ELSEVIER

Leaf Optical Properties with Explicit Description of Its Biochemical Composition: Direct and Inverse Problems

Th. Fourty,[*,†] *F. Baret,*[*] *S. Jacquemoud,*[‡] *G. Schmuck,*[§] *and J. Verdebout*[§]

This study presents a methodology to estimate the leaf biochemical compounds specific absorption coefficients and to use them to predict leaf biochemistry. A wide range of leaves was collected including variations in species and leaf status. All the leaves were dried out. The biochemical composition was measured using classical wet chemistry techniques to determine lignin, cellulose, hemicellulose, starch, and protein contents. Concurrently, leaf reflectance and transmittance were measured with a high spectral resolution spectrophotometer in the 800–2500 nm range with approximately 1 nm spectral resolution and sampling interval. In addition, infinite reflectance achieved by stacking leaves was also measured. The PROSPECT leaf optical properties model was first inverted over a selection of wavebands in the 800–2400 nm domain to provide estimates of the scattering characteristics using leaf reflectance, transmittance, and infinite reflectance data. Then, the model was inverted again over all the wavelengths to estimate the global absorption coefficient, using the previously estimated scattering properties. The global absorption coefficient was eventually explained using the measured biochemical composition by fitting the corresponding specific absorption coefficients after substraction of the measured contribution of the residual structural water absorption. Results show that the derived specific absorption coefficients are quite robustly estimated. Further, they are in good agreement with known absorption features of each biochemical compound. The average contribution of each biochemical compound to leaf absorption feature is also evaluated. Sugar, cellulose, and hemicellulose are the main compounds that contribute to absorption. Results demonstrate the possibility of modeling leaf optical properties of dry leaves with explicit description of leaf biochemistry. Estimates of the detailed biochemical composition obtained by model inversion over the 1300–2400 nm spectral domain show poor predictive performances. In particular, the protein content is very poorly retrieved. The retrieval performances of several combinations of the biochemical compounds are investigated. Results show that the total amount of dry matter per unit leaf area is the only variable to be accurately retrieved. Possible improvements of these results are discussed.

INTRODUCTION

Knowledge of the canopy biochemical composition may provide critical information to describe and predict vegetation productivity, litter decomposition processes, or nutrient cycles within an ecosystem (Running et al., 1985). The understanding of biogeochemical cycles involved in an ecosystem requires the measurement of canopy biochemical composition generally estimated by ground level sampling techniques; however, these techniques are tedious and time-consuming. For that reason, the representativeness of ground measurements, as well as the fine spatial and temporal distribution of the biochemical, are generally poorly achieved, particularly when large ecosystems are investigated.

Remote sensing allows large and continuous radiometric measurements from which the biophysical and biochemical characteristics of canopy may be derived.

[*]INRA Bioclimatologie, Site Agroparc, Avignon, France
[†]GEOSYS, Parc Technologique, Ramonville, France
[‡]Department of Land, Air and Water Resources, University of California, Davis
[§]IRSA / AT, Joint Research Centre, Ispra, Varese, Italy

Address correspondence to Th. Fourty, INRA, Station de Bioclimatologie, Domaine Saint Paul, Site Agroparc, 84914 Avignon Cedex 9, France.

Received 2 March 1995; revised 5 September 1995.

REMOTE SENS. ENVIRON. 56:104–117 (1996)

655 Avenue of the Americas, New York, NY 10010

0034-4257 / 96 / $15.00
SSDI 0034-4257(95)00234-0

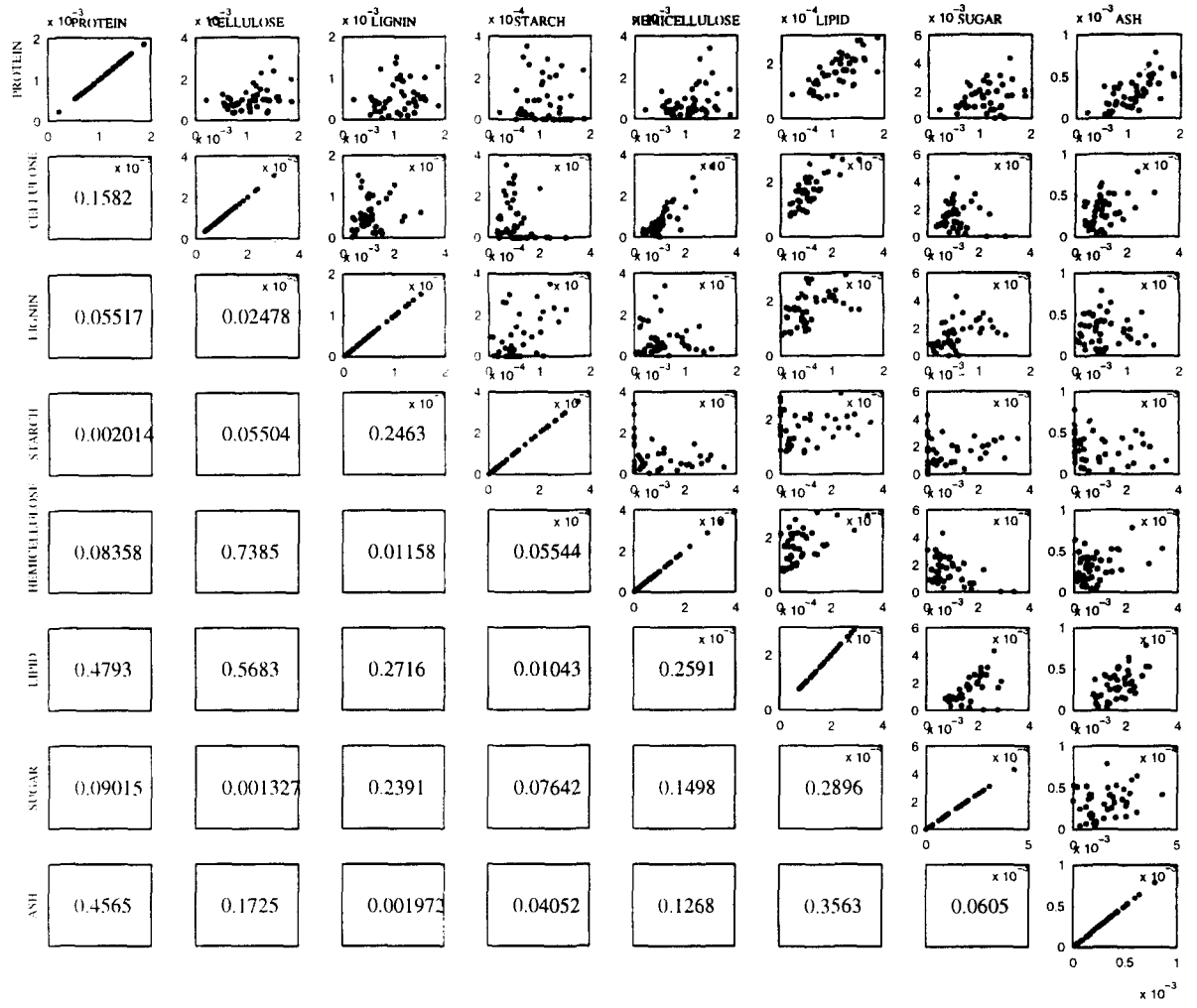


Figure 1. Correlations between the concentrations observed over the 43 leaves. The concentrations are expressed in g cm^{-2}. The number in the lower diagonal matrix is the associated R^2 value.

