*International Journal of Remote Sensing*
Vol. 27, No. 5, 10 March 2006, 885–902

Taylor & Francis
Taylor & Francis Group

# Estimation of forest leaf area index from SPOT imagery using NDVI distribution over forest stands

H. DAVI*†‡, K. SOUDANI†, T. DECKX†, E. DUFRENE†, V. LE DANTEC‡ and C. FRANÇOIS†

†Ecologie, Systématique et Évolution (ESE), Université Paris Sud, Bât 362, 91405, Orsay, France

‡Centre Etudes Spatiales de la Biosphere (CESBIO) [CNES/CNRS/UPS] 18 Ave. E. Belin 31401, Toulouse, France

(*Received 10 December 2004; in final form 12 June 2005*)

Leaf area index (LAI) is a key parameter of atmosphere–vegetation exchanges, affecting the net ecosystem exchange and the productivity. At regional or continental scales, LAI can be estimated by remotely-sensed spectral vegetation indices (SVI). Nevertheless, relationships between LAI and SVI show saturation for LAI values greater than 3–5. This is one of the principal limitations of remote sensing of LAI in forest canopies. In this article, a new approach is developed to determine LAI from the spatial variability of radiometric data. To test this method, *in situ* measurements for LAI of 40 stands, with three dominant species (European beech, oak and Scots pine) were available over 5 years in the Fontainebleau forest near Paris. If all years and all species are pooled, a good linear relationship without saturation is founded between average stand LAI measurements and a model combining the logarithm of the standard deviation and the skewness of the normalized difference vegetation index (NDVI) ($R^2$=0.73 rmse=1.08). We demonstrate that this relation can be slightly improved by using different linear models for each year and each species ($R^2$=0.82 rmse=0.86), but the standard deviation is less sensitive to the species and the year effects than the mean NDVI and is therefore a performing index.

## 1. Introduction

The leaf area index (LAI), defined as one half of the total surface of leaves by unit ground surface area, corresponds to the main surface exchange of carbon and water between the forest canopy and the atmosphere. Biophysical characteristics of the forest stand such as LAI, gap size and distribution show a strong spatial variability from stand to regional scales. Some optical (LI-COR LAI-2000, hemispherical photographs, Sunscan, Demon) and direct or semi-direct methods (litter collection and allometric methods) exist for LAI estimation at stand level (Albrekston 1984, Dufrêne and Bréda 1995). The estimation of LAI variability at regional and global scales is essential but not feasible using *in situ* methods such as those described above. For these reasons, remote sensing of LAI is the subject of numerous works due to the growing demand of this parameter for ecological modelling (Justice *et al.* 1998, Lucas and Curran 1999, Running *et al.* 1999).

*Corresponding author. Email: hendrik.davi@avignon.inra.fr

*International Journal of Remote Sensing*
ISSN 0143-1161 print/ISSN 1366-5901 online 
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160500227896

The approaches used, generally based on Bidirectional Reflectance Distribution Function (BRDF) model inversion procedures or on empirical relationships, illustrate the difficulties with LAI remote sensing due to the complexity of the radiometric behaviour of forest canopies (Peterson *et al.* 1987, Asner 1998). Several factors such as 3-dimensional canopy structure heterogeneity, shrubs and herbaceous species of the understorey, plus exogenous factors such as atmospheric effects, solar illumination and view geometry conditions strongly affect the spectral properties of the canopy (Myneni and Asrar 1994, Gastellu-Etchegorry *et al.* 1999). These factors make interpretation difficult, particularly for multi-temporal remote sensing images.

LAI–radiometric data relationships are very scattered (Turner *et al.* 1999, Colombo *et al.* 2003) and show a well known plateau termed saturation at which the increase of LAI has no effect, neither on single nor combined spectral bands (Tian *et al.* 2000). The level of saturation is variable and species-dependent (Chen *et al.* 2002) but generally reached for mid-LAI values around 3–5 (Turner *et al.* 1999). An LAI value greater than 5 is frequent and one-third of the terrestrial land surface is occupied by such forests (Turner *et al.* 1999). Moreover, local measurements may lead to LAI values less than 1 and greater than 10 in the same plot. Thus, forest LAI variability could not be estimated by LAI versus radiometric data relationships with saturation plateau.

In this paper, to overcome the saturation problem, we try to estimate stand LAI using not only averaged radiometric data but also their variability over the stand. First, we thought that the LAI distribution could bring additional valuable information linked to the LAI itself. Second, and somewhat paradoxically, the saturation of the normalized difference vegetation index (NDVI)–LAI relationship provides in theory a way to estimate the LAI: the higher the LAI, the larger the number of saturating pixels, and the smaller the variability of NDVI. Third, an increase in stand LAI implies a decrease in canopy openness and a lower contribution of the soil reflectance to the whole canopy reflectance; this may cause a decrease in the spatial variability of reflectance. For these reasons we have studied the shape of the NDVI distribution (i.e. first-order texture) to see if an improvement in LAI estimation from remote sensing images was possible.

## 2. Materials and methods

### 2.1 *Site description*

Measurements were taken in the Fontainebleau forest, located south-east of Paris, France (48°25′ N, 2°40′ E). This large mixed deciduous forest extends over 17 000 ha at an average elevation of 120 m. Climate is temperate with an average annual temperature of 10.2°C and an average annual precipitation of 720 mm.

The dominant species are oaks (*Quercus petraea* (Matus) Liebl., *Quercus robur* (Matus) Liebl.), beech (*Fagus sylvatica* L.) and Scots pine (*Pinus sylvestris* L.). The main understorey species are hornbeam (*Carpinus betulus* L.) and beech. The herbaceous understorey species are mainly bramble (*Rubus fruticosus* L.), bracken (*Pteridium aquilinum* (L.) Kuhn) and purple moor-grass (*Molinia caerulea* (L.)). The Fontainebleau forest is dominated by oaks, which represent 50%, beech 10% and Scot pines 40%.

Most of the deciduous stands are located on flat ground (i.e. on windborne sands), while the coniferous stands are found on the hilly parts of the forest (i.e. on

the sandy Stampian or the sand-stone, with a shallow soil). Almost all of the humus types are found in the Fontainebleau forest from mull to mor.

All stands are managed by the French forest service, Office National des Forêts. Regular thinning and other forestry practices are carried out and appear to be the main factors of spatial and annual variations of LAI and other canopy structure characteristics (Le Dantec *et al.* 2000).

### 2.2 *LAI measurements*

Leaf area index measurements were made using a ground-based optical instrument, the Plant Canopy Analyser LAI-2000 (LI-COR Inc., Lincoln, NE, USA), named LAI-2000 hereafter. This instrument uses a fisheye optical sensor, comprising five detectors arranged in concentric rings, to measure diffuse radiation (below 490 nm) from different sections of the sky. The ratio of below-canopy and above-canopy radiation, measured for the five zenith angles, corresponds to the transmittance or the gap frequency. Estimation of LAI is based on the inversion of the standard Poisson model of gap frequency distribution (Nilson 1971).

LAI-2000 measurements were taken from the end of June to the middle of July during clear sky days to avoid rapid changes in sky conditions. All measurements were made during periods of very low solar elevation: less than 2 hours after sunrise or before sunset (Le Dantec *et al.* 2000). To prevent direct radiation on the sensor, measurements were made in the opposite direction of the sun, using a view restrictor of 180°. Above-canopy incident diffuse radiation measurements have been taken in open areas without edge effects close to the stands studied. For a better agreement with direct measurements, the three upper rings of LAI-2000 are used to estimate LAI (Welles and Norman 1991, Dufrêne and Bréda 1995, Olthof and King 2000).

