

Original article

## Evaluation of large-scale unsupervised classification of New Caledonia reef ecosystems using Landsat 7 ETM+ imagery

### Évaluation d'une classification à grande échelle avec une classification non-supervisée dans les écosystèmes récifaux de Nouvelle-Calédonie en utilisant les images Landsat 7 ETM+

Guénolé Bouvet <sup>a,\*</sup>, Jocelyne Ferraris <sup>a</sup>, Serge Andréfouët <sup>b</sup>

<sup>a</sup> Institut de la recherche pour le développement, UR128 (CoRéUs), BP A5, 98848 Nouméa cedex, Nouvelle-Calédonie, France

<sup>b</sup> Institute for Marine Remote Sensing, College of Marine Science, University of South Florida, St Petersburg, FL 33701, USA

Received 13 June 2002; revised and accepted 12 November 2002

#### Abstract

The capacity of the Landsat 7 Enhanced Thematic Mapper Plus sensor to classify the shallow benthic ecosystems of New Caledonia (South Pacific) is tested using a novel unsupervised classification method. The classes are defined by using a set of multiple spectral decision rules based on the image spectral bands. A general model is applied to the entire Southwest lagoon (5500 km<sup>2</sup>) and tested on three representative sites: a section of the barrier reef, a cay reef flat rich in corals, and a cay reef flat rich in algae and seagrass beds. In the latter one, the classification results are compared with a locally optimized model, with aerial color photographs and extensive ground-truthed observations. Results show that a reconnaissance of the main benthic habitats in shallow areas (<5 m depth) is possible, at a geomorphological scale for coral reef structure and at a habitat scale for seagrass beds. However, results directly issued from the model must be cautiously interpreted according to empirical spatial rules, especially to avoid confusion between coral slopes and shallow dense seagrass.

© 2003 Éditions scientifiques et médicales Elsevier SAS and Ifremer/CNRS/IRD. All rights reserved.

#### Résumé

Le but de cette étude est de tester la capacité des images Landsat 7 Enhanced Thematic Mapper Plus à discriminer les principales classes d'habitats benthiques rencontrées dans les parties peu profondes du système récifal et lagonaire de Nouvelle-Calédonie (Pacifique Sud). Une méthode originale de classification non-supervisée est proposée. Les habitats benthiques correspondent à une combinaison de plusieurs règles de décision établies à partir des bandes radiométriques Landsat. Cette modélisation statistique des habitats benthiques est appliquée au lagon sud-ouest de Nouvelle-Calédonie (5500 km<sup>2</sup>). Les résultats sont testés sur trois sites témoins contrastés: un platier de récif barrière, un platier d'îlot riche en corail et un platier d'îlot riche en herbiers/algues. Pour ce dernier, le résultat est comparé à celui d'un modèle optimisé, construit à échelle locale et validé à partir de photographies aériennes et d'observations de terrain. Les résultats montrent qu'une reconnaissance des différentes classes benthiques est possible pour des fonds peu profonds (< 5 m de profondeur), à l'échelle géomorphologique pour les structures récifales et à l'échelle des habitats pour les herbiers. Toutefois, les résultats bruts du modèle doivent être interprétés en fonction de critères spatiaux pour corriger les confusions entre certaines classes, notamment entre les pentes coralliennes et les herbiers denses.

© 2003 Éditions scientifiques et médicales Elsevier SAS and Ifremer/CNRS/IRD. All rights reserved.

**Keywords:** Remote sensing; Statistical modeling; Indo-Pacific; Seagrass; Coral

**Mots clés :** Télédétection ; Modèle statistique ; Indo-Pacifique ; Herbier ; Corail

\* Corresponding author.

E-mail address: [bouvet@geolitto.com](mailto:bouvet@geolitto.com) (G. Bouvet).

## 1. Introduction

The heavy anthropogenic pressure along the coasts of the Indo-Pacific region severely threatens coastal marine ecosystems, which results in a quick degradation and loss of fragile biota such as coral reefs and seagrass beds (Wilkinson, 2000). Countless reports document locally the dramatic consequences of pollution, overfishing, urban development and coral mining (Wilkinson, 2000). Moreover, a global-scale climatic change seems to induce new threats, even for pristine ecosystems that are not directly under the human influence (Kleypas et al., 2001). Therefore, large-scale quantitative assessments are urgently required to establish quickly and accurately the status of these crucial marine ecosystems, in order to monitor the general trends and discriminate between human influences and natural factors.

The coastal marine ecosystems of New Caledonia (161–167 E, 18–23 S, Fig. 1) present an ideal configuration to conduct such a large-scale assessment (Labrosse et al., 2000). The mainland, extending over 19,100 km<sup>2</sup>, presents large coastal plains on the western side and abrupt shores covered by rainforests on the eastern side. This isolated land is surrounded by a very large deep lagoon (23,400 km<sup>2</sup>) rimmed by an almost continuous shallow barrier reef. This reef complex is one of the largest in the world. It could be added in the short future to the list of Unesco World heritage areas that already includes the Great barrier reef in Australia. Most of the New Caledonia reef system is considered to be pristine without significant human-induced threats (Labrosse et al., 2000), except in the vicinity of the capital Nouméa (Fig. 1).

Nouméa and the southern province of New Caledonia are increasing their industrial (nickel extraction), aquaculture, fisheries and tourist activities sometimes at the detriment of mangroves and fringing reefs. As a consequence, most of the research and environmental monitoring activities on lagoon, reef and coastal systems takes place in the large Southwest lagoon (Fig. 1). These investigations focus on reef and soft-

bottom fish communities (Rossier and Kulbicki, 2000), influence of marine reserves on fish populations (Wantiez et al., 1997), water and sediment chemistry, hydrodynamic modeling and patterns of sedimentation within the lagoon (Douillet et al., 2001). Soft-bottom communities in deep areas have been studied and mapped using discrete sampling (Garrigue, 1995) and acoustic technologies are currently under evaluation for seagrass, sediments and coral mapping (Chevillon, 2001). However, there are still no exhaustive and comprehensive inventories of the various marine habitats present in the lagoon and reef flats. The shallow barrier reefs and reef flats surrounding cays have been neglected and information on their structures remain scarce, despite several optical remote sensing studies in the mid-eighties. Satellite (real or simulated) multi-spectral images were interpreted to map islands and barrier reef flats (Thomassin, 1985; De Vel and Bour, 1990), to assess coral reef health (Bour et al., 1993) and to estimate marine resources stocks (Bour et al., 1986). Despite this pioneer work, no further studies were pursued in New Caledonia during the 1990s, while in the same time, both sensors and analytical methods were improved worldwide.

