

# A SVM-based model for predicting distribution of the invasive tree *Miconia calvenscens* in tropical rainforests

Robin Pouteau<sup>a,\*</sup>, Jean-Yves Meyer<sup>b</sup>, Benoît Stoll<sup>a</sup>

<sup>a</sup> Laboratoire Géosciences du Pacifique Sud, Université de Polynésie française, BP 6570, Faa'a, French Polynesia

<sup>b</sup> Délégation à la Recherche, Gouvernement de Polynésie française, BP 20981, Papeete, French Polynesia

## ARTICLE INFO

### Article history:

Received 5 October 2010

Received in revised form 12 March 2011

Accepted 30 April 2011

Available online 9 June 2011

### Keywords:

Distribution model

Digital elevation model (DEM)

Support vector machines (SVM)

Vegetation mapping

Invasive species

Rainforest

## ABSTRACT

Advances in spatial and spectral resolution of sensors can be useless to detect directly understory and subcanopy species but make remote sensing of some ecosystems components increasingly feasible. We propose to use support vector machines (SVM) to integrate multisource-derived biophysical descriptors (overstory plant species, physiography and climate) for the indirect detection of the small invasive tree *Miconia calvenscens* in tropical rainforests on the island of Tahiti (South Pacific). Our model consists in classifying overstory plant species from an optical Quickbird scene, with the output then used in a subsequent fusion process with digital elevation model (DEM) extracted physiographic and climatic descriptors. A range of accuracy metrics was calculated to assess the SVM-based model which widely outperforms the commonly used GARP model. Biophysical descriptors alone are necessary and sufficient to explain *M. calvenscens* distribution and suggest that the potential invaded area is currently saturated in our study site on Tahiti. Rainfall, elevation and slope steepness are the major variables explaining the species distribution. In addition, our results show that morning insolation plays a critical role on *M. calvenscens* height whether it is restricted to the subcanopy or reaches the forest surface. The model can be used to map the potential distribution of *M. calvenscens* in areas where it has been recently introduced and rapidly spreading, such as in the Hawaiian islands, New Caledonia and Australia or in other French Polynesian islands. It also may be adapted to detect other species (plants or animals, alien invasives or rare endemics) in the understory and subcanopy of forest ecosystems.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Invasive species are recognized as a primary cause of global biodiversity loss and an agent of global change (Millennium Ecosystem Assessment, 2005). Because of their isolation and small areas, island ecosystems are particularly vulnerable to biological invasions (Loope and Mueller-Dombois, 1989; Paulay, 1994; Kueffer et al., 2010). As reported by Atherton et al. (2007), invasive alien species represent the first threat to the insular native biota in the “Pacific Ecoregion”. They underlined the need for “more information on the distribution and impact of invasive species in sensitive sites and the identification of alien-free habitats” as their first strategic direction.

Advances in the spatial and spectral resolutions of sensors now allow a direct detection at species or community levels (Xie et al., 2008). As an example, Asner et al. (2008a) mapped a selected group of native and invasive species in a Hawaiian rainforest using

hyperspectral Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data, identifying species by their differences in canopy spectral signature. Asner et al. (2008b) also mapped the location and impacts of five invasive plant species in Hawaii based on Light Detection And Ranging (LiDAR) imagery by identifying 3D transformation of the structure of native rainforests. Walsh et al. (2008) studied the positive synergism between Quickbird and Hyperion data at the pixel- and object-levels to characterize the location, density and spatial structure of *Psidium guajava* (Myrtaceae), a small alien invasive tree in the Galápagos islands.

As stated in Joshi et al. (2004), an increasing number of studies deal with the use of remote sensing for mapping invasive species using various remotely sensed data and remote sensing techniques from global to local scales. They divided these past studies into four classes: (i) applications in detecting canopy dominating species; (ii) mixed canopy dominant species; (iii) invaders influencing canopy dominant species; and (iv) understory and subcanopy species.