For each stand and according to its size, 40–150 measurements (71 on average) were taken from several transects at distance intervals ranging from 5 to 10 m. The mean LAI over the stand is therefore calculated and hereafter named simply LAI. From these local measurements, it is also possible to have the within-stand distribution of LAI measured *in situ*.

The measurements were made within 40 stands (12 are dominated by beech, 17 by oak and 11 by Scots pine) and over 5 years (1994–1998). Selection criteria based on stand age, tree density and biomass are considered and this sample is representative of the main types of the Fontainebleau forest stands (Le Dantec *et al.* 2000). Note that LAI is not measured every year for all stands. For that reason, we have only 162 plots at our command (and not 200). The number of stands used in this study and the dominated species are summarized in table 1.

### 2.3 *Remote sensing stand data acquisition and processing*

Five Satellite pow l'Observation de la Terre (SPOT) High Resolution Visible (HRV) images, taken between June and August, from 1994 to 1998 were used. They had a 20 m pixel size and we used the red (XS1 610–680 nm), green (XS2 500–590 nm) and near-infrared (XS3 790–890 nm) bands. Images were rectified and geo-referenced using ground control points and integrated in a GIS database of the Fontainebleau forest. Digital counts (grey tone) were converted to at-satellite (top of atmosphere (TOA)) radiance ($Wm^{-2}\,sr^{-1}\,\mu m^{-1}$) using the gains and offsets contained in the image headers. The images were then calibrated to scaled surface reflectance after

Table 1. Statistical summary of leaf area index and NDVI per species and per year.

| | Scots pine dominated-stands | | | | | Oaks dominated-stands | | | | | Beech dominated-stands | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | 1994 | 1995 | 1996 | 1997 | 1998 | 1994 | 1995 | 1996 | 1997 | 1998 | 1994 | 1995 | 1996 | 1997 | 1998 |
| Number of stands | 3 | 9 | 10 | 5 | 10 | 14 | 18 | 18 | 8 | 16 | 8 | 11 | 11 | 10 | 11 |
| LAI | 2.30 | 2.31 | 2.89 | 2.67 | 2.61 | 4.66 | 4.62 | 4.96 | 5.51 | 5.27 | 5.52 | 5.00 | 5.39 | 4.82 | 5.34 |
| $std_{LAI}$ | 0.98 | 0.75 | 0.99 | 0.60 | 0.80 | 1.03 | 1.13 | 1.15 | 1.17 | 1.17 | 1.04 | 1.09 | 1.28 | 1.15 | 1.39 |
| NDVI | 0.792 | 0.762 | 0.779 | 0.701 | 0.673 | 0.881 | 0.862 | 0.873 | 0.828 | 0.812 | 0.908 | 0.875 | 0.898 | 0.812 | 0.838 |
| $std_{NDVI}$ | 0.021 | 0.023 | 0.023 | 0.030 | 0.038 | 0.018 | 0.016 | 0.022 | 0.016 | 0.021 | 0.011 | 0.011 | 0.014 | 0.016 | 0.014 |
| $skew_{NDVI}$ | −0.319 | −0.059 | −0.207 | 0.340 | 0.118 | −0.618 | −0.616 | −0.881 | −0.086 | −0.649 | −1.158 | −0.676 | −1.651 | −0.319 | −0.682 |
| $kurt_{NDVI}$ | 0.536 | 0.493 | 0.819 | 0.943 | −0.160 | 1.575 | 1.738 | 2.979 | 0.381 | 1.590 | 3.393 | 1.521 | 7.064 | 0.233 | 2.361 |
| $R_1$ | 0.99 | 0.74 | 0.69 | $0.17^{ns}$ | 0.88 | 0.87 | 0.84 | 0.96 | $-0.49^{ns}$ | 0.96 | 0.90 | 0.89 | 0.81 | $0.09^{ns}$ | 0.93 |
| $R_2$ | $-0.92^{ns}$ | −0.65 | $-0.52^{ns}$ | $0.11^{ns}$ | −0.72 | −0.92 | −0.94 | −0.90 | −0.81 | −0.94 | −0.84 | −0.96 | −0.60 | $-0.41^{ns}$ | −0.68 |

std, average within-stand standard deviation. $R_1$, the correlation coefficient between mean NDVI and LAI. $R_2$, the correlation coefficient between logarithm of standard deviation of NDVI and LAI. Statistical analysis: ns, not significant ($p>0.05$).

atmospheric corrections using a dark object subtraction (DOS) approach (Song *et al.* 2001).

Using ENVI software (Environment of Visualizing Images, Research System Inc., Boulder, CO), limits of each stand and for each year were located and vectorized taking special care not to include pixels situated on the borders. Geographic limits and measurements taken within each stand are well documented and the high resolution of SPOT images shows sufficient details to allow for their accurate positioning. Geolocation errors of the stands are thus inferior to one pixel (20 m). For each stand and in each band, all grey tone pixels were extracted, corrected for atmospheric effects and processed using SAS software (SAS Institute Inc., Cary, NC).

### 2.4 *Methodology and statistical analysis*

Among different radiometric indices, we have tested NDVI and simple ratio (SR), which are the most widely used indices. Several studies showed high correlations of these indices to canopy cover (Chen and Cihlar 1996), LAI (Fassnacht *et al.* 1997), above-ground biomass (Dong *et al.* 2003) and other environmental variables. Note that the idea of our study is to demonstrate that the distribution of spectral vegetation indices can be used not to search for the best indices. The SR index has given less accurate results for predicting LAI and is therefore not presented in this paper.

A multiple least square linear regression analysis with LAI as the dependent variable and $\text{mean}_{NDVI}$, $\text{std}_{NDVI}$ (or logarithm of $\text{std}_{NDVI}$), $\text{skew}_{NDVI}$, $\text{kurt}_{NDVI}$ as the explicative variables, were then computed. Only explicative variables, which contribute significantly to the linear regression model, were retained. The first objective is to find predictive relationships as general as possible (by incorporating all the data). Then we attempt to analyse if the year and species effects can explain residuals of the general model, and how much the accuracy of the model increases if these two effects are taken into account.

The experimental design of the statistical analysis is described in figure 1:

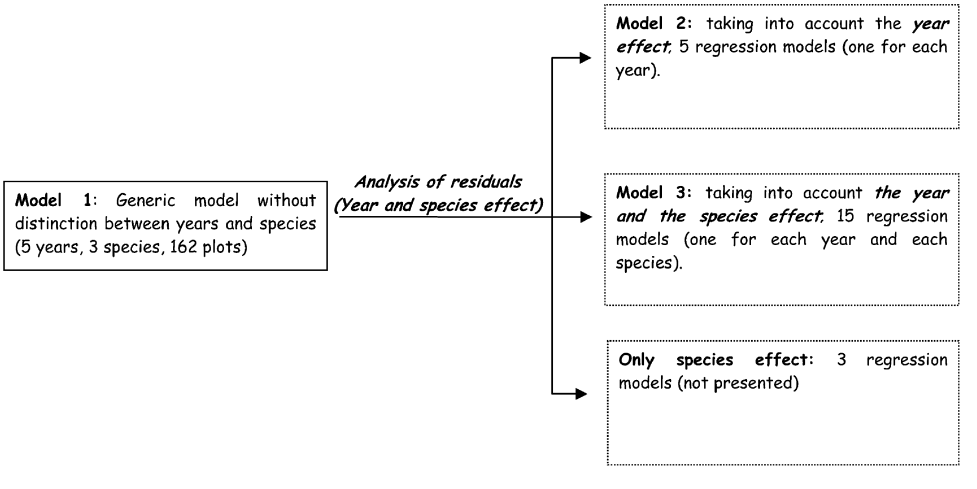


Figure 1. Experimental design of the different models tested to relate NDVI distribution and measured LAI according to the year or the species.