The potential of remote sensing data to assess tropical coastal ecosystems is now quite well understood (Green et al., 2000; Hochberg and Atkinson, 2000; Andréfouët et al., 2001; Chauvaud et al., 2001). Inventories and mapping of coastal systems described in broad 4–6 broad classes are more cost-effective when they use Landsat or similar satellite sensors (Green et al., 2000). For more detailed assessments (e.g. pigmentation) costly airborne hyperspectral data may be required (Andréfouët et al., 2003).

Our long-term objectives are to map the reef flats of New-Caledonia's Southwest lagoon using the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor that was launched in April 1999. So far, this study is one of the very few that used Landsat ETM+ data for a large-scale assessment of shallow marine communities in coral reef environments. We limit this study to the reef flats and slopes of approximately 5 m depth maximum. To our knowledge, only the Great barrier reef (GBR) in Australia has been studied intensively at large-scale using the Landsat multi-spectral scanner (MSS) and the thematic mapper (MS) data (e.g. Jupp et al., 1985), while most of the other studies have only considered one particular reef or a small area of interest, where supervised classification, ground-truthing and accuracy assessment could be performed in a reasonable amount of time. Here, our constraints are different and the exploration of the variability of the large Southwest lagoon requires a different approach. Since few data exist on these reef-flats, we have only used unsupervised techniques and radiometric criteria to cluster the pixels. This paper aims to critically review the unsupervised classification of three different sites that are representative of the Southwest lagoon. Therefore, this study provides: (1) one of the first assessments of ETM+ data for a large-scale coastal study, and (2) a classification scheme for a New Caledonian reef flats mapping.

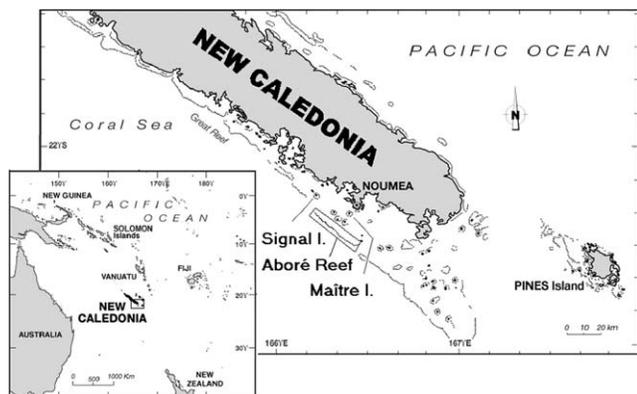


Fig. 1. Southwest lagoon of New Caledonia (South Pacific) and location of the three validation sites: Signal Islet, Maître Islet and Aboré Reef.

## 2. Materials and methods

### 2.1. Study area

The New Caledonian southwest lagoon extends over 5500 km<sup>2</sup> (Fig. 1), with an average depth of 17.5 m (Douillet et al., 2001). Many sand cay islands are present in the lagoon, associated with large shallow flats. Two pilot sites representative of the cay morphology were considered for this study (Maître and Signal, Fig. 1). Signal is characterized by a reef flat with live and dead coral and algal communities. Conversely, the Maître reef flat is characterized by seagrass beds, frequently mixed with algae populations. Both sites display extensive and healthy coral populations, especially on the outer slopes exposed to dominant trade winds. A ridge of coral rubble at the periphery of the flats, emergent at low tide, is present on both the sites.

The lagoonal system is closed by a large barrier reef system 15 km offshore. The shallowest part of the barrier reef is characterized by two main environments. First, a coral-rich flat extends from the breaking wave area out to an internal spur and grooves system. The transition to the second zone is sharp. It leads to a sand zone, with rubble pools and isolated coral patches along a gentle slope. The Aboré reef is our third test site (Fig. 1).

### 2.2. Remote sensing data

For this study, we used an ETM+ image that was acquired on August 11, 1999, centered by 23.11°S/174.41°W. The image covered a 180 × 180 km area, and included a large part of the southern New Caledonia lagoon and reef system. Specifications of ETM+ data were described elsewhere (Thome, 2001). Briefly, the ETM+ image provided four spectral bands relevant for underwater targets (ETM+ 1–4) at 30 m resolution data. They measured the total radiance in the blue (450–520 nm), green (520–600 nm), red (630–690 nm) and in the near-infrared part of the spectrum (760–900 nm), the latter can be of interest for extremely shallow conditions (Menges et al., 1998).

The ETM+ data were complemented by a collection of scanned low altitude aerial orthophotographs of several reef flats acquired in June, 1999. This very high resolution data set (resolution 2 m) was used to stratify the reconnaissance of shallow habitats in the Maître and Signal cay islands and assess the accuracy of the classifications.

### 2.3. Image processing

The ETM+ image was imported into Savane. Savane is a Geographical Information System (GIS) software used to merge remote sensing data with a geographical database, named LagonNC. The classification of the radiometric data (explained hereafter) was achieved with SPADN, a statistical package.

The ETM+ data processing was primarily constrained by the initial lack of information on the structure and composi-

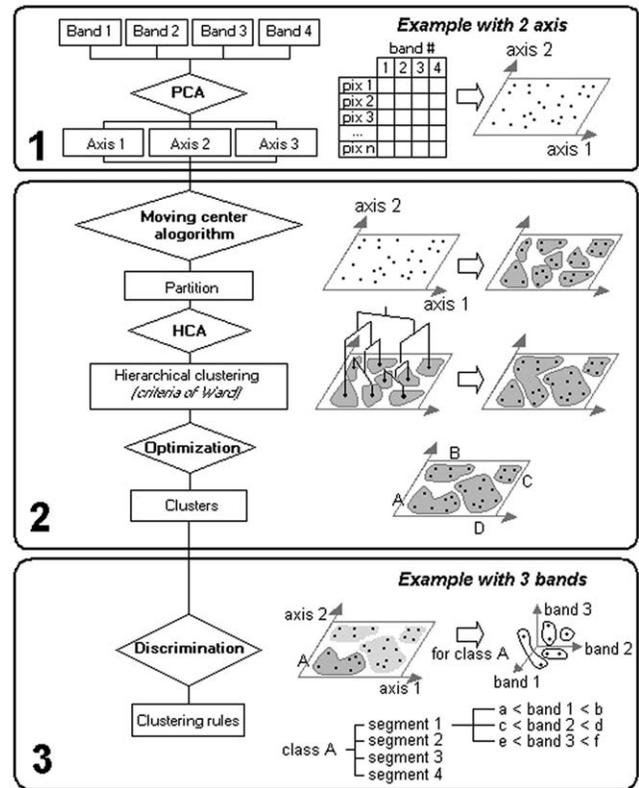


Fig. 2. Flowchart for unsupervised classification/segmentation of Landsat imagery. Rectangular boxes refer to objects, while diamonds refer to actions.