While many plant invaders do not reach the canopy surface, the fourth class is the least represented in terms of case studies, because understory and subcanopy species are inherently more difficult to detect and there is a need for extensive ground truth surveys. Joshi et al. (2006) mapped cover percentage and potential areas

\* Corresponding author. Tel.: +689 866 438; fax: +689 803 842.

E-mail address: [robin.pouteau@upf.pf](mailto:robin.pouteau@upf.pf) (R. Pouteau).

of *Chromolaena odorata* (Asteraceae), an herbaceous weed commonly found in a range of habitats, including rainforest understory. They extracted six ancillary variables from a Landsat ETM+ image and integrated using an artificial neural network: grazing intensity, distance from forest ecotone, distance from the road, understory biomass, forest canopy density and light intensity reaching the understory.

All fourth class studies need to use predictive mapping which involves making correspond meaningful environmental layers, known a priori to be a proxy of the presence/absence of the tracked species, with known localities of the tracked species for extrapolation to the entire study site. Most of the species predictive mapping approaches are reviewed in Franklin (1995). Principal integration methods include rule-based models, generalized linear models, classification (and regression) trees and more recently machine learning algorithms such as artificial neural networks (Linderman et al., 2004; Joshi et al., 2006), genetic algorithms (Peterson et al., 2003; Roura-Pascual et al., 2004; Loo et al., 2007; LaRosa et al., 2007; Sánchez-Flores, 2007; Ebeling et al., 2008; Roura-Pascual et al., 2008) and support vector machines (Guo et al., 2005; Drake et al., 2006). Input data are usually ancillary data only but some studies also couple ancillary data with remotely sensed imagery (Strahler, 1981; Hutchinson, 1982; Linderman et al., 2004; Joshi et al., 2006).

Before the present study, several attempts to map the distribution of *Miconia calvescens* (Melastomataceae) were conducted in the Society islands of French Polynesia (Tahiti and Moorea) and in the Hawaiian islands where the species is also invasive.

In 1990, the Station Polynésienne de Télédétection (Polynesian Remote Sensing Station) made the very first attempt to map *M. calvescens* comparing its spectral properties on two SPOT scenes with spectro-radiometric data taken by helicopter and in the field (Wibaux, 1990). In 1997, a collaboration between the University of California at Berkeley's Richard B. Gump Research Station and the Ministry of the Environment of the Government of French Polynesia attempts to map *M. calvescens* using aerial photographs taken during helicopter surveys over Tahiti and Moorea (Krantz, 1997). These two studies identified limitations of aerial imagery when *M. calvescens* was confined to the understory or the subcanopy. Indeed, *M. calvescens* is a shade-tolerant small tree (<15 m tall) that regenerates in rainforest understory and grows more rapidly in forest gaps and open areas, becoming dominant in canopy where it shades out other plants, except the higher trees (>15–20 m) which density determine whether it is directly discernible or not.

In 1993, Florence (1993) delimits the distribution of *M. calvescens* mainly based on personal field observations as *M. calvescens* was indiscernible on aerial photograph when located in the understory and subcanopy of rainforest (Florence, personal communication cited in Meyer, 1996). Despite the temporal bias between today and 1993, that map is still satisfactory at the considered scale which suggests that the stretch of the invasion has already reached its maximum. This saturation hypothesis is important to produce an actual distribution map and not a potential distribution one. The limit of Florence's study is the coarse spatial resolution which is not relevant for an effective in situ monitoring effort.

The most recent study was conducted by the "Big Island Invasive Species Committee" in the Hawaiian islands (LaRosa et al., 2007). The niche-distribution model Genetic Algorithm for Rule Set Production (GARP) (Stockwell and Peters, 1999) was used to project potential habitat of *M. calvescens* using GIS-based information, namely elevation, precipitation, slope steepness and windwardness and the ecological range of *M. calvescens* on Tahiti from Meyer (1998). Unfortunately, the accuracy of this method was not really assessed.

