ELSEVIER

Remote Sensing of Environment 80 (2002) 418–434

Remote Sensing of Environment

www.elsevier.com/locate/rse

Assessing the information content of multiangle satellite data for mapping biomes I. Statistical analysis

Y. Zhang*, Y. Tian, R.B. Myneni, Y. Knyazikhin, C.E. Woodcock

Department of Geography, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

Received 9 January 2001; received in revised form 4 September 2001; accepted 15 September 2001

Abstract

The insights gained from present land cover classification activities suggest integration of multiangle data into classification attempts for future progress. Land cover types that exhibit distinct signatures in the space of remote sensing data facilitate unambiguous identification of cover types. In this first part, we develop a theme for consistency between cover type definitions, uniqueness of their signatures, and physics of the remote sensing data. The idea of angular signatures in spectral space is proposed to provide a cogent synthesis of information from spectral and angular domains. Three new metrics, angular signature slope (ASSI), length (ASLI), and intercept indices, are introduced to characterize biome signatures. The statistical analyses with these indices confirm the idea that incorporation of the directional variable should improve biome classification result. The consistency principle is tested with the Multiangle Imaging SpectroRadiometer (MISR) leaf area index (LAI) algorithm by examining retrievals when both unique and nonunique signatures are input together with a land cover map. It is shown that this requirement guarantees valid retrievals. Part II provides a theoretical basis for these concepts [Zhang et al., Remote Sens. Environ., in press.].

1. Introduction

The identification and classification of global vegetation into cover types and biomes is valuable for at least two reasons. First, land cover and land use changes inferred from vegetation maps is not only a direct evidence of the human and climate impact on the land but is also a key piece of information required for the study of global biogeochemical cycles (Vitousek, Mooney, Lubchenoco, & Melillo, 1997). Second, most climate and biogeochemical models and algorithms that estimate surface biophysical variables from remote sensing data utilize vegetation maps to assign certain key parameters to reduce the number of problem unknowns (e.g., Bonan, 1998; Knyazikhin, Martonchik, Diner, et al., 1998; Knyazikhin, Martonchik, Myneni, Diner, & Running, 1998; Potter et al., 1993). In short, a fundamental description of global vegetation is simply characterizing the cover type at a certain location and a given time.

Remote sensing data, especially satellite-measured reflected radiation magnitudes at broad wavelength bands in the solar spectrum, have been used to segregate global vegetation into a finite set of spectrally similar classes using cluster analysis, decision trees, neural networks, etc. (Carpenter et al., 1999; DeFries & Townshend, 1994; Friedl & Brodley, 1997; Gopal, Woodcock, & Strahlar, 1999; Hansen, DeFries, Townshend, & Sohlberg, 2000; Moody & Strahler, 1994; Running, Loveland, Pierce, Nemani, & Hunt, 1995). The accuracy of the produced vegetation maps varies widely depending on the training data, input remote sensing data, and the classification method. The insights gained from these activities suggest integration of multiangle data into land cover classification attempts for future progress in this area.

The theme of this two-part series is the following. Land cover types that exhibit distinct signatures in the space of the remote sensing data facilitate unambiguous identification of the cover types. This implies: (a) cover definitions consistent with physics of the data, (b) number of classes dependent on the number of distinct signatures, and (c) spatial resolution of remote sensing data consistent with cover type definitions to minimize the problem of mixtures.

* Corresponding author. Tel.: +1-617-353-8845; fax: +1-617-353-8399.

E-mail address: yuzhang@crsa.bu.edu (Y. Zhang).

0034-4257/01/$ – see front matter
PII: S0034-4257(01)00322-4

The discussion in these papers is limited to spectral and angular dimensions of the optical remote sensing data.

The consistency mentioned above is the following. The physics of processes operative in the generation of optical remote sensing data, namely, radiative transfer, admits certain variables that characterize the transport and interaction of photons with the host medium (Myneni et al., 1995). These variables are determined by the structure and optics of vegetated land surfaces. Land cover definitions, if given in terms of these surface properties, provide consistency with the signal generating processes of the measurement. Possibly, such a consistency principle constitutes a theoretical basis for land cover identification.

This two-part series is organized as follows. The relationship among cover type definitions, spatial resolution of the data, and cover mixtures is addressed first. The spectral and angular signatures of different cover types as recorded in the Polarization and Directionality of the Earth's Reflectances (POLDER) data are presented next. Information from spectral and angular domains is synthesized as angular signatures in spectral space. Two new metrics, angular signature slope and length indices (ASSI and ASLI, respectively), are then introduced and evaluated. The consistency requirement is tested with the Multiangle Imaging SpectroRadiometer (MISR) leaf area index (LAI) algorithm by examining retrievals when both unique and nonunique signatures are input together with a land cover map. It is shown that the consistency requirement guarantees valid retrievals. The second part of this series provides a theoretical basis for these concepts (Zhang, Shabanov, Knyazikhin, & Myneni, in press).

2. Land cover definitions and spatial resolution of the data

Land cover or biome classification is typically based on identifying the spectral signature (and its temporal evolution) of a cover type and utilizing this signature to classify a large region. A common problem with land cover and biome maps is one of misclassification; that is, a pixel is wrongly designated as a certain cover type. However, even when a pixel is designated as the correct cover type, there is still the possibility that the designated cover type is just the dominant cover type, and other cover types can potentially exist at that location. This problem of land cover mixtures generally arises when the chosen set of land cover classes and spatial resolution of the data or the classification map are incompatible. In some instances, however, mixtures occur naturally even at high spatial resolutions, as for example, needle leaf and broadleaf forest mixtures, in which case the mixture should be treated as a separate cover or biome type. Since land cover information is routinely used in retrievals of LAI, fraction of incident photosynthetically active radiation absorbed by vegetation (FPAR) and albedo (Knyazikhin, Martonchik, Myneni, et al., 1998; Lucht, Schaaf, & Strahler, 2000) and in global climate and biogeochemistry models (e.g., Bonan, 1998; Potter et al., 1993), it is important to characterize the degree of land cover heterogeneity. This is facilitated here by the availability of a higher resolution biome map (1 km) compared to the resolution of the POLDER data (6.17 km).

2.1. POLDER data

The POLDER instrument is a wide field of view imaging radiometer designed to measure the directionality and polarization of the sunlight scattered by the Earth/atmosphere system (Deschamps et al., 1994). The instrument acquired eight months of data globally from November 1996 to June 1997 at 6.17-km resolution and eight wavelengths (443, 490, 565, 670, 763, 765, 865, and 910 nm). Polarization of the incident light is measured at three channels, 443, 670, and 865 nm. The Level 2 POLDER multiangle surface reflectance products are geocoded, calibrated, cloud screened, and partially atmosphere-corrected data (Leroy et al., 1997). For

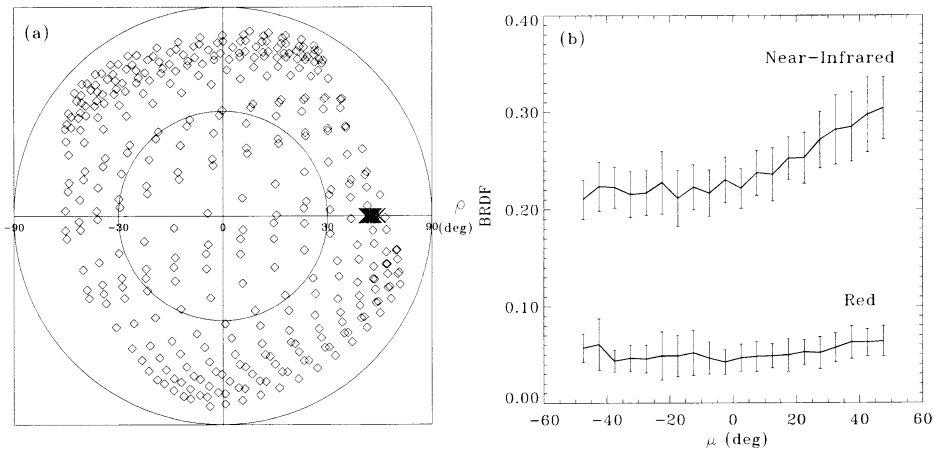

Fig. 1. Angular variation of the BRDFs of a homogeneous patch of needle leaf forest as measured by POLDER in June 1997 (55.36°N, 92.20°W) and the corresponding sun–view geometries. The symbols ◊ and 🌡 denote view and sun directions. The horizontal axis in panel (b) is expressed in terms $\sin^{-1}\mu$, where parameter μ is defined in Fig. 6.

operational reasons, the surface reflectances were not corrected for the tropospheric aerosol effects (Hautecœur & Leroy, 1998). The data thus contain some residual noise, possibly from subpixel clouds and aerosol effects, which can be gauged from the reflectance at 443-nm channel (Bicheron & Leroy, 1999). The POLDER multiangle surface reflectances (BRDF—the bidirectional reflectance distribution function) at red (670 nm) and near-infrared (NIR, 865 nm) wavelengths over North America for the period June 1, 1997 to June 30, 1997 were used in this study. The maximum view angle is approximately 60° and there are up to 14 looks for each multiangle measurement, but the typical number is 12.