Since Peterson et al. (1988), many experimental results showed statistical correlations between the canopy biochemical composition and the corresponding reflectance spectra recorded by spectroimaging systems (Gastellu-Etchegory et al. 1994; 1995; Johnson et al., 1994; Wessman et al., 1988; Zagolski, 1994; Martin and Aber, 1993; 1994b; Smith and Curran, 1992). However, recent developments demonstrate that the relationships elaborated on one site on fresh leaves had very poor predictive performances when applied to another site or even to another year (Grossman et al., 1994). The lack of consistency and robustness of these empirical approaches forces to propose a more analytical way to describe the possible relationships between canopy reflectance and its biochemical composition.

Table 1. Minimum, Average and Maximum Mass Fraction (% of Total Dry Mass) of the Main Biochemical Compounds Observed over the 43 Leaf Types Studied

Biochemical Compound	*Minimum*	*Average*	*Maximum*
Cellulose	9.1	18.9	33.8
Hemicellulose	0.3	14.3	38.7
Lignin	1.1	9.93	27.5
Protein	7.1	20.13	36.8
Starch	0	1.7	6.5
Sugar	0	26.5	48.6
Lipid	2	2	2
Ash	0.9	5.6	13.6

Deformations of the biochemical bonds (stretching, rotations, or vibrations) between light atoms (C, H, O, N) absorb at specific fundamental frequencies and their harmonics (Curran, 1989). In the near-infrared domain (800–2500 nm), the absorption features result from the combination of harmonics and overtones of the fundamental frequencies of each chemical bond. Since many years, near-infrared spectroscopists have developed a successfully large body of knowledge based on statistical relationships between the biochemical composition of dried and powder materials and their reflectance spectra

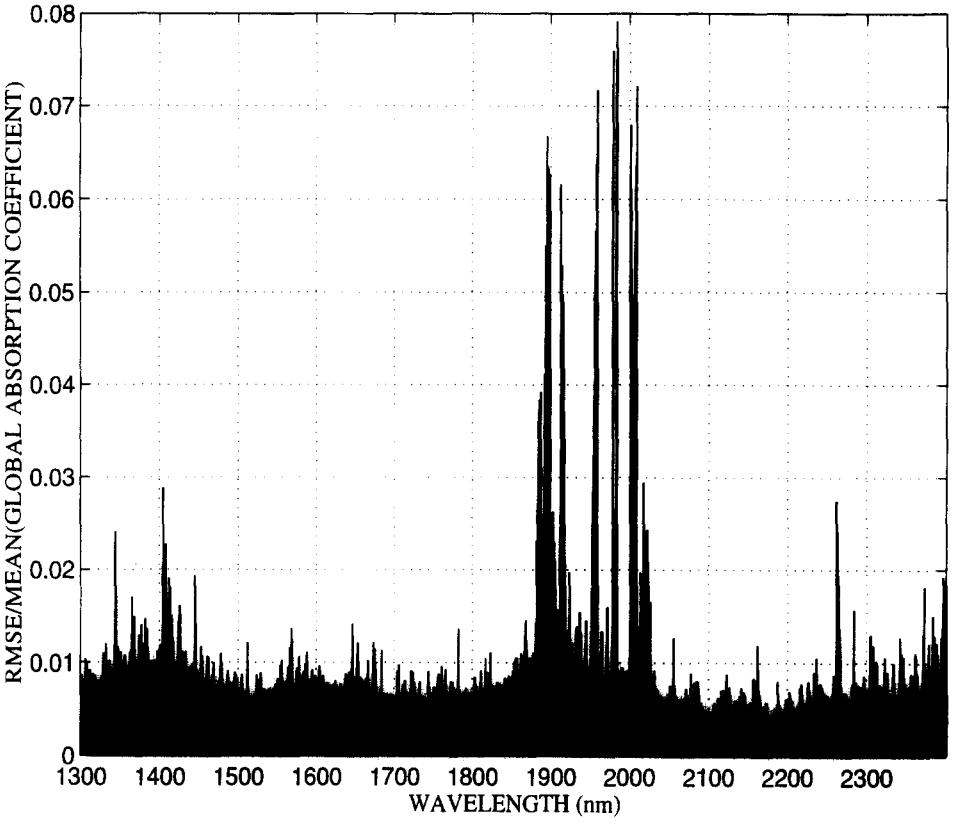
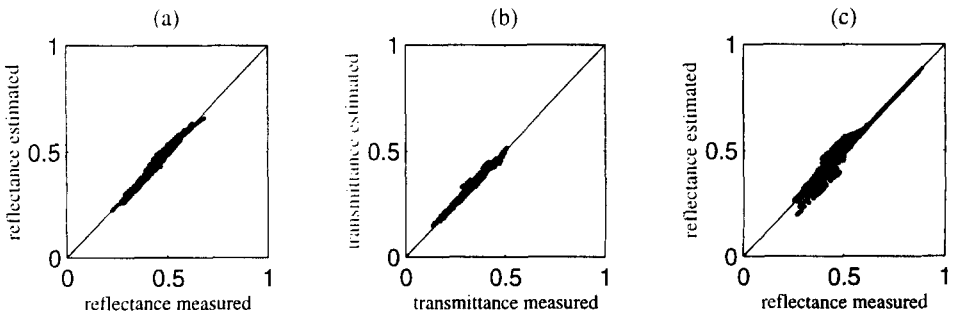


Figure 2. RMSE values associated to the comparison between the absorption coefficients evaluated at a 10-nm spectral sampling interval after interpolation at a 1-nm interval and that obtained originally at a 1-nm spectral sampling interval. This is computed over the 43 leaves investigated.

analysis (Williams and Norris, 1987; Norris et al., 1976; Marten et al., 1989; Weyer, 1985). However, due to the statistical nature of these relationships, their robustness is always questionable. Further, when applied to intact leaves or to canopies, retrieval of the biochemical composition is much more complex due to the strong water absorption that masks the weakest absorption features of compounds such as lignin, protein, cellulose, and starch. Canopy structure, soil background, or the atmosphere disturbances act as additionally confounding factors that complicate the interpretation of the radiometric signal.

Canopy reflectance results from elementary scattering processes that take place at leaf or soil levels. Thus, modeling leaf optical properties will allow one to investigate canopy reflectance sensitivity to its biochemical

Figure 3. Comparison between the measured reflectance (a), transmittance (b), and infinite reflectance (c) and the corresponding values simulated using the PROSPECT model using the retrieved values of $K(\lambda)$, N, and a parameters.



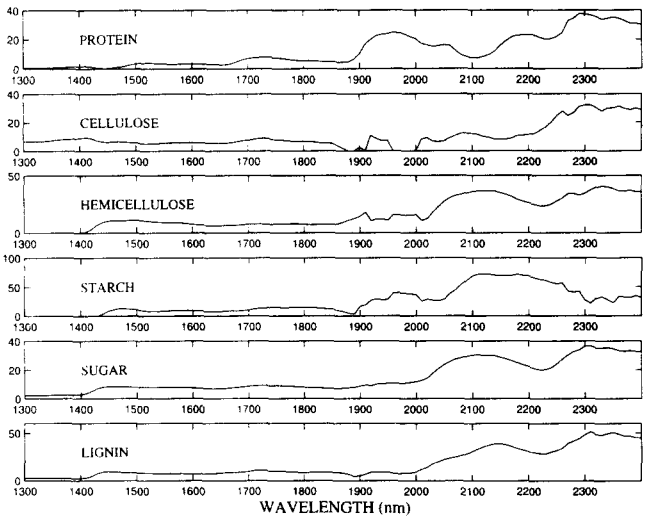


Figure 4. Specific absorption coefficients ($cm^2\ g^{-1}$) of each biochemical compound.

composition and will potentially provide a tool to design algorithms dedicated to the retrieval of canopy biochemical composition.

