

# Comparing and classifying one-dimensional spatial patterns: an application to laser altimeter profiles

S. Ollier<sup>a,\*</sup>, D. Chessel<sup>a</sup>, P. Couteron<sup>b</sup>, R. Pélissier<sup>c</sup>, J. Thioulouse<sup>a</sup>

<sup>a</sup>UMR CNRS 5558, Laboratoire de Biométrie et Biologie Evolutive, Université Claude Bernard Lyon 1, 69 622 Villeurbanne Cedex, France

<sup>b</sup>ENGREF/UMR AMAP, Boulevard de la Lironde, TA40/PS2, 34 398 Montpellier Cedex 05, France

<sup>c</sup>IRD/UMR AMAP, Boulevard de la Lironde, TA40/PS2, 34 398 Montpellier Cedex 05, France

Received 19 July 2002; received in revised form 29 January 2003; accepted 2 February 2003

## Abstract

Numerical analyses of remotely sensed data may valuably contribute to an understanding of the vegetation/land surface interface by pointing out at which scales a given variable displays a high level of spatial variability. Thus, there is a need of methods aimed at classifying many one-dimensional signals, such as airborne laser profiles, on the basis of their spatial structure. The present paper proposes a theoretical framework ensuring a consistent combination of a multi-scale pattern characterization, based on the Haar wavelet variance (also called in ecology Two Terms Local Variance, TTLV), with two multivariate techniques such as principal components analysis (PCA) and hierarchical cluster analysis. We illustrate our approach by comparing and classifying 257 laser profiles, with a length of 64 measurements (448 m), that were collected by the BRGM in French Guiana over three main landscape units with distinct geomorphological and ecological characteristics. We calculate for each profile a scalogram that summarized the multi-scale pattern and analyze the structural variability of profiles via a typology and a classification of one-dimensional patterns. More than 80% of the variability between spatial patterns of laser profiles has been summarized by two PCA axes, while four classes of spatial patterns were identified by cluster analysis. Each landscape unit was associated with one or two dominant classes of spatial patterns. These results confirmed the ability of the method to extract landscape scaling properties from complex and large sets of remotely sensed data.

© 2003 Elsevier Science Inc. All rights reserved.

*Keywords:* Multi-scale pattern analysis; Haar wavelet variance; Two Terms Local Variance; Classification of spatial patterns; Laser altimeter profile; Landforms; Forest canopy; Tropical rain forest

## 1. Introduction

Concepts of spatial pattern and scale play a central role in the study of both ecological and land surface processes (Levin, 1992; Scheidegger, 1991). The uneven distribution through space of resource, material and biomass is, indeed, both a determinant and a result of dynamic processes taking place at the interface between vegetation and landforms. Biotic as well as abiotic processes generally exert an influence on a broad range of scales, which does not mean that all scales are equally relevant to the study of a given phenomenon or land system. Hence, pattern characterization via numerical analyzes of remotely sensed data may valuably contribute to an understanding of the vegetation/land

surface interface by pointing out at which scales a given variable (e.g., total biomass) displays a high level of spatial variability.

Several methods, using autocorrelation, fractals, variograms, or wavelet variance have been proposed to investigate scales of spatial heterogeneity in remotely sensed data (Pachepsky, Ritchie, & Gimenez, 1997; Parker, Lefsky, & Harding, 2001; Rango et al., 2000). All these methods characterize a pattern observed in a particular sampling unit, for instance along a given laser transect or within a portion of a satellite scene, by studying the relationship between variance and scale (Dale, 1999; Dale et al., 2002; Ver\_Hoef, Cressie, & Glenn-Lewin, 1993). What has been largely missing, until now, is a general approach enabling multi-scale comparisons between a large number of spatial patterns, i.e., between many sampling units in which patterns are observed and quantified via one of the above methods. A reciprocal question would be how to assess the relative

\* Corresponding author. Fax: +33-4-78-89-27-19.

E-mail address: [ollier@biomserv.univ-lyon1.fr](mailto:ollier@biomserv.univ-lyon1.fr) (S. Ollier).

importance of scales on the basis of a large set of sampling units.

This is the very scope of the present paper which proposes a theoretical framework for a systematic comparison of spatial patterns observed in one-dimensional sampling units (i.e., transects). By revisiting an established method for pattern quantification, namely the Haar wavelet variance (Bradshaw & Spies, 1992; Burrus, Gopinath, & Guo, 1998; Dale & Mah, 1998)—also known in ecological sciences as Two-Terms-Local-Variance (Hill, 1973)—we propose a multivariate approach allowing to consistently compare and classify one-dimensional spatial patterns on a multi-scale basis. We shall also consider an application of this approach to airborne laser profiles obtained for a tropical forested landscape in French Guiana. The objective of that application is to compare patterns of laser altimeter profiles with the geomorphological and ecological context to determine whether pattern analysis of laser data could be a useful means for quantifying scaling properties of tropical forested landscapes.

## 2. Methods

### 2.1. Quantifying patterns via the Haar wavelet variance

Wavelet transforms are becoming increasingly popular for signal analysis, denoising and compressing (Burrus et al., 1998). One way to produce a wavelet transform,  $W(b, x_i)$ , from one-dimensional  $y=[y_i]_{1 \leq i \leq n}$ , is to compare the data sequence with a particular template or wavelet function,  $w(t)$ , centered at successive locations  $x_i$ , and expressing the scale,  $b$ , of the analysis (Bradshaw & Spies, 1992; Dale & Mah, 1998):

$$W(b, x_i) = \frac{1}{\sqrt{b}} \sum_{j=1}^n y_j w\left(\frac{x_j - x_i}{b}\right).$$

This formulation corresponds to a continuous wavelet transform (Misiti, Misiti, Oppenheim, & Poggi, 1993), for which the analyzing template is shifted smoothly over the full signal. Wavelet transforms generate results in terms of scale and position, i.e., a two-dimensional array from a one-dimensional transect. To facilitate comparisons between transects, the more synthetic wavelet variance function,  $V(b)$ , can be computed by averaging all squared values  $W(b, x_i)^2$  obtained at a particular scale  $b$ . In so doing, emphasis is put on scales while analytical information related to position is dropped to enable a systematic comparison of spatial patterns observed in many sampling units.