tion of the reef ecosystems throughout the lagoon. The approach is a 4-stage process. Fig. 2 summarizes the statistical analysis. A large set of pixels (more than 200,000) issued from nine representative sites lagoon-wide was used to explore the spectral variability of the Southwest lagoon area. Land, clouds and cloud shadows were manually masked using the NIR bands. The 4-stage classification strategy, inspired by Lebart et al. (1984), was:

1. Ordination: A principal component analysis (PCA) was applied to the bands 1–4. Only the three main axes summarizing the most important part of the total variance are used in the following steps.
2. Partition: A moving center partitioning algorithm (also called “des nuées dynamiques” algorithm (Diday, 1971)) is applied twice using two sets of 10 random initial centers. Calculations are performed on the coordinates of the individuals on the third main axes of the PCA. Pixels are assigned to the most similar center. By cross-tabulating the two resulting partitions, one can identify stable pixels grouped together in both partitions. These stable groups are aggregated pairwise using a hierarchical clustering algorithm (HCA) with the minimum variance criterion of Ward (Ward, 1963), that, at each step, minimizes the variance within clusters and maximizes the variance between clusters. Discontinuities in the criterion of aggregation emphasize successive optimal partitions with  $k$  classes. We keep the first optimal partition with at least 10 classes to get a number

of classes not too high (Green et al., 2000), but sufficient to likely capture the various main reef biota. To consolidate the partition, a final moving center algorithm is applied, but instead of a random selection of 10 initial centers, we use the *k* centers of the *k* classes to start the process.

At this stage, the unsupervised classification was finished. In most of the unsupervised mapping exercises (Green et al., 2000), the posterior field work allows to assign a thematic meaning to the computed classes. Eventually, after the field trip, a final supervised process is conducted using decision rules that depend on the type of the selected algorithm (Maximum Likelihood, Mahalanobis Distance, Minimum Distance or others) and the adequate statistical parameterization of the classes (a priori probabilities, covariance matrix, standard deviation). Most of these methods assumed that the shape of the class in the spectral space is simple (parallelepiped, ellipsoids, etc.). This assumption may not be verified if the class is composed of different objects with contrasting spectral signatures. For example, a class "seagrass" may include low density seagrass with a bright signal (due to the dominant bright sandy floor) and dense seagrass with a much darker signal (due to pigments absorption). A unique parameterization is thus inadequate and two (or more) sets of parameters for the same class may be more appropriate. Such observations have led to various fuzzy clustering methods (e.g. using fuzzy covariance matrix (Gustafson and Kessel, 1979)) to allow complex representation in the spectral space. Thus, we have proposed a method to accommodate complex classes, by simply allowing for more complex decision rules.

3. Discrimination: the optimized partition obtained in stage two undergoes a segmentation (or discrimination) based on the initial variables, i.e. the radiometric bands ETM+1-4, and not the PCA-derived neo-bands. We follow the classification and regression tree (CART) algorithm (Breiman et al., 1984). This segmentation hierarchically splits the initial pixel population into child-segments. It minimizes intrasegment variance and maximizes intersegment variance. The two segments have to be more homogeneous than the parent-segment or the process stops. The result is a sequence of  $N_c$  segments for each class *c*. The statistical parameters of each segment are then used as decision rules to assign a pixel to one of the segments (then to a class). Here, we used a criteria similar to the parallelepiped distance algorithm (Jensen, 1996), i.e. a segment is defined by its lower and upper spectral boundaries,  $DN_{\min,j}$  and  $DN_{\max,j}$ , in each ETM+ band *j*. A pixel *p* belongs to the segment *S* (and the corresponding class *c*) if

$$DN_{\min,j}(S) \leq DN_j(p) \leq DN_{\max,j}(S),$$

with  $j = 1, 2, 3, 4$

Classification rules thus depend on the digital counts (*DN*) of the different spectral bands. Of course, similar methods can be applied on digital counts, total radiance, total reflectance, water leaving radiance, or remote sensing reflectance, according to the type of normalization and corrections applied to the image.

4. Validation: The set of decision rules obtained from nine representative sites defined what we call as "model I". The model was built using 200,000 pixels lagoon-wide, but did not include the test sites, the Maître, Signal or Aboré reefs. To estimate the relevance of this global model, we followed two approaches (Fig. 3). First, we created an optimized model using the same previously described three stages, but at a local scale on one site (Maître reef flat). We assessed the final classification using aerial orthophotographs and ground-truth data. Orthophotographs and ground-truth data were integrated in the GIS database. GIS tools were used to aggregate segments, merge classes and field stations, obtain statistics by class, calculate areas, inquire DN of each pixel and output layouts. This model II was compared to model I for the same set of Maître reef flat pixels. Second, the composition of the classes and segments identified with the decision rules of the model I on the Maître, Signal and Aboré reefs was compared between each site to identify possible confusions between benthic habitats.

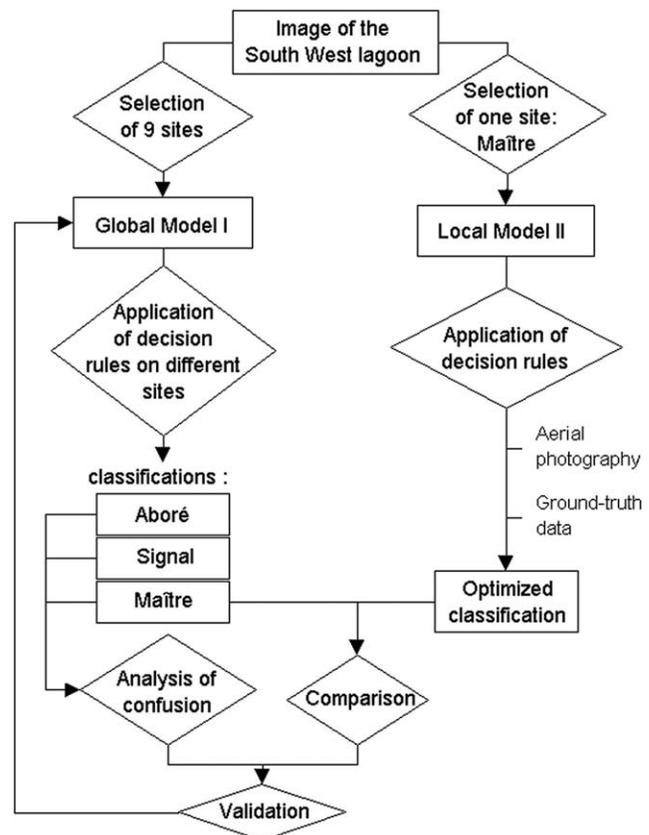


Fig. 3. Flowchart for the validation of large-scale unsupervised classification based on the general model I.

