metadata fields are set accordingly.

3. Next, the MODAGAGG aggregate QC bits are parsed and analyzed. If the cloud bit equals 1 (is too cloudy), or the LandWater bit is != 1, (e.g. is NOT land) the FPAR, LAI QC bits are set to IsQaWorstQuality, and the pixel is not processed.

4. Next, the pixel is checked to see if it meets the minimum quality threshold. This is true is any of the following conditions are met:
   a) the cloud bit is >= 1 (e.g. mixed or assumed not set)
   b) the cirrus bit is >=2 (e.g. average or high amount of cirrus)
   c) the cloud shadow bit ==1 (e.g. pixel is in cloud shadow)
   d) the aerosol bit ==3 (e.g. a high amount of aerosol has been encountered)
   e) the snow/ice bit must be 0 (e.g. snow/ice ***does not*** occupy a significant portion of the pixel)

5. The 1KM aggregated surface reflectances for all channels used (2 or 4) are then unpacked, along with their pixel level QC bits, and are then further analyzed. If the MODLAND bits are > 1 or the Atmosphere bit is >= 13, the pixel is rated IsQaWorstQuality and processing is blocked for the pixel. A check is then made on how many of the individual bands used (2 or 4) supplied inputs at the level of IsQaMinAcceptQuality or better. For the main method to be used, all 2 or 4 bands must have supplied this level of quality. If even one band supplied inferior quality inputs, then the main method is blocked, and a flag is set indicating that the backup method should be tried instead, provided that the first (2) bands of input --the minimum required for the backup method-- were scored at the level of IsQaMinAcceptQuality or better.

6.3.3 Exception Handling

Exception handling for the MOD15A1 (PGE-33) and MOD15A2 (PGE-34) is performed using the standard ECS compliant SDTPK SMF software layer. In our implementation of this method, we define (4) classes of exceptions according to their severity. A common SMF message file is used for all our algorithms (PGS_MODIS_37150.t/.h). The single letter severity codes {U,W,E,F} were adapted from the SDPTK User's Guide, Section 6.128. The table below summarizes these:

Univ. Montana SCF Exception Handling (PGE_MODIS_37150.t symbols)		
Message Code	***Severity***	***Comments***
MODIS_U_MUM_ADVISORY	Advisory	These messages are used to passively inform the operator or user about a given condition.
MODIS_W_MUM_ADVISORY	Warning	These messages indicate that an out of the ordinary condition has occurred, that may require monitoring further.
MODIS_E_MUM_ADVISORY	Error	These messages indicate that a (non-fatal) program error has occurred which should be investigated as soon as possible.
MODIS_F_MUM_ADVISORY	Fatal exception	These messages indicate that a fatal program error or condition has been encountered. The PGE will halt shortly after performing as much damage-containment as it can.

Generally, when an exception of class MODIS_F_MUM_ADVISORY is encountered, this will necessitate re-running the PGE against the indicated tile, once the cause has been identified

and a solution has been determined. A single "collector" exception handling call is made whenever any of these exceptions is encountered. This call routes the text of the message to the (3) standard SDPTK session log files, e.g. (MOD15_StatusLog.log, etc). An example of such a call is shown below:

```
mum_message(MODIS_U_MUM_ADVISORY,
        "lai_main.c","lai_driver","Pixels successfully processed %ld\n",n_total);
```

6.3.4 Output Product Details

The MOD15A1 (PGE-33) daily executable produces IS tile (NASA HDFEOS v.2.4 format) output files that are formally identical to the archived 8-day product generated by MOD15A2. The full baselined file format for these may normally be found at:

ftp://modis-xl.nascom.nasa.gov/modisbaselinedcode/COMMON/filespec

To summarize the contents of the archived MOD15A2 FPAR, LAI product file, it contains (4) spatially defined 2D gridfields, the full set of ECS Core, Archive, and Struct metadata fields in Parameter-Value-Language (PVL)/Object Data Language (ODL) format blocks, as well as a small set of gridfield (or SDS) attributes. Note that HDF and HDF-EOS format files use an underlying "xdr" based numeric representation for data, which allows data of any numeric data type to be ported to virtually any compute platform, regardless of the "endian" byte-ordering of the native word on the platform. The common datatype the main gridfields are stored in (e.g. DFNT_UINT8) is a platform independent, unsigned, 8-bit integer type capable of representing numeric values in the range {0 <= dn <= 255} inclusive. The (4) main gridfields are summarized further in the following table:

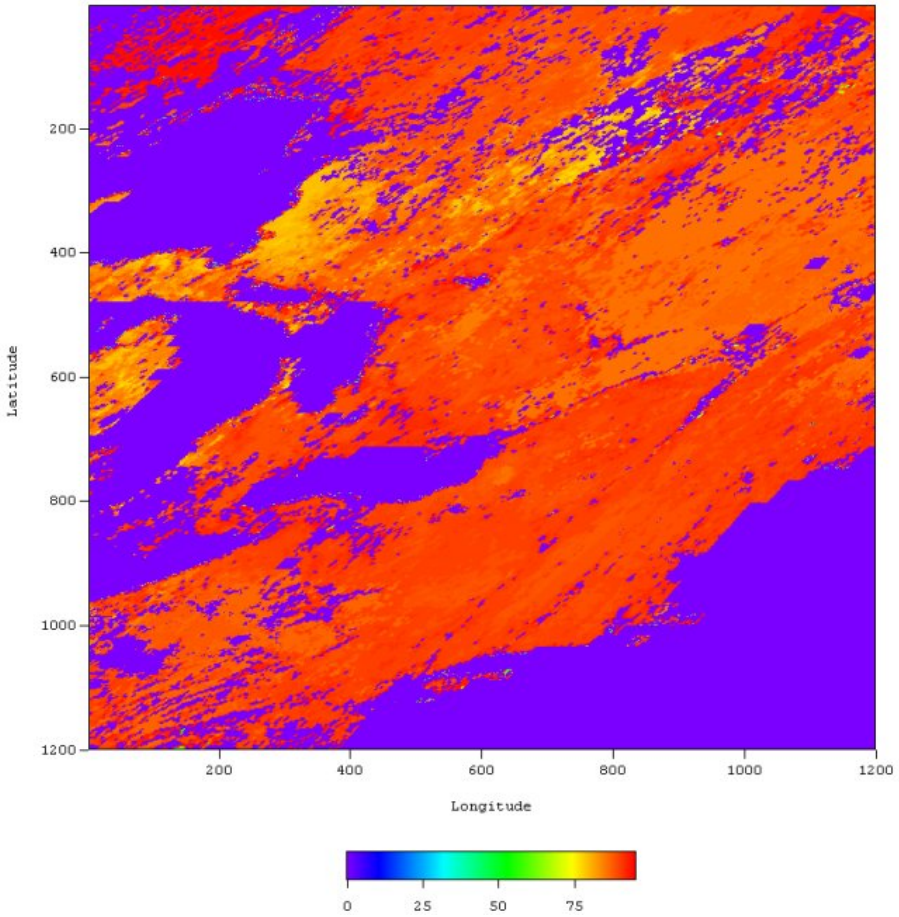
MOD15A2 FPAR, LAI 8-day composite archive file summary			
Grid Field Name	Datatype	Dimensions	Description
Fpar_1km	DFNT_UINT8	1200 x 1200	spatial FPAR field
Lai_1km	DFNT_UINT8	1200 x 1200	spatial LAI field
Fpar_1km_QC	DFNT_UINT8	1200 x 1200	Pixel-wise quality control for the modeled FPAR measure.
Lai_1km_QC	DFNT_UINT8	1200 x 1200	Pixel-wise quality control flags for the LAI measure.

Users may refer to a very complete information base on the ECS metadata concepts may be found at: http://ecsinfo.hitc.com/metadata/metadata.html