With the increasing availability and accuracy of remotely sensed data and the concomitant development of predictive mapping

methods, it is now possible to precisely describe invasive species distribution, a crucial step to improve policy, legislation and procedures implemented against them, and also to manage landscapes and natural risks, improve natural resources monitoring and maintain sustainable development.

The aim of this study is twofold: (i) to introduce a set of biophysical predictors in relation with the ecology of *M. calvescens*; and (ii) to compare our SVM-based model with the widely used GARP model for *M. calvescens* predictive mapping.

## 2. Material and methods

### 2.1. Study area and species

Tahiti, located 17°38' South and 149°30' West in the South Pacific Ocean, is a high volcanic tropical oceanic island, 0.3–1.7 Myr old (Duncan et al., 1994). It is the largest (1045 km<sup>2</sup>), highest (2241 m elevation) and most populated (ca. 180,000 inhabitants in 2007) island of French Polynesia. It has also the richest vascular flora with 467 native plant species, 212 (ca. 45%) of which are endemic (Meyer and Florence, 1996).

As in most of the small Pacific islands, the native forests of French Polynesia, and Tahiti in particular, have been invaded by alien plants, accidentally or purposely introduced by humans (Meyer, 2004). A total of 593 plant species is currently naturalized, with 57 considered highly invasive (Fourdrigniez and Meyer, 2008). *M. calvescens* is probably the worst plant invader on Tahiti, being also one of the world's 100 worst invasive species (ISSG, 2010). This small tree, commonly 4–12 m tall (up to 15 m), is native to rainforests of Central and South America. It was introduced in Tahiti in 1937 as an ornamental plant because of its large leaves with purple undersides and has invaded about two-thirds of the forested area on the island (between 70,000 and 80,000 ha) and became the dominant canopy tree over large areas (Meyer, 1998). This spectacular invasion is probably due to the combination of its rapid vegetative growth and its reproductive competitiveness (early sexual maturity, self pollination, trimodally reproduction each year, active dispersal over long distances, high rate of germination, large size and persistence of the soil seed bank) (Meyer et al., 2008).

Our study site is the Papenoo valley, the largest valley in Tahiti and French Polynesia (18 km long, covering about 9000 ha) ranging from sea level to the Mont Orohena (2241 m), the highest peak of French Polynesia and the South Pacific. It was selected because of its accessibility (by road) and its ecological importance. Indeed, the Papenoo valley includes the only protected natural area in Tahiti, "Te Faaiti Natural Park", a 750 ha rainforest (Meyer, 2007), as well as the unique Anaor'i'i swampy mid-elevation plateau, an extremely rare habitat in French Polynesia (Meyer, 2009). Both areas are partially invaded by *M. calvescens*.

Zonal vegetation types occurring in the Papenoo valley are a function of a climatic gradient (precipitation and temperature) controlled by the topography. Vegetation of the study site can be divided into three types, according to their preferential ecological conditions (Papy, 1951–1954; Florence, 1993; Florence and Lorence, 1997; Meyer, 2007):

- a low- to mid-elevation mesic vegetation where precipitation ranges from 1500 to 3000 mm per year and elevation is lower than 900 m. This vegetation type is commonly found on steep slopes and ridge crests where native *Dicranopteris linearis* fernlands and stands of the large alien tree *Falcataia moluccana* occur;
- a low- to mid-elevation wet forest for which precipitation exceeds 3000 mm per year and elevation is lower than 900 m, found in valleys and dominated by the small native tree *Hibiscus tiliaceus*, the large alien tree *Spathodea campanulata*, and the Polynesian introduced bamboo *Schizostachyum glaucifolium*;

- and a montane cloud forest which is found on the upper slopes (up to 900 m) where temperatures are cool, precipitation is above 3000 mm per year, with almost persistent diurnal cloud cover. About 60% of the endemic vascular plant species of French Polynesia are found in cloud forests and 25% of these endemics are confined to this habitat type (Meyer, 2010).