(i) A first linear regression model, including all species and all years ($n$=162), was computed. The residuals, which represent the unexplained variance, were investigated by means of a two-way analysis of variance to test the year and species effects.
(ii) Five linear regression models, one for each year without distinction between the three species, were computed to improve the estimation of *in situ* LAI and to isolate the year effect.
(iii) Fifteen linear regression models, one for each species and for each year, were computed to obtain the most accurate predictions of *in situ* LAI and to evaluate the year and species combined effects.

The species effect on the LAI measurements was tested by an analysis of variance over all the stands and for the 5 years. Since LAI measurements were not made on all the stands each year, the year effect is only tested over 15 stands (7 dominated by oak, 7 by beech, 1 by Scots pine), where LAI was measured each year. The species and the year effects on the NDVI are tested with the same methodology as for the *in situ* LAI measurements. Therefore, all the plots are used for the analysis of species effect and only 15 stands for the year effect.

To estimate the NDVI variability in each stand, we used the standard deviation ($std_{NDVI}$), the skewness ($skew_{NDVI}$), which quantifies the degree of asymmetry of a given distribution and the kurtosis ($kurt_{NDVI}$). If a symmetrical distribution is considered to have a centre, two shoulders, and two tails, kurtosis describes the proportions found in the centre and in the tails in relation to those in the shoulders (Sokal and Rohlf 1995). The symmetry of the LAI distribution is examined because the LAI distribution is not symmetrical and the NDVI–LAI relationship is not linear. Skewness and kurtosis were therefore expected to be pertinent parameters, which could bring additional information linked to the LAI itself. The LAI measurements were not spatially explicit (they were measured on several randomly distributed transects). Therefore it would have been very difficult to explain and interpret an empirical relationship between LAI and the second-order statistical measures of image texture (Haralick *et al.* 1973). For that reason, the second-order statistical measures of image texture are not examined.

The relationships between *in situ* LAI and the mean NDVI ($mean_{NDVI}$) and descriptive statistics of NDVI dispersion were first analysed on the basis of the Pearson correlation coefficients to highlight the potential correlations.

Concerning the analysis of variance, the level of significance was calculated by analysis of variance (ANOVA) followed by a Bonferroni test for multiple comparisons. The bonferroni test allows the significance of the difference between all pairs of means (pairwise comparison of means) to be tested, while ANOVA tests the significance of the global effect. For all the tests, a probability level at $p \leqslant 0.05$ was considered to be significant. The normality of model residuals is tested by the Shapiro–Wilk W test (Shapiro *et al.* 1968). All statistical analyses were performed using SAS software.

The predictive ability and the stability of the three models were also assessed by means of a cross-validation procedure. Cross-validation is a technique to estimate the forecast skill of a statistical forecasting model (Michaelsen 1987). Each member of a given dataset is excluded in turn from the prediction algorithm process and then predicted using the algorithm or fit or relationship derived without it. This is done for each member (i.e. the prediction algorithm is computed $n$ times, if $n$ is the size of the dataset). This avoids separating the dataset into a calibration and test dataset,

and allows testing the predictive ability of the algorithm on each member of the dataset. The procedure was repeated $n$ times and the Pearson correlation coefficients and the root mean square error (rmse) were calculated between the predicted and the observed values to assess the accuracy of the model. The stability was evaluated by the coefficient of variation (CV) of the different coefficients associated to the regression variables.

## 3. Results

### 3.1 In situ *LAI and remote sensing characteristics of stands*

**3.1.1** ***In situ*** **LAI of stands.** The statistical summary of NDVI and LAI is given in table 1.

Concerning all the stands, the mean *in situ* LAI over the five years is 4.49 and the standard deviation is 2.07 ($4.95 \pm 2.21$ for oak, $5.25 \pm 1.57$ for beech and $2.60 \pm 1.00$ for Scots pine).

The temporal variation of LAI over the five years is estimated by the mean value of only 15 stands, where LAI was measured each year (figure 2). The differences between years are not significant (ANOVA $p=0.59$). The slight decrease of the mean LAI in 1997 may be due to the effect of drought in 1996 (Le Dantec *et al.* 2000).

**3.1.2 Radiometric characteristics of stands.** For the NDVI and for all the stands, the descriptive statistics of the radiometric data are presented in table 1. Differences are significant between species ($p<0.05$) for the mean, the standard deviation and the skewness, but not significant for kurtosis.

The year effect on the radiometric parameters on the 15 stands, where LAI was measured all the years, has been tested (figure 3). The same sample used for the *in situ* LAI analysis was considered, to allow a comparison with the LAI results. We note that there is a significant decrease in the mean NDVI in 1997 and 1998 compared to 1994, 1995 and 1996 and a significant increase in the standard deviation in 1997 and 1998 compared to 1994 and 1995. On the other hand, there is no effect of the year on kurtosis and for the skewness of the NDVI only 1996 differs significantly from 1997.

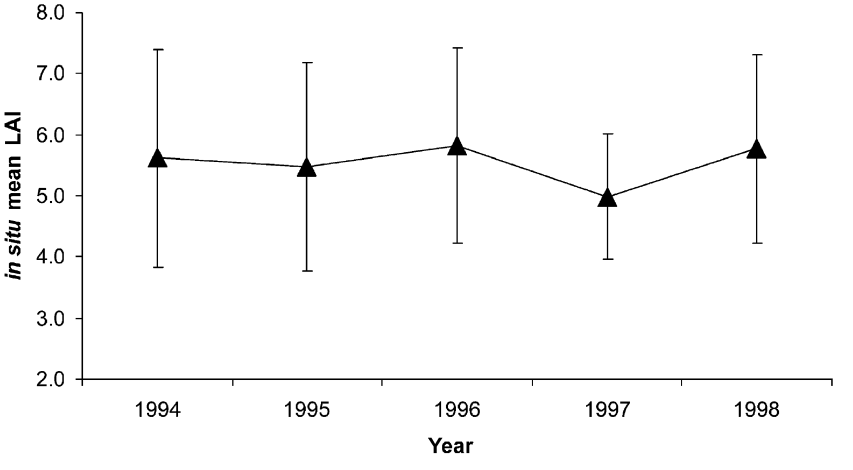


Figure 2. Evolution of the mean *in situ* leaf area index (LAI) for 15 stands (error bars show LAI standard deviation).

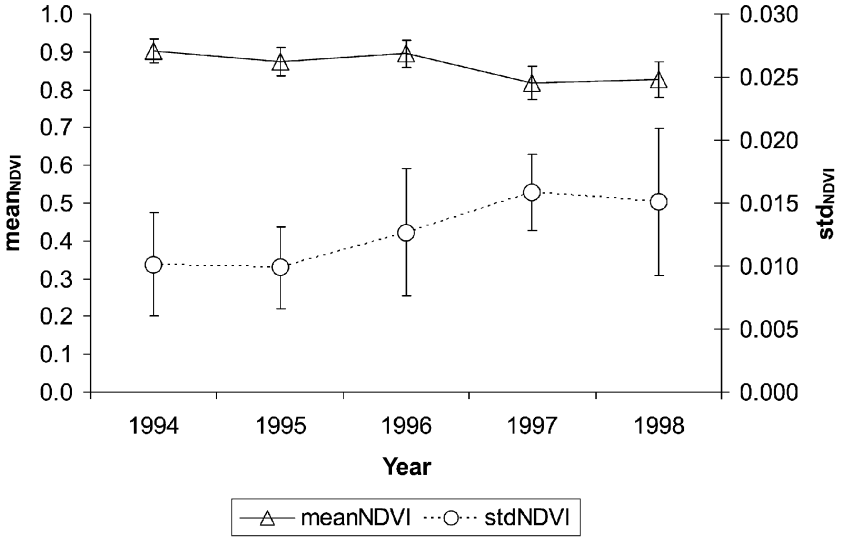


Figure 3. Evolution of mean NDVI ($mean_{NDVI}$) and standard deviation of NDVI ($std_{NDVI}$) for 15 stands (error bars show standard deviation).