The accuracy of surface reflectances after removal of atmospheric effects essentially determines the accuracy of surface parameter retrievals (Martonchik, 1994). Uncertainties of a multiangle measurement can be evaluated from several records of directional reflectances of the same pixel (as in Fig. 1 denote view and sun directions, respectively. The horizontal axis in panel (b) is expressed in terms $\sin^{-1}\mu$, where parameter μ is defined in Fig. 6).The average shape of BRDF and its directional spatial variation, namely BRVF (bidirectional reflectance variance function), can be used to study the typical signatures among different land cover types (Ni & Jupp, 2000; Ni, Woodcock, & Jupp, 1999). The mean BRDF, its variance BRVF for an individual pixel or a group of pixels, are calculated as follows:

$$\overline{\mathrm{BRDF}}(\theta_v,\varphi_v)=\frac{1}{N}\sum_{k=1}^{N}\mathrm{BRDF}_k(\theta_v,\varphi_v) \tag{1}$$

$$\mathrm{BRVF}(\theta_v,\varphi_v)=\frac{1}{N-1}\sum_{k=1}^{N}\left[\mathrm{BRDF}_k(\theta_v,\varphi_v)-\overline{\mathrm{BRDF}}(\theta_v,\varphi_v)\right]^2 \tag{2}$$

where $\overline{\mathrm{BRDF}}$ denotes mean BRDF, BRVF is the variance, and N is number of available records over direction (θ_v,φ_v). Specifically, we evaluate the mean BRDF, BRVF, and the relative uncertainty ε for each pattern of POLDER configuration by substituting (θ_v,φ_v) with (μ_i) for the above formulas (the definition of μ is explained later; cf. Fig. 5). Therefore, the BRDF(μ_i) and BRVF(μ_i) denote the average of BRDF and BRVF values over the interval $[\mu_{i-1}\ \mu_i)$ (see Table 1). The relative uncertainty ε can be calculated as follows (Eq. (3)):

$$\varepsilon(\mu)=\frac{\sqrt{\mathrm{BRVF}(\mu)}}{\overline{\mathrm{BRDF}}(\mu)}. \tag{3}$$

Table 1 shows variation in $\varepsilon(\mu)$ derived from all available observations of a selected needle leaf forest pixel and for each of the three patterns of observation configuration (these are explained later; cf. Fig. 5). The uncertainties of the multiangle records are about 31% at red and 11% at NIR, irrespective of the observation geometry. The higher uncertainty at red is due to a stronger atmospheric effect and smaller vegetation signal (Kaufman & Sendra, 1988). Variations in BRDF corresponding to oblique view angles exhibit similar behavior, i.e., they are maximal at off-nadir directions and minimal at near-nadir looks. Again, this is due to a larger atmospheric effect at off-nadir directions compared to near-nadir views.

Table 1
Mean uncertainties $\varepsilon(\mu)$ in POLDER BRDFs

	Pattern 1 (%)		Pattern 2 (%)		Pattern 3 (%)	
$\sin^{-1}\mu(°)$	Red	NIR	Red	NIR	Red	NIR
−45.0	25.5	9.03	32.4	10.4	9.67	9.01
−37.5	30.9	6.51	46.4	8.77	56.0	19.6
−30.0	30.5	10.1	20.3	9.78	53.8	14.0
−22.5	45.1	7.99	17.9	10.4	43.1	10.9
−15.0	34.3	8.25	22.3	4.38	33.6	9.77
−7.5	31.9	8.70	16.7	9.11	16.5	11.5
0.0	**39.6**	**7.08**	**14.9**	**5.07**	**27.3**	**10.3**
7.5	32.1	8.27	14.2	2.33	22.0	11.2
15.0	42.8	7.61	18.8	5.62	15.6	7.04
22.5	29.0	5.25	19.9	10.9	25.0	10.7
30.0	37.8	8.51	17.9	8.10	12.7	7.82
37.5	23.3	10.3	23.1	12.8	16.5	8.84
45.0	33.2	3.65	13.5	10.3	14.9	8.57

Parameter μ is defined in Fig. 6. Bold numbers indicate the symmetric center of each pattern.

2.2. Biome classification map

A decision tree classification algorithm was used to generate a six-biome North American land cover map from AVHRR normalized difference vegetation index (NDVI) and ancillary data sources at 1-km resolution (Lotsch, Tian, Friedl, & Myneni, submitted). The biome classification scheme segregates global vegetation into six major biome types depending on the vegetation structure, optical properties, and backgrounds (Myneni, Nemani, & Running, 1997). These biome types are: grasses and cereal crops, shrubs, broadleaf crops, savannas, broadleaf forests, and needle leaf forests. The site-based accuracy of this map is 73%. When compared with maps generated from the same data but classified using the International Geosphere Biosphere Program (IGBP) classification scheme (e.g., the EDC map, Loveland et al., 1995) and the UMD map (Hansen et al., 2000)), the six biomes were mapped with approximately 5% higher overall accuracies (Lotsch et al., submitted). This improvement is possibly due to fewer biome classes than IGBP classes.

2.3. Land cover mixtures

Each 6.17-km POLDER pixel encompasses about 36 pixels from the 1-km biome map. Let HF(biome) be the fractional coverage of the various biomes or the homogeneity factor. We assign the predominant biome to the larger

POLDER pixel. Thus, the 6-km map provides information on the large-scale biome distribution and also on the degree of cover heterogeneity within each pixel. The fractional coverage of the dominant type varies considerably and can be less than 50% in the case of severely mixed pixels. If one insists on highly homogeneous pixels, say HF = 90% or 100%, then only 27–10% of the 6-km pixels satisfy this requirement (Table 2). About 20% of the pixels cannot be assigned a biome type because the homogeneity factor is less than 50% at 6-km resolution.

We shall assume the homogeneity factor to be an approximate index of biome patch size. For pixels with HF = 100%, the patch size is 36 km^2. Likewise, the patch size is 18 km^2 for HF = 50%. Strictly speaking, this definition of patch size is not valid when HF is less than 100%, as the smaller 1-km^2 pixels of a particular cover type are not necessarily spatially contiguous within the patch, although the degree of contiguity should be higher for large values of the homogeneity factor. For this reason, this analysis is restricted to homogeneity factors greater than 50%. The relative proportion amongst the six biomes of patches greater than a certain size is shown in Fig. 2. This proportion in the case of broadleaf and needle leaf forests is independent of the patch size, and is approximately 10% and 30%, respectively. In the case of shrubs, there is a skewness towards larger patches, while the opposite is seen in the case of grasses. This is further illustrated with a plot of the patch size distribution in Fig. 3a. Grasses, crops, and savannas show smaller patches with increasing frequency, possibly indicating their prevalence everywhere. Shrubs on the other hand show just the opposite. The shrubs are generally found in arid regions where it is hot and dry, and the location of such regions is usually spatially not fragmented. Hence, not all patch sizes of shrubs are to be found. In the case of forests, the frequency of very large patches is high, and the frequency for patches of all other sizes is constant. The merit of this discussion is not in its rigor, but as an approximate guide for taking the first steps towards addressing the issues of scale and mixtures in this context.

To minimize land cover mixtures, and thus restrict most errors to misclassification only, patches of all sizes must be uniquely assigned a land cover or biome type. While this demands very high-resolution data, one can aim for minimizing the problem of mixtures in, say 60%, of the patches. The critical resolution of the remote sensing data required is now defined by the patch size at which the cumulative size

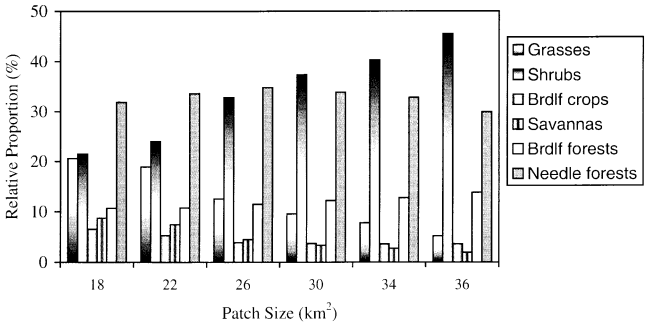

Fig. 2. Relative proportion amongst the six biomes of patches greater than a certain size.

distribution crosses this threshold when cumulated along decreasing patch size (Fig. 3b). For shrubs, which show a high frequency of larger patches, the critical resolution is lower, about 5.7 km. On the other hand, grasses, crops, and savannas, which show smaller patches with increasing frequency, the critical resolution is about 5 km. Thus, the critical resolution is biome dependent. Again, we emphasize the lack of rigor in these arguments. The results are provided for illustrative purposes only.

3. Spectral signatures

The location of reflectance data in the spectral space is the basic source of information about the vegetation canopy conveyed by single-angle multispectral satellite data. This information is typically used for biome identification. The movement of data in this space characterizes changes in canopy properties (Shabanov, Zhou, Knyazikhin, Myneni, & Tucker, in press). This is exploited in the design of vegetation indices (Huete, 1988). The biome spectral signatures, defined as the location of canopy BRDFs in the spectral space, are shown in Fig. 4, which depicts the POLDER BRDF densities in the red–NIR and red–blue spectral spaces as a function of biome type and homogeneity factors. Each contour in Fig. 4 separates an area in the spectral space of high data density containing 50% of the pixels from a given biome. Thus, the density contour shows the most probable location of the biomes in the spectral space. This location depends on canopy structure, optical properties of the leaves and ground, sun–view geometry, and spatial resolution of the data.

Table 2
Distribution of biomes based on homogeneity factors

Homogeneity factor (%)	Biome type						
	Grasses and cereal crops	Shrubs	Broadleaf crops	Savannas	Broadleaf forests	Needle leaf forests	Total
100	0.50	4.39	0.35	0.18	1.34	2.88	9.64
90	**2.55**	**9.95**	**0.98**	**0.88**	**3.24**	**9.00**	**26.60**
50	15.09	19.12	4.21	5.93	8.57	26.75	79.67
All data	20.61	21.55	6.56	8.74	10.68	31.86	100

Distribution values at homogeneity level of 90% are in boldface.

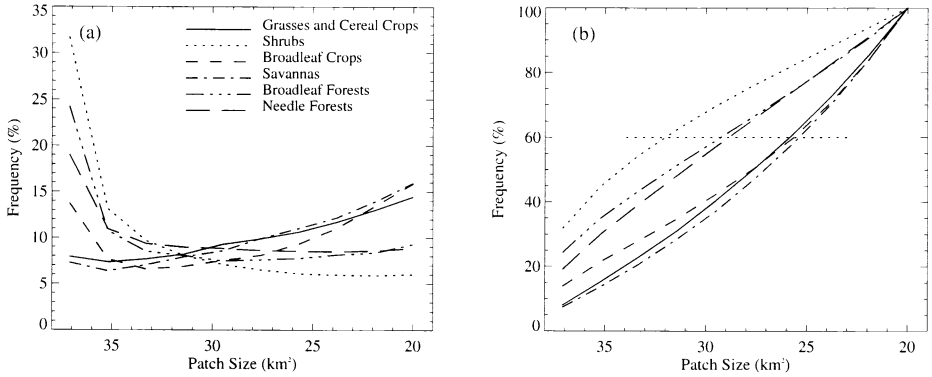

Fig. 3. Biome patch size distributions (a) and the corresponding cumulative distribution functions (b).