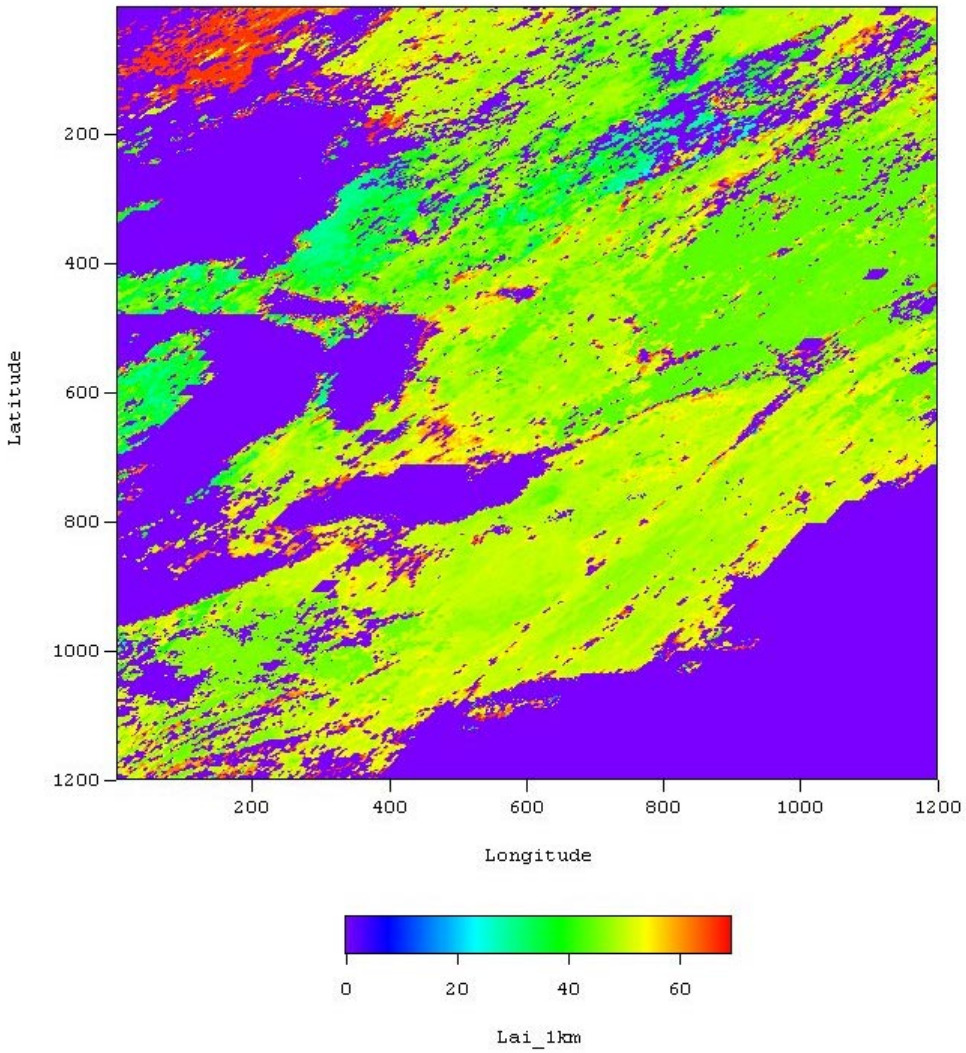
A set of ECS compliant Product Specific Attributes (PSA) is also included in each tile file, within the CoreMetadata.0 block. These provide users with a very coarse but quick "tile-level" quality assessment for the product file. The table below summarizes these:

MOD15A2 (PGE 34) Archive Product Specific Attributes	
PSA Metadata Field Name	Field Description
N_DAYS_COMPOSITED	PGE33 always 1; PGE34 this is the number of days contributing to the final composite.
QAPERCENTGOODFPAR	Percent {0<=p<=100} of FPAR pixels rated at good (e.g. IsQaMinAcceptQuality) or better
QAPERCENTGOODLAI	Percent {0 <= p <= 100} of LAI pixels rated at good (IsQaMinAcceptQuality) or better.
QAPERCENTMAINMETHOD	Percent {0 <= p <= 100} of pixels calculated using main method
QAPERCENTEMPIRICALMETHOD	Percent {0 <= p <= 100} of pixels calculated using backup method (reciprocal of above)
TILEID	IS Tile ID code, an 8-digit integer that identifies the map projection used, the tile's size code (quarter, half, or full-tile) and the tile's horizontal and vertical position in the grid.

A one-tile sample image illustrating how the FPAR spatial field will look is shown below, for test file H12V04 in the 10 degree IS grid.



A one-tile sample image illustrating how the LAI spatial field will look is shown below, for test file H12V04 in the 10 degree IS grid.



REFERENCES

Anisimov, O. A., and G. V. Menzulin, The problem of modeling the radiative radiation regime in the plant cover, *Sov. Meteorol. and Hydrol., 10*, 70-74, 1981.