This study aims at developing a model of leaf optical properties that takes explicitly into account leaf biochemical composition. We will use the PROSPECT model (Jacquemoud and Baret, 1990) to describe the radiative transfer in the leaf. This simple model assumes that the leaf is made up of a pile of elementary layers separated by air spaces. The number of layers (N) mimics the scattering processes within the leaf internal structure. Each layer is characterized by a refraction index (n) and an absorption coefficient $K(\lambda)$. Assuming random spatial distribution of each compound, this global absorption coefficient $[K(\lambda)]$ is determined by the leaf biochemical composition:

$$K(\lambda) = \sum_{i=1}^{i=n} k_i(\lambda) \cdot C_i, \qquad (1)$$

where k_i and C_i are respectively the specific absorption coefficient and concentration (in weight per unit leaf area) of compound i. Thus, the modeling reduces to the estimation of the specific absorption coefficients for each compounds. This will be achieved through inversion of the PROSPECT model. To minimize the strong masking effect of water absorption, we will mostly consider dry leaf materials. To enhance the absorption features of weak absorbers, we measured the reflectance of a semi-infinite medium made up of stacked leaves.

In the first part of this article, we describe the data set. The second part presents the methodology used to estimate the specific absorption coefficients. The third part investigates the performances of model inversion to estimate the leaf biochemical composition. Discussion and conclusion then follow with a special emphasis to the possible remote detection of canopy biochemical compositions.

THE DATA SETS

Forty-three leaf types corresponding to a wide range of species and biophysical status were collected by the Joint Research Centre at Ispra (Italy) (Hosgood et al., 1995). The leaves were subjected to a gentle drying to remove most of the water without changing too much the biochemical composition. Two kinds of measurements were performed on each leaf type: leaf optical properties and biochemical composition analyses.

Optical Properties Measurements

Reflectance and transmittance measurements were performed in the 400–2500 nm range using a Perkin-Elmer spectrophotometer equipped with an integrating sphere. This type of measurement configuration minimizes problems related to the nonlambertian properties of plant leaves. This instrument provides a spectral resolution around 2 nm depending on the wavelength, with a 1-nm sampling interval. The output signal was calibrated into absolute directional / hemispherical reflectance or transmittance using spectralon references. The reflectance and transmittance of five leaves for each leaf type were acquired and then averaged with calibration. The noise was very small, close to 0.05% of the signal. Additional measurements of infinite reflectance were collected over a pile of more than 20 leaves for each leaf type.

Biochemical Composition Analyses

For each leaf type, a subsample of leaf materials was sent for analyses to a laboratory (C.R.A. Gembloux, 100 Serpont Road, B-6800 Libramont-Chevigny, Belgium). The following were measured:

cellulose	Weende (1985)
hemicellulose	NDF-ADF (Van Soest and Masson, 1967; Van Soest and Wine, 1991)
Lignin	Van Soest (Van Soest and Masson, 1967; Van Soest and Wine, 1991)
protein	Kjeldhal (AOAC, 1970)
starch	Ewerts (1985)
ash	elemental microanalysis

The remaining dry matter was explained as being made of 2% lipid (average value taken from the literature (Penning de Vries et al., 1976; Williams et al., 1987) and of cell solubles. For convenience, we used

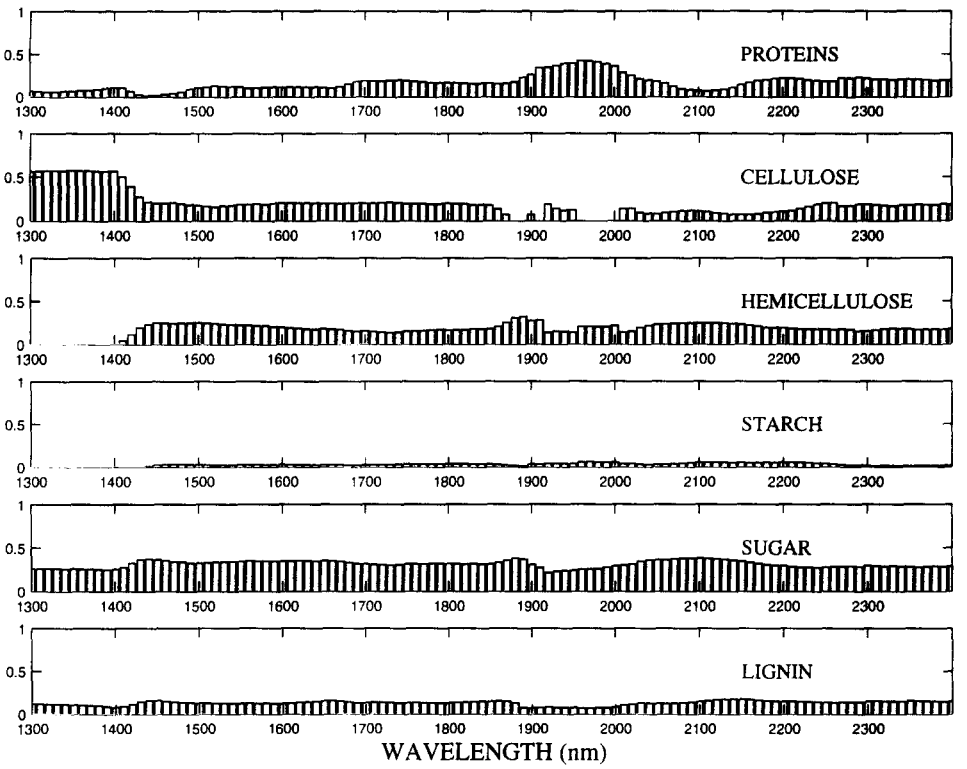


Figure 5. Contribution (λ_i) of each biochemical compound to the global absorption coefficient. The contribution of each compound is computed as the product of the specific absorption coefficient [$K_i(\lambda)$] and the average concentration ($\overline{C_i}$) divided by the average global absorption coefficient ($\overline{K}$): $\gamma_i \frac{k_i(\lambda) \cdot \overline{C_i}}{\overline{K}}$.

the generic name "sugar" for cell solubles, sugar representing most of it.

The remaining structural water content was evaluated after drying the samples at 85°C for 48 h. It represented an average of 4.15% of the dry mass. The biochemical concentration originally expressed in mass of compound per unit leaf dry mass was transformed into mass of compound per unit leaf area using measurements of the specific leaf area. We eventually came up with the average biochemical composition as presented in Table 1. Almost no correlation between the concentrations of the biochemical compounds were observed, except for cellulose and hemicellulose and for lipid (Fig. 1). The correlation with the lipid concentration can be explained by the fact that it was not actually measured but was set to 2% of the dry mass.