The locations of biome data in the spectral space, that is, their spectral signatures, are more distinct for the case of homogeneous patches (HF = 100%) as expected. This is especially true in the red–NIR space than in the red–blue space for the following reason. The canopy-leaving radiation is a function of canopy structure, which is wavelength independent, and optical properties of the leaves and the canopy background. The optical properties of the foliage are similar at blue and red wavelengths. Therefore, the spectral properties of a sufficiently dense canopy are comparable in magnitude at these wavelengths.

The signatures of broadleaf forests, needle leaf forests, shrubs, and to some extent grasses are distinct in the red–NIR space (Fig. 4a). The other two biomes, savannas and broadleaf crops, tend to overlap with broadleaf forests, grasses, and shrubs. Such confusion will lead to misclassification, if only information on location in the red–NIR space is used. Therefore, additional information, preferably angular as it characterizes canopy structure, is needed for biome identification and parameter retrieval.

4. Angular signatures

A vegetated surface scatters shortwave radiation into an angular reflectance pattern, known as the BRDF, the mag-

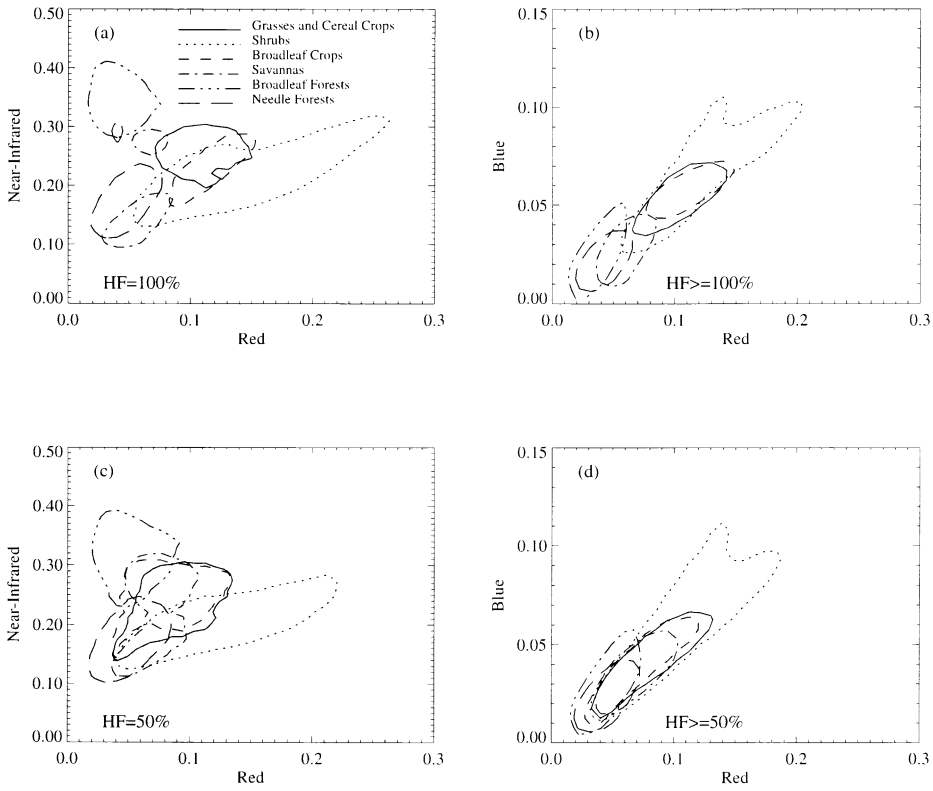

Fig. 4. Contour plot of POLDER BRDF density distribution for different values of the homogeneity factor. Each contour separates an area in the spectral space of high data density containing 50% of the pixels from a given biome and represents all sun and view geometries of the data.

nitude and shape of which is governed by the composition, density, optical properties, and geometric structure of the vegetation canopy. The BRDF is defined as the directional radiance emanating from a target divided by the irradiance (incident flux density) illuminating the target along a single incident angle (Nicodemus, Richmond, Hsia, Ginsberg, & Limperis, 1977). Its shape is informative of canopy structure and ground beneath the canopy.

The angular signatures of vegetation canopies are often demonstrated as variations in reflectance with respect to view polar angle in the principal plane (Brèon et al., 1997). These signatures typically show a sharp peak about the retrosolar direction, the so-called hot-spot effect (Kuusk, 1983), because of the absence of shadows. Strong backscattering is another characteristic feature of vegetation angular signatures. Deriving such signatures from single-look instruments requires compositing many days of data, which introduces considerable uncertainties due to changing atmospheric and surface conditions. However, with instruments such as POLDER and MISR, near-simultaneous looks of the same target are possible and this facilitates the study of angular signatures.

The angular variation of the BRDFs of a homogeneous patch of needle leaf forest as measured by POLDER is shown in Fig. 1. The BRDFs were assembled from 30 different POLDER views of the same pixel during June of 1997, which then were averaged over the records. The corresponding sun–view geometries are also shown in Fig. 1. The variation about the mean reflectance values can be largely attributed to changing measurement geometry and atmospheric conditions because the needle leaf forest can be assumed to be a reasonably stable target during this month. Such variations are indicative of the degree of uncertainty in the measurements, which is a valuable source of information for inverse problems (Wang et al., in press). The geometrical variables required to express the angular signatures are developed in the next section.

The general shape of the BRDFs for different biomes and the corresponding directional spatial variance information (BRVF) can be reconstructed from a group of multiangle measurements by Eqs. (1) and (2). The average BRDF and BRVF shapes are demonstrated in Fig. 5 from homogeneous pixel data of the six biomes. These BRDF shapes for the six biomes show certain differences and in general, the hot-spot can be clearly observed. BRVF shapes appear to be similar to the corresponding BRDFs, which indicates that a variation peak exists in hot-spot area and minimum value at forward scattering. This shape is similar to the BRVF measurements and models (Ni et al., 1999). The vegetation cover is an important determinant of spatial variance, especially when the background is brighter than the vegetation canopy. The contrast between sunlit and shaded tree crowns in the case of forests also contributes to the spatial variance (Ni & Jupp, 2000). Therefore, we can expect the BRVF to also peak in hot-spot direction. The reason that the BRVFs of POLDER data do not show an apparent peak value about the nadir direction, compared to measurements, is probably because of the coarse resolution of the data. For fine resolution data, the nadir view has more probability to capture the bare soil or dense vegetation. The BRVFs exhibit higher value at the edge and lower values in the middle. The high value of BRVF indicates either the directions have more information or the uncertainties of the measurements are high.

5. Geometry of angular signatures

POLDER view directions define a curve on the reference horizontal plane in cylindrical coordinates (Fig. 1a). All curves corresponding to the various multiangle records have similar orientation as determined by the POLDER observation configuration. One of these curves is a near straight line passing through the pole and represents a multiangle record with an almost constant view azimuth. Let this be denoted as the reference curve or line. The other curves are nearly symmetric with respect to the reference line. The angle between the reference line and the horizontal OX axis, measured counterclockwise, is used to parameterize the reference line (Fig. 6). This angle is denoted as the slope of the reference line. The slope angle (φ_s) can change with latitude of the pixel. Its value varies between 30° and 60° for North America. Each direction $\Omega_v=(\theta_v,\varphi_v)$ of a multiangle record can be projected onto the reference line and parameterized in terms of the distance μ between the pole O and its projection (Fig. 6), and assigned a "+" sign if $\mu>0$ and "−" otherwise.

Let $\Omega_d\sim(d,\varphi_d)$ be the direction from a multiangle record closest to the nadir (Fig. 6). Here $\sin^{-1} d$ and φ_d are the corresponding polar angle and azimuth, respectively. A negative value is assigned to d if the curve is located below the reference line, i.e., the distances d and $-d$ correspond to $\Omega_d\sim(d,\varphi_d)$ and $\Omega_{-d}\sim(d,\ 180°+\varphi_d)$, respectively. In this manner, the multiangle record is characterized in terms of the slope, φ_s, the distance, d, and its view directions expressed in terms of the distance μ. The horizontal axis in Fig. 1b shows variation in μ. Based on the foregoing, "typical patterns" of sun–view configurations can be identified by holding φ_s and d constant. For example, we split the set of multiangle records into three patterns. Two of these contain off-nadir directions (corresponding to $+d$ and $-d$) and the BRDF variation is mostly due to changes in the view azimuth. The third represents records close to the reference line and captures variations in the BRDF due to changes in view polar angle.

The above parameterization depends on the slope φ_s distance d and the solar zenith angle (SZA). The probable illumination conditions are summarized in Table 3 (upper part). For a given biome type, probable SZA range and φ_d are selected to include majority of available data (shown as "Data Used" in Table 3). Then three patterns can be defined depending on the value of d; for example, $0.5<d\leq 1$