## 2.4. Fieldwork

Detailed fieldwork was conducted on the Maître reef flat in August 2000, 1 year after the acquisition of the Landsat image, but in similar season. Fieldwork occurred after the unsupervised processing of the ETM+ image. Stations were located along transects positioned to cross as many classes as possible. In each station (total = 50) depth and nature of substrate were assessed. To categorize the substrate, we followed a simple typology (sand, bare pavement, rubble, seagrass, algae, coral, heterogeneous) inspired by other classification schemes from other tropical coral reef areas (Mumby and Harborne, 1999; Andréfouët et al., 2000). For seagrass beds, cover density was estimated on a semi-quantitative scale (1: sandy bottom with less than 10 blades/m<sup>2</sup>, to 5: compact dense seagrass cover without sediment visible beneath).

A second field campaign occurred in December 2001 and January 2002 to identify the macrophytes and seagrass communities and to assess the accuracy of the classes identified on the Maître, Aboré and Signal reefs.

## 3. Results

Model I was elaborated with more than 200,000 pixels from nine sites corresponding to different lagoonal zones (reef flats, barrier reef, bay) throughout the southwest lagoon. We remind that the Maître, Signal and Aboré reefs were not included in the training sites. The first three axes of the PCA explained 98.2% of the total variance (49, 28.3 and 20.9%, respectively). The last axis explained 1.8%. The partitioning and segmentation provided 11 classes split into 68 segments. One or a few dominant segments generally described each class.

For Maître, the decision rules of the model I provided 10 classes, with 8 relevant classes (K1-K8) located on the reef flat and slopes (Fig. 4). In comparison, the model II, optimized for Maître, described the reef flat and slopes in seven relevant classes (C1-C7) segmented into 58 segments. In contrast with model I, the first three axes of the PCA explained 72.8, 23.9 and 2.6% of the total variance, respectively. The high variance explained by the first axis confirmed that one reef alone was spectrally more homogeneous than the assemblage of the many different reefs used to build model I. Generally, the spatial patterns visible on Maître reef flat using model I and model II were consistent (Fig. 4) with some variations. This was especially true on the reef flat. Table 1 presents a cross-table of pixel occurrences within each class from model I and model II.

Ground-truth data revealed that the Maître reef flat was characterized by seagrass beds, monospecific (*Halophila ovalis*, *H. uninervis*) or heterogeneous (mix of *H. uninervis*, *C. serrulata*, *Syringodium isoetifolium*), or frequently mixed with algae populations (including Phaeophyta *Padina* sp., *Sargassum* spp., *Lobophora variegata*, *Cystoseira trinodis*, *Dictyota bartayresiana*, *Hormophysa triquetra*, Rhodophyta

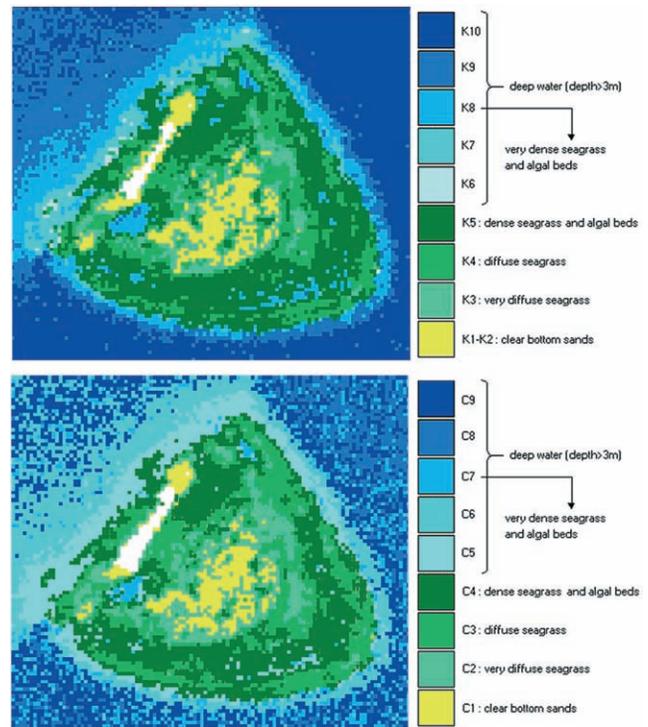


Fig. 4. Comparison of the classifications of the Maître Islet obtained from global model I (top) and local model II (bottom).

*Acantophora spicifera*, *Digenea simplex*, *Galaxaura lapidescens*, *Spyridia filamentosa*, Chlorophyta *Halimeda macroloba*, *H. cylindracea*, *H. opuntia*, *H. sinulans*, cyanobacteria). The site displayed outer slopes with extensive and healthy coral populations especially on the eastern and southern sides that were exposed to dominant trade winds (mostly branching *Acropora* spp. on the top and the slope, and highly diverse formations on the floor).

Shallow sandy bottoms corresponded to the class C1 in model II and both K1 and K2 in model I, with an average depth of 0.7 m. A similar average depth was also found for other shallow classes, i.e. C2, C3, C4 and C7 in model II and K3, K4, K5 and K8 in model I. Each deep class (C5-C6, K6-K7) integrated sand channels, coral zones and heterogeneous areas (made of isolated coral colonies and patches, seagrass, rubble and algae), but model II provided results far less noisy than model I (Fig. 4). None of the models clearly separated the rubble zone at the periphery of the reef flats. The classes C2, C3 and C4 in model II (respectively, K3, K4 and K5 in model I) characterized well the gradient of seagrass densities as evidenced by the overlays of ground-truth data, aerial photographs and classification results (Figs. 5, 6).