## 2.2. Remote sensing methods

### 2.2.1. Ancillary biophysical variables

LaRosa et al. (2007) propose four data fields to map *M. calvescens* in Hawaiian islands. We keep them except windwardness since, working at an entire island-scale, windwardness appears pertinent but not at a single valley-scale:

elevation (id. the DEM) affecting temperature. Considering an environmental lapse rate of  $0.0058^{\circ}\text{C m}^{-1}$  as observed in Hawaii (Baruch and Goldstein, 1999), there is a shift of  $13^{\circ}\text{C}$  between sea-level and the highest summit of Tahiti. Air temperature is one of the most important factors controlling vegetation zonation and key processes such as evapotranspiration, carbon fixation and decomposition, plant productivity and mortality in mountain ecosystems (Chen et al., 1999; Nagy et al., 2003).

precipitations. Our precipitations map is an interpolation of data from a network of meteorological stations set up by Météo France in French Polynesia and the Groupement d'Etude et de Gestion du Domaine Public (Laurent et al., 2004). Precipitation refers to valley-scale pattern of water availability but excludes effect of micro-topography on water flow and concurrence at the community scale.

slope steepness can be considered as a proxy of overland and subsurface flow velocity and runoff rate, effect of micro-topography on precipitation, geomorphology, soil water content (Wilson and Gallant, 2000), mechanical effect on plant rooting and seed dispersion.

In this study, in addition to those introduced by LaRosa et al. (2007), we propose four additive biophysical variables potentially impacting the presence/absence of *M. calvescens*:

- annual potential insolation (API) calculated by the Solar Analysis tool, an ArcView GIS extension. It equates to the amount of radiative energy received over the ground (Fu and Rich, 2002);
- slope exposure as a proxy of solar insolation, evapotranspiration (Wilson and Gallant, 2000) and trade wind exposure;
- compound topographic index (CTI) at a 5% threshold, quantifying water drainage by micro-topography and explaining geomorphology (Moore et al., 1993; Gessler et al., 2000);
- and overstory plant species map, indeed forest overstory can affect temperature and resources availability (light, water, nutrients) by competition (Riegel et al., 1992). Vegetation map is obtained using a support vector machine classification.

Redundancy between biophysical variables is inevitable, above all between primary physiographic attributes (elevation, slope steepness, slope exposure) and secondary physiographic attributes (annual potential insolation and compound topographic index) that are computed from primary physiographic attributes (Moore et al., 1991). In consequence, correlation between biophysical variables is tested to avoid overweighting problem. Elevation, slope steepness, API, slope exposure and CTI are extracted from a 5 m-resolution digital elevation model (DEM) projected in the system WGS 84 UTM 6 South.

### 2.3. Genetic algorithm for rule-set production (GARP)

The niche-distribution model GARP used in LaRosa et al. (2007) for mapping *M. calvescens* was also successfully used for other

invasive species mapping in many studies (Peterson et al., 2003; Roura-Pascual et al., 2004; Loo et al., 2007; Sánchez-Flores, 2007; Ebeling et al., 2008; Roura-Pascual et al., 2008). GARP is a machine learning-based analytical package developed by Stockwell and Peters (1999). More details on the program are available at the URL <http://www.nhm.ku.edu/desktopgarp/index.html> where the Desktop version used in this study can also be downloaded. Due to unstable traits of the GARP algorithm producing various solutions, we choose to keep the map producing the best overall accuracy after 100 runs.

### 2.4. Support vector machines (SVM)

SVM is widely used in classification of remotely sensed data such as optical data (e.g. Foody and Mathur, 2006), synthetic aperture radar (SAR) data (e.g. Lardeux et al., 2009), hyperspectral data (e.g. Lennon et al., 2002) and multisensor data (Hill et al., 2005; Camps-Valls et al., 2006; Waske and Benediktsson, 2007). Although SVM is seldom used in the literature for biogeographic modelling (see Guo et al., 2005; Drake et al., 2006 for applications), it can be an optimal statistical tool for our study case since it is adapted to numerous and heterogeneous variables and mixtures of qualitative and quantitative variables. In Waske and Benediktsson (2007), a range of the most commonly used classifiers are compared for multisource classification, namely maximum likelihood, decision trees, boosted decision trees and support vector machines (SVM). According to their results, SVM gives the best accuracy. Artificial neural networks is another widely used classifiers which is not compared in the latter study but shows lacks face to SVM in other studies (e.g. Foody and Mathur, 2004; Pal and Mather, 2004; Nemmour and Chibani, 2006).