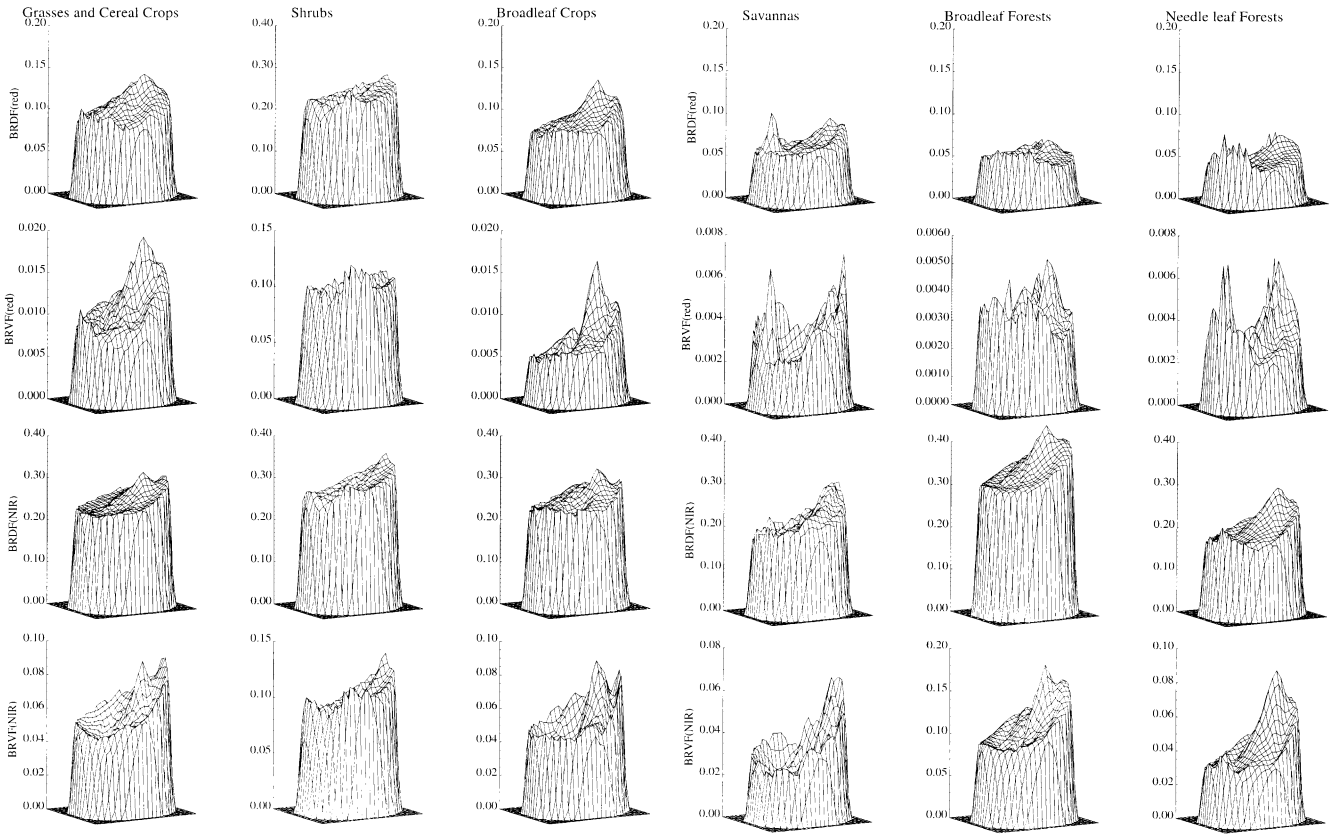

Fig. 5. POLDER data mean BRDFs and BRVFs for six biomes in red and NIR spectral bands.

("Pattern 1, $+d$"), $-0.34 < d \leq 0.5$ ("Pattern 2, 0d"), and $-1 < d \leq -0.34$ ("Pattern 3, $-d$"). The intervals are chosen such that they have an approximately equal probability of occurrence (Table 3). The variation of d within each class is ignored and represented by its mean value. The directional variation is expressed in terms of the distance μ. In the case of Patterns 1 and 3, view polar angles are approximated by the corresponding mean value of $\sin^{-1} d$. Variations in μ, therefore, are due to changing of azimuth. In the case of Pattern 2, $d = 0$ and thus variations in μ are due to changing view polar angle. The three patterns cover more than 70–80% of the available data. Table 3 also shows the mean and standard deviation of d and φ_s for all the patterns. The range of slope angles is, in general, less than 30°. The angular signatures can now be represented as BRDF variations of Patterns 1, 2, and 3.

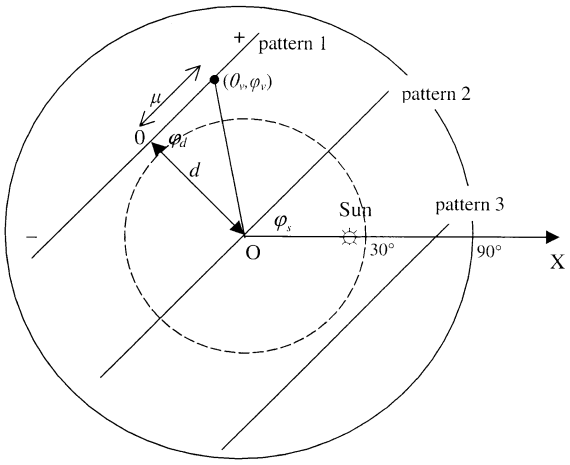

Fig. 6. Parameterization of POLDER observation configurations on a horizontal plane in cylindrical coordinates. A multiangular record through the pole O is taken as a reference line and parameterized in terms of the angle φ_s between this line and OX axis; d is the shortest distance between the multiangle record and the pole O; φ_d is the azimuth of the direction $\Omega_d = (\sin^{-1} d, \varphi_d)$ closest to the nadir. Each view direction $\Omega_v = (\theta_v, \varphi_v)$ from the multiangle record is projected onto the reference line and expressed by the distance $\mu = \sin\theta_v \cdot \cos(\varphi_v - \varphi_s)$ between the projection and the pole O. Thus, the POLDER record is parameterized in terms of φ_s, d, and μ.

The biome-specific BRDF averaged over pixels with homogeneity factors greater than 90% is shown in Fig. 7 for the three angular geometry patterns. Two features of the angular signature can be used to specify the angular variation of the reflected radiation: the magnitude of the BRDF and its shape. Typically, the reflectance of a vegetation canopy tends to vary between 0 and 0.2 at the red band and between 0.1 and 0.4 at the NIR band. Broadleaf and needle leaf forests have the lowest red reflectances, while their NIR reflectance varies about the minimum (needle leaf forests) and maximum (broadleaf forests) values of canopy reflectances. This feature permits identification of these biomes using the magnitude feature. Shrubs can be regarded as the brightest biome, exhibiting almost the highest reflectances at

Table 3
Geometrical characterization of POLDER data

All data	Grasses and cereal crops	Shrubs	Broadleaf crops	Savannas	Broadleaf forests	Needle leaf forests
Sun angle (°)						
Range	17–32	38–53	17–32	23–42	17–32	23–38
Mean	23.7 (5.9)	46.2 (3.2)	24.4 (3.3)	33.0 (6.5)	23.9 (4.0)	30.3 (7.0)
Percent	86.9	69.7	99.8	89.9	95.5	84.9
Slope angle (°)						
Range	25–65	5–45	25–65	5–45	20–60	10–50
Percent	78.0	71.8	92.5	80.2	79.2	80.7
Data used (%)	**70.5**	**68.0**	**85.2**	**74.1**	**74.6**	**79.3**
Pattern 1 ($+d$)						
$\sin^{-1} d$ (°)	42.6 (10.8)	45.3 (11.2)	38.0 (9.0)	45.1 (10.0)	42.6 (10.4)	45.7 (11.4)
Slope (°)	39.9 (9.6)	23.1 (10.2)	35.4 (5.4)	22.0 (6.2)	34.0 (7.1)	24.8 (8.1)
Percent	30.0	28.2	32.8	26.4	30.5	25.6
Pattern 2 (0d)						
$\sin^{-1} d$ (°)	−2.9 (13.5)	−1.6 (14.5)	−6.2 (13.8)	2.0 (13.6)	−2.5 (15.0)	−1.6 (14.8)
Slope (°)	52.5 (9.1)	26.9 (4.6)	50.9 (5.2)	34.1 (6.0)	50.7 (6.3)	37.2 (7.2)
Percent	38.2	32.3	46.7	26.9	31.8	28.2
Pattern 3 ($-d$)						
$\sin^{-1}$ d (°)	−42.6 (8.4)	−43.2 (8.4)	−43.3 (8.6)	−45.3 (8.0)	−41.8 (7.7)	−43.3 (8.3)
Slope (°)	56.7 (7.8)	33.5 (3.5)	60.1 (4.0)	42.4 (3.5)	59.2 (4.0)	44.4 (5.2)
Percent	31.8	39.5	20.5	46.7	37.7	46.2

The numbers in the parenthesis are standard deviations. The numbers in bold indicate the data selected for further analysis in this study.

both red and NIR. On the other hand, savannas, which are mixtures of grasses and woody vegetation, tend to exhibit the lowest reflectances at these bands. The remaining biomes, grasses and crops, have intermediate reflectance magnitudes, that is, between shrubs and savannas.

As for the shape of the BRDF, shrubs have a distinct shape with a sharp jump at about the upper and lower bounds of μ (Pattern 1 of POLDER observation configuration). This corresponds to low sun ($\mu_0 \approx \sin(50°)$) and view ($\mu \approx \sin(40°) \sim \sin(60°)$) polar angles. The view of the forward scattering direction as determined by the distance $\mu \approx \sin(-50°)$ (corresponding to azimuth $\approx 180°$) lies on the principal plane, and thus, the BRDF takes on its local maximum. The distance $\mu \approx \sin(50°)$ (azimuth $\approx 70°$) specifies a back scattering direction close to the hot-spot direction. The values of the BRDF vary about its global maximum in this case. These two extreme situations are responsible for the observed jumps in the canopy angular signature. In the case of grasses and crops, the angular signatures are nearly identical, thus their identification is difficult. In the case of Pattern 2, the backscattering directions lie in the planes close to the principal plane (cf. Table 3), and thus, the hot-spot, although not well discerned, is recorded in the POLDER data. It appears that there are biome-specific features in both the magnitude and shape of the BRDF but there is also considerable variation within a biome type from changing geometry of view and illumination.

6. Angular signatures in spectral space

The angular signatures at different spectral bands are not independent. To demonstrate their correlation, we treat a multiangular record, which contains the angular variation of BRDF as a function of view directions, as a curve on the red–NIR spectral plane. Such curves are shown in Fig. 8 depicts the point corresponding to $\sin^{-1} \mu = -45°$. for the six biomes from the three geometry patterns introduced earlier. These signatures show four features that can be used to distinguish the biomes: (1) location in the spectral space, (2) inclination, (3) length, and (4) intercept. With the exception of grasses and crops that tend to have identical angular signatures in the red–NIR space, the signatures are unique in terms of the above three metrics. The angular signatures are nearly linear in the spectral space. In fact, statistical analysis indicates a significant linear relation between the red and NIR reflectances (Table 4).