Asrar, G., M. Fuchs, E.T. Kanemasu, and J.L.Hatfield, Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat, *Agron. J., 76*, 300-306, 1984.

Baret, F., S. Jacquemoud, and J.F. Hanocq, The soil concept in remote sensing, *Remote Sens. Rev., 7*, 65-82, 1993.

Barnsley, M. F., *Fractals Everywhere*, 531 pp., Academic, San Diego, Calif., 1993.

Bass, L.P., A.M.Volocsvhenko, and T.A.Germogenova, *Methods of Discrete Ordinates in Radiation Transport Problems* (in Russian, with English abstract), 231 pp., Inst. Appl. Math., Russ. Acad. of Sci., Moscow, 1986.

Belward, A.S. and Loveland, T.R., The IGBP-DIS 1-km land cover project. Proceedings, RSS'95 Remote Sensing in Action, 21st Annual Conference of the Remote Sensing Society, Southhampton, UK, 11-14 September 1995 (Nottingham: Remote Sensing Society), pp. 1099-1106, 1995.

Borel, C.C., S.A., Gerstl, and B.J. Powers, The radiosity method in optical remote sensing of structured 3-D surfaces, *Remote Sens. Environ., 36*, 13-44, 1991.

Chen. J. M, Canopy architecture and remote sensing of the fraction of photosynthetically active radiation absorbed by boreal conifer forests, *IEEE Trans. Geosci. Remote Sens., 34*, 1353-1368, 1996.

Chen. J. M and Jose Cihlar, Retrieving leaf area index of boreal conifer forests using landsat TM images, *Remote Sens. Environ., 55*, 153-162, 1996.

Cihlar, J, J. Chen, and Z. Li, Seasonal AVHRR multichannel data sets and products for studies of surface-atmosphere interactions, J. Geophys. Res., 102, 29625-29640, 1997.

Clevers, J. G. P. W, The application of a weighted infrared-red vegetation index for estimating leaf area index by correcting for soil moisture, *Remote Sens. Environ., 29*, 25-37, 1989.

Deschamps, P.Y., F.M. Bréon, M. Leroy, A. Podaire, A. Bricaud, J.C. Buriez, and G. Sèze, The POLDER mission: Instrument characteristics and scientific objectives, *IEEE Trans. Geosci. Remote Sens., GE-32*, 598-615, 1994.

Diner, D.J., J.V.Martonchik, C. Borel, S.A.W. Gerstl, H.R. Gordon, Y. Knyazikhin, R. Myneni, B. Pinty, and M.M. Verstraete, MISR: Level 2 surface retrieval algorithm theoretical basis, *JPL Internal Doc. D-11401, Rev. C*, Calif. Inst. of Technol., Jet Propul. Lab., Pasadena, 1998a.

Diner, D.J., W.A. Abdou, H.R. Gordon, R.A. Kahn, Y. Knyazikhin, J.V.Martonchik, S. McMuldroch, R.B. Myneni, and R.A. West, Level 2 ancillary products and data sets algorithm theoretical basis, *JPL Internal Doc. D-13402, Rev. A*, Calif. Inst. of Technol., Jet Propul. Lab., Pasadena, 1998b.

Eisen, M., *Introduction to Mathematical Probability Theory*, 535 pp., Prentice Hall, Englewood Cliffs, New Jersey, 1969.

GCOS/GTOS, GHOST-global hierarchical observing strategy, GCOS-33, WMO-No. 862, World Meteorol. Org., Geneva, 1997.

Germogenova, T.A., *The Local Properties of the Solution of the Transport Equation* (in Russian), 272 pp., Nauka, Moscow, 1986.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G. and Sohlberg, R., Global land cover classification at 1km spatial resolution using a classification tree approach, International Journal of Remote Sensing, Submitted, 1999.

Holben, B. N, Characteristics of maximum-value composite images from temporal AVHRR data, *Int. J. Remote Sens.*, *7*, 1417-1434, 1986.

Jacquemoud, S., F. Barret, and J.F. Hanocq, Modeling spectral and bidirectional soil reflectance, *Remote Sens. Environ.*, *41*, 123-132, 1992.