In ecological sciences, wavelet variances have long been used to analyze single transects under the name of Two Terms Local Variance or TTLV (Hill, 1973). However, Dale & Mah (1998) recognized that the TTLV can be seen as a

variance function of the oldest and simplest wavelet template, which was due to Haar (1910). With this template we have:

$$\begin{cases} w(t) = 1 & \text{if } (0 \leq t < \frac{1}{2}) \\ w(t) = -1 & \text{if } (\frac{1}{2} \leq t < 1) \\ w(t) = 0 & \text{otherwise} \end{cases}$$

Thus, using the Haar wavelet at a particular scale  $b$  is equivalent to computing the difference between successive blocks of  $b$  values observed along the transect. The wavelet variance is then computed from the average of the squared difference between blocks, which can be written for a finite data sequence,  $y=[y_i]_{1 \leq i \leq n}$  of  $n$  values, as:

$$V(b) = \text{average}(W(b, x_i)^2) = \frac{1}{K(b)} \times \sum_{i=1}^{n+1-2b} \left[ \sum_{j=i}^{i+b-1} (y_j - y_{j+b}) \right]^2$$

where  $K(b)$  is a scaling coefficient. Its choice may depend on the objective of the analysis. Our aim being to achieve a multi-scale comparison of spatial patterns, there is a need of a standardization of  $V(b)$  values, which would allow consistent comparisons of pattern intensity for different values of  $b$ . Indeed, large scales features (e.g., landforms) are likely to have a higher variance than the features at smaller scales (e.g., emanating from forest canopy), thereby determining most of the results of a multi-scale pattern comparison if a priori standardization is omitted.

### 2.2. Using matrix formulation to standardize $V(b)$ values

$V(b)$  is a quadratic form and, as such, can be written as:  $V(b) = (\mathbf{y}^t \mathbf{A}_b \mathbf{y}) / (K(b))$  where  $\mathbf{y}$  stands for the vector of observations and where  $\mathbf{A}_b$  denotes the square matrix (or metric) defining the quadratic form at scale  $b$ . This matrix is easily computable on the basis of the relationships linking observational units and blocks. Let  $\mathbf{P}$  be the matrix expressing the affectation of the  $n$  observational units to the  $N_b = n + 1 - b$  blocks defined at scale  $b$ .  $\mathbf{P}(l, m)$  is one if unit  $l$  belongs to block  $m$  and zero otherwise. Let  $\mathbf{M}$  stand for the  $N_b \times N_b$  matrix expressing the neighbourhood relationship between blocks. Only adjacent blocks that do not overlap are considered as neighbours, for instance blocks  $[x_i, x_i + b - 1]$  and  $[x_i + b, x_i + 2b - 1]$ .  $\mathbf{M}(l, m)$  is one if blocks  $l$  and  $m$  are neighbours and zero otherwise. Finally, let  $\mathbf{N}$  be the diagonal matrix, with  $\mathbf{N}(m, m)$  being the number of neighbours of block  $m$  and with  $\mathbf{N}(l, m) = 0$  for  $l \neq m$ . Matrix  $\mathbf{A}_b$  can be rewritten as  $\mathbf{A}_b = \mathbf{P}^t (\mathbf{N} - \mathbf{M}) \mathbf{P}$  (Lebart, 1969; see Appendix A for an example).

Note that the sum of all diagonal terms of  $\mathbf{A}_b$ , i.e.,  $\text{trace}(\mathbf{A}_b)$  is  $2b(n+1-2b)$ , a value which has been proposed by Hill (1973) for standardizing the TTLV. Indeed, taking  $K(b)=\text{trace}(\mathbf{A}_b)$  (Euclidean standardization) allows to consider  $V(b)$  as an estimate of variance at scale  $b$ . However, dividing by  $\text{trace}(\mathbf{A}_b)$  does not ensure that  $V(b)$  values are comparable at different scales. Furthermore, matrices  $\mathbf{A}_b$  are not diagonal matrices due to the use of overlapping blocks.

This problem can be solved using what has been defined by Chatelin (1988) as a “spectral standardization”, which relies on a singular value decomposition (SVD) of  $\mathbf{A}_b$ . (All matrices  $\mathbf{A}_b$  are semi-definite positive matrices and can be decomposed by SVD.) With this approach, the spatial pattern defined by the first eigenvector of  $\mathbf{A}_b$  is taken as reference for pattern analysis at scale  $b$ , and the associated eigenvalue,  $\lambda_1(b)$ , is used as reference variance for standardization, i.e.,  $K(b)=\lambda_1(b)$ . For instance, considering a transect length of  $n=64$  and scales of  $b=\{1, 2, 4, 8, 16, 24\}$  yields the reference spatial patterns displayed in Fig. 1. Using these eigenvectors as references means that, for a given transect, pattern intensity at a particular scale  $b$ , will be equal to 1 if and only if the transect values,  $y_{1 \leq i \leq n}$ , are proportional to the reference pattern. Thus, comparing transects on the basis of their underlying multi-scale structure, as well as assessing the relative importance of scales on the basis of a large set of transects can be both conceived in a consistent way.

### 2.3. Pattern ordination and classification

Let us consider the  $p \times q$  table  $\mathbf{Z}$ , for which each of the  $p$  rows represent the  $V(b)$  values of a given transect (the latter are called “scalogram” sensu Strang & Nguyen, 1996), and each of the  $q$  columns contain the scalogram values related to a given scale  $b \in \{1, 2, 3, \dots, (n/2)\}$ . Thus, in this table, the transects correspond to statistical units, while scales (i.e., sizes of blocks) are quantitative variables characterizing the transects in terms of spatial patterns. Note that Couteron (2002) used an analogous table for texture-based comparisons of air photographs, with the difference that Fourier spectra were used instead of Haar scalograms. Fourier spectra do not require any standardization prior to comparison, since reference functions (sine and cosine) are orthogonal, but Fourier decomposition is reputed to be less robust to non-stationarity than its wavelet analogues (Burrus et al., 1998).