Canopy reflectances having the same NDVI value lie on a single line passing through the origin of the red–NIR plane. The NDVI of all biomes, with the exception of shrubs, is insensitive to the view directions. That is, the intercept values from the regression are identical to zero. This is further evidence for the insensitivity of NDVI to view, and from reciprocity arguments, sun angle changes in the case of dense vegetation canopies. The fact that shrubs tend to have a nonzero intercept indicates NDVI sensitivity to view and sun angle variations. This is because the surface reflectances are nonlinear combinations of vegetation (shrub) and ground reflectances. Theoretical arguments further elucidating these concepts are presented in the second part of this series (Zhang et al., in press; also see Kaufmann et al., 2000). The view direction-averaged NDVI values of the different biomes, shown in Table 4, indicate the unique inclination of the angular signatures in the spectral space, with the exception of grasses and crops. The location of the biome data in the spectral space is also

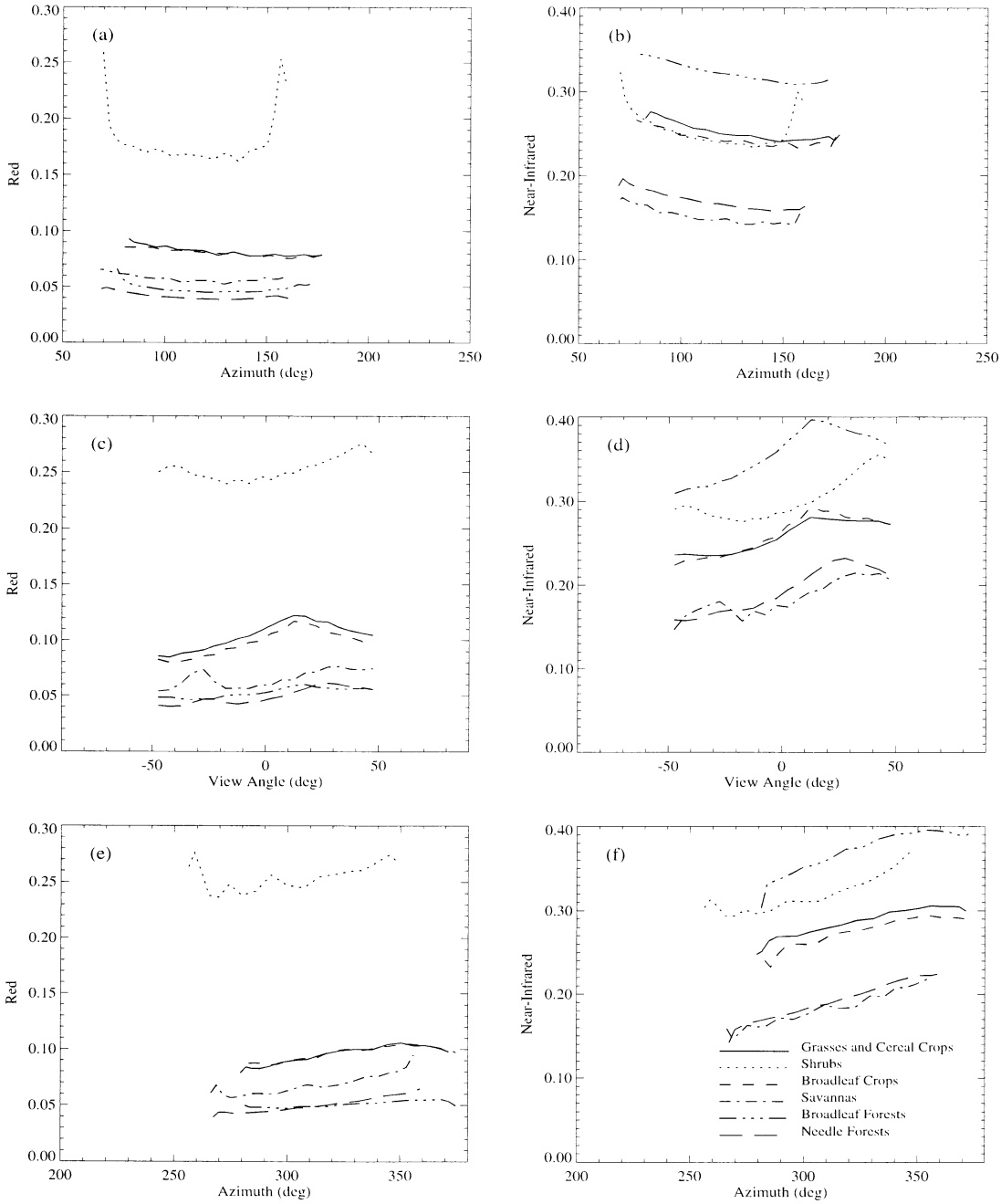

Fig. 7. Biome-specific BRDF signatures obtained from averaging over pixels with HF>90% for the three angular geometry patterns show in Fig. 6. Here, view angle is defined as $\sin^{-1}\mu$ and azimuth is defined φ_v.

distinct as it can be ascertained from the mean red and NIR reflectance values shown in Table 4. A methodology for quantifying the three metrics characterizing the angular signatures in spectral space is given in the next section.

The six biome types proposed by Myneni et al. (1997) were defined in terms of vegetation structural, optical, and background attributes that define variables admitted by the radiative transfer. They argued that this linkage was needed to establish a theoretical basis for the biome identification with remote sensing data. This is similar to the consistency theme mentioned previously. We note from Fig. 7 that the grasses and cereal crops have a signature similar to broadleaf crops. Therefore, potentially five biomes can be identified with multispectral multiangle reflectance data. It is, however, possible that with additional information, say in the form of temporal variation of remote sensing data, this confusion can be resolved, possibly by seasonality and ground cover differences.

This is not to argue that global vegetation can be classified into five biomes only with optical remote sensing

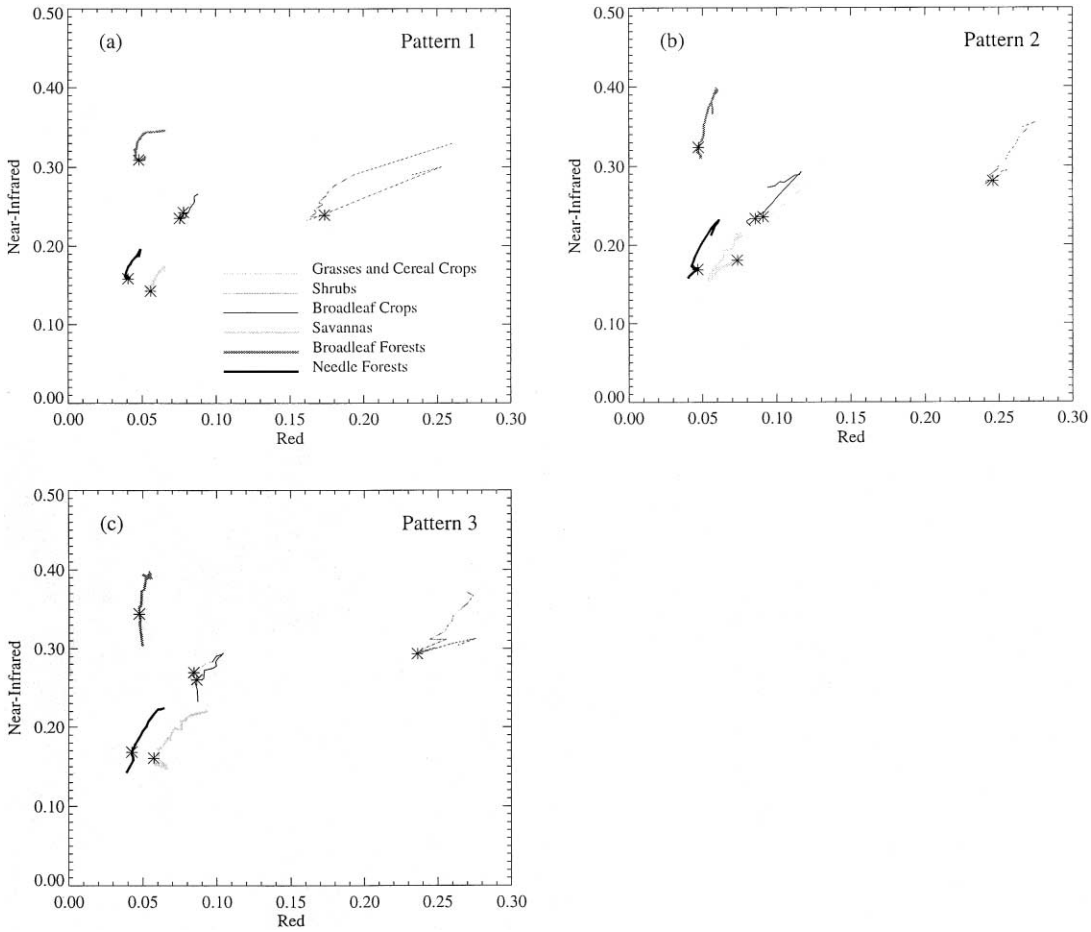

Fig. 8. Angular signatures in the red–NIR spectral space for the three angular geometry patterns. The symbol 🌡 depict the point corresponding to $\sin^{-1}\mu = -45°$.

data. It is possible that other unique signatures exist in the data space that we have not identified. Therefore, we examine below the angular signatures in spectral space of vegetation cover types from a more detailed classification (Hansen et al., 2000).

A 1-km land cover map (Hansen et al., 2000) was used to derive land cover specific angular signatures from the POLDER data. This classification includes 10 vegetation cover types: evergreen needle leaf forest (ENF), evergreen broadleaf forest (EBF), deciduous broadleaf forest (DBF), mixed forest (MF), woodland (WL), wooded grassland (WG), closed shrubland (CS), open shrubland (OS), cropland (CL), and grassland (GL). Deciduous needle leaf forest is not predominant in North America; so, we exclude this cover type. Thus, the major forest covers are ENF, EBF, and DBF. The EBF is located mainly in Central America. The OS represents vegetation in the high-latitude areas. The other land covers are generally distributed in the middle of the continent. The 1-km map was aggregated to 6-km resolution as per homogeneity considerations discussed earlier.

The angular signatures of the 10 land covers are shown in Fig. 9 in the red–NIR space. Fig. 9a depicts the signatures of the forests. Fig. 9b shows the angular signatures of the shrublands, grasslands, and crops. In both panels, the signatures of WLs and WGs are included for better visualization of forest signatures vs. the herbaceous vegetation covers. The signatures of DBF and EBF have similar location and slope in the spectral space. In the six-biome classification, these land covers are aggregated into one biome type, broadleaf forests. The biome and land cover signatures are identical; for instance, compare the broadleaf forest biome signature in Fig. 8 to the signature of EBF and DBF in Fig. 9a. Likewise, the signature of the needle leaf forest biome is comparable to the signature of the ENF. The MFs are located in between the broadleaf and needle leaf biomes. The location depends on the proportion of the two biome types in the mixture. If that were the case, the signature of the MF may be treated as a superimposition of the signatures of the two forest biomes, but this needs further investigation.