To assess the accuracy of the Maître reef flat model, we compared the results with higher resolution aerial photographs classified by photo-interpretation. At this stage, we lacked enough control points to establish a formal site-specific confusion matrix (Foody, 2002), thus we used non-site-specific accuracy assessment based on the comparison of computed and reference surface areas (Congalton and Green, 1999). The reference was provided by aerial photographs

Table 1

Cross-table between clusters of global model I and local model II on Maître, showing the good agreement between models on the reef flats. K1-K2 contains soft-bottom sand, K3-K4-K5 regroup seagrass and algal beds with a growing density from K3 to K5, K6-K10 are clusters of deep water with a depth growing from K6 to K10

| Model II classes |     | C1  | C2  | C3   | C4   | C5  | C6   | C7  | C8  | C9   | Σ    |
|------------------|-----|-----|-----|------|------|-----|------|-----|-----|------|------|
| Model I classes  | K1  | 57  |     |      |      |     |      |     |     |      | 57   |
|                  | K2  | 298 | 78  | 3    |      |     |      |     |     |      | 379  |
|                  | K3  | 19  | 374 | 6    |      |     |      |     |     |      | 399  |
|                  | K4  | 3   | 195 | 794  | 27   |     |      |     |     |      | 1019 |
|                  | K5  | 4   | 9   | 340  | 1340 |     |      |     |     |      | 1693 |
|                  | K6  |     |     |      | 15   | 6   |      |     |     |      | 21   |
|                  | K7  |     |     |      | 10   | 391 |      |     |     |      | 401  |
|                  | K8  |     |     | 3    | 87   | 366 | 564  | 40  |     |      | 1060 |
|                  | K9  |     |     |      |      |     | 1057 | 178 |     | 342  | 1577 |
|                  | K10 |     |     |      |      |     | 3    | 486 | 740 | 1393 | 2622 |
|                  | Σ   | 381 | 656 | 1146 | 1479 | 763 | 1624 | 704 | 740 | 1735 | 9228 |



Fig. 5. Location of classes C1 (clear bottom sand), C2 (very diffuse seagrass), C3 (diffuse seagrass) and C4 (dense seagrass or algal beds) obtained with model II on the Maître reef flat. The accuracy of the boundaries is provided by comparison with aerial photograph and ground-truthing that validate the presence of the different bottom types.

interpretation. We acknowledge that there may be a bias since the photograph interpretation may not be perfect. As a measure of accuracy, we reported the ratio between the areas correctly classified and the total area of the reef flat. Details are provided in Table 2. The agreement for model I is 0.688. The aerial photograph and model I provided very similar class boundaries (Fig. 5), though they did not tightly match in some cases. This is more likely a question of precise georeferencing in the GIS different spatial resolution, and less likely a problem of aerial photographs interpretation. In such conditions, a precision of 0.688 is satisfactory.

Applied to the Aboré reef, model I revealed spatial patterns consistent with the structure of the barrier reefs (Fig. 7). Ground-truthing revealed that Aboré is covered by coral communities dominated by branching and plating Acroporids on a pavement floor. Coral cover differed greatly from place to place (15–80%) but the general trend was an increase in coral cover from the crest to the internal spur and grooves, where the density, size and variety of growth forms

of coral colonies are very high (cover >80%). Fleshy algae were scarce on the coral-flat, but encrusting coralline algae cover were high. Cyanobacteria also frequently covered the sedimentary zone in several places. Spatial arrangement between the classes rendered well the geomorphological succession from the sedimentary internal slope between 3 and 7 m (mostly K3), the shallow (<2 m) coral-rich reef flats (K4 and K5), and outer slope (down to 7–10 m).

The Signal reef flat was characterized by a reef flat with live and dead coral and algal communities. (A southwestern reef flat dominated by live (cover >50%, essentially branching acroporids), a western flat made of a dead coral framework, and a center and eastern flat dominated by sediments, algal communities (*Padina* spp., *Acantophora spicifera*, *Actinotrichia* sp., *Galaxaura* sp., cyanobacteria) and sparse seagrass beds (*Halodule uninervis*, *Cymodocea rotundata* and *C. serrulata*). Applied to this site, model I provided good results and zonations (not presented), though, as in the Maître reef flat, the rubble ridge did not appear separated.

Table 2

Model I evaluation based on cross-tabulated areas. Classes of the model are vectorized and intersected with polygons coming from the visual interpretation of the Maître aerial photograph. Areas are in hm<sup>2</sup> (10,000 m<sup>2</sup>). vdS: very diffuse seagrass; dS: diffuse seagrass; DS: dense seagrass and algal bed; VDS: very dense seagrass and algal bed

| Model I classes |      | sand | vdS  | dS    | DS    | VDS | Σ     |
|-----------------|------|------|------|-------|-------|-----|-------|
| Photograph      | sand | 33   | 16   | 10    | 4     | 0   | 63    |
|                 | vdS  | 4    | 45   | 31    | 15    | 0   | 95    |
|                 | dS   | 1.5  | 5    | 105   | 52    | 0.7 | 164.2 |
|                 | DS   | 2    | 4    | 19    | 188   | 2   | 215   |
|                 | VDS  | 1.8  | 0.6  | 0.3   | 1.8   | 5   | 9.5   |
|                 | Σ    | 42.3 | 70.6 | 165.3 | 260.8 | 7.7 |       |

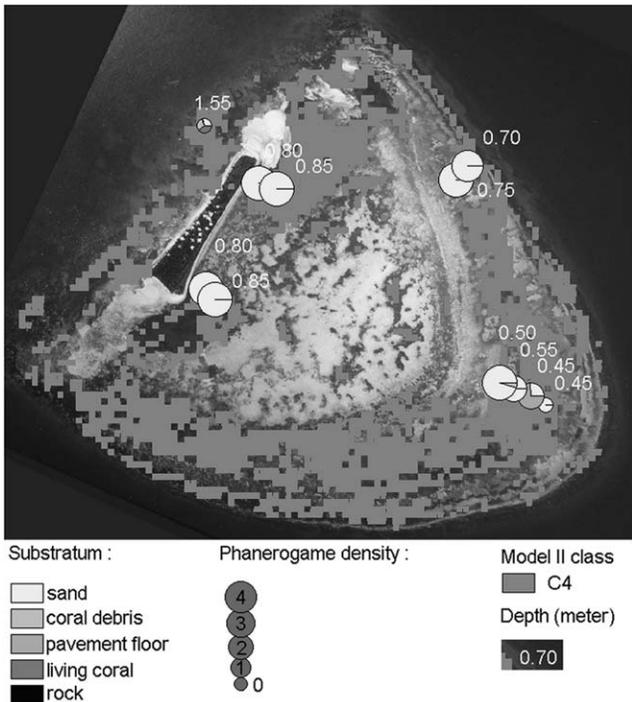


Fig. 6. Overlays of ground-truth data, aerial photograph and class C4 issued from model II. Macrophytes density is expressed on a 5-level qualitative scale: a density of "0" corresponds to a bottom without macrophyte; a density of 4 corresponds to a bottom entirely masked by plants. The density computed with GIS tools in class C4 is about 3 and corresponds to the class of dense seagrass and algal beds.

K7, K5, K4/K3 and K2/K1 corresponded principally to zones with high live coral cover, dead coral zones, algal zones and sandy zones, respectively. K3 and K4 reflected different densities of algal coverage.