We submit the  $(p \times q)$  table  $\mathbf{Z}$  to a Principal Components Analysis (PCA) on correlation matrix (Manly, 1994) that looks for a limited number of synthetic new variables accounting for a substantial share of the variability between scalograms. These new variables are linear combinations of initial variables and are called principal components. As scalograms have been preliminarily standardized by the “spectral” norm, each of the scale variables has the same weight in the analysis. PCA results have been displayed

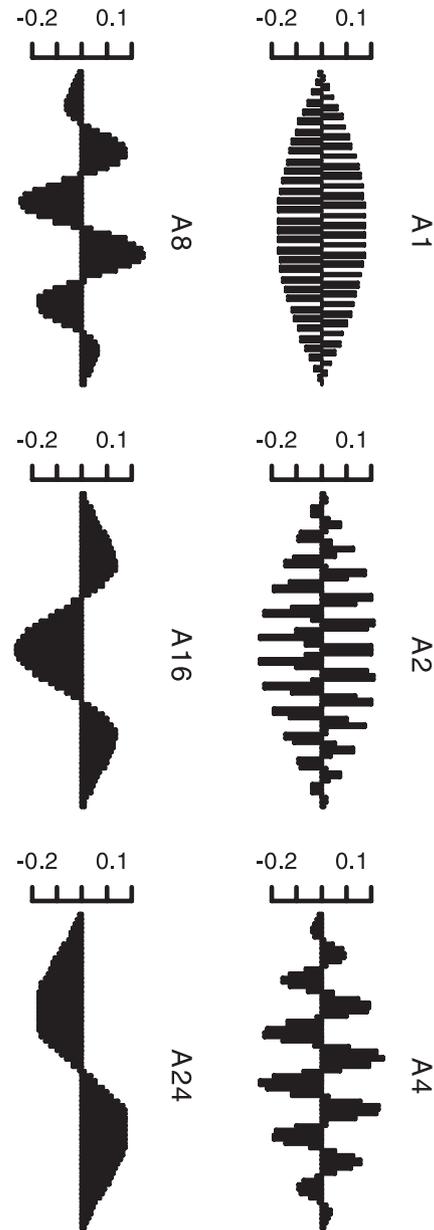


Fig. 1. Example of reference patterns, used for standardization, for a transect length of  $n=64$  and scales  $b=\{1, 2, 4, 8, 16, 24\}$ . Each of these patterns is deduced from the first eigenvector of the corresponding matrix  $\mathbf{A}_b$ .

using a particular graphical method called biplot (Gabriel, 1981).

As a complement to PCA ordination, a hierarchical cluster analysis has been carried out to classify scalograms. The analysis used the Euclidean distance between two scalograms  $i$  and  $j$ ,  $d_{ij}^2 = \sum_{k=1}^q (z_{ik} - z_{jk})^2$ , which is consistent with PCA, and the Ward’s minimum variance criterion that yields compact spherical clusters (Gordon, 1981).

All computations involved in the preparation of this paper were carried out using R (Ihaka & Gentleman, 1996), ADE-4 (Thioulouse, Chessel, Dolédec, & Olivier,

1997) and ArcView® as softwares, with both pre-programmed and personal routines.

### 3. Application to laser data

To illustrate the method, we used airborne laser altimetry data extracted from a very large data set of ca.  $10^5$  km of cumulated laser profiles, that were recorded in 1996 throughout French Guiana by the BRGM (Delor, Perrin, Truffert, Asfirane, & Rossi, 1998).

#### 3.1. Study site

We have restricted ourselves to a study site of 15,000 ha located at about 143 km towards the northwest of the main town Cayenne, between  $5^{\circ}28'$  and  $5^{\circ}38'$  North latitude and  $-53^{\circ}17'$  and  $-53^{\circ}28'$  West longitude (Fig. 2). It is situated in an unlogged rain forest called Counami forest. Climate is wet tropical with annual rainfall ranging between 2750 and 3000 mm and scattered over 9 months (Blancaneaux, 2001).

Within the Counami site, experienced geomorphologists have identified three main landscape units (Fig. 2) that differ in altitude and morphological complexity (Hutter, 2001). The first unit, i.e., “alluvial plains” (A), is characterized by a very simple and flat relief, and encompasses the valleys of the three main rivers (Counamama, Counami and Iracoubo) flowing through the study site. The second unit (B) stands for complex relief forms, having maximal altitudes below 60 m above sea level, and corresponding to tabular landforms that gently slopes toward talwegs. The last unit (C) is characterized by an altitude generally exceeding 60 m above sea level, associated with complex landforms that present steep slopes with small round-off summits.

#### 3.2. Laser data

An airborne laser altimeter was used to measure the distance from airplane to landscape surface. The profiling laser altimeter was a pulsed gallium–arsenide diode laser operating at a frequency of 195 Hz and a wavelength of 905 nm. The field of view of the laser was 5 rad giving a