The savanna biome is represented by WL and WG in the Hansen et al. (2000) classification. These cover types have identical signatures and are comparable to the signature of the savanna biome, shown in Fig. 7. The signatures of GL, CS, and crop cover types are indistinguishable and this compares to an identical situation with respect to the signatures of grasses and crops in the biome classification scheme. The signature of the OS cover type is comparable to the signature of the shrubs biome. Thus, with the

Table 4
Characteristics of angular signatures in spectral space

Pattern	Grasses and cereal crops	Shrubs	Broadleaf crops	Savannas	Broadleaf forests	Needle leaf forests
Slope						
1	2.287	0.782	2.590	2.648	1.448	3.291
2	1.395	2.456	1.870	2.390	6.391	3.582
3	2.074	1.489	2.596	1.969	7.343	3.186
Intercept						
1	0.065	0.115	0.036	− 0.001	0.252	0.032
2	0.113	− 0.322	0.078	0.028	0.017	0.015
3	0.089	− 0.059	0.025	0.047	− 0.006	0.029
Correlation coefficient						
1	.938	.893	.952	.885	.525	.883
2	.906	.964	.949	.902	.962	.969
3	.953	.747	.923	.889	.747	.971
Mean red reflectance						
1	0.082	0.186	0.081	0.058	0.050	0.042
2	0.104	0.253	0.097	0.065	0.053	0.050
3	0.094	0.256	0.096	0.069	0.051	0.050
Mean NIR reflectance						
1	0.253	0.260	0.245	0.153	0.324	0.171
2	0.258	0.303	0.260	0.184	0.355	0.193
3	0.285	0.321	0.273	0.183	0.368	0.190
Mean NDVI						
1	0.510	0.168	0.505	0.449	0.733	0.605
2	0.426	0.088	0.455	0.476	0.741	0.591
3	0.503	0.113	0.481	0.453	0.757	0.581
Standard deviation						
1	0.007	0.034	0.005	0.012	0.018	0.011
2	0.024	0.022	0.018	0.020	0.005	0.012
3	0.013	0.024	0.011	0.027	0.011	0.012

exception of the MF cover type, the other nine cover types have five distinct angular signatures in the spectral space, and these map to the five biome types proposed by Myneni et al. (1997). This also implies that there would be even fewer unique spectral signatures. It is, however, possible that the cover types have distinct signatures in the spectral–temporal space, as suggested by Hansen et al. (2000), but this needs to be investigated.

7. Slope and length indices

As mentioned previously, the biome-specific angular signatures in the red–NIR spectral space can be characterized by three metrics: (a) their location in the spectral space, which is mainly determined by the biome type (Fig. 4); (b) inclination (slope and intercept) of the signature, which is determined by leaf and soil optical properties and the

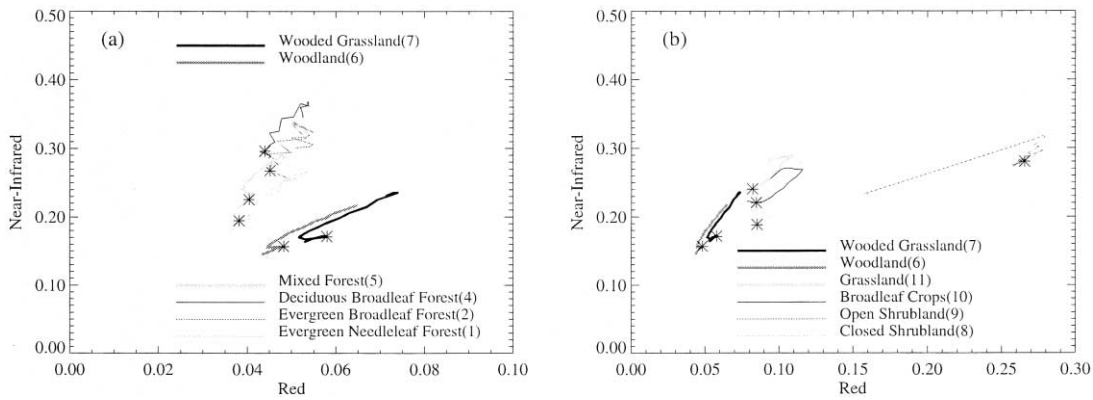

Fig. 9. Angular signatures in the red–NIR spectral space of the 10 land covers from Hansen et al. (2000) 1 km land cover map of North America.

structure of the canopy; and (c) the length of the signature, which describes spectral variation in the shape of the BRDF. A rigorous derivation of these interpretations using radiation transport theory is given in the second part of this series.

The slope of the angular signature in spectral space is quantified by the ASSI (Eq. (4)), defined as

$$\mathrm{ASSI} = \frac{1}{\mu_{\max} - \mu_{\min}} \int_{\mu=\mu_{\min}}^{\mu_{\max}} \left| \frac{\partial \mathrm{BRDF}_{\mathrm{NIR}}(\mu)}{\partial \mathrm{BRDF}_{\mathrm{red}}(\mathrm{m})} \right| \mathrm{d}\mu \approx \sum_{v=1}^{N-1} \left| \frac{\mathrm{BRDF}_{\mathrm{NIR}}(\mu_{v+1}) - \mathrm{BRDF}_{\mathrm{NIR}}(\mu_v)}{\mathrm{BRDF}_{\mathrm{red}}(\mu_{v+1}) - \mathrm{BRDF}_{\mathrm{red}}(\mu_v)} \frac{\mu_{v+1} - \mu_v}{\mu_N - \mu_1} \right| \tag{4}$$

where μ is the distance defined in Fig. 6. This index was evaluated for all pixels with homogeneity factors of 100%. The mean and standard deviations for the six biomes are shown in Table 5. Small variations in the red band and large variations in the NIR band result in large values of ASSI. Therefore, the forest biomes have larger ASSI values compared to the other biomes although there is considerable variability within a biome type. The ASSI can also be taken as the slope of the mean angular signature in the spectral space, if it is a linear or near-linear function. For the pixel-mean signatures shown in Fig. 8, the slope and intercept values are given in Table 4. The forest biomes show a larger slope, consistent with detailed pixel level calculations presented in Table 5. Shrubs show the largest intercept, which as mentioned previously, is indicative of the sensitivity of NDVI to view angle changes.

The length of the angular signature in spectral space can be characterized by the ASLI (Eq. (5)), defined as

$$\mathrm{ASLI} = \frac{2}{\mu_{\max} - \mu_{\min}} \int_{\mu=\mu_{\min}}^{\mu_{\max}} \sqrt{\left(\frac{\mathrm{dBRDF}_{\mathrm{red}}(\mu)}{\mathrm{d}\mu}\right)^2 + \left(\frac{\mathrm{dBRDF}_{\mathrm{NIR}}(\mu)}{\mathrm{d}\mu}\right)^2} \mathrm{d}\mu \approx \frac{2}{\mu_N - \mu_1} \sum_{v=1}^{N-1} \times \sqrt{\left(\mathrm{BRDF}_{\mathrm{red}}(\mu_{v+1}) - \mathrm{BRDF}_{\mathrm{red}}(\mu_v)\right)^2 + \left(\mathrm{BRDF}_{\mathrm{NIR}}(\mu_{v+1}) - \mathrm{BRDF}_{\mathrm{NIR}}(\mu_v)\right)^2} \tag{5}$$

where μ is the distance defined in Fig. 6. This index measures the degree of anisotropy in the reflected radiation field that is dependent on the heterogeneity of the medium.

Table 5
ASSI of biomes

	Grasses and cereal crops	Shrubs	Broadleaf crops	Savannas	Broadleaf forests	Needle leaf forests
Pattern 1 (+ d)						
Mean	2.111	1.996	2.163	2.024	2.569	2.484
S.D.	1.097	1.086	1.080	0.953	1.626	1.363
Pattern 2 (0d)						
Mean	1.682	1.957	2.046	2.398	4.480	3.161
S.D.	0.896	1.023	0.839	1.313	2.590	1.703
Pattern 3 (− d)						
Mean	1.963	2.113	2.323	2.463	4.100	2.990
S.D.	1.080	1.201	1.056	1.368	2.749	1.317

Table 6
ASLI of biomes

	Grasses and cereal crops	Shrubs	Broadleaf crops	Savannas	Broadleaf forests	Needle Leaf forests
Pattern 1 (+ d)						
Mean	0.119	0.226	0.109	0.106	0.134	0.121
S.D.	0.055	0.119	0.047	0.063	0.068	0.068
Pattern 2 (0d)						
Mean	0.162	0.281	0.170	0.208	0.209	0.201
S.D.	0.070	0.146	0.063	0.153	0.095	0.115
Pattern 3 (− d)						
Mean	0.137	0.271	0.146	0.140	0.187	0.151
S.D.	0.048	0.145	0.047	0.066	0.072	0.055

For the case of a homogeneous medium, defined as an isotropic reflector, the ASSL is zero because the angular signature in spectral space is a point. Vegetation canopies exhibit structural features such as preferred orientation of foliage elements, mutual shadowing, vertical layering, and spatial discontinuities, which contribute to the anisotropy of the scattered radiation field.

POLDER measurement configuration Patterns 1 and 3 characterize BRDF variations with respect to view azimuth, and similarly Pattern 2 with respect to view polar angle (Fig. 7). Correspondingly, the length indices evaluated from BRDF variations due to view azimuth and polar angle changes capture lateral and vertical heterogeneity of the medium. The mean length indices for the three measurement configuration patterns are shown in Table 6 for the six biomes. These were evaluated from BRDF data of all pixels with homogeneity factors of 100%. Shrubs have the highest length index values, relative to other biomes, in all three measurement configurations, thus indicating a high degree of both lateral and vertical heterogeneity. The forest and savanna biomes exhibit larger Pattern 2 length index magnitudes compared to grasses and crops. This indicates a higher degree of vertical heterogeneity in the case of the former.