MODELING LEAF OPTICAL PROPERTIES

Modeling leaf optical properties amounts to estimate the specific absorption coefficients, used in the PROSPECT model as described earlier. This will be achieved in two steps: The first one consists in estimating for each sample a global absorption coefficient $K(\lambda)$ through inversion of the PROSPECT model. In the second step, we will compute the specific absorption coefficients for each biochemical compound using the retrieved $K(\lambda)$ spectra and the measured concentrations of each biochemical compound.

Estimation of the Global Absorption Coefficient

The PROSPECT model computes the leaf optical properties as a function of a structural parameter (N) that mimics the scattering process and a global absorption coefficient $K(\lambda)$:

$$[\rho(\lambda),\tau(\lambda)] = \text{PROSPECT}(K(\lambda),N), \quad (2)$$

where $[\rho(\lambda),\tau(\lambda)]$ are respectively leaf spectral reflectance and transmittance. It also simulates the reflectance of semiinfinite medium by assigning large values of the N parameter. In that case, the radiative transfer theory shows that only an unscaled absorption coefficient can be retrieved. For example, the simple two-stream Kubelka and Munk (1931) model states that infinite reflectance is a simple function of the ratio between the absorption and the scattering coefficients. Infinite reflectance depends obviously on the mixing ratio between the scatterers and the absorbers. To mimic this process, the PROSPECT model can be run using a large value of the N parameter (N_∞) and a

multiplicative coefficient (a) that accounts for the mixing ratio between scatterers and absorbers:

$$\rho_\infty(\lambda) = \text{PROSPECT}(K(\lambda), a, N_\infty), \quad (3)$$

where $\rho_\infty(\lambda)$ is the infinite reflectance. After some tests, we fixed $N_\infty = 300$ that provides accurate simulation of the infinite reflectance while avoiding numerical problems occurring with greater values. Inverting the PROSPECT model consists in finding the set of parameters $K(\lambda)$, N, and a that minimizes the distance between the measured and simulated values of reflectance, transmittance, and infinite reflectance of the leaves. The inversion is performed using the simplex algorithm (Nelder and Mead, 1963), a very good compromise between the robustness of the solution and the computing time. However, since an inversion applied simultaneously over the 800–2400 nm domain would be very computer time-consuming, we decided to select a few wavelengths for the estimation of the structural parameters N and a. Ten wavelengths were chosen to represent well the scattering and the absorption features: four of them located in the near-infrared plateau where minimum absorption occurs (780–820–880–920 nm); the other ones along the near-infrared region (1400–1500–2100–2200–2300–2400 nm). We noticed that this approach provide more robust estimates of the structural parameters N and a than the inversion applied only over the near-infrared plateau where absorption is minimum. Subsequently, we used the N and a values previously estimated to invert the model a second time and retrieve the global absorption coefficient $K(\lambda)$ at each wavelength.

The effect of the spectral sampling interval was investigated on the retrieved global absorption coefficients values, by comparing the original 1-nm sampling interval data with those obtained by linear interpolation between 10-nm spectral sampling interval data. Figure 2 shows that the corresponding root mean square errors (RMSEs) values are very small. It mostly corresponds to the instrumental noise and to that associated with the inversion process for the retrieval of the absorption coefficient, especially in the water absorption bands. These RMSE values being globally very small, we concluded that a spectral resolution lower than 10 nm would not provide any significant additional information. To reduce the computing time, we thus used a 10-nm sampling interval instead of the original 1 nm.

To evaluate the performances of the inversion process, we reconstructed the reflectance, transmittance, and infinite reflectance spectra using the retrieved values of the absorption coefficient and structural parameters N and α and compared them with the original spectra. The little scatter around the 1:1 line observed mainly for the infinite reflectance (Fig. 3) is due to the higher sensitivity of infinite reflectance to slight inaccuracies in the absorption coefficient. The RMSE values associated with leaf reflectance, transmittance, and infinite reflectance reconstruction are very small and are, respectively, 0.011, 0.012, and 0.01544. This demonstrates that the PROSPECT model describes quite accurately the optical properties of dry leaves. In conclusion, this inversion process provides 43 global absorption coefficients spectra corresponding to the 43 leaf types investigated. Using the measured biochemical composition, we will now derive the specific absorption coefficients of each leaf biochemical compound.

Determination of the Specific Absorption Coefficients for Individual Biochemical Compound

For each leaf sample, the contribution of the remaining structural water absorption to the global absorption coefficient $K(\lambda)$ was removed to provide the actual dry leaf global absorption coefficient $K_d(\lambda)$:

$$K_d(\lambda) = K(\lambda) - k_w(\lambda)C_w, \quad (4)$$

where C_w is the remaining structural water equivalent thickness of the softly dried leaf and $K_w(\lambda)$ is the specific absorption coefficient of water taken from Curcio and Petty (1951). The resulting 43 absorption coefficients spectra (l wavelengths) were stored in a matrix $K_d[l,43]$ corresponding to the biochemical composition matrix $C[n,43]$, where n is the number of biochemical compounds. Solving Eq. (1) to derive the specific absorption coefficients $k_i(\lambda)$ reduces to

$$[K_d(\lambda) = \sum_{i=1}^{n} k_i(\lambda) \cdot C_i] \Leftrightarrow [K_d = k \cdot C] \Rightarrow [k = K_d \cdot C^{-1}], \quad (5)$$

where the matrix $k[l,n]$ is the matrix of the specific absorption coefficients. The lack of significative correlations between the concentrations of each biochemical compound make the inversion of matrix $[C]$ possible. The values of k were constrained to be positive to avoid a physically meaningless negative value. We used the nonnegative linear least-square algorithm to perform this matrix inversion (Lawson and Hanson, 1974).

Figure 4 shows the specific absorption coefficient spectra of six leaf biochemicals except lipid and ash. As a matter of fact, the lipid fraction is small, and its concentration is highly correlated with the other ones as explained earlier; the ash fraction is also very small. Also, because of the mineral nature of ash, no well-defined and strong absorption features are expected. It follows that taking into account these constituents might have decreased the retrieval performances for the other constituents.

We present only the results corresponding to the 1300–2400 nm domain. Brown pigments (polyphenols) appeared during the drying process of the leaves and for wavelengths below 1300 nm; their absorption features overlapped the absorption features of the other biochemical constituents (results not presented).