### 3.2 *Relationships between* in situ *LAI and radiometric data*

The relationship between *in situ* LAI and the radiometric data is tested over the 5 years and for all the 162 sampled stands on which LAI was measured. Significant correlations between the LAI and single bands, near-infrared (NIR) and visible (VIS), were found. These correlations are consistent with the expectations based on previous works (Spanner *et al.* 1990): LAI is positively correlated to the NIR band ($r=0.71$, $p<0.0001$) and negatively correlated to the red ($r=-0.52$, $p<0.0001$) and green bands ($r=-0.5$, $p<0.0001$). Significant inverse curvilinear relationships are also found between LAI and the standard deviation of radiometric data in the VIS band but not in the NIR band. Correlation coefficients are $-0.75$ ($p<0.0001$) in the green band and $-0.65$ ($p<0.0001$) in the red band. Saturation occurs at LAI$\sim$2.5 for the relationship between LAI and red or green SPOT bands, while no saturation is observed for the NIR band. Nevertheless, all these relationships are very scattered.

Figure 4(*a*) shows a significant relationship ($p<0.0001$) between LAI and NDVI in agreement with other studies. Over all years and all species, the correlation coefficient between *in situ* LAI and $mean_{NDVI}$ is 0.70. This relation is very scattered, particularly for low LAI values, year and species-dependent, and shows saturation when *in situ* LAI is greater than $\sim$4.

Between *in situ* LAI and the standard deviation of the NDVI ($std_{NDVI}$), the correlation coefficient is $-0.77$ ($p<0.0001$), and there is a strong logarithmic relationship ($r=-0.83$, $p<0.0001$) between these two variables (figure 4(*b*), (*c*)).

The correlation coefficients between $mean_{NDVI}$ and LAI or between the logarithm of $std_{NDVI}$ and LAI are also calculated by year and by species (table 1). In both cases, the correlation coefficients are higher for the oak stands than for the beech stands and higher for the beech stands than for the pine stands. By distinguishing between species and year, we note that $mean_{NDVI}$ is as correlated with LAI as the logarithm $std_{NDVI}$, while it is less correlated when all the years and species are pooled. This indicates a smaller effect of species and year on correlation based on $std_{NDVI}$ than on those based on $mean_{NDVI}$.

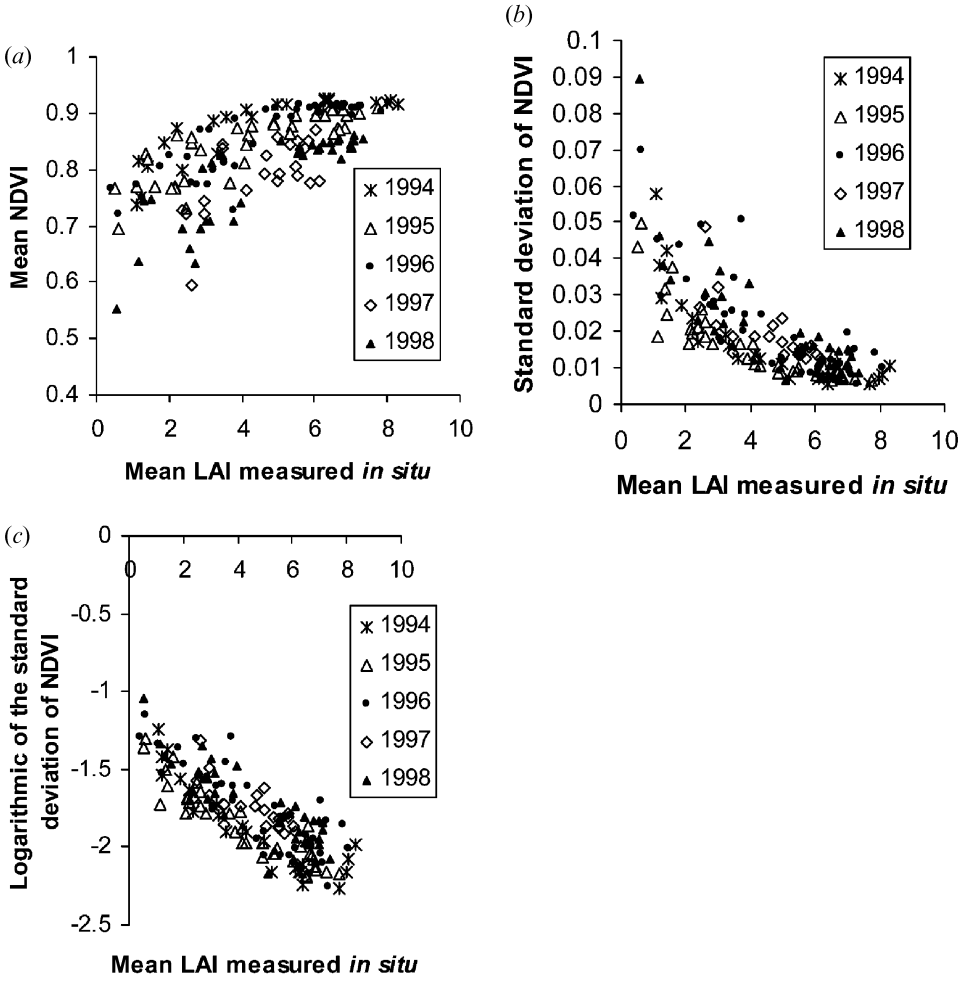


Figure 4. Relationship between mean LAI measured *in situ* and (*a*) mean NDVI, (*b*) standard deviation of NDVI, (*c*) logarithm of the standard deviation of NDVI. Results are presented by year from 1994 to 1998.

### 3.3 *Empirical model*

**3.3.1 Model 1: Generic model (all species, all years).** Only the intercept of the relationship, the logarithm of $std_{NDVI}$ and the $skew_{NDVI}$ contribute significantly to the regression model (table 2). The other explicative variables ($mean_{NDVI}$ or $kurt_{NDVI}$) are therefore not used to compute the model. The cross-validation correlation coefficient between measured and predicted values is 0.85 ($R^2$=0.73), and the root mean square error (rmse) is 1.08 (figure 5(*a*)).