8. Statistical analysis of the information content of POLDER data

Central to the idea of using directional data from POLDER is the information content of the various directional variables. To assess the value and dimensionality of the directional variables, a number of statistical analyses were conducted. In particular, we were curious to know if the

Table 7
MANOVA for six biomes

	Red	NIR	S1	L1	I1	S2	L2	I2	S3	L3	I3
F_{17031}^{5}	1289	2937	5.74	982	4723	589	562	2759	377	229	2821
F_{56497}^{21}	2505		F value at level α=.001; F_{17031}^{5} = 4.2; F_{56497}^{21} = 2.2								

Table 8
Correlation matrix for six biomes

	Red	NIR	S1	L1	I1	S2	L2	I2	S3	L3
NIR	**.57**									
S1	−.06	−.011								
L1	**.59**	.30	−.06							
I1	−.16	**.54**	.005	−.06						
S2	−.30	.02	.09	−.18	.22					
L2	.47	.31	−.02	**.54**	−.03	−.21				
I2	−.02	**.56**	.02	−.05	**.71**	.07	.04			
S3	−.28	−.05	.04	−.15	.15	.20	−.09	.10		
L3	.26	.16	−.01	.27	−.04	−.04	.25	−.01	.26	
I3	.14	**.70**	.004	.05	**.70**	.16	.13	**.65**	.01	.06

Bold numbers indicate significant correlations.

directional signal of the six biomes of interest could be captured by a simpler subset of the directional variables.

The multiangle measurements from each POLDER pixel can be represented as the spectral mean BRDF values and three patterns of angular signatures, each of which are represented by three angular signature characteristics — slope (*S*), length (*L*), and intercept (*I*). Therefore, a total of 11 variables measure the spectral and angular signature of a pixel. Multivariate analysis of variance (MANOVA) tests the null hypothesis that the means of the six biomes are the same. Each variable is tested individually and then all variables are tested together. The results shown in Table 7 indicate a high significance level for all variables (*F* value). This result allows rejection of the highly conservative null hypothesis that the means of the six biomes are the same. Table 7 shows that the two spectral bands and the three intercepts have higher *F* values than the other variables, and thus are likely to be the most useful in separating these biomes. The slope for Pattern 1 (S1) has the lowest *F* value and may not prove as useful as the other variables. Given highly significant results for each variable, the question of the dimensionality of the directional data and possible redundancy between the nine directional variables arises.

To determine if strong correlation exists among the variables, a correlation matrix was calculated (Table 8). In general, it indicates that the correlation between the variables is not very high. This result indicates little redundancy among the directional variables. One possible exception is that all three intercept indices are correlated, indicating that it may not be necessary to use all three in image classification efforts. Examination of the correlation matrix for the individual biomes indicates that for shrubs (not shown here), there are higher correlations among almost all variables, which indicate that the dimensionality in data for shrubs is less than other biomes. The result of little correlation between variables was also confirmed by a principal component analysis (PCA) conducted for each biome, individually and for all the biomes combined. This zero-correlation rotational transformation is frequently used to remove correlation among variables and reduce the number of variables used in analyses. Fig. 10 shows the variance distribution among all the 11 principal components for the different biomes and the aggregated PCA for all biomes. The pattern of eigenvalues for the six biomes indicates that there is considerable dimensionality in the data. The fifth components still contain as much as 10% of the total variance in the dataset, a much larger fraction that would occur in a dataset with 11 variables if extensive correlation existed between variables. While all nine directional variables may not be necessary to characterize the BRDF for improving image classification, it does imply that BRDF shapes are quite complex and may require several variables. Further examination of the pattern of eigenvectors for all six biomes in Table 9 indicates several interesting patterns. The first component, and by definition the component that explains the most variance, is dominated by the intercept variables and the NIR band. This result is not surprising as the MANOVA showed these four parameters to have the highest *F* values, and the correlation matrix indicates they are correlated. The implication is that there is considerable information content in the intercept

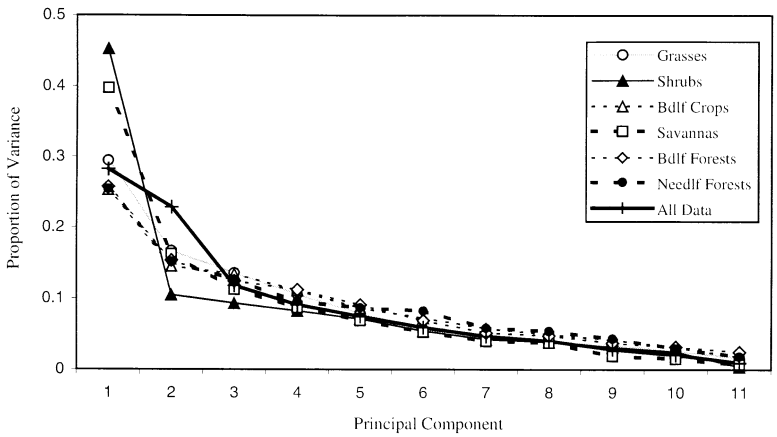

Fig. 10. The variance distribution among principal components for different biomes.

Table 9
Eigenvectors of PCA

Component	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
Red	.27	−.47			.11	.36	.38	.17	.15		.60
NIR	.52					.25	.28	.18		.28	−.69
S1			.11	−.96	−.23						
L1	.21	−.44	.11		.22	−.30	.13	−.73		−.20	−.13
I1	.40	.36				−.17	−.13	−.29	−.22	.65	.34
S2		.32	.25	−.17	.86				.21		
L2	.23	−.38	.12			−.66	−.27	.51			
I2	.41	.28			−.26				.75	−.30	
S3		.22	.68	.15	−.22	−.20	.60				
L3	.13	−.20	.65		−.11	.45	−.55				
I3	.47	.20					−.11	.13	−.56	−.60	.13
Proportion of variance	.28	.23	.12	.09	.07	.06	.05	.04	.03	.02	.01
Cumulative variance	.28	.51	.63	.72	.79	.85	.90	.94	.97	.99	1.00

values, but all three may not be necessary. The second most striking pattern is that Components 3, 4, and 5 are virtually devoid of contribution from the spectral data and hence are essentially directional variables. Each of these components is dominated by different combination of kinds of directional data (e.g., slope, length, and intercept) and the three patterns.

One overall interpretation of these results is that the shapes of BRDFs are complex and it takes many variables to effectively parameterize them and capture their information content. The idea of using three variables (slope, length, and intercept) as well as multiple chords crossing the BRDF is supported by the high dimensionality of the data as indicated in the PCA results and the significance of each of the variables in the MANOVA. In summary, the statistical analyses indicate that both spectral and angular variables are significantly different among six biomes and they all convey information valuable for identifying biomes. Overall, these results strongly confirm the idea that incorporation of the directional variable should improve biome classification.

9. Test for the consistency argument

The estimation of LAI from reflectance measurements requires accomplishing a solution of the radiation transport problem in the inverse mode, which is a well-known ill-posed problem (Kimes, Knyazikhin, Privette, Abuelgasim, & Gao, 2000). It is essential to reduce the dimensionality of this problem, that is, to reduce the number of unknowns. This is done by assuming knowledge of some parameters. These can be either conservative parameters (e.g., leaf optical properties) or those that are difficult to measure (e.g., leaf normal orientation). These parameters are assumed to vary by biome type only and a global distribution of biome types is used as a surrogate for the global distribution of these parameters. Hence, the use of biome maps in retrieval algorithms (Knyazikhin, Martonchik, Diner, et al., 1998; Knyazikhin, Martonchik, Myneni, et al., 1998), and also in many global models of land surface processes (Bonan, 1998; Sellers et al., 1996) and biogeochemical cycles (Potter et al., 1993).

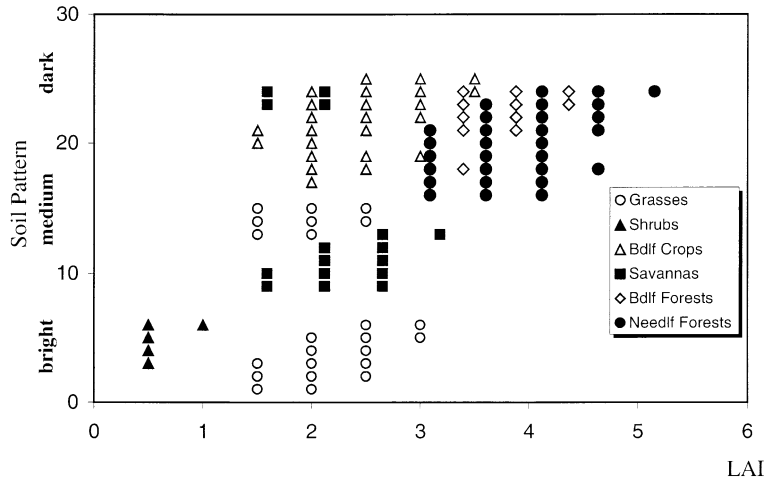

Fig. 11. LAI and soil patterns constituting the solution distribution function of the MISR LAI algorithm.

The problem of retrieving LAI from vegetated surface reflectances (BRDF) can be formulated as follows (Knyazikhin, Martonchik, Diner, et al., 1998): evaluate LAI from information on measurement geometry, multispectral multiangle reflectances, and their uncertainties. The algorithm compares observed and modeled canopy reflectances for a suite of canopy structures and soil patterns that represent a range of expected natural conditions. All canopy/soil patterns for which the modeled and observed BRDFs differ by an amount equivalent to or less than the corresponding uncertainty are considered as acceptable solutions. This is called the solution distribution function. The mean values of LAI obtained from averaging over all acceptable solutions and their dispersions are taken as the final solution and retrieval uncertainty (Knyazikhin, Martonchik, Myneni, et al., 1998; Wang et al., in press; Zhang et al., 2000). Vegetation structural attributes are parameterized in terms of variables that transport theory admits, i.e., this algorithm is compatible with the biome definitions. The model reflectances are evaluated from solution of the three-dimensional radiative transfer equation and expressed as a function of sun–view geometry, canopy/soil pattern, and biome type.

The angular signatures shown in Fig. 8b represent BRDF variations as a function of view polar angle in the red–NIR POLDER band space. The biome-specific signatures were obtained by averaging over all pixels with homogeneity factors of 100%. We note that the signatures of grasses and crops are indistinguishable. The signatures of the other biomes can be treated as distinct (cf. statistical analysis above). The following question tests the consistency required between biome definitions with the physics of remote sensing problem. Can distinct (needle leaf and broadleaf forests) and similar (grasses/cereal crops and broadleaf crops) biome signatures result from canopies with identical LAI values?