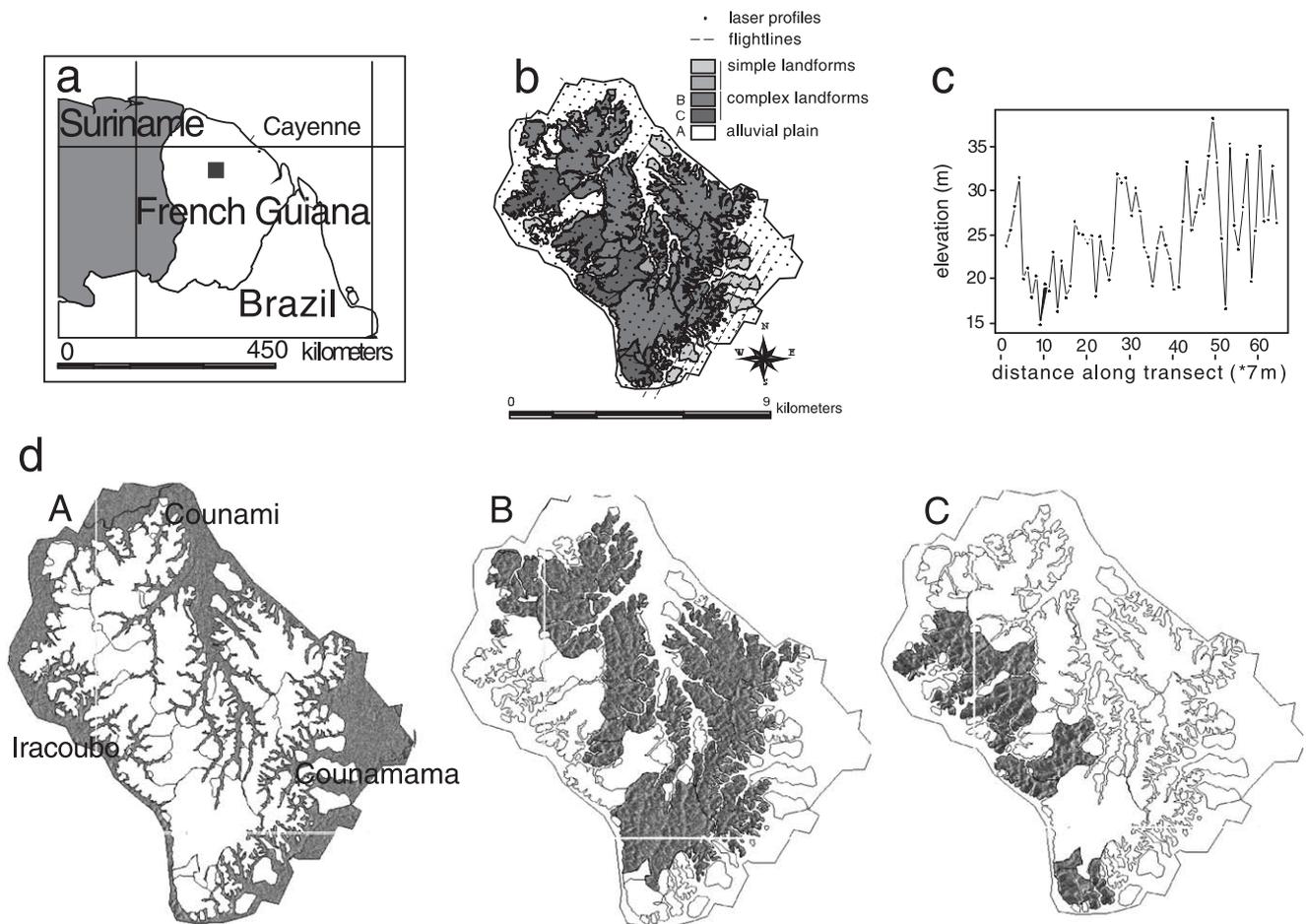


Fig. 2. Maps of general situation and sampling design of the Counami study site. (a) General situation. (b) Location of sampled transects (represented by the position of their midpoint). (c) Example of a unitary laser profile. (d) Maps of the three landscape units (A, B, C): for each unit, the texture of the relief is made apparent via an excerpt of a radar image, while the rest of the map is left blank.

“footprint” on the ground equal to 12 cm for the nominal altitude used during these flights. The timing mechanism in the laser receiver enabled a vertical resolution of 10 cm for a single measurement. The altitude of the airplane was approximately 120 m above the forest canopy with a nominal ground speed of 260 km/h. Data were pre-processed during recording by combining the number of shots so that a laser measurement occurs at intervals of 7 m along the flight line (Fig. 2). The geo-location of laser footprint was established with a spatial resolution of 2–3 m using a combination of ancillary data recorded simultaneously to laser measurements (video frames and GPS data). Landscape surface elevation was finally calculated for each measurement by using ground elevations along a flight line (GPS data) to convert the relative data into absolute elevations. Each measurement relates to the first obstacle encountered by the laser, and may correspond to vegetation items (leaf, branch) or to the forest floor itself.

Flight lines were oriented 30 °N and separated by 500 m. We extracted a set of  $p=556$  transects with a length of

$n=64$  points of measurements (448 m) from the 30 flight lines that intersected the study site. Each transect was represented in the geographical space by a point corresponding to its middle (Fig. 2). We eliminated transects that intersected several geomorphological types and kept only  $p=257$  homogeneous transects. We computed the scalogram for each of these 257 transects, using the following set of block sizes  $b \in \{1, 2, 4, 6, 8, 12, 16, 24\}$ . Then the scalograms were compared and classified to provide a typology of transects based on their spatial pattern.

### 3.3. Results

The two main PCA axes accounted for 81% of the variability between scalograms, with 48% and 33% for the first and the second axes, respectively (Fig. 3a). The first axis displayed a high negative correlation with small block sizes and a high positive correlation with large block sizes (Fig. 3b). The second axis was negatively correlated with intermediate block sizes. In terms of transect scores (Fig. 3b), it

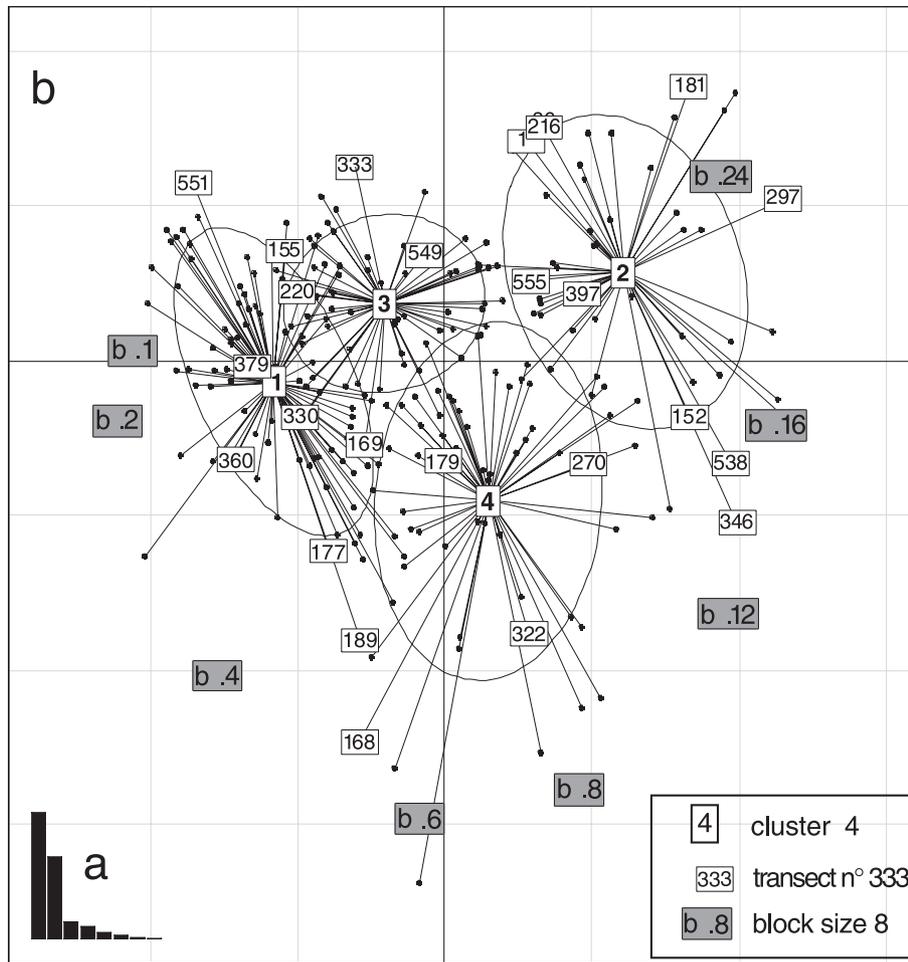


Fig. 3. Ordination and classification of the 257 scalograms. (a) Histogram of eigenvalues. (b) Biplot based on axes 1 and 2 of the principal components analysis (PCA) of the scalograms table Y. The biplot technique allows a simultaneous plotting of individuals (transects, denoted by points) and variables (block sizes, represented by grey labels). The identifying number of some transects, evoked in the text, is mentioned on the plot (white labels). Straight lines are linking transects to the gravity center of the cluster to which they belong (clusters are labelled from 1 to 4). Ellipses are based on the clusters' standard deviations.