Table 2. Parameters of regression model 1 (generic model).

| Variable | Intercept | $\log(std_{NDVI})$ | $skew_{NDVI}$ |
|---|---|---|---|
| Parameter estimate | −6.825 | −2.685 | −0.484 |
| Standard error | 0.597 | 0.146 | 0.095 |
| $p>\|t\|$ | <0.0001 | <0.0001 | <0.0001 |

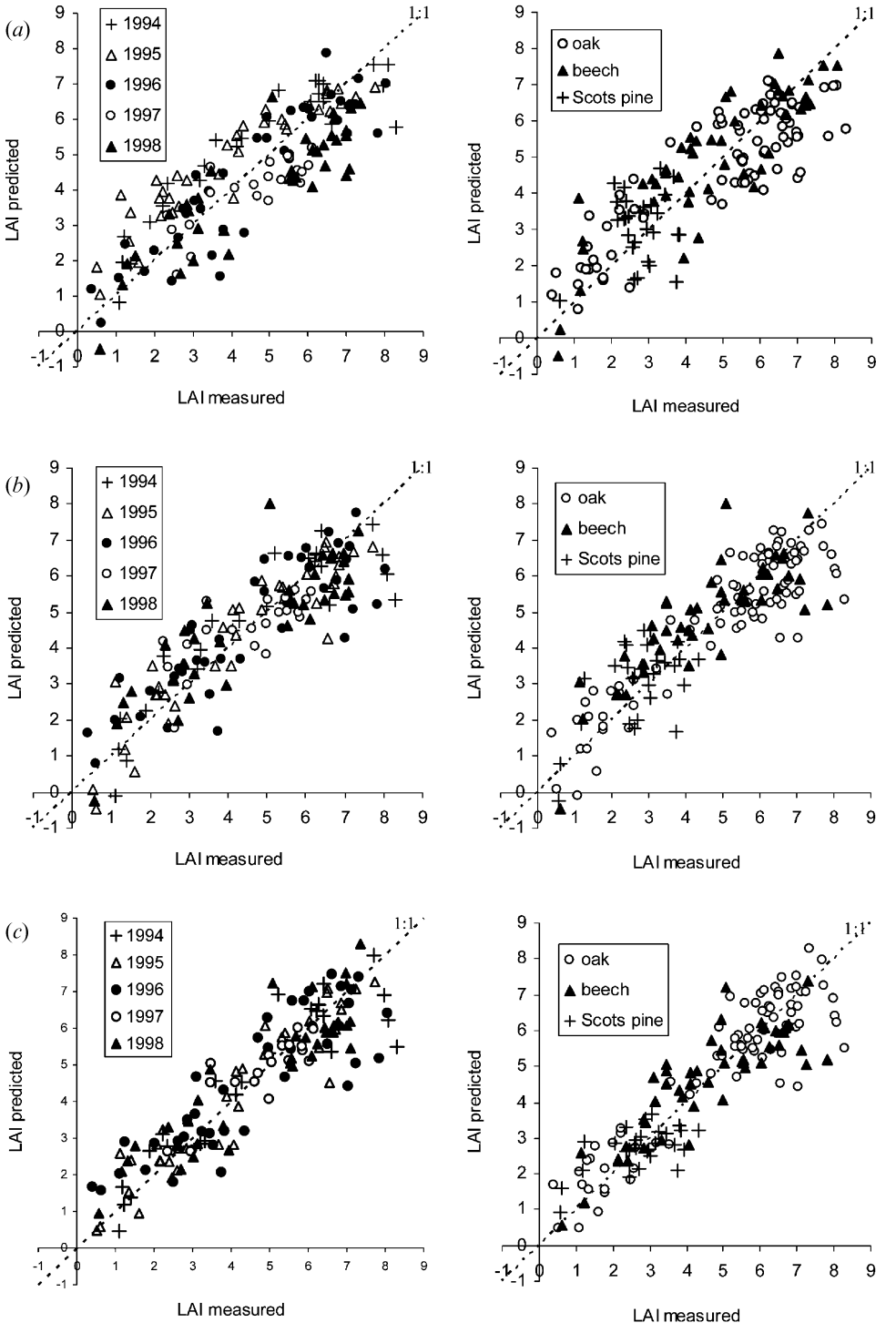


Figure 5. Relationships between *in situ* LAI and LAI predicted by linear regression: for generic model (*a*), for model year dependent (*b*), for model year- and species-dependent (*c*). Results are presented by year (left) or by species (right).

By distinguishing results between different species and years, we note that there is a year effect. Indeed in 1997 and 1998, predicted LAI is slightly lower. By a two-way ANOVA, we conclude that the effects of year ($p<0.0001$) and species ($p<0.0072$) on residuals are significant, nevertheless there is no significant interaction effect ($p<0.599$). As there is a significant year effect on NDVI and not on LAI, a year effect on the generic model was predictable.

This empirical model can therefore be improved by distinguishing between different years or between different species.

To validate this model, a cross-validation was carried out. The correlation coefficient between predicted and measured values is 0.85 and the rmse is 1.13. The stability of the model is estimated by the coefficient of variation of the regression parameters (see table 3). For the three variables the coefficient of variation is very low (<2%) and the regression model is quite stable.

**3.3.2 Model 2: Model year dependence.** In this case, there are five models, one per year from 1994 to 1998. Only the intercept of the relationship and the logarithm of $std_{NDVI}$ significantly contribute to the regression model. The correlation coefficient between measured and predicted values is 0.88 and the root mean square error (rmse) is 0.98 (figure 5(*b*)).

By a two-way ANOVA, we conclude that the species-effect on residuals remains significant ($p<0.0016$). The species-effect is not only a cover type effect (between resinous and deciduous species): the residuals also differ significantly between beech and oak (Bonferoni test $p<0.05$).

The correlation coefficient of the cross-validation for the model year dependent is 0.86 and the rmse is 1.09. The coefficient of variation of parameters for each model is lower than 3% except in 1997 (table 3). The fitted parameters significantly differ according to the year ($p<0.05$), which confirms the year effect.

**3.3.3 Model 3: Model year- and species-dependence.** In this case, there are 15 models, one per year and per species. Due to the small size of samples, we kept only the logarithm of $std_{NDVI}$ as an explicative variable. Each model is used independently to predict LAI. The correlation coefficient between measured and predicted values is 0.91, and the rmse is 0.86 (figure 5(*c*)).

Scots pine stands are removed from cross-validation analyses for the years 1994 and 1997, because there is not enough data (3 and 5 stands, respectively, see table 1). The correlation coefficient of the cross-validation is 0.74 and rmse is 2.10. The coefficient of variation of fitted parameters for each model might be very high: e.g.

Table 3. Coefficient of variation (%) of coefficients associated to the explicative variables in models 1 and 2 in the cross-validation.

| Model 1 | Intercept | $\log(std_{NDVI})$ | $skew_{NDVI}$ |
|---|---|---|---|
| | 0.58 | 1.84 | 0.36 |
| Model 2 | Intercept | $\log(std_{NDVI})$ | |
| 1994 | 2.96 | 1.97 | |
| 1995 | 1.67 | 1.18 | |
| 1996 | 1.90 | 1.18 | |
| 1997 | 28.70 | 10.24 | |
| 1998 | 2.94 | 1.92 | |

higher than 400% for beech in 1996 and oaks in 1997. Even if the two models associated to these higher values are removed, the mean variation coefficient remains high: 8% for parameters associated with the logarithm of $std_{NDVI}$ and 16% for parameters associated with intercept. The regression models are thus quite unstable because of the small size of samples used (see table 1). These regression models are therefore unlikely to be useful for estimating LAI.

## 4. Discussion

### 4.1 *Causes of the relationship between average stand LAI and within-stand variability of radiometric data*

Our results show that the main radiometric variable correlated to *in situ* LAI is the logarithm of $std_{NDVI}$. This negative correlation indicates that at the stand level, the stand average LAI is inversely related to the variability of NDVI. This negative correlation can be attributed to several factors.