The specific absorption coefficients range from 0 to

Table 2. Interpretation of the Absorption Peaks According to Curan (1989) and Himmelsbach et al. (1988)

λ (nm) Measured	λ (nm) Literature	Absorption Mechanisms	Absorbing Compounds
		Protein	
1420	?	?	?
1520	1510	Stretch NH	Protein, nitrogen
1730	1730	Stretch CH	Protein
1940	1940	Stretch OH, deformation OH	Water, lignin, protein, nitrogen, starch, cellulose
1960	1980	Asymmetric NH	Protein
2060	2060	Rotation N = H, NH, stretch NH	Protein, nitrogen
2200	2180	Rotation NH, stretch CH, CO, C = O, NH	Protein, nitrogen
2270	2240	Stretch CH	Protein
2290	2300	Stretch NH, C = O, rotation CH	Protein, nitrogen
2350	2350	Rotation CH^2, deformation CH	Cellulose, protein, nitrogen
		Cellulose	
1410	?	?	?
1470	1490	Stretch OH	Cellulose, sugar
1550	1540	Stretch OH	Starch, cellulose
1730	1736	Stretch OH	Cellulose
1770	1780	Stretch CH, OH, deformation HOH	Cellulose, sugar, starch
1820	1820	Stretch OH, stretch CO	Cellulose
1920	1924	Stretch OH, deformation OH	Cellulose
1950	1950	Stretch OH, deformation OH	Water, lignin, protein, nitrogen, starch, cellulose
2020	?	?	?
2090	2100	Rotation OH, deformation OH, stretch COC, CH	Starch, cellulose
2260	2270	Rotation CH, CH^2, stretch OH, CH^2	Starch, cellulose
2300	2280	Stretch CH, deformation CH^2	Starch, cellulose
2380	?	?	?
		Sugar	
1450	1450	Stretch OH, stretch CH	Starch, sugar, lignin, water
1560	1580	Stretch OH	Starch, sugar
1720	?	?	?
1750	1780	Stretch CH, stretch OH, deformation HOH	Cellulose, sugar, starch
1910	?	?	?
1950	1960	Stretch OH, bond OH	Sugar, starch
2110	?	?	?
2140	?	?	?
2280	2270	Rotation CH, CH^2, stretch OH, CH^2	Cellulose, sugar, starch
2310	?	?	?
2340	?	?	?
2380	?	?	?
		Lignin	
1450	1450	Stretch OH, stretch CH	Starch, sugar, lignin, water
1670	1690	Stretch OH	Lignin, starch protein
1710	?	?	?
1760	1754		Lignin
1950	1940	Stretch OH, deformation HOH	Water, lignin, protein, starch cellulose
1980	?	?	?
2050	?	?	?
2150	?	?	?
2270	2262	Stretch CH, stretch C = C	Lignin
2310	?	?	?
2350	2232	Stretch C	?
2380	2380	Stretch OH, aromatic deformation	Lignin
2340	2340	Stretch OH, CH, deformation CH, OH	Cellulose
2360	2350	Rotation CH^2, deformation CH	Cellulose, protein, nitrogen
2390	?	?	?

Continued

Table 2. (Continued)

λ (nm) Measured	λ (nm) Literature	Absorption Mechanisms	Absorbing Compounds
		Hemicellulose	
1460	?	?	?
1500	1490	Stretch OH	Cellulose
1540	1540	Stretch OH	Starch, cellulose
1580	1580	Stretch OH	Starch, sugar
1780	1780	Stretch OH, CH, deformation HON	Cellulose, starch, sugar
1830	1820	Stretch OH, CO	Cellulose
1910	?	?	?
1960	1960	Stretch OH, bond OH	Sugar, starch
2000	2000	Deformation OH, CO	Starch
2120	2100	Rotation O = H, stretch CO, C = O	Starch, cellulose
2140	?	?	?
2280	2280	Stretch OH, deformation CH^2	Starch, cellulose
2330	2340	Stretch OH, CH, deformation OH, CH	Cellulose
2380	?	?	?
		Starch	
1470	1450	Stretch OH, CH	Starch, sugar, lignin, water
1450	1450	Stretch OH, stretch CH	Starch, sugar, lignin, water
1670	1690	Stretch CH	Lignin, starch, protein, nitrogen
1760	?	?	?
1950	1940	Stretch OH, deformation OH	Water, lignin, protein, nitrogen, starch, cellulose
1980	?	?	?
2100	?	?	?
2180	?	?	?
2270	?	?	?
2350	?	?	?

50 $cm^2\ g^{-1}$, except for starch that reaches 70 $cm^2\ g^{-1}$. However, these high values do not compensate enough the small mass fraction of starch in the leaf (Table 1) to provide a significant contribution to leaf absorption (Fig. 5).

From a general point of view, in the 1300–1850 nm region, absorption is low and smooth, without strong spectral features. Conversely, between 1850 nm and 2400 nm, absorption is more important and spectral features are more marked, probably because this spectral domain is closer to the fundamental frequencies domain. Strong analogies are observed among the specific absorption coefficients of cellulose, hemicellulose, starch, and sugar, especially around 2100 nm. All these compounds are made of C—O, C—H, and O—H bonds. Table 4 shows that most of the absorption features in this spectral domain correspond to the O—H bonds. The specific absorption coefficient of lignin differs from that of the other compounds by the low level of absorption observed around 2050 nm. Proteins have quite distinct absorption features with the highest absorption levels in the 1900–2000 nm range that corresponds to N—H bonds. Figure 5 shows the contribution (γ_i) of each biochemical compound to the global absorption coefficient. It has been computed as the product of the specific absorption coefficient [$k_i(\lambda)$] and the average concentration ($\overline{C_i}$) divided by the average global absorption coefficient ($\overline{K}$):

$$\gamma_i = \frac{k_i(\lambda) \cdot \overline{C_i}}{\overline{K}}. \tag{6}$$

Sugar has the highest contribution over the whole wavelength range considered. It is also the compound that contributes the more to the leaf mass (Table 1). Its specific absorption coefficient pattern confirms that assumption. The contributions of starch, lignin, sugar, and hemicellulose are very smooth and constant over the 1300–2400 nm range. They do not exhibit important characteristic peaks. In contrast, cellulose shows a strong peak in the 1300–1400 nm domain, related to the higher specific absorption coefficients values observed over the same spectral domain. The same observation applies for the protein in the 1900–2000 nm.

Table 2 shows the correspondence between the absorption peaks reported in the literature for pure and mixed material (Curran, 1989; Himmelsbach et al., 1988) and those derived as described previously. We interpreted hemicellulose absorbing peaks according to the cellulose, sugar, and starch specific absorbing features, these three compounds having similar chemical bonds. We find generally a good agreement within a 10-nm window between the position of peaks reported

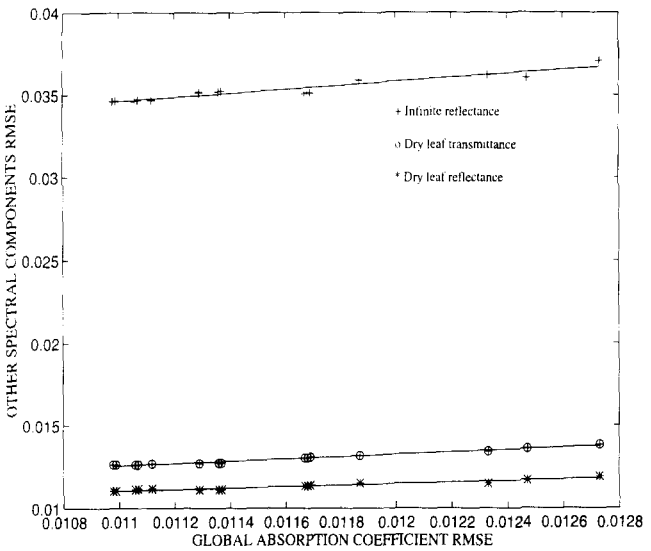


Figure 6. Sensitivity of the reconstruction performances of single leaf reflectance, transmittance, and infinite reflectance to the accuracy with which the global absorption coefficient is reconstructed. Each data point corresponds to one of the biochemical composition presented in Table 3.