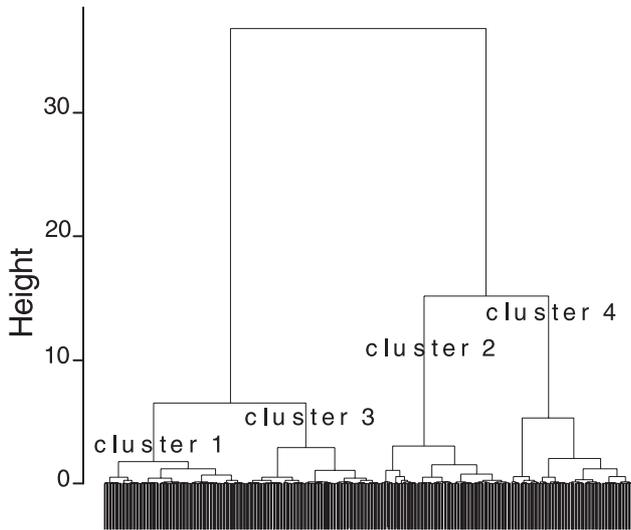


Fig. 4. Dendrogram yielded by the hierarchical cluster analysis of the scalograms table Y.

was particularly difficult to define clusters of transects by visual delineation. However, the hierarchical cluster analysis yielded a dendrogram containing four well-defined clusters (Fig. 4).

Centers of gravity of the clusters identified through the previous dendrogram are represented on the biplot in Fig. 3b. For each cluster, the average scalogram is displayed in Fig. 5a along with the laser profile having the smallest Euclidian distance to the average scalogram of the cluster. Cluster 1 was typical of fine-grained patterns, with scalograms that were exclusively skewed towards small scales. Laser profiles belonging to that cluster displayed a substantial variability, that was mainly observable along axis 2 in Fig. 3b, with a gradation of small scales from  $b=1, 2$  (transect 551, 379) to  $b=4$  (transect 360) and 6 (transect 189), i.e., from structures of size 7–42 m. As opposed to that cluster, the second one (cluster 2) was characterized by patterns which mainly presented large-scale variation between 100 ( $b=16$ ; transects 152, 346) and 200 m ( $b=24$ ; transects 181, 297). Scale

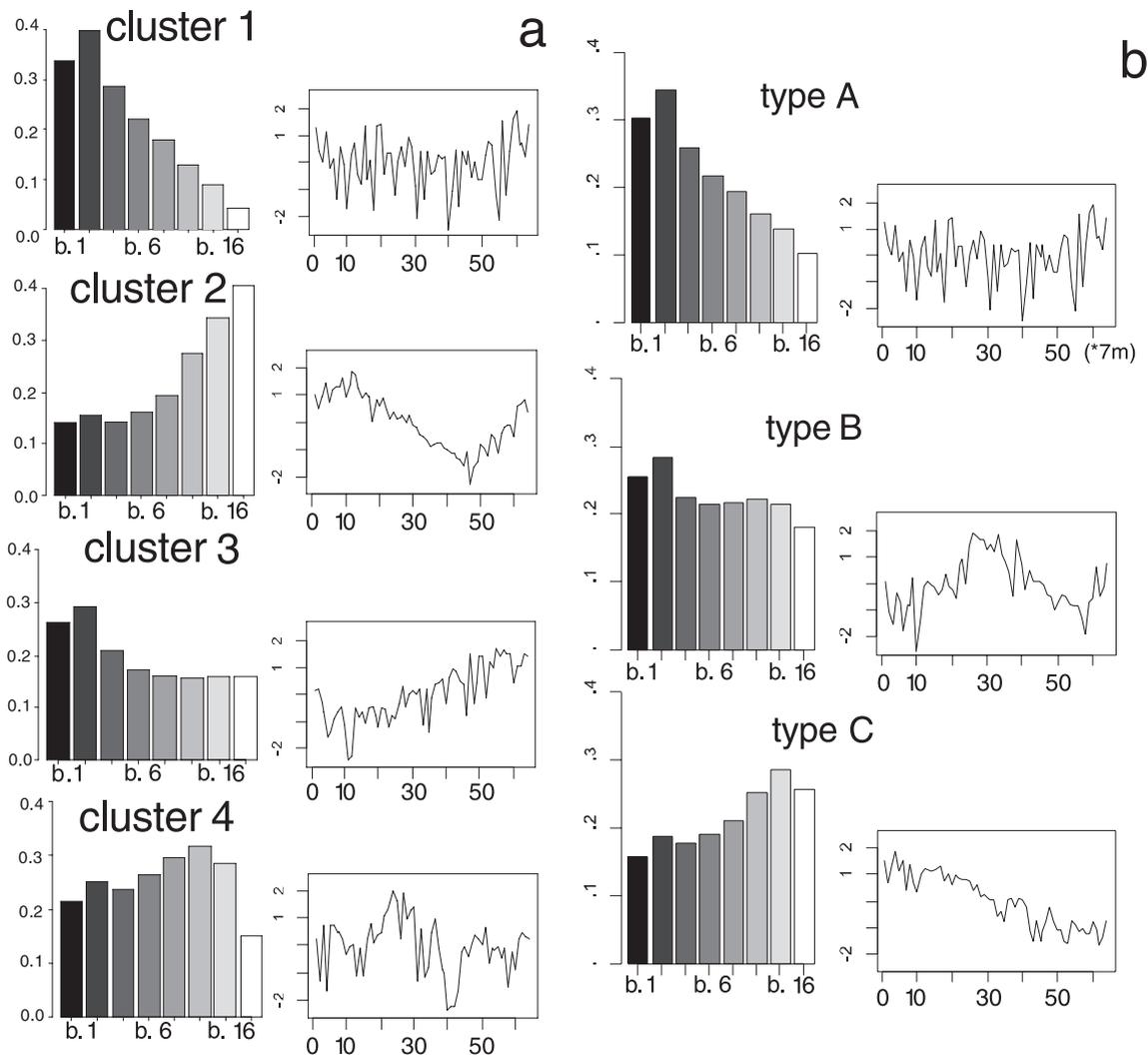


Fig. 5. Average scalograms and typical laser profiles for clusters (a) and landscape units (b). The raw laser profile of the transect showing the smallest Euclidian distance to the average scalogram is displayed for each cluster and each landscape unit.