First, due to the saturation of the relationship between NDVI and LAI, similar spatial variation of LAI leads to much higher spatial variability of NDVI for low average stand LAI values than for higher values. In other words, the higher the LAI, the larger the number of saturating pixels and the smaller the variability of NDVI. To illustrate this, we compare histograms of the distribution of *in situ* LAI and NDVI for two oak stands ($C_3$ and $C_{21}$) having two different average LAI values but similar standard deviations of LAI (figure 6(*a*)). The stand $C_3$ is dense with an average LAI of 6.3. The LAI values within this stand range between 3.5 and 8.5, which reflect the absence of large gaps. The stand $C_{21}$ is more open with gaps sufficiently large to have LAI values close to zero and an average stand LAI is 1.61. The statistical distribution of LAI of this stand is clearly asymmetrical and positively skewed with the mode smaller than the mean.

Consequently, for a mean LAI of 6.3, as in $C_3$, NDVI varies little (figure 6(*d*)). Oppositely, for a mean LAI of 1.61, the NDVI does not saturate and as a consequence NDVI distribution is wider (figure 6(*c*)). This phenomena can partly explain the relationship found between LAI and $\log(std_{NDVI})$. As the relationship between LAI and NDVI saturates, the relationship between LAI and $std_{NDVI}$ also saturates, but for higher LAI than in the case of the LAI–NDVI relationship. Indeed, even for high stand mean LAI (6 or 7), there is an area inside the stand, where local LAI is lower and not in the saturation range.

Analysis of statistical distribution of within-stand local LAI show that these distributions are variable as shown for the stands $C_3$ and $C_{21}$ (figure 6(*e*)). We found that these distributions can be fitted to the Weibull probability density function with different parameters depending on average stand LAI (data not shown). Knowing a stand LAI, we can thus reproduce theoretically its distribution. Then using an observed relationship between LAI and NDVI, we can generate several theoretical NDVI distributions and theoretically find if there is a relationship between LAI and $std_{NDVI}$. In figure 6(*c*), we show the simulations of several distributions of LAI and NDVI for different mean LAI. As intuitively expected we find that the distribution of NDVI becomes more compact (i.e $std_{NDVI}$ is smaller) the higher the LAI.

Note that other factors could have an effect on NDVI distribution. Among them the roughness of the top of the canopy determines the spatial distribution of shaded and sunlit vegetation and the soil can also contribute to the whole canopy reflectance variability. The real contribution of these factors is difficult to evaluate

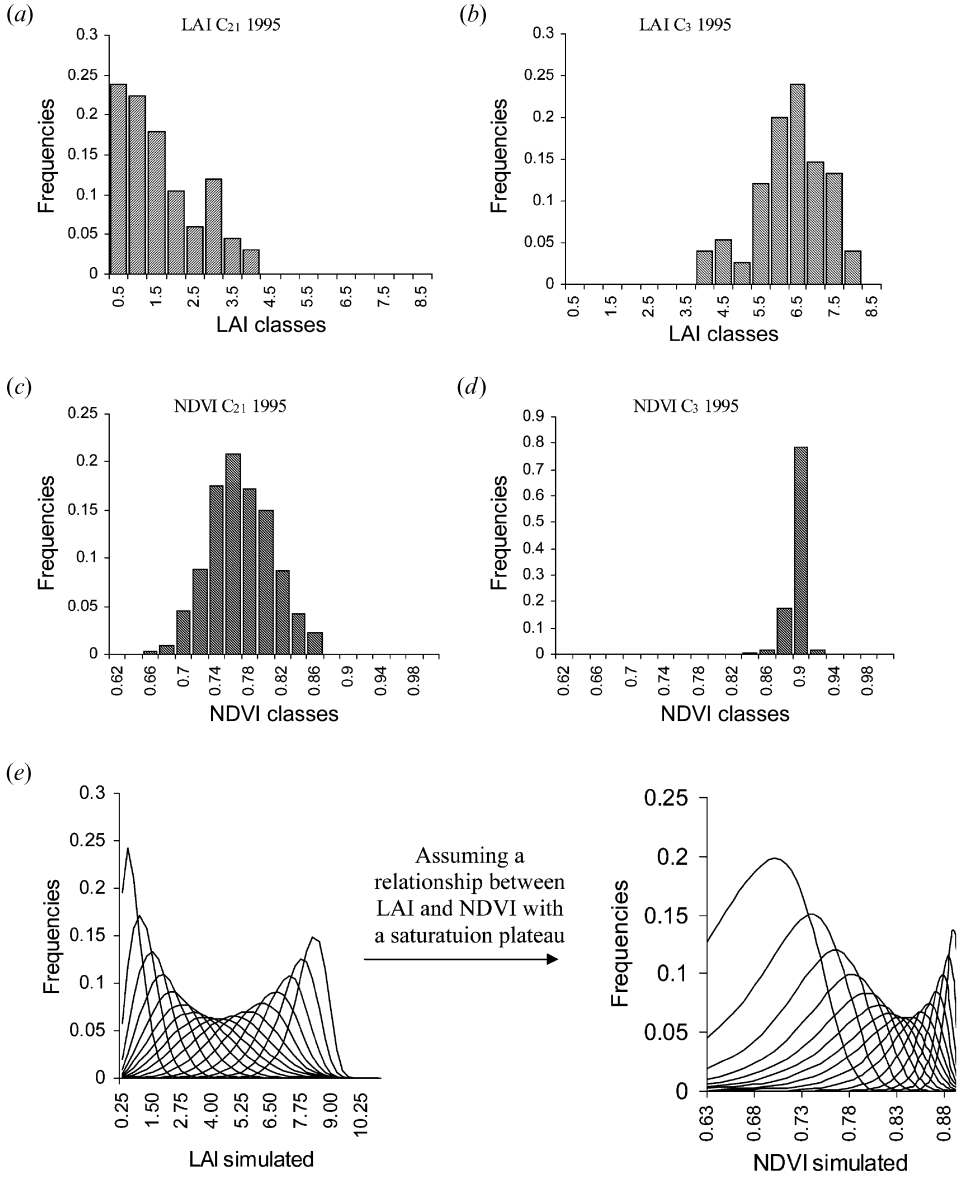


Figure 6. Histograms of LAI values on two contrasted stands (*a*) $C_{21}$ LAI=1.61+/−1.00 and (*b*) $C_3$ LAI=6.29+/−0.94 in 1995. (*b*) Historgrams of NDVI values on (*c*) $C_{21}$ and (*d*) $C_3$. (*e*) Sixteen modelled distributions of LAI computed for 44 LAI classes. LAI distributions are estimated from the Weibull function of $\text{mean}_{\text{LAI}}$. (*f*) Sixteen modelled distributions of NDVI distributions are calculated by assuming a simple logarithmic relationship between LAI and NDVI.

because of the complex relationship between LAI, the roughness of the top of the canopy, the soil reflectance and the radiation/vegetation interactions.

Even when year and species effects are taken into account, the variation of the relationship between LAI and NDVI spatial variability remains strong (figure 5(*c*)). The residual variations are difficult to reduce. Indeed, for the same LAI many configurations of the canopy structure are possible. For example, the roughness of

the top of the canopy or the stratification, can strongly affect the reflectance in the visible and infrared bands and the NDVI (Leblanc *et al.* 1997). These architectural differences in canopy structure are among the main factors explaining the variation in LAI–NDVI relationships (Chen *et al.* 1999). Furthermore, there may be a residual effect of the understorey on the reflectance, which may vary across the years and with the dominant species.

### 4.2 *Dependency of the regression results on stand selections*

Obviously, an estimation of LAI from the NDVI variability requires a judicious selection of the area over which the reflectance variability is estimated. Here stands were set out by photointerpretation and stands are 'relatively' homogeneous in species composition, age and forestry practices. This method requires therefore a preliminary work in order to classify the areas.