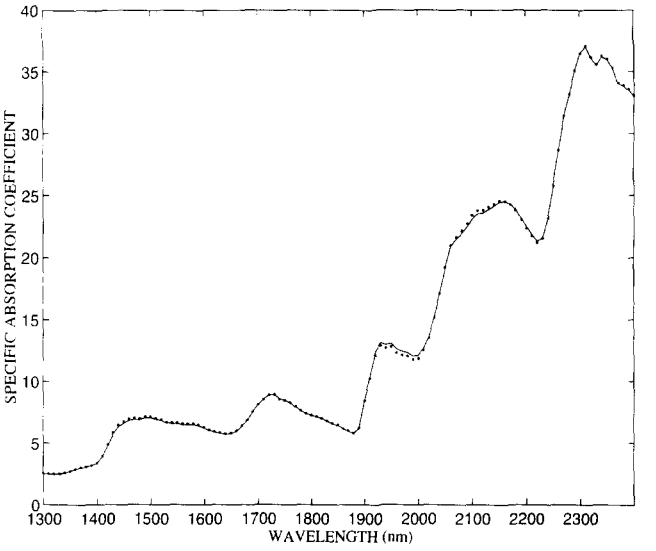
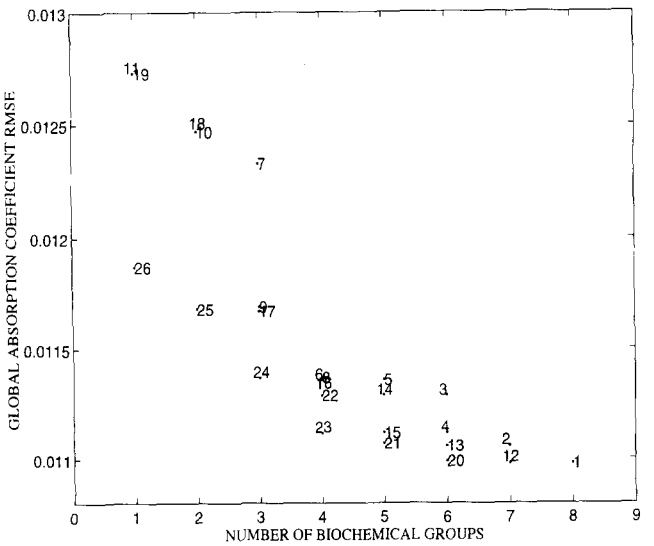


Figure 8. Comparison between the specific absorption coefficient computed for dry matter [all biochemicals (except lipid and ash) grouped together] and that computed using individual specific absorption coefficients and the average biochemical composition.

in the literature and ours. The 10-nm window roughly corresponds to the accuracy with which the absorption peaks can be located, in view of the 10-nm spectral sampling interval used. Despite this general good agreement, some observed absorption peaks do not correspond to well identified absorption peaks.

These results demonstrate that modeling dry leaf optical properties with explicit description of the biochemical composition is possible. We will now investigate the performances of model inversion for the retrieval of leaf biochemical composition.

Figure 7. RMSE values associated with the global absorption coefficient reconstruction as a function of the number of compounds or biochemical groups taken into account (described by a label explained in Table 3).



RETRIEVAL OF LEAF BIOCHEMICAL COMPOSITION

Retrieval of the Detailed Biochemical Composition

The PROSPECT model using the set of specific absorption coefficients previously determined is inverted to retrieve the leaf biochemical composition. The inversion is performed using the same set of leaf reflectance and transmittance spectra as described earlier. The inversion is achieved again in two successive steps: the first one over the 10 spectral bands selected to get the leaf structural parameters N and the mixing ratio a. The model is then inverted in the second step over each wavelength, using the previous N and a values to estimate the global absorption coefficient. The concentration of each biochemical compound is computed using the linear equation (5). The retrieval performances associated to each biochemical compound are characterized by the following T statistic:

$$T = 1 - \frac{\text{var}(C - \hat{C})}{\text{var}(C)}, \qquad (7)$$

where C and $\hat{C}$ are, respectively, the measured and estimated concentrations. The usually computed R^2 statistic characterizes well the relationship between two variables according to a linear model, whereas the T statistic is more appropriate to evaluate the equality of two variables that are in our case the measured and estimated values. T is a normalized measure of the

distance between the estimated and the observed values. In contrast to the R^2, T values can be negative when $\text{var}(C - \hat{C}) > \text{var}(C)$, indicating very poor retrieval performances. To better evaluate the predictive performances of the inversion, we used a cross validation technique (Wallach and Goffinet, 1987), and computed the corresponding T values. Results indicate that none of the biochemical compound is accurately retrieved (columns 1–8 of Table 4). Nevertheless, the orders of magnitude are preserved. This can be explained by the weakness of the absorption observed for each individual biochemical compound, as compared to the relatively small range of variation we would like to be sensitive to. To bypass these dramatic limitations, we will investigate the possibility of grouping some biochemical compounds together to enhance their absorption features. This will possibly lead to define a simplified way to describe leaf biochemical composition.

Optimizing the Way To Describe the Biochemical Composition for Its Retrieval

The estimation of the specific absorption coefficients was performed for several combinations of biochemical groupings. We started from the more detailed description of the leaf biochemical composition using individual compounds, and then progressively degraded the description by grouping similar compounds together. For instance, we started with cellulose and hemicellulose that are well correlated (Fig. 1) and have similar chemical structures and specific absorption coefficients. We keep protein and lignin separate from the other compounds because their chemical structures and absorption coefficients are different. This process finally ends with the simplest description of the biochemical composition that corresponds to the sum of all compounds, thus leaf dry weight per unit leaf area, that is the specific leaf area. Table 3 presents the several groupings investigated and the associated reconstruction performances for the absorption coefficient, single leaf reflectance, and transmittance as well as infinite reflectance spectra. We also investigated the cases where lipid and ash are taken explicitly into account, although that was not done for the detailed composition formerly analyzed.

The accuracy of the reconstructions of the spectral variables such as reflectance, transmittance, and infinite reflectance depends obviously on the accuracy of the global absorption coefficient reconstruction. Figure 6 shows that infinite reflectance is significantly more sensitive to errors in the global absorption coefficient than are single leaf reflectance and transmittance. This can be explained by multiple scattering that enhances the absorption features (Baret et al., 1994). As illustrated in Figure 7, the accuracy of the reconstruction depends also on the number of biochemical groups accounted for. The more groups or individual compounds are separately taken into account, the lower the associated RMSE. However, the increase in the accuracy becomes very small when the number of groups exceeds 4–5. Moreover, the accuracy is higher when lipid or ash are ignored, which is in good agreement with our previous observations when studying the detailed biochemical composition.

For each grouping investigated, we characterized the retrieval performances of the concentration using the T statistics. As in the previous section, we used the same cross validation technique to get a predictive value and to compute the associated T statistic. The more detailed biochemical composition are associated with very small T values corresponding to very poor retrieval performances. Good results ($T > 0.7$) appear when cellulose, hemicellulose, and sugar are grouped together. This group of biochemicals represents from 50% to 60% of the whole leaf dry mass and almost all the carbonated molecules. We note that adding the starch fraction to these three last compounds does not improve the T values, while adding the lignin fraction increases T up to 0.88. Taking the other compounds into consideration does not lead to significant improvements.