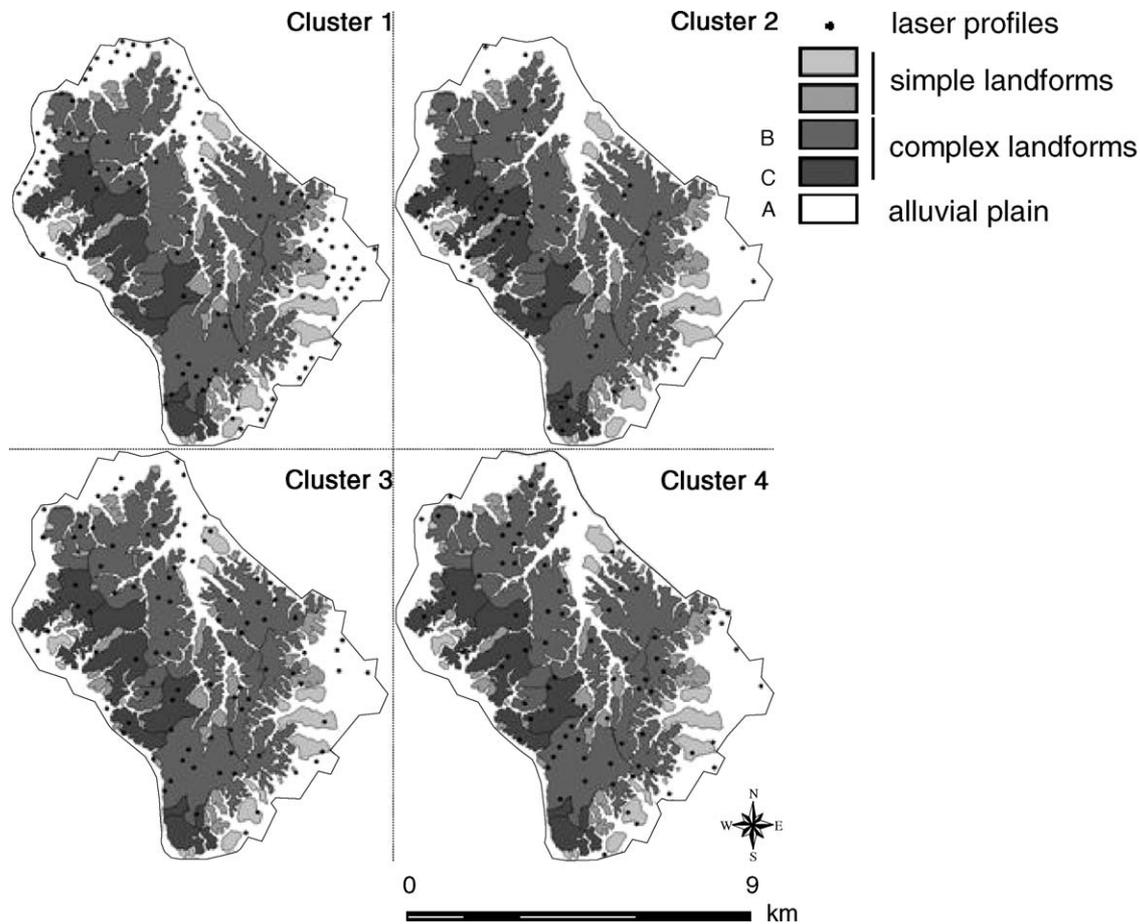


Fig. 6. Spatial relationship between clusters and landscape units. Transects belonging to the focal cluster are represented by their midpoint.

variability within this cluster was mainly apparent along axis 2. Cluster 3 gathered patterns characterized by the importance of both small and large scales (i.e., <40 and >100 m, respectively). It was typically a class of multi-scale patterns. Intensity attached to small scales was quite lower for this cluster than for cluster 1. Cluster 4 featured all the patterns characterized by the dominance of intermediate scales, ranging from about 40 to 100 m.

Laser transects belonging to the same cluster were not evenly distributed throughout the study area (Fig. 6). In particular, transects belonging to clusters 1 and 2 appeared clearly aggregated. This means that neighbouring transects displayed similar spatial patterns of laser response. The local heterogeneity of landscape surface roughness was thus quite limited. Moreover, laser patterns were related to landscape units, since the contingency table based on the cross-classification of the three landscape units and the four clusters of patterns (Table 1) departed significantly from the null hypothesis of independence ( $\chi^2$  test,  $P=2e-12$ ). Indeed, each landscape unit had most of its transects belonging to a particular cluster. Patterns defining cluster 1 were generally found in the alluvial plain (landscape unit A), while patterns making clusters 3 and 4 characterized the landscape unit B, and patterns of cluster 2 mostly belonged

to landscape unit C. The differences between landscape units were thus explained simultaneously by large and small scales patterns. Large scale differences were not surprising since they reflected large scale topographical fluctuations that characterized the landscape units B and C. The scale gradient observed from the laser profiles, that stretched from the alluvial plain to the complex landforms was also consistent with the visual aspect of the radar scenes (see Fig. 2c). Results at smaller scales (large intensity for the alluvial plain, lower intensity for the complex forms) were less expected, but may be interpreted as differences in forest

Table 1  
Contingency table expressing the cross-classification of the 257 transects into the three landscape units and the four clusters of patterns (row percent)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
(A) Alluvial plain (altitude < 20 m)	55	10	14	21	108
(B) Relief with complex forms (altitude < 60 m)	23	16	31	30	105
(C) Relief with complex forms (60 m > altitude)	9	52	23	16	44
Total	88	51	58	60	257

cover: small-scaled structures may correspond to convex units of canopy surface topography formed by the crown of emergent trees, isolated or juxtaposed with a minimal inter-crown space (Birnbaum, 2001; Pitman, Terborgh, Silman, & Nuñez, 1999). Other forms of forest organization can result in large-scale variations due to the presence of canopy gaps (Bradshaw & Spies, 1992; Birnbaum, 2001). These forms of canopy organization are liable to vary in relation to topographical and geological factors. Moreover, distinct floristic compositions were found for parts of the Counami forest corresponding roughly to the three landscape units (Couteron, Pélissier, Mapaga, Molino, & Teillier, 2002). Here, the correspondence between landscape units and clusters remained fuzziest, because the variability of laser profiles within landscape units was large (Table 1). For example, in the alluvial plain, in spite of a clear dominance of fine-grained patterns (cluster 1), large gaps in the forest cover may occasionally determine the presence of laser profiles belonging to the other clusters.