Justice, C., D. Starr, D. Wickland, J. Privette and T. Suttles, EOS land validation coordination: an update, *Earth Observer*, 10(3):55-60, 1998.

Kimes, D.S., Radiative transfer in homogeneous and heterogeneous vegetation canopies, in *Photon-Vegetation Interactions: Applications in Plant Physiology and Optical Remote Sensing*, pp. 339-388, edited by R. B. Myneni and J. Ross, Springer-Verlag, New York, 1991.

Knyazikhin, Y., and A. Marshak, Fundamental equations of radiative transfer in leaf canopies and iterative methods for their solution, in *Photon-Vegetation Interactions: Applications in Plant Physiology and Optical Remote Sensing*, pp. 9-43, edited by R. B. Myneni and J. Ross, Springer-Verlag, New York, 1991.

Knyazikhin, Y., J. Kranigk, G. Miessen, O. Panfyorov, N. Vygodskaya, and G. Gravenhorst, Modelling three-dimensional distribution of photosynthetically active radiation in sloping coniferous stands, *Biomass Bioenerg.*, *2/3*, 189-200, 1996.

Knyazikhin, Y., G. Mießen, O. Panfyorov, and G. Gravenhorst, Small-scale study of three-dimensional distribution of photosynthetically active radiation in a forest, *Agric. For. Meteorol.*, *88*, 215-239, 1997.

Knyazikhin, Y., J.V., Martonchik, R.B. Myneni, D.J. Diner, and S.W. Running, Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data, *J. Geophys. Res.*, *103*, D24, 32,257-32,275, 1998a.

Knyazikhin, Y., J.V. Martonchik, D.J.Diner, R.B. Myneni, M.M. Verstraete, B.Pinty, and N. Gobron, Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data, *J. Geophys. Res*., *103*, D24, 32,239-32,256, 1998b.

Knyazikhin, Y., J. Kranigk, R.B. Myneni, O.Panfyorov, and G.Gravenhorst, Influence of small-scale structure on radiative transfer and photosynthesis in vegetation cover, *J. Geophys. Res.*, *103*, 6133-6144, 1998c.

Kolmogorov, A.M., *Foundations of the Theory of Probability*, 71 pp., Chelsea, New York, 1950.

Kondratyev, K. Y., *Radiation in the Atmosphere*, 912 pp., Academic, San Diego, Calif., 1969.

Kranigk, J., *Ein Model für den Strahlungstransport in Fichtenbeständen*, 127 pp., Cuvillier, Göttingen, Germany, 1996.

Krein, S.G (Ed.), *Functional Analysis*, 352 pp., Foreign Technol. Div., Wright-Patterson Air Force Base, Ohio, 1967.

Kuusk, A., The hot spot effect of a uniform vegetative cover, *Sov. J. Remote. Sens.*, *3*, 645-658, 1985.

Lebesgue, H., Intégrale, Longueur, Aire, *Thése (Paris 1902) ou Annali Mat Pura e Appl*., vol. 3, Nr. 7, 231-359, 1902.

Li, X., and A.H. Strahler, Geometrical-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing, *IEEE Trans. Geosci. Remote Sens*., *30*, 276-292, 1992.

Li, X., and A. Strahler, Geometric-optical modeling of a coniferous forest canopy, *IEEE Trans. Geosci. Remote Sens., GE-23*, 207-221, 1985.

Li, X., A. Strahler, and C. E. Woodcock, A hybrid geometric optical-radiative transfer approach for modelling albedo and directional reflectance of discontinuous canopies, *IEEE Trans. Geosci. Remote Sens., 2*, 466-480, 1995.

Liou, K.-N., *An Introduction to Atmospheric Radiation*, 392 pp., Academic, San Diego, Calif., 1980.

Loveland, T. R., Merchant, J. W., Brown, J. F., Ohlen, D. O., Reed, B. C., Olsen, P. and Hutchinson, J., Seasonal land cover of the United States, Annals of the Association of American Geographers 85(2), 339-355,1995.

Loveland, T.R. and Belward, A.S., The IGBP-DIS global 1km land cover data set, DISCover: first results, International Journal of Remote Sensing 18, 3289-3295, 1997.

Marshak, A., Effect of the hot spot on the transport equation in plant canopies, *J. Quant. Spectrosc. Radiat. Transfer, 42*, 615-630, 1989.