The best retrieval performance of concentration is observed when grouping protein, cellulose, hemicellulose, sugar, starch, and lignin together. The associated T value reaches 0.9. This way of describing simply the biochemical composition corresponds to the total dry matter per unit of leaf area.

The specific absorption coefficient corresponding to the dry matter (protein + cellulose + hemicellulose + sugar + starch + lignin) results mostly from the absorption features of each individual biochemical compound. The comparison between the specific absorption coefficient of the dry matter and that computed as a weighted sum of the specific absorption coefficients of each individual compound is excellent (Fig. 8). This means that it is possible to interpret accurately the absorption peaks of the specific absorption coefficient derived for the dry matter according to the absorption peaks of the detailed composition.

CONCLUSION

This study demonstrates that it is possible to use a physically based approach to describe the effects of the biochemical composition on the leaf optical properties. We propose to incorporate the specific absorption coefficients of most leaf biochemical families such as cellulose, hemicellulose, sugar, starch, lignin, and protein into the PROSPECT leaf optical properties model. These specific absorption coefficients features are in good agreement with those described in the literature, proving the soundness of the approach used. Further, the wide range of leaf types used in this study ensures

Table 4. *T* Values Associated with the Retrieval of the Biochemical Composition Using the Model Inversion and the Previously Estimated Specific Absorption Coefficients[a]

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
Protein	−0.07	−0.06	−0.06	−0.02	−0.03	−0.04	−0.05	−0.03	−0.07	−0.09	
Cellulose	0.27										
Hemicellulose	−0.30										
Starch	−0.72	−0.51		−0.26							
Sugar	−0.13	−0.03	−0.47								
Lignin	−1.42	−1.02	−0.70	−0.7	−0.61	−0.60		−0.61			
Lipid	0.00	0.00	0.00	0.00	0.00						
Ash	−33.3	−27.3	−21.3	−33.7	−22.4			−23.5	−31.6		
Cellulose + Hemicellulose		0.35									
Lipid + Ash						−19.3					
Cellulose + Hemicellulose + Starch			0.25								
Cellulose + Hemicellulose + Sugar				0.73							
Lignin + Ash + Lipid							−0.43				
Cellulose + Hemicellulose + Sugar + Starch					0.73	0.73	0.71				
Cellulose + Hemicellulose + Sugar + Starch + Lipid								0.74			
Cellulose + Hemicellulose + Sugar + Starch + Lignin											
Cellulose + Hemicellulose + Starch + Lignin + Lipid									0.88		
Protein + Cellulose + Hemicellulose + Sugar + Starch + Lignin											
Cellulose + Hemicellulose + Sugar + Starch + Lignin + Ash											
Cellulose + Hemicellulose + Starch + Sugar + Lignin + Lipid + Ash										0.87	
Protein + Cellulose + Hemicellulose + Sugar + Starch + Lignin + Ash											
Protein + Cellulose + Hemicellulose + Sugar + Starch + Lignin + Lipid + Ash										0.88	

[a] The model was inverted over single leaf reflectance $\rho_s(\lambda)$ and transmittance $\tau_s(\lambda)$ and the infinite reflectance $\rho_\infty(\lambda)$. Several ways to describe the leaf biochemical composition are compared. They are described by the number presented in the first column which was explained previously in Table 3. Italic numbers correspond to biochemical composition whose lipids are not accounted for. Bold numbers correspond to biochemical composition whose lipid and ash are not accounted for.

the result to be applied over most of the leaves encountered over the Earth's surface.

A spectral resolution lower than 10 nm does not bring any significant amount of information since the absorption features observed for the biochemical compounds are quite broad.

The inversion of the PROSPECT model using the retrieved specific absorption coefficients spectra to estimate the detailed biochemical composition leads to very poor performances. The order of magnitude of the composition is roughly approximated, but the fine variations of the biochemical composition observed among the various leaves are not well estimated. This may be due to the low sensitivity of leaf optical properties to minor variations of the biochemical composition. Further, cellulose, hemicellulose, starch, sugar, and lignin have similar absorption features. Only protein has distinct absorption patterns with an important peak around the 1900–2000 nm corresponding to the N—H bonds. However, even for this compound, the retrieval performances are very poor. This is quite unfortunate since protein (nitrogen) is a key factor governing many physiological processes at the leaf level such as photosynthesis (Aber and Federer, 1992).

The poor retrieval performances observed for the detailed biochemical composition disagree with results issued from near-infrared spectrometric (NIRS) techniques commonly applied on dried ground or fresh samples (Norris et al. 1976; Jacquemoud et al., 1994; Wessman et al., 1988; Peterson et al., 1988; Gastellu-

Table 4. (Continued)

12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
−0.07	*−0.06*	*−0.06*	*−0.02*	*−0.03*	*−0.06*	*−0.08*		**−0.06**	**−0.07**	**−0.07**	**−0.02**	**−0.05**	**−0.08**	
0.27								**0.05**						
−0.30								**−0.28**						
−0.73	*−0.51*		*−0.25*					**−1.05**	**−0.42**		**−0.19**			
0.13	*−0.03*	*−0.47*						**0.32**	**−0.2**	**−0.71**				
−1.42	*−1.02*	*−0.70*	*−0.70*	*−0.61*				**−1.72**	**−1.44**	**−0.98**	**−1.06**	**−0.80**		
−33.3	*−27.3*	*−21.3*	*−33.7*	*−22.4*	*−29.6*									
	0.35								**0.21**					
		0.25								**0.02**				
			0.72								**0.65**			
				0.73								**0.67**		
					0.88								**0.88**	
														0.90
						0.87								
							0.88							

Etchegory et al., 1995; Grossman et al., 1994; Martin and Aber, 1994a; Curran, 1989; 1992), some of them being offical methods for biochemistry agricultural products analysis (Williams and Norris, 1987; Norris et al., 1976; Marten et al., 1989; Weyer, 1985). This apparent contradiction is mostly explained by the difference between our physically based approach and the statistical / empirical approach associated with NIRS techniques. While improvements in the estimation of leaf biochemistry may be achieved by using an empirical approach, there is no certainty about the predictive performances of that approach when applied to independent data sets (Grossman et al., 1994). Another way to improve the biochemical composition retrieval performances consists in grouping similar compounds together to strengthen their absorption features. We thus investigated various combinations to describe the biochemical composition by grouping progressively the biochemicals together according to their absorption similarities. Results show that the highest accuracy of the concentration estimates is achieved when all the compounds are grouped together. It amounts to estimate the specific leaf area, a very interesting variable that participates in many physiological processes occurring in canopy functioning models.

The results presented in this study were obtained over dry leaves. To apply these findings to fresh leaves, one obviously has to take into account the absorption of water inside the leaf. This was beyond the objectives of this article but will be the next issue to be investigated before going up to canopy level.

We thank very much the agencies that participated in the funding of this study: European Space Agency (Europe) under the "Plant Canopy Biochemistry" Contract 10661 / 93 / NL / SF, Centre National de la Recherche Spatiale (France), and NASA

through the LEAF MOD Project. We thank also the Joint Research Centre (Europe) for kindly providing the data set used in this study. We are grateful to G. Andreoli, B. Hoosgood, A. Pedrini, and S. E. Horning for their assistance during the JRC experiment. Special thanks also to Lee Johnson, Bary Ganopol, D. Peterson, Vern Vanderbilt, and D. Himmelsbach, who provided constructive comments on this study.