Moreover, the notion of homogeneity is scale dependent. Accurate definition of this notion is given in Raffy *et al.* (2003). The radiometric homogeneity depends on the spatial resolution of the image and on the different scales of variation in canopy structure. The spatial variability of radiometric data decreases when the image spatial resolution becomes coarser. For these reasons, the obtained relationships certainly depend on satellite spatial resolution and stand size. There is probably an optimum resolution for which the correlation may be improved (Marceau *et al.* 1994). Despite the fact that significant correlations were found, more investigations are needed to enhance our understanding of the dependency between the image spatial resolution and the forest canopy structure for the purposes of LAI estimation.

The model stability also requires enough samples, as shown by the variability of regression parameters during the cross-validation in our species and year dependent models. In future works, the model stability should always be tested, if one wants to use this method to estimate LAI at a regional scale. In our case, a good compromise would be to pool the stands dominated by deciduous species and to increase the number of stands dominated by Scots pine.

It is obviously necessary to test and validate this methodology on other sites. In each case, before calculation of NDVI variability and application of this method for LAI determination, preliminary work is necessary to define the homogenous stands. This task is relatively easy in forest regions, where inventory tables and thematic maps are available and easy to use via GIS systems. On the contrary, photo-interpretation, image classification and segmentation techniques may be used to partition the images taken by airborne and satellite sensors into homogenous regions (Welch *et al.* 2002, Burnett and Blaschke 2003).

### 4.3 *Improving the generality of regression models over the years*

We found that there is a year effect on the relationships between LAI and the NDVI variability. This effect limits the generality of the model. It is probably due to an effect of image acquisition date. In 1997 and 1998, satellite measurements were taken late in August while in 1994, 1995 and 1996 they were taken in July. The $std_{NDVI}$ and the $mean_{NDVI}$ measured in 1997 and 1998 significantly differ from those measured in 1994, 1995 and 1996 ($p=0.0019$ and $p=0.0001$, respectively). Moreover, over the 15 test stands, a good correlation was found between the mean of $std_{NDVI}$ and the date of reflectance measurements (figure 7).

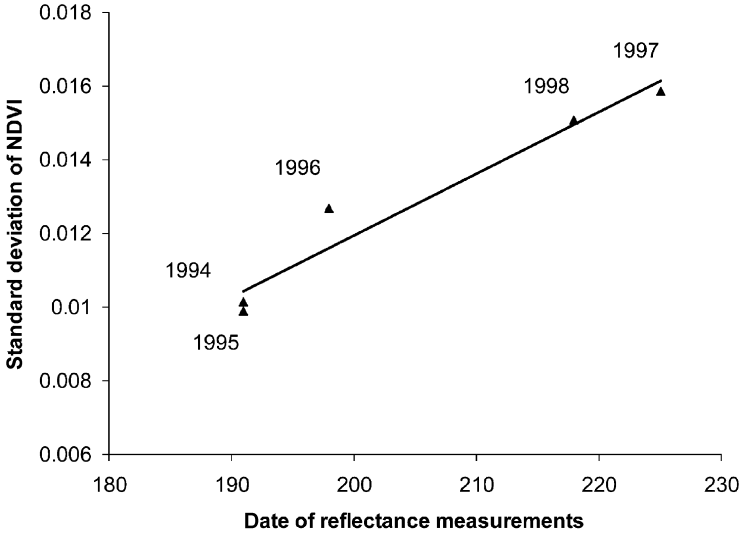


Figure 7. Mean NDVI standard deviation (over the same 15 stands) as a function of date of reflectance measurements.

Different authors have also measured temporal variation of mean reflectances during the leafy season: between July and August for Japanese beech, Kodani *et al.* (2002) found a decrease of near-infrared reflectance, while the NDVI remained stable; for European beech, Blackburn and Milton (1995) found a slow increase in red reflectance and a decrease in near-infrared reflectance. Thus, these date effects probably depend on a complex relationship between sun elevation, atmospheric conditions, biochemical composition (increase in chlorophyll content in Kodani *et al.* (2002)), canopy structure, soil reflectance and understorey variations.

Finally, several improvements could potentially enhance the accuracy of our results:

1. We can improve the standardization of radiometric data acquisitions (accurate corrections for the effects of atmospheric and satellite-sun geometry conditions). Nevertheless, we have carried out the same work without atmospheric correction and the results are still quite good. Using the standard deviation, which is based on the differences between local values and the mean, may indeed reduce the atmospheric effects. So the radiometric correction essentially highlights the year effect more rigorously.
2. We can also take into account the ground, herb and shrub layers, which affect the integrated canopy reflectance particularly in sparse forest canopies. Spectral Mixture Analysis (SMA) offers the possibility to achieve these aims. Based on the assumptions that the spectrum measured by a sensor is a linear or non-linear mixture of spectrally distinct endmembers (Roberts *et al.* 1997), the SMA/modelling may be relevant to 'unmix' these pixels. The spatial and temporal dynamics of understorey and soil reflectances can be measured in order to apply a spectral mixture analysis.
3. To improve our understanding of the year effect, it is also necessary to study the seasonal and spatial patterns of foliar biochemistry and canopy structure in the Fontainebleau forest, with simultaneous above-canopy reflectance measurements.

Simulations with BRDF models might also be useful for understanding interactions between these different factors and the relationship between radiometric variability and LAI.

## 5. Summary and conclusions

In this study, we investigated the potential of using the within stand spatial variability of NDVI to improve the estimation of mean stand LAI. A large database was available: 162 *in situ* mean LAI estimations (i.e. 13265 LAI-2000 measurements) on 40 homogeneous stands across a large region, over 5 years, with three dominant species (oak, beech, Scots pine). There is a good relationship between measurements and a linear model combining standard deviation and the skewness of NDVI ($R^2$=0.73) valuable for 5 years of data. We showed that this relationship is year and species dependent. The accuracy of model estimations can therefore be improved by using different linear models for each year. But the year effect on LAI-std$_{NDVI}$ relationship is very slight in comparison with the year effect on LAI-mean$_{NDVI}$. Moreover, we showed that this year effect is probably due to an acquisition date effect, so this 'year' effect is in fact a seasonal effect and the relationship is probably stable from year to year. Finally, it is noteworthy that no saturation occurs in any of the different statistical models tested, allowing forest LAI estimations with the same accuracy over the entire range of observations. Even if the year and the species effects could be further studied, we have shown by cross-validation that the predictive ability of the generic model is accurate enough (cross-validation rmse=1.13). This methodology may be applied as it is on other forest stands, where homogeneous stands are mapped or previously delimited by remote sensing. The causes of the relationship between LAI and NDVI variability will have to be theoretically further understood to improve the generality of that methodology.

**Acknowledgments**

We would like to thank the Office National des Forêts (ONF), Paris, France, for giving us facilities for *in situ* LAI measurements. This work was supported by the National Centre of Scientific Research (CNRS), Paris, France. We also thank all the people who helped us to collect the ground data. We acknowledge Guerric le Maire and Nick Barrett for their comments on an earlier version of the manuscript.

## References

ALBREKSTON, A., 1984, Sapwood basal area and needle mass of Scots Pine (Pinus sylvestris L.) trees in Central Sweden. *Forestry*, **57**, pp. 35–43.

ASNER, G.P., 1998, Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, **64**, pp. 234–253.

BLACKBURN, G.A. and MILTON, E.J., 1995, Seasonal variations in the spectral reflectance of deciduous tree canopies. *International Journal of Remote Sensing*, **16**, pp. 709–720.