Martonchik, J.V., D.J. Diner, B. Pinty, M.M. Verstraete, R.B. Myneni, Y. Knyazikhin, and H.R. Gordon, Determination of land and ocean reflective, radiative and biophysical properties using multi-angle imaging, *IEEE Trans. Geosci. Remote Sens., 36*, 1266-1281, 1998.

Myneni, R.B., Modeling radiative transfer and photosynthesis in three-dimensional vegetation canopies, *Agric. For. Meteorol*., *55*, 323-344, 1991.

Myneni, R. B., and G. Asrar, Photon interaction cross sections for aggregations of finite dimensional leaves, *Remote Sens. Environ*., *37*, 219-224, 1991.

Myneni, R. B., and D. L. Williams, On the relationship between FAPAR and NDVI, *Remote Sens. Environ*., *49*, 200-211, 1994.

Myneni, R.B., J. Ross, and G. Asrar, A review on the theory of photon transport in leaf canopies in slab geometry, *Agric. For. Meteorol., 45*, 1-165, 1989.

Myneni, R.B., A.L. Marshak, and Y. V. Knyazikhin, Transport theory for a leaf canopy of finite-dimensional scattering centers, *Quant. Spectrosc. Radiat. Transfer, 46*, 259-280, 1991.

Myneni, R.B., F.G. Hall, P.J. Sellers, and A.L. Marshak, The interpretation of spectral vegetation indexes, *IEEE Trans. Geosci. Remote Sens., 33*, 481-486, 1995a.

Myneni,R.B., S. Maggion, J. Iaquinta, J.L.Privette, N. Gobron, B. Pinty, D.S. Kimes, M.M. Verstraete, and D.L. Williams, Optical remote sensing of vegetation: Modeling, caveats and algorithm, *Remote Sens. Environ*, *51*, 169-188, 1995.

Myneni, R.B., R.R. Nemani, and S.W. Running, Estimation of global leaf area index and absorbed PAR using radiative transfer model, *IEEE Trans. Geosci. Remote Sens., 35*, 1380-1393, 1997.

Myneni,R.B., C.D. Keeling, C.J. Tucker, G.Asrar, and R.R. Nemani, Increased plant growth in the northern high latitudes from 1981 to 1991, *Nature, 386*, 698-702, 1997a.

Myneni, R.B., R.R. Nemani, and S.W. Running, Estimation of global leaf area index and absorbed PAR using radiative transfer model, *IEEE Trans. Geosci. Remote Sens., 35*, 1380-1393, 1997b.

Nilson, T., A theory of radiation penetration into non-homogeneous plant canopies, in *The Penetration of Solar Radiation into Plant Canopies* (in Russian), pp. 5-70, Estonian Acad. of Sci., Tartu, 1977.

Norman, J.M., and J. M. Welles, Radiative transfer in an array of canopies, *Agron. J., 75*, 481-488, 1983.

Oker-Blom, P., J. Lappi., and H. Smolander, Radiation regime and photosynthesis of coniferous stands, in *Photon-Vegetation Interactions: Applications in Plant Physiology and Optical Remote Sensing*, pp. 469-499, edited by R. B. Myneni and J. Ross, Springer-Verlag, New York, 1991.

Peterson, D. L, M. A. Spanner, S. W. Running, and L. Band, Relationship of Thematic Mapper Simulator data to leaf area index, *IEEE Trans. Geosci. Remote Sens., 22*, 323-341, 1987.

Pinty, B., and M.M. Verstraete, Modeling the scattering of light by vegetation in optical remote sensing, *J. Atmos. Sci., 55*, 137-150, 1998.

Pinty, B., M.M.Verstraete, and R.E.Dickenson, A physical model for predicting bidirectional reflectance over bare soils, *Remote Sens. Environ., 27*, 273-288, 1989.

Pokrowski, G.I., Über die Helligkeitsverteilung am Himmel (in German), *Phys. Z., 30*, 697-700, 1929.

Privette, J. L., R. B. Myneni, W. L. Emery, and C. J. Tucker, Invertibility of a 1D discrete ordinates canopy reflectance model, *Remote Sens. Environ., 48*, 89-105, 1994.