#### 4. Discussion

Using laser data, we have illustrated a new method for classifying one-dimensional patterns. In spite of a coarse resolution of the laser signal, results regarding geomorphology and to some extent forest canopy structure are interesting and nontrivial since they were related to an independent mapping of landforms. This is in agreement with results of several studies that have demonstrated the usefulness of various kinds of laser data to characterize vegetation cover and landforms (Drake & Weishampel, 2000; Pachepsky & Ritchie, 1998; Pachepsky et al., 1997; Weishampel, Blair, Knox, Dubayah, & Clark, 2000), and to investigate scaling properties of landscapes. Indeed as for other studies, we have found spatial patterns expressing themselves at several distinct scales. Of course, the characteristic of the laser information that was used did not allow a clear distinction of landform vs. canopy cover structure at intermediate scales but other kinds of laser techniques can allow such a distinction thanks to a higher spatial resolution, or to new sensors like the LVIS generation (Blair, Rabine, & Hofton, 1999), that yield separate signals for canopy cover and forest floor. Corresponding data sets will, nevertheless, represent a large amount of information that will increase the need of methods allowing consistent comparison and classification of spatial patterns. Thus, whatever the limitation of the data set we used for illustration, the present study demonstrated the efficiency of our new method to summarize and classify a large number of laser profiles.

Indeed, the method was able to summarize a rich set of data into a limited number of ordination axes that are linear combinations of spatial scales. More than 80% of the variability between scalograms, and thus between spatial patterns, was accounted for by the two main PCA axes. As

laser data contain pertinent information on landscape scaling properties, both the main axes of the PCA and the cluster dendrogram constitute valuable tools to explore, summarize and monitor landforms and/or vegetation covers. This kind of analysis is actually multi-scale at least for the range of scales that can be considered via a sampling window having a particular length (here 448 m).

Our approach is not restricted to a particular method of spatial pattern quantification, since pattern classification may stem from any method that produce an expression of variance with respect to scale. Indeed, the variance at a particular scale,  $b$ , can always be represented as a quadratic form ( $\mathbf{x}^t \mathbf{A}_b \mathbf{x}$ ) of a data vector  $\mathbf{x}$ . Consequently, the analysis of matrices  $\mathbf{A}_b$ , via singular value decomposition, provides explicit representations of the reference spatial patterns that a given method consider at a given scale (Fig. 1). This opens the way to a standardization of variance vs. scale functions (scalograms).

In spite of this unifying principle, the extent to which the choice of a particular method of pattern quantification (via a family of matrices  $\mathbf{A}_b$ ) may have some influence on the final result of the pattern classification is still an open question. For the present paper, we used the Haar wavelet variance (also known as Two Terms Local Variance) which is closely related to the variogram (Ver Hoef et al., 1993). The use of other simple wavelet templates (e.g., the ‘French Top Hat’; Dale & Mah 1998) is also thinkable. Another alternative is Fourier decomposition, which has been already used for classifying spatial patterns that display a strong periodic component (Couteron, 2002). Since our analysis only deals with scales and not with positions, the advantage of wavelets over Fourier decomposition may, in fact, be questioned. From a theoretical standpoint, it is well known that Fourier analysis is very good at detecting genuine periodicity, though non-stationarity may have some blurring effects. Taking short portions of the signal, as our ‘transects’ of 64 observations, may limit non-stationarity, but is also detrimental to the resolution of the Fourier spectrum (Kumaresan, 1993). Hence, even if only scales are dealt with, using wavelet alternatives to Fourier decomposition retain a substantial appeal if one has no a priori reason to suspect the presence of periodic features in the data. From a practical standpoint, however, the sensibility of a pattern classification to the Fourier vs. Haar wavelet choice is a question, which is still largely to be explored from both real-world and computer-generated data. This is, anyway, beyond the scope of the present paper, which intend mainly to underline that, thanks to spectral standardization, consistent multi-scale comparisons of one-dimensional patterns are possible from various families of scale vs. variance functions.

#### Acknowledgements

This work was undertaken thanks to a research agreement between the “Bureau de Recherches Géologiques et Minières” (BRGM) and the “Institut de Recherche pour le

Développement” (IRD) that allowed Sébastien Ollier to use the laser data from the 1996 airborne geophysical project of BRGM in French Guiana. We are particularly grateful to Catherine Truffert from BRGM-Orléans, who actively supported this collaboration.

**Appendix A**

The goal of this appendix is to illustrate on an example how the wavelet variance function at scale  $b$ ,  $V(b) = \frac{1}{K(b)} \times \sum_{i=1}^{n+1-2b} [\sum_{j=i}^{i+b-1} (y_j - y_{j+b})]^2$ , can be rewritten under a quadratic form  $V(b) = (\mathbf{y}^t \mathbf{A}_b \mathbf{y}) / K(b)$  where  $\mathbf{A}_b = \mathbf{P}^t (\mathbf{N} - \mathbf{M}) \mathbf{P}$ .

Let us consider a transect  $\mathbf{y}^t = (y_1, y_2, y_3, y_4, y_5, y_6)$  of length  $n = 6$ , and a block size  $b = 2$ .