#### 4. Discussion

Overall, spatial zonations were correct on the three sites (Maître, Aboré and Signal). However, the same class could include different bottom types. For instance, the class K4 included large sections of the coral-rich reef flat in Aboré and moderately dense seagrass beds in Maître. Merging such different bottom types in one class is clearly not acceptable. This confusion is of course explained by the unsupervised

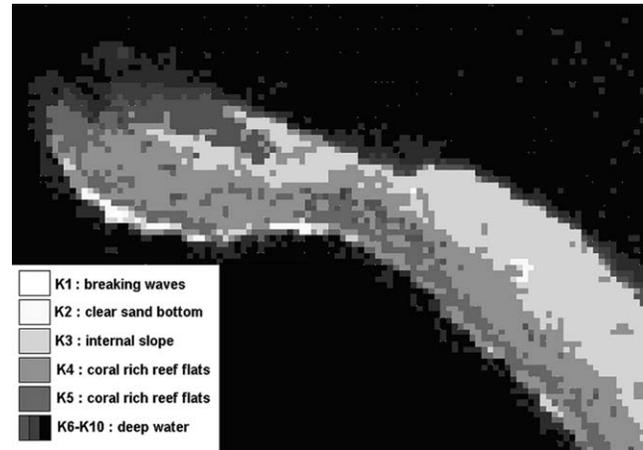


Fig. 7. Result of the classification of the Aboré Reef site with global model I: spatial arrangement between the classes reflected the geomorphological succession from the sedimentary internal slope in the range 7-3 m (mostly K3), the shallow (<2 m) coral-rich reef flats (K4 and K5), and outer slope (till 7-10 m).

approach applied to build a model including all reef formations in the Southwest lagoon, by the similar optical signature of these bottom types, and by the range of depth in each site. The example of Maître suggests that locally optimized model (model II) may not perform significantly better than the general model (model I). But the confusion revealed by multisites comparison points to the strict obligation to interpret the results locally with the help of contextual knowledge (Groom et al., 1996).

In several cases, the confusion between two different bottom types merged in the same class can be explained at the level of the segments. For instance, the class K3 in model I was made of two segments. One of the segments (no. 2, black in Fig. 8) collected pixels from the rubble/sand zones of the Aboré reef, while the other segment (no. 1, white in Fig. 8) collected pixels from shallow diffuse seagrass in the Maître and Signal flats (Fig. 8). Table 3 presents the decision rules of these two segments. Segment 1 represents 89% of K3 on the Maître reef flat, less than 71% on the Signal reef flat and only 2% on the Aboré reef. This shows the interest of working by segments. They were created to allow for some flexibility in the type of cover within one class (e.g. diffuse seagrass and dense seagrass), but in case of true thematic confusion within

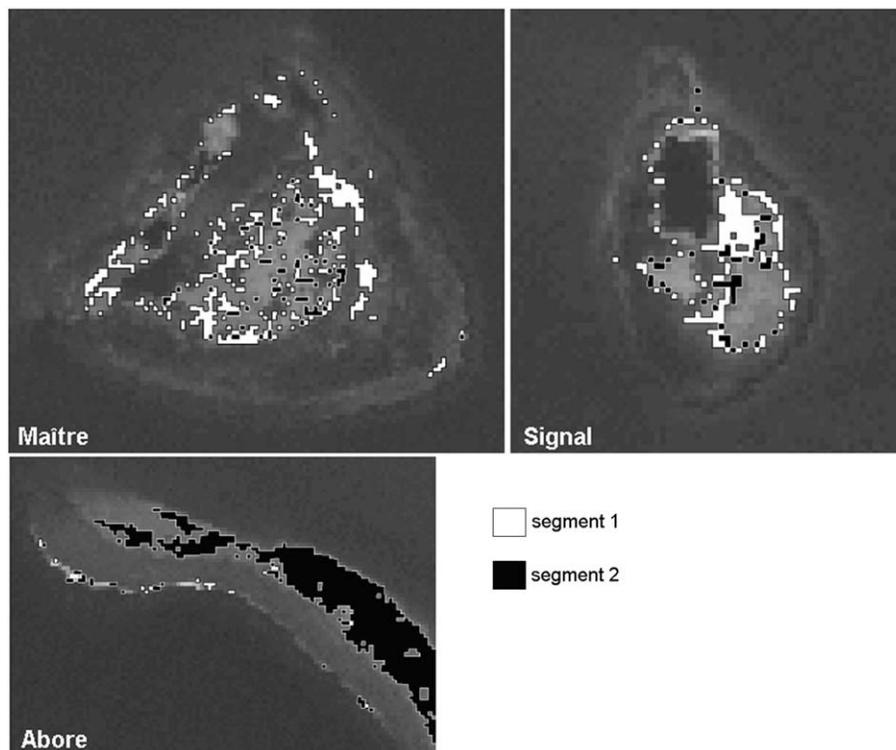


Fig. 8. Location of segments no. 1 (white) and no. 2 (black) from class K3, on Maître, Aboré and Signal. The segment no. 2 collected pixels from the rubble/sand zones of the Aboré reef, while the segment no. 1 collected pixels from shallow diffuse seagrass in the Maître and Signal flats.

a class (e.g. coral slope and dense seagrass), it is then possible to reallocate a misclassified segment to the correct class.

Incorrect segments are easily identified if there is a topological problem (e.g. dense seagrass beds are generally inside shallow flats, but appear frequently in confusion with outer slopes, outside the barrier reef). It seems then logical to introduce in the processing, contextual decision rules based on spatial information between segments and between classes. For instance, an a priori discrimination between barrier reef flats and cay flats is very possible. However, integration of spatial knowledge in classification schemes in a rigorous formal way (and not just by masking out zones of potential confusion by photo-interpretation) is a complex task that requires tools derived from artificial intelligence. For instance, Andréfouët et al. (2000) used fuzzy membership functions to integrate the distance to the shore into a spectral classifier. This was a simple contextual rule. Description of more complex interclass topological relationships (inside, outside, between, among) still requires many conceptual developments to define the terms, grammar and

fuzzy logic algorithms necessary to tackle these objectives (Suzuki et al., 2001).