REFERENCES

Aber, J. D., and Federer, C. A. (1992), A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems, *Oecologia* 92:463–474.

AOAC (1970), *Official Method of Analysis*, 11th ed., Association of Official Analytical Chemists, Washington, DC.

Baret, F., Vanderbilt, V. C., Steven, M. D., and Jacquemoud, S. (1994), Use of spectral analogy to evaluate canopy reflectance sensitivity to leaf optical properties. *Remote Sens. Environ.* 47:1–25.

Curran, P. J. (1989), Remote sensing of foliar chemistry: review, *Remote Sens. Environ.* 30:271–278.

Curran, P. J., Dungan, J. L., Macler, B. A., Plummer, S. E., and Peterson, D. L. (1992), Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration, *Remote Sens. Environ.* 39:153–166.

Curcio, J. A., and Petty, C. C. (1951), The near infrared absorption spectrum of liquid water, *J. Opt. Soc. Am.* 41(5): 302–304.

Ewerts (1985), 3ème Directive avec rectification (J.O.C.E. du 27 novembre 1980), in *Recueil des Normes Françaises-AFNOR.*

Gastellu-Etchegory, J. P., Zagolski, F., Mougin, E., and Marty, G. (1995), Assessment of canopy chemistry using AVIRIS, a case study in the landes forest, south-west France, *Int. J. Remote Sens.* 16(3):487–501.

Gastellu-Etchegory, J. P., Zagolski, F., Pinel, V., et al. (1994), Assessment of forest canopy chemistry with ISM. *Int. J. Remote Sens.* (in press).

Grossman, Y. L., Ustin, S. L., Sanderson, E., Jacquemoud, S., Schumck, G., and Verdebout, J. (1994), Examination of regression and correlation approaches for extraction of leaf biochemistry information from leaf reflectance data, *Remote Sens. Environ.* (in press).

Himmelsbach, D. S., Boer, J. D., Akin, D. E., and Barton, F. E., II (1988), Solid-state carbon-13 NMR, FTIR and NIR spectroscopic studies of ruminant silage digestion, in *Analytical Applications of Spectroscopy* (C. S. Creaser and A. M. C. Davies, Eds.), Royal Society of Chemistry, London, pp. 410–413.

Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, G., and Schmuck, G. (1995), Leaf Optical Properties Experiment 93 (LOPEX 93), Report EUR-16095-EN, European Commission, Joint Research Centre, Institute For Remote Sensing Applications, Ispra, Italy.

Jacquemoud, S., and Baret, F. (1990), PROSPECT: a model of leaf optical properties, *Remote Sens. Environ.* 34:75–91.

Jacquemoud, S., Verdebout, J., Schmuck, J., Andreoli, G., and Hosgood, B. (1995), Investigation of leaf biochemistry by statistics, *Remote Sens. Environ.* 54:180–188.

Johnson, L. F., Hlavka, C. A., and Peterson, D. L. (1994), Multivariate analysis of AVIRIS data for the canopy biochemical estimation along the Oregon transect, *Remote Sens. Environ.* 47:216–230.

Kubelka, P., and Munk, F. (1931), Ein Beitrag zur Optik de Farbanstriche, Z. *Tech. Phys.* 12:593.

Kupiec, J., Smith, G. M., and Curran, P. J. (1993), Aviris spectra correlated with the chlorophyll concentration of a forest, in *Summaries of the Fourth Annual JPL Airborne Geoscience Workshop* (R. O. Green, Ed.), 25–29 October, Pasedana, California, pp. 105–108.

Lawson, C. L., and Hanson, R. J. (1974), *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, Chap. 23.

Marten, G. C., Shenk, J. S., and Barton, F. E. (1989), Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality, USDA Handbook 643, U.S. Department of Agriculture, Washington, DC.

Martin, M. E., and Aber, J. D. (1993), Measurements of canopy chemistry with 1992 aviris data at Blackhawk Island and Harvard Forest, in *Summaries of the Fourth Annual JPL Airborne Geoscience Workshop* (R. O. Green, Ed.), 25–29 October, pp. 113–116.

Martin, M. E., and Aber, J. D. (1994a), Analyses of forest foliage III: Determining nitrogen, lignin and cellulose in fresh leaves using near infrared reflectance data, *J. Near Infrared Spectrosc.* 2:25–32.

Martin, M. E., Aber, J. D. (1994b), High spectral resolution remote sensing of canopy chemistry, in *Proceedings of the 6th International Symposium, Physical Measurements and Signatures in Remote Sensing,* Val d'Isere, France, 17–21 January, pp. 201–207.

Nelder, J. A., and Mead, R. (1963). A simplex method for function optimization, *Comput. J.* 7:308–313.

Norris, K. H., Barnes, R. F., Moore, J. E., and Shenk, J. S. (1976), Predicting forage quality by infrared reflectance spectroscopy. *J. Animal Sci.* 43:839–878.

Pening de Vries, F. W. T., Brunsing, A. H. M., and Van Laar, H. H. (1976), Products, requirements and efficiency of biosynthesis: A quantitative approach. *J. Theor. Biol.* 45: 339–377.

Peterson, D. L., Aber, J. A., Matson, P. A., et al. (1988), Remote sensing of forest canopy and leaf biochemical contents, *Remote Sens. Environ.* 24:85–108.

Running, S. W., Aber, J. D., Peterson, D. L., Matson, P. A., and Vitousek, P. M. (1985), A simulated model integrating carbon, water and nitrogen cycles in forest, in *Proc. Symp. IUFRO Whole Plant Physiology,* Oak Ridge, TN.

Smith G. M., and Curran, P. J. (1992), Exploring the remote sensing of foliar biochemical concentrations with AVIRIS data, in *Summaries of the Third Annual JPL Airborne Geoscience Workshop,* (R. O. Green, Ed.), 1–5 June, pp. 50–52.

Van Soest, P. J., and Masson, V. C. (1967), The influence of the Maillard reaction upon the nutritive values of fibrous feed, *Feed Sci. Technol.* 32:45–53.

Van Soest, P. J., and Wine, R. (1991), Use of detergent in the analysis of fibrous feeds. IV. Determination of plant cell wall constituents. *J. Assoc. Off. Agric. Chem.* 50:50–55.

Wallach, D., and Goffinet, B. (1987), Mean squared error of prediction in models for studying ecological and agronomic systems, *Biometrics* 43:561–573.

Weende, P. (1985), NF V03-400, in *Recueil des Normes Françaises-AFNOR.*

Wessman, C. A., Aber, J. D., Peterson, D. L., and Mellilo, J. M., (1988), Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems, *Nature* 335:154–156.

Weyer, L. G. (1985), Near infrared spectroscopy of organic substances, *Appl. Spectrosc. Rev.* 21:1–43.

Williams, K., Percival, F., Merino, J., and Mooney, H. A. (1987), Estimation of tissue construction cost from heat of combustion and organic nitrogen content, *Plant Cell Environ.* 10:725–734.

Williams, P., and Norris, K. H., Eds. (1987), *Near-Infrared Technology in the Agricultural and Food Industries.* American Association of Cereal Chemists, St. Paul, MN.

Zagolski, F. (1994), Evaluation des concentrations chimiques d'un couvert forestier à l'aide de la haute résolution spectrale, Thesis, Université Paul Sabatier, Toulouse, France, 173 pp.