Privette, J., R. Myneni, J. Morisette and C. Justice, Global validation of EOS LA1 and FPAR products, Earth Observer, 10(6):39-42, 1998.

Privette, J.L. and G.P. Asner, The prototype validation exercise (PROVE) for EOS land and atmosphere products, *IGARSS'99 Proceedings* (99.110 1), 1999.

Ross, J., *The Radiation Regime and Architecture of Plant Stands*, 391 pp., Dr. W. Junk, Norwell, Mass., 1981.

Ross, J.K., and A. L. Marshak, Calculation of the solar radiation reflection from plant cover using the Monte-Carlo method, *Sov. J. Remote. Sens., 5*, 58-67, 1984.

Ross, J., and T. Nilson, A mathematical model of radiation regime of plant cover, in *Actinometry and Atmospheric Optics* (in Russian), Valgus, Tallin, Estonia, 1968.

Ross, J., Y. Knyazikhin, A. Kuusk, A. Marshak, and T. Nilson, *Mathematical Modeling of the Solar Radiation Transfer in Plant Canopies* (in Russian, with English abstract), 195 pp., Gidrometeoizdat, St. Peterburg, Russia, 1992.

Running, S.W., R.B. Myneni, R. Nemani, J. Glassy, MOD15 LAI/FPAR algorithm theoretical basis document, MODIS LAI (leaf area index) and MODIS FPAR (fraction of absorbed photosynthetically active radiation), *Tech. Rep.*, Sch. of For., Univ. of Mont., Missoula, 1996.

Sellers, P.J., J.A. Berry, G.J. Gollatz, C.B.Field, and F.G. Hall, Canopy reflectance, photosynthesis and transpiration, III, A reanalysis using improved leaf models and a new canopy integration scheme, *Remote Sens. Environ., 42*, 187-216, 1992.

Sellers, P.J., et al., Modeling the exchange of energy, water, and carbon between continents and atmosphere, *Science, 275*, 602-509, 1997.

Simmer, C., and S.A. Gerstl, Remote sensing of angular characteristics of canopy reflectances, *IEEE Trans. Geosci. Remote Sens., GE-23*, 648-658, 1985.

Titov, G.A., Statistical description of radiative transfer in clouds, *J. Atmos. Sci.,47*, 24-38, 1990.

Tucker, C.J., Red and photographic infrared linear combination for monitoring vegetation, *Remote Sens. Environ., 8*, 127-150, 1979.

Tucker, C. J, P. J. Sellers, Satellite remote sensing of primary production, *Int. J. Remote Sens., 7*, 1395-1416, 1986.

Tucker, C.J., Y.Fung, C.D. Keeling, and R.H. Gammon, Relationship between atmospheric CO_2 variations and a satellite-derived vegetation index, *Nature*, *319*, 195-199, 1986.

Vainikko, G.M, Equation for mean intensity in broken clouds (In Russian), *Meteorol. Issled., 21*, 28-37, 1973.

Verma, S. B, P. J. Sellers, C. L. Walthall, F. G. Hall, J. Kim and S. J. Goetz, Photosynthesis and stomatal conductance related to reflectance on the canopy scale, *Remote Sens. Environ., 44*, 103-116, 1993.

Vermote, E., L. A. Remer, C. O. Justice, Y. J. Kaufman, and D. Tanré, Atmospheric correction algorithm: Spectral reflectances (MOD09), version 2.0, Algorithm technical background document, *NASA EOS-ID 2015 Doc*., 42 pp., 1995.

Verstraete, M.M., and B. Pinty, Designing spectral indexes for remote sensing applications, *IEEE Trans. Geosci. Remote Sens*., *34*, 1254-1265, 1996.

Verstraete M.M., B. Pinty, and R.E. Dickenson, A physical model of the bidirectional reflectance of vegetation canopies, *J. Geophys. Res*., *95*, 11765-11775, 1990.

Vladimirov, V.S., Mathematical problems in the one-velocity theory of particle transport, *Tech. Rep. AECL-1661*, At. Energy of Can. Ltd., Chalk River, Ontario, 1963.

Vygodskaya, N.N., and I.I. Gorshkova, *Theory and Experiment in Vegetation Remote Sensing* (in Russian, with English abstract), 248 pp., Gidrometeoizdat, St. Petesburg, Russia, 1987.