Our classification is initially unsupervised. We have shown that this unsupervised approach provides good results to determine regions optically and thematically homogeneous within one reef, but it does not reflect the real complexity since we obtain classes that include very different bottom types encountered on different reefs. Because of the spectral confusion between different bottom types on different reef formations, the unsupervised model needs to evolve into a supervised model. The general model I in 11 classes can benefit from field information to evolve into two or three different optimized models in 5–8 classes well labeled and described for each reef formation. This is a very flexible process that will be pursued in the future when new ground-truth data will be available for both training and accuracy assessment. Since we are at the stage of assessing an unsupervised scheme, we lack here final confusion matrices to compare with results achieved by various authors in other regions of the world. According to these studies, for moderately complex classification schemes (4–8 classes) the overall

Table 3  
Decision rules for segments no. 1 and no. 2 of the model I, class K3

| Band | Segment no. 1 |     | Segment no. 2 |     |
|------|---------------|-----|---------------|-----|
|      | Min           | Max | Min           | Max |
| 1    | 101           | 129 | 130           | 203 |
| 2    | 131           | 151 | 97            | 151 |
| 3    | 0             | 75  | 0             | 75  |
| 4    | 0             | 110 | 0             | 110 |

accuracy is around 65–75% using Landsat TM or SPOT data. The only exceptions are the SPOT classifications of Guadeloupe and Martinique (Chauvaud et al., 1998, Chauvaud et al., 2001). On Guadeloupe, these authors reach 95.7% overall accuracy for a classification scheme including up to 32 classes. This unusual high value, obtained for a highly heterogeneous coastal system, can be explained by several factors that optimistically biased the results (e.g. training sites were also used to control the classification) (Hammond and Verbyla, 1996). Qualitatively, we report confusions between benthic classes (e.g. class K4) similar to those encountered in the Caribbean and South Pacific reefs (Mumby et al., 1997a; Capolsini et al., 2003). Final maps (from a supervised classification) should provide a 65–75% overall accuracy on New Caledonia reefs.

The most interesting results achieved here are obtained for mapping and characterization of seagrass areas. The ETM+ classification results on the Maître reef flat, overlaid on the aerial photographs, demonstrate an excellent agreement in the spatial patterns between data acquired at very different spatial resolution (Figs. 5 and 6). Other authors have also presented good results in mapping tropical seagrass beds with various satellite and airborne sensors (Mumby et al., 1997b). Our results strongly suggest that monitoring shallow seagrass beds using remote sensing data is feasible in New Caledonia. Time-series of SPOT images on Ile des Pins (data not shown) show that seagrass beds may be highly evolutive in the shallow New Caledonian waters and a remote sensing approach would likely be a cost-effective management tool. This is strengthened by the fact that New Caledonia benefits from regular coverage by the long-term acquisition plan of the Landsat 7 program (Arvidson et al., 2001) with four to six cloud-free images planned every year. Conversely, our results suggest that for coral and heterogeneous areas, this is not true. The accuracy achieved with Landsat data in mapping coral zones is not sufficient to detect changes in coral cover for instance, a critical information to assess coral reef health. However, Landsat data are adequate to identify and estimate the areas covered by coral-rich reef flats (Aboré and Signal for instance). These basic information are still not available for New Caledonia (Labrosse et al., 2000). Eventually, ETM+ data should identify areas of prime interest that deserve in situ investigations or higher spectral and spatial resolution acquisitions.

The work conducted here was performed without any type of correction applied to the image. This is acceptable for a work conducted on one single image. However, if one wishes to apply the model I decision rules to other areas on other images, the models should be recomputed and calibrated in remote sensing reflectance, after calibration of the images from DN to radiance, normalization of the radiance by solar irradiance, and correction of the atmospheric effects. If not the decision rules would be ineffective. Variations in water depth due to tide variations are also likely and will require some corrections. Because accurate atmospheric and water quality corrections are difficult to conduct accurately in

coastal environments (Hu et al., 2001), it may be possible to apply instead empirical techniques to normalize different images, using stable uniform areas such as sandy bottoms. This technique is very forgiving and provides change detection results similar in quality than more sophisticated methods requiring absolute corrections (Andréfouët et al., 2001).

## 5. Conclusion

This study provides the foundation to design an effective tool to assess shallow coastal ecosystems of New Caledonian waters using moderate resolution remote sensing data in general and Landsat ETM+ data in particular. This is the first time that Landsat data (ETM+ on-board the Landsat 7 satellite, or Thematic mapper, TM, on-board the Landsat 5 satellite) are used to map marine benthic habitats in New Caledonia. It is also one of the first studies that uses Landsat ETM+ data for a large-scale assessment of shallow marine communities in coral reef environments. Exploration of the variability of the large Southwest lagoon implied an unsupervised approach that proved to be efficient to recognize the location of the main biotopes. The unsupervised classification scheme presented here is original, since it is based on a consecutive partitioning/segmentation that allows for a better tracking of the confusion between different biota types. However, this process requires local validation and future optimization to avoid misclassification and define accurately the different bottom types. The first results described here clearly highlight the strong potential for monitoring shallow seagrass beds and the potential to map the geomorphology of the poorly known but extensive Caledonian barrier reef system.

## Acknowledgements

This study was supported by the Programme National Environment Côtier and by NASA's grant NAG5-10908 to SA and by a NASA Carbon Cycle Science and Applications Grant to Julie Robinson and Kamlesh Lulla, Johnson Space Center. We thank the Latical (IRD Nouméa) and Pascal Dumas for providing hardware and software. The Direction des Ressources Naturelles of Province Sud provided the aerial photographs and logistics for field work. Dr Claude Payri, Université de la Polynésie Française is acknowledged for courageously identifying the flora taxa despite the presence of painful stingers. This is IMaRS contribution 042.

## References

- Andréfouët, S., Payri, C., MaoChe, L., Hochberg, E.J., Atkinson, M.J., 2003. Airborne hyperspectral detection of microbial mats pigmentation in Rangiroa atoll (French Polynesia). *Limnol. Oceanogr.* 48 (1), 426–430.