The biome signatures shown in Fig. 8b were input to the MISR LAI algorithm. Band and angle-averaged uncertainties in POLDER BRDFs were assumed to be equal to 20%; this corresponds to the mean uncertainty (Table 1, Pattern 2). The resulting LAI and soil pattern distributions for each of the six retrievals are shown in Fig. 11. The mean model reflectance signatures of these solutions are shown in Fig. 12 as angular signature in spectral space for comparability. The LAI values and soil patterns constituting the solution distribution of the two forest biomes are comparable although their reflectance signatures were distinct. This indicates that the information provided by the biome map (biome definitions) and that embedded in the model calculations (physics of the problem) was critical to associating these identical solutions with distinct reflectance patterns. The solution distribution function of grasses/cereal crops and broadleaf crops indicates similar LAI values but different soil patterns, in spite of the fact the reflectance signatures input to the algorithm in this instance were nondistinguishable. The broadleaf crops are

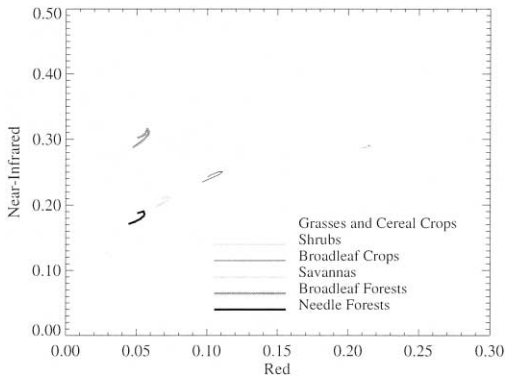

Fig. 12. Modeled angular signatures in the red–NIR spectral space from the MISR LAI algorithm.

associated with darker soils and the grasses with brighter soils. Again, we conclude that the information provided by the biome map and that embedded in the model calculations was critical to associate these different solutions with similar reflectance patterns. These results argue for the consistency required between biome definitions and their signatures with the physics of the spectral–angular reflectance data, for were it not for this consistency, the retrievals would be inaccurate.

10. Concluding remarks

The insights gained from present land cover classification activities suggest integration of multiangle data into classification attempts for future progress. Land cover types that exhibit distinct signatures in the space of remote sensing data facilitate unambiguous identification of cover types. In this first part, we develop a theme for consistency between cover type definitions, uniqueness of their signatures, and physics of the remote sensing data. Angular signatures in spectral space (Fig. 8) provide a cogent synthesis of information from spectral and angular domains (Figs. 3 and 7). These signatures can be characterized in terms of their (a) location in the spectral space, which is mainly determined by the biome type (Fig. 4); (b) inclination (slope and intercept), which is determined by leaf and soil optical properties and the structure of the canopy; and (c) length, which describes spectral variation in the shape of the BRDF. The statistical analyses indicate that both spectral and angular variables are significantly different among six biomes and they all convey information valuable for identifying biomes. The consistency requirement guarantees valid biophysical retrievals because the information provided by the biome map is consistent with the physics of the problem and data (Figs. 11 and 12). Our approach was tested with POLDER data. The availability of multiangle data of higher resolution, less uncertainties, and wider angular coverage can reduce the effect of biome mixture and increase the information content of angular signatures in

spectral space that, in turn, will result in a more accurate biome classification and can lead to more biome parameters that are better specified. Part II provides a theoretical basis for this consistency requirement (Zhang et al., in press).

Acknowledgments

This research was funded by the NASA MODIS and MISR programs. We thank Marc Leroy of CESBIO/CNES for providing the POLDER data. The UMD classification made available to the community is due principally to Matt Hansen, Ruth DeFries, and John Townshend. We thank John Martonchik and Alan Strahler for comments on the manuscript.

References

Bicheron, P., & Leroy, M. (1999). A method of biophysical parameter retrieval at global scale by inversion of a vegetation reflectance model. *Remote Sensing of Environment*, *67*, 251–266.

Bonan, G. B. (1998). The land surface climatology of the NCAR land surface model coupled to the NCAR community climate model. *American Meteorological Society*, *11*, 1307–1327.

Brèon, F. M., Vanderbilt, V., Leroy, M., Bicheron, P., Walthall, C. L., & Kalshoven, J. E. (1997). Evidence of hot spot directional signature from airborne POLDER measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 479–484.

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., & Franklin, J. (1999). A neural network method for efficient vegetation mapping. *Remote Sensing of Environment*, *70*, 326–338.

DeFries, R. S., & Townshend, J. R. G. (1994). NDVI-derived land cover classification at global scales. *International Journal of Remote Sensing*, *15*, 3567–3586.

Deschamps, P. Y., Brèon, F. M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J. C., & Seze, G. (1994). The POLDER Mission: instrument characteristics and scientific objectives. *IEEE Transactions on Geoscience and Remote Sensing*, *32*, 598–615.

Friedl, M., & Brodley, C. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, *61*, 399–409.

Gopal, S., Woodcock, C. E., & Strahlar, A. H. (1999). Fuzzy neural network classification of global land cover from a 1 degrees AVHRR data set. *Remote Sensing of Environment*, *67*, 230–243.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*, 1331–1364.

Hautecœur, O., & Leroy, M. (1998). Surface bidirectional reflectance distribution function observed at global scale by POLDER/ADEOS. *Geophysical Research Letters*, *25*, 4197–4200.

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*, 295–309.

Kaufman, Y. J., & Sendra, C. (1988). Algorithm for automatic atmospheric corrections to visible and near-IR satellite imager. *International Journal of Remote Sensing*, *9*, 1357–1381.

Kaufmann, R. K., Zhou, L., Knyazikhin, Y., Shabanov, N. V., Myneni, R. B., & Tucker, C. J. (2000). Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 2584–2597.

Kimes, D., Knyazikhin, Y., Privette, J., Abuelgasim, A., & Gao, F. (2000). Inversion methods for physically-based models. *Remote Sensing Reviews*, *18*, 381–439.

Knyazikhin, Y., Martonchik, J. V., Diner, D. J., Myneni, R. B., Verstraete, M. M., Pinty, B., & Gobron, N. (1998a). Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data. *Journal of Geophysical Research*, *103*, 32239–32256.

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Diner, D. J., & Running, S. (1998b). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, *103*, 32257–32275.

Kuusk, A. (1983). The hot-spot effect of a uniform vegetative cover. *Earth Research Space*, *4*, 90–99 (in Russian, English translation in *Soviet Journal of Remote Sensing, 3,* 645–658, 1985).

Leroy, M., Deuzè, J. L., Brèon, F. M., Hautecœur, O., Herman, M., Buriez, J. C., Tanrè, D., Bouffies, S., Chazette, P., & Roujean, J. L. (1997). Retrieval of atmospheric properties and surface bidirectional reflectances over land from POLDER/ADEOS. *Journal of Geophysical Research*, *102*, 17023–17037.

Lotsch, A., Tian, Y., Friedl, M. A., & Myneni, R. B. (2000). Land cover mapping in support of LAI/FAPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors. *International Journal of Remote Sensing*, (submitted).

Loveland, T. R., Merchant, J. W., Brown, J. F., Ohlen, D. O., Reed, B. C., Olsen, P., & Hutchinson, J. (1995). Seasonal land cover of the United States. *Annals of the Association of American Geographers*, *57*, 1453–1463.

Lucht, W., Schaaf, C. B., & Strahler, A. H. (2000). An algorithm for the retrieval of albedo from space using semiempirical BRDF models. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 977–998.

Martonchik, J. V. (1994). Retrieval of surface directional reflectance properties using ground level multiangle measurements. *Remote Sensing of Environment*, *50*, 303–316.

Moody, A., & Strahler, A. (1994). Characteristics of composited AVHRR data and problems in their classification. *International Journal of Remote Sensing*, *15*, 3473–3491.

Myneni, R. B., Maggion, S., Iaquinta, J., Privette, J. L., Gobron, N., Pinty, B., Verstraete, M. M., Kimes, D. S., & Williams, D. L. (1995). Optical remote sensing of vegetation: modeling, caveats and algorithms. *Remote Sensing of Environment*, *51*, 169–188.

Myneni, R. B., Nemani, R. R., & Running, S. W. (1997). Estimation of global leaf area index and absorbed PAR using radiative transfer models. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 1380–1393.

Ni, W., & Jupp, L. B. (2000). Spatial variance in directional remote sensing imagery recent developments and future perspectives. *Remote Sensing Reviews*, *18*, 441–479.

Ni, W., Woodcock, C. E., & Jupp, L. B. (1999). Variance in bi-directional reflectance over discontinuous plant canopies. *Remote Sensing of Environment*, *69*, 1–15.

Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., & Limperis, T. (1977). *Geometrical considerations and nomenclature for reflectance* (NBS Monograph No. 160, 52 pp.). National Bureau of Standards, U.S. Department of Commerce.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*, 811–841.

Running, S. W., Loveland, T. R., Pierce, L. L., Nemani, R. R., & Hunt Jr., E. R. (1995). A remote sensing based vegetation classification logic for global land cover analysis. *Remote Sensing of Environment*, *51*, 39–48.

Sellers, P. J., Los, S. O., Tucker, C. J., Justice, C. O., Dazlich, D. A., Collatz, G. J., & Randall, D. A. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs: Part II. the generation of global fields of terrestrial biophysical parameters from satellite data. *Journal of Climate*, *9*, 706–737.

Shabanov, N. V., Zhou, L., Knyazikhin, Y., Myneni, R. B., & Tucker, C. J. (2001). Analysis of interannual changes in northern vegetation activity

observed in AVHRR data during 1981 to 1994. *IEEE Transactions on Geoscience and Remote Sensing* (in press).

Vitousek, P. M., Mooney, H. A., Lubchenoco, J., & Melillo, J. M. (1997). Human domination of Earth's ecosystems. *Science*, *277*, 494–499.

Wang, Y., Tian, Y., Zhang, Y., El-Saleous, N., Knyazikhin, Y., Vermote, E., & Myneni, R. B. (2001). Investigation of product accuracy as a function of input and model uncertainties: case study with SeaWiFS and MODIS LAI/FPAR Algorithm. *Remote Sensing of Environment* (in press).

Zhang, Y., Shabanov, N., Knyazikhin, Y., & Myneni, R. B. (2001). Assesing the information content of multiangle satellite data for mapping biomes. 2: theory. *Remote Sensing of Environment* (accepted September 2001).

Zhang, Y., Tian, Y., Knyazikhin, Y., Martonchik, J. V., Diner, D. J., Leroy, M., & Myneni, R. B. (2000). Prototyping of MISR LAI and FPAR algorithm with POLDER data over Africa. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 2402–2418.