BURNETT, C. and BLASCHKE, T., 2003, A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling*, **168**, pp. 233–249.

CHEN, J.M. and CIHLAR, J.J., 1996, Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, **55**, pp. 153–162.

CHEN, J.M., LEBLANC, S.G. and WHITE, H.P., 1999, Compact Airborne Spectrographic Imager (CASI) used for mapping biophysical parameters of boreal forests. *Journal of Geophysical Research. Atmospheres*, **104**, pp. 27945–27958.

CHEN, J.M., PAVLIC, G., BROWN, L., CIHLAR, J., LEBLANC, S.G., WHITE, H.P., HALL, R.J., PEDDLE, D.R., KING, D.J., TROFYMOW, J.A., SWIFT, E., VAN DER SANDEN J. and PELLIKKA, P.K.E., 2002, Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sensing of Environment*, **80**, pp. 165–184.

COLOMBO, R., BELLINGERI, D., FASOLINI, D. and MARINO, C.M., 2003, Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sensing of Environment*, **86**, pp. 120–131.

DONG, J., KAUFMANN, R.K., MYNENI, R.B., TUCKER, C.T., KAUPPI, P.E., LISKI, J., BUERMANN, W., ALEXEYEV, V. and HUGHES, M.K., 2003, Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sensing of Environment*, **84**, pp. 393–410.

DUFRÊNE, E. and BRÉDA, N., 1995, Estimation of deciduous forest leaf area index using direct and indirect methods. *Oecologia*, **104**, pp. 156–162.

FASSNACHT, K.S., GOWER, S.T., MACKENZIE, M.D., NORDHEIM, E.V. and LILLESAND, T.M., 1997, Estimating the leaf area index of north central Wisconsin forest using the Landsat Thematic Mapper. *Remote Sensing of Environment*, **61**, pp. 229–245.

GASTELLU-ETCHEGORRY, J.P., GUILLEVIC, P., ZAGOLSKI, F., DEMAREZ, V., TRICHON, V., DEERING, D. and LEROY, M., 1999, Modeling BRF and radiation regime of boreal and tropical forests: I. BRF. *Remote Sensing of Environment*, **68**, pp. 281–316.

HARALICK, R.M., SHANMUGAN, K. and DINSTEIN, I., 1973, Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, **3**, pp. 610–621.

JUSTICE, C., VERMOTE, E., TOWNSHEND, J.R.G., DEFRIES, R., ROY, D.P., HALL, D.K., SALOMONSON, V.V., PRIVETTE, J., RIGGS, G., STRAHLER, A., LUCHT, W., MYNENI, R., KNJAZIHHIN, Y., RUNNING, S., NEMANI, R., WAN, Z., HUETE, A., VAN LEEUWEN, W., WOLFE, R., GIGLIO, L., MULLER, J-P., LEWIS, P. and BARNSLEY, M., 1998, The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, pp. 1228–1249.

KODANI, E., AWAYA, Y., TANAKA, K. and MATSUMARA, N., 2002, Seasonal patterns of canopy structure, biochemistry and spectral reflectance in a broad-leaved deciduous Fagus crenata canopy. *Forest Ecology and Management*, **167**, pp. 233–249.

LEBLANC, S.G., CHEN, J.M. and CIHLAR, J., 1997, NDVI directionality in boreal forests. A model interpretation of measurements. *Canadian Journal of Remote Sensing*, **23**, pp. 368–379.

LE DANTEC, V., DUFRÊNE, E. and SAUGIER, B., 2000, Interannual and spatial variation in maximum leaf area index of temperate deciduous stands. *Forest Ecology and Management*, **134**, pp. 78–81.

LUCAS, N.S. and CURRAN, P.J., 1999, Forest ecosystem simulation modelling: the role of remote sensing. *Progress in Physical Geography*, **23**, pp. 391–423.

MARCEAU, D.J., HOWARTH, P.J. and GRATTON, D.J., 1994, Remote sensing and the measurement of geographical entities in a forested environment. Part 1: The scale and spatial aggregation problem. *Remote Sensing of Environment*, **49**, pp. 93–104.

MICHAELSEN, J., 1987, Cross-validation in statistical climate forecast models. *Journal of Climate and Applied Meteorology*, **26**, pp. 1589–1600.

MYNENI, R.B. and ASRAR, G., 1994, Atmospheric effects and spectral vegetation indices. *Remote Sensing of Environment*, **47**, pp. 390–402.

NILSON, T., 1971, A theoretical analysis of the frequency of gaps in plant stands. *Agricultural and Forest Meteorology*, **8**, pp. 25–38.

OLTHOF, I. and KING, D.J., 2000, Development of a forest health index using multispectral airborne digital camera imagery. *Canadian Journal of Remote Sensing*, **26**, pp. 166–176.

PETERSON, D., SPANNER, M., RUNNING, S. and TEUBER, K., 1987, Relationship of Thematic Mapper data to leaf area index of temperate coniferous forests. *Remote Sensing of Environnement*, **22**, pp. 323–341.

RAFFY, M., SOUDANI, K. and TRAUTMANN, J., 2003, On the variability of the LAI of homogeneous covers with respect to the surface size. *International Journal of Remote Sensing*, **24**, pp. 2017–2035.

ROBERTS, D.A., GREEN, R.O. and ADAMS, J.B., 1997, Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS. *Remote Sensing of Environment*, **62**, pp. 223–240.

RUNNING, S.W., BALDOCCHI, D.D., TURNER, D.P., GOWER, S.T., BAKWIN, P.S. and HIBBARD, K.A., 1999, A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. *Remote Sensing of Environment*, **70**, pp. 108–127.

SHAPIRO, S.S., WILK, M.B. and CHEN, H.G., 1968, A comparative study of various tests for normality. *Journal of the American Statistical Association*, **63**, pp. 1343–1372.

SOKAL, R.R. and ROHLF, F.J., 1995, *Biometry: The principles and practice of statistics in biological research*, 3rd edn (New York: W.H. Freeman and Co.).

SONG, C., WOODCOCK, C.E., SETO, K.C., LENNEY, M.P. and MACOMBER, S.A., 2001, Classification and change detection using Landsat TM data. When and how to correct atmospheric effects? *Remote Sensing of Environment*, **75**, pp. 230–244.

SPANNER, M., PIERCE, L.L., PETERSON, D.L. and RUNNING, S.W., 1990, Remote sensing of temperate coniferous forest leaf area index: the influence of canopy closure, understory vegetation and background reflectance. *International Journal of Remote Sensing*, **11**, pp. 95–111.

TIAN, Y., ZHANG, Y., KNYAZIKHIN, Y., MYNENI, R.B., GLASSY, J.M., DEDIEU, G. and RUNNING, S.W., 2000, Prototyping of MODIS LAI and FPAR algorithm with LASUR and Landsat data. *IEEE Transactions on Geoscience and Remote Sensing*, **38**, pp. 2387–2401.

TURNER, P.D., COHEN, W.B., KENNEDY, R.E., FASSNACHT, K.S. and BRIGGS, J.M., 1999, Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, **70**, pp. 52–68.

WELCH, R., MADDEN, M. and JORDAN, T., 2002, Photogrammetric and GIS techniques for the development of vegetation databases of mountainous areas: Great Smoky Mountains National Park. *ISPRS Journal of Photogrammetry and Remote Sensing*, **57**, pp. 53–68.

WELLES, J.M. and NORMAN, J.M., 1991, Instrument for indirect measurement of canopy architecture. *Agronomy Journal*, **83**, pp. 818–825.