- Andréfouët, S., Roux, L., Chancerelle, Y., Bonneville, A., 2000. A fuzzy possibilistic scheme of study for objects with indeterminate boundaries: application to french polynesian reefs. *IEEE Trans. Geosci. Remote Sens.* 38 (1), 257–270.
- Andréfouët, S., Muller-Karger, F., Hochberg, E., Hu, C., Carder, K., 2001. Change detection in shallow coral reef environments using Landsat 7 ETM+ data. *Remote Sens. Environ* 79, 150–162.
- Arvidson, T., Gasch, J., Goward, S.N., 2001. Landsat 7's long term acquisition plan—an innovative approach to building a global imagery archive. *Remote Sens Environ.* 78 (1-2), 13–26.
- Bour, W., Dupont, S., Joannot, P., 1993. Establishing a SPOT thematic neo-channel for the study of hard-of-access lagoon environments. Example of application on the growth areas of the New Caledonian reefs. *Geocarto Int.* 11 (1), 29–39.
- Bour, W., Loubersac, L., Rual, P., 1986. Thematic mapping of reefs by processing of simulated SPOT satellite data: application to the *Trochus niloticus* biotope on Tetembia Reef (New Caledonia). *Mar. Ecol. Prog. Ser.* 34, 243–249.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. Wadsworth, Belmont.
- Capolsini, P., Andréfouët, S., Rion, C., Payri, C., 2003. A comparison of Landsat ETM+, SPOT HRV, Ikonos, ASTER and airborne MASTER data for coral reef habitat mapping in South Pacific islands. *Can. J. Remote Sens.* 29 (2) (in press).
- Chauvaud, S., Bouchon, C., Manière, R., 1998. Remote sensing techniques adapted to high resolution mapping of tropical coastal marine ecosystems (coral reefs, seagrass beds and mangrove). *Int. J. Remote Sens.* 19 (18), 3625–3639.
- Chauvaud, S., Bouchon, C., Manière, R., 2001. Cartographie des biocénoses marines de Guadeloupe à partir des données SPOT (récifs, coralliens, phanérogames marines, mangroves). *Oceanol. Acta* 24, S3–S16.
- Chevillon, C., 2001. Caractérisation des types de fonds et habitats benthiques par classification hydro-acoustique dans le lagon sud-ouest de Nouvelle-Calédonie. Rapport d'activité Convention ZoNéCo-IRD.
- Congalton, R., Green, K., 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Lewis-Publishers, New York, 137 p.
- De Vel, O., Bour, W., 1990. The structural and thematic mapping of coral reefs using high resolution SPOT data: application to the Tetembia Reef (New Caledonia). *Geocarto Int* 2, 27–34.
- Diday, E., 1971. La méthode des Nuées Dynamiques. In: Diday, E. (Ed.), *Rev. Statist. Appl.* 19 (2), 19–34.
- Douillet, P., Ouillon, S., Cordier, E., 2001. A numerical model for fine suspended sediment transport in the Southwest Lagoon of New Caledonia. *Coral Reefs* 20, 361–372.
- Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote Sens. Environ* 80, 185–201.
- Garrigue, C., 1995. Macrophyte associations on the soft-bottoms of the Southwest Lagoon of New-Caledonia: description, structure and biomass. *Bot. Mar* 38, 481–492.
- Green, E.P., Mumby, P.J., Edwards, A.J., Clark, C.D., 2000. *Remote sensing handbook for tropical coastal management*. UNESCO, Paris, 316 p.
- Groom, G.B., Fuller, R.M., Jones, A.R., 1996. Contextual correction techniques for improving land cover mapping from remotely sensed data. *Int. J. Remote Sens* 17, 69–89.
- Gustafson, E.E., Kessel, W.C., 1979. Fuzzy clustering with a fuzzy covariance matrix. *Proc. IEEE Conf. Decision and Control*, San Diego, CA, IEEE Press, Pictaway, NJ, 761–766.
- Hammond, T.O., Verbyla, D.L., 1996. Optimistic bias in classification accuracy assessment. *Int. J. Remote Sens.* 17 (6), 1261–1266.
- Hochberg, E., Atkinson, M., 2000. Spectral discrimination of coral reef benthic communities. *Coral Reefs* 19, 164–171.
- Hu, C., Muller-Karger, F., Andréfouët, S., Carder, K., 2001. Atmospheric correction and cross-calibration of LANDSAT-7/ETM+ imagery over aquatic environments: multi-platform approach using SeaWiFS/MODIS. *Remote Sens. Environ.* 78 (1-2), 99–107.
- Jensen, J.R., 1996. *Introductory digital image processing: a remote sensing perspective*, second ed. Geogr Inf. Sci., Prentice Hall, Upper Saddle River, NJ, 316 p.
- Jupp, D.L.B., Mayo, K.K., Kuchler, D.A., Claasen, D., Kenchington, R., Guerin, P., 1985. Remote sensing for planning and managing the Great barrier reef of Australia. *Photogrammetria* 40, 21–42.
- Kleypas, J.A., Buddemeier, R.W., Gattuso, J.P., 2001. The future of coral reefs in an age of global change. *Int. J. Earth Sci* 90, 426–437.
- Labrosse, P., Fichez, R., Farman, R., Adams, T., 2000. New Caledonia. In: *Seas at the Millenium: An Environmental Evaluation*. In: Sheppard, C. (Ed.), 2. Elsevier Science, Amsterdam, pp. 723–736.
- Lebart, L., Morineau, A., Warwick, K.M., 1984. *Multivariate Descriptive Statistical Analysis*. In: Lebart, L., Morineau, A., Warwick, K.M. (Eds.). John Wiley & Sons, New York, 231 p.
- Menges, C.H., Hill, G.J.E., Ahmad, W., 1998. Landsat TM data and potential feeding grounds for threatened marine species turtle in northern Australia. *Int. J. Remote Sens.* 19 (6), 1207–1221.
- Mumby, P.J., Green, E.P., Clark, C.D., Edwards, A.J., 1997a. Coral reef habitat mapping: how much detail can remote sensing provide? *Mar. Biol.* 130 (1), 193–202.
- Mumby, P.J., Green, E.P., Edwards, A.J., Clark, C.D., 1997b. Measurement of seagrass standing crop using satellite and digital airborne remote sensing. *Mar. Ecol. Prog. Ser.* 159 (2), 51–60.
- Mumby, P.J., Harborne, A.R., 1999. Development of a systematic classification scheme of marine habitats to facilitate regional management and mapping of Caribbean coral reefs. *Biol. Conserv* 88, 155–163.
- Rossier, O., Kulbicki, M., 2000. A comparison of fish assemblages from two types of algae beds and coral reefs in the Southwest Lagoon of New Caledonia. *Cybiurn* 24 (1), 3–26.
- Suzuki, H., Matsakis, P., Andréfouët, S., Desachy, J., 2001. Satellite image classification using expert structural knowledge: a method based on fuzzy partition computation and simulated annealing. Annual Conference of the International Association for Mathematical Geology. IAMG, Cancun, Mexico CDROM.
- Thomassin, B., 1985. Cartographie thématique des récifs coralliens par télédétection (simulation des données du satellite SPOT) : problèmes méthodologiques en relation avec les stades évolutifs des platiers. Rapport ATP CNRS Télédétection 1985/AIP (06931) 1196.
- Thome, K., 2001. Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method. *Remote Sens. Environ.* 78 (1-2), 27–38.
- Wantiez, L., Thollot, P., Kulbicki, M., 1997. Effects of marine reserve on the communities of five islands in the lagoon of New Caledonia. *Coral Reefs* 16, 215–224.
- Ward, J.H., 1963. Hierarchical grouping to optimise an objective function. *J. Amer. Stat. Assoc.* 58 (301), 236–244.
- Wilkinson, C., 2000. Status of the coral reefs of the world. *Aust. Inst. Mar. Sci.*, Townsville, 363 p.