There are  $N_2 = n + 1 - b = 5$  blocks of length 2 in the transect:  $[(y_1, y_2); (y_2, y_3); (y_3, y_4); (y_4, y_5); (y_5, y_6)]$ . This information is summarized in the matrix

$$\mathbf{P}_{(N_2, n)} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

The neighbourhood relationships between blocks is expressed by the matrix

$$\mathbf{M}_{(N_b, N_b)} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

and the number of neighbours of each blocks is represented on the diagonal of the matrix

$$\mathbf{N}_{(N_b, N_b)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The Haar wavelet variance at scale,  $b = 2$ , is

$$V(2) = \frac{[(y_1 + y_2 - y_3 - y_4)^2 + (y_2 + y_3 - y_4 - y_5)^2 + (y_3 + y_4 - y_5 - y_6)^2]}{K(b)}$$

$$= \frac{\begin{pmatrix} y_1 + y_2 \\ y_2 + y_3 \\ y_3 + y_4 \\ y_4 + y_5 \\ y_5 + y_6 \end{pmatrix}^t \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 + y_2 \\ y_2 + y_3 \\ y_3 + y_4 \\ y_4 + y_5 \\ y_5 + y_6 \end{pmatrix}}{K(b)},$$

i.e.,

$$V(2) = \frac{(\mathbf{P}\mathbf{y})^t (\mathbf{N} - \mathbf{M})(\mathbf{P}\mathbf{y})}{K(b)} = \frac{\mathbf{y}^t (\mathbf{P}^t (\mathbf{N} - \mathbf{M}) \mathbf{P}) \mathbf{y}}{K(b)}$$

$$= \frac{\mathbf{y}^t \mathbf{A}_b \mathbf{y}}{K(b)} \quad \text{with } \mathbf{A}_b = \mathbf{P}^t (\mathbf{N} - \mathbf{M}) \mathbf{P}.$$

**References**

Birnbaum, P. (2001). Canopy surface topography in a French Guiana forest and the folded forest theory. *Plant Ecology*, 123, 293–300.

Blair, J. B., Rabine, D. L., & Hofton, M. A. (1999). The laser vegetation imaging sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 115–122.

Blancaneaux, P. (2001). Le climat guyanais. In J. Barret (Ed.), *Atlas illustré de la Guyane* (pp. 46–49). CNES, IESG, IRD, Région Guyane.

Bradshaw, G. A., & Spies, T. A. (1992). Characterizing canopy gap structure in forests using wavelet analysis. *Journal of Ecology*, 80, 205–215.

Burrus, C. S., Gopinath, R. A., & Guo, H. (1998). *Introduction to wavelets and wavelet transforms*. Upper Saddle River, NJ: Prentice-Hall.

Chatelin, F. (1988). *Valeurs propres de matrices*. Masson, Paris.

Couteron, P. (2002). Quantifying change in patterned semi-arid vegetation by Fourier analysis of digitised aerial photographs. *International Journal of Remote Sensing*, 23(17), 3407–3425.

Couteron, P., Pélissier, R., Mapaga, D., Molino, J. F., & Teillier, L. (2002). Drawing ecological insights from a management-oriented forest inventory in French Guiana. *Forest Ecology and Management*, 172(1), 89–108.

Dale, M. R. T. (1999). *Spatial pattern analysis in plant ecology*. Cambridge: Cambridge Univ. Press.

Dale, M. R. T., Dixon, P., Fortin, M. J., Legendre, P., Myers, D., & Rosenberg, M. S. (2002). The conceptual and mathematical relationships among methods for spatial analysis. *Ecography*, 25, 558–577.

Dale, M. R. T., & Mah, M. (1998). The use of wavelets for spatial pattern analysis in ecology. *Journal of Vegetation Science*, 9, 805–814.

Delor, C., Perrin, J., Truffert, C., Asfirane, F., & Rossi, P. (1998). Images géophysiques dans le socle guyanais. *Géochronique*, 67, 7–12.

Drake, J. B., & Weishampel, J. F. (2000). Multifractal analysis of canopy height measures in a longleaf pine savanna. *Forest Ecology and Management*, 128, 121–127.

Gabriel, K. R. (1981). Biplot display of multivariate matrices for inspection of data and diagnosis. In V. Barnett (Ed.), *Interpreting multivariate data* (pp. 147–174). New York: Wiley.

Gordon, A. D. (1981). *Classification*. London: Chapman & Hall.

Hill, M. O. (1973). The intensity of spatial pattern in plant communities. *Journal of Ecology*, 61, 225–235.

- Hutter, S. (2001). *Etude geomorphologique du massif forestier de Counami*. Technical report, CIRAD.
- Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5, 299–314.
- Kumaresan, R. (1993). Spectral analysis. In S. K. Mitra, & J. F. Kaiser (Eds.), *Handbook for digital signal processing* (pp. 1143–1169). New York: Wiley.
- Lebart, L. (1969). Analyse statistique de la contiguïté. *Publication de l'Institut de Statistiques de l'Université de Paris*, 28, 81–112.
- Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73, 1943–1967.
- Manly, B. F. (1994). *Multivariate statistical methods*. London: Chapman & Hall.
- Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. M. (1993). Ondelettes en statistique et traitement du signal. *Revue de Statistique Appliquée*, XLI, 33–43.
- Pachepsky, Y. A., & Ritchie, J. C. (1998). Seasonal changes in fractal landscape surface roughness estimated from airborne laser altimetry data. *International Journal of Remote Sensing*, 19, 2509–2516.
- Pachepsky, Y. A., Ritchie, J. C., & Gimenez, D. (1997). Fractal modelling of airborne altimeter laser altimetry data. *Remote Sensing of Environment*, 61, 150–161.
- Parker, G. G., Lefsky, M. A., & Harding, D. J. (2001). Light transmittance in forest canopies determined using airborne laser altimetry and in-canopy quantum measurements. *Remote Sensing of Environment*, 76, 298–309.
- Pitman, N. C. A., Terborgh, J., Silman, M. R., & Nuñez, P. (1999). Tree species distributions in an upper Amazonian Forest. *Ecology*, 80, 2651–2666.
- Rango, A., Chopping, M., Ritchie, J. C., Havstad, K., Kustas, W., & Schmugge, T. (2000). Morphological characteristics of shrub coppice dunes in desert grasslands of southern New Mexico derived from scanning LIDAR. *Remote Sensing of Environment*, 74, 26–44.
- Scheidegger, A. E. (1991). *Theoretical geomorphology*. New York: Springer-Verlag.
- Strang, G., & Nguyen, T. (1996). *Wavelets and filter banks*. Wellesley, MA: Wellesley-Cambridge Press.
- Thioulouse, J., Chessel, D., Dolédec, S., & Olivier, J. M. (1997). ADE-4: a multivariate analysis and graphical display software. *Statistics and Computing*, 7, 75–83.
- Ver\_Hoef, J. M., Cressie, N. A. C., & Glenn-Lewin, D. C. (1993). Spatial models for spatial statistics: some unification. *Journal of Vegetation Science*, 4, 441–452.
- Weishampel, J. F., Blair, J. B., Knox, R. G., Dubayah, R., & Clark, D. B. (2000). Volumetric lidar return patterns from an old-growth tropical rainforest canopy. *International Journal of Remote Sensing*, 21, 409–415.