Fig. 1 presents the SVM, originally introduced as a binary classifier (Vapnik and Chervonenkis, 1998). SVM uses two classes of training samples within a multi-dimensional feature space to fit an optimal separating hyperplane (in each dimension, vector components is a biophysical variable). In this way, it tries to maximize the margin that is the distance between the closest training samples, or support vectors, and the hyperplane itself.

SVM consists in projecting vectors into a high dimension feature space by means of a kernel function then fitting the optimal hyperplane that separates classes using an optimization function (1). For a generic pattern  $\mathbf{x}$ , the corresponding estimated label  $\hat{y}$  is given by (1).

$$\hat{y} = \text{sign}[f(\mathbf{x})] = \text{sign} \left( \sum_{i=1}^N a_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right) \quad (1)$$

wherein  $N$  is the number of training points, the label of the  $i$ th sample is  $y_i$ ,  $b$  is a bias parameter,  $K(\mathbf{x}_i, \mathbf{x})$  is the chosen kernel function and  $a_i$  denotes the Lagrangian multipliers.

SVM, introduced by Vapnik and Chervonenkis (1998), is extensively described by Burges (1998), Schölkopf and Smola (2002) and Hsu et al. (2009). As suggested by the latter and supported by many other papers, we use the radial basis function (RBF) as the kernel. The equation is (2).

$$K(\mathbf{x}_i, \mathbf{x}) = \exp \frac{\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2} \quad (2)$$

Noise in the data can be accounted for by defining a distance tolerating the data scattering, thus relaxing the decision constraint. This regularization parameter  $C$  as well as  $\sigma$  is found by cross-validation.

SVM have been created for two-class problem but extensions have been developed to deal with  $Q$  class problem ( $Q > 2$ ). Among



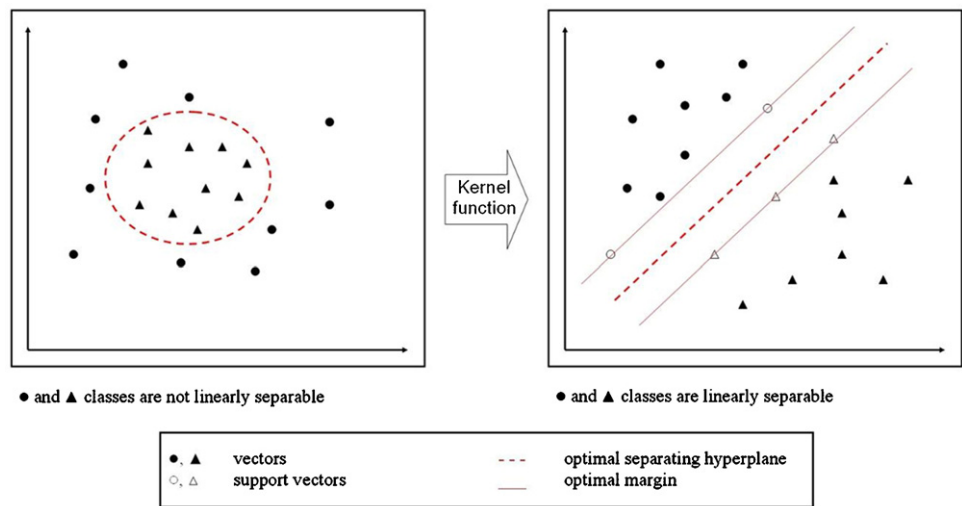


Fig. 1. SVM classifier scheme.

them, we choose the One-Against-One algorithm which consists in the construction of  $Q(Q - 1)/2$  hyperplanes which separate each pair of classes.

2.5. Proposition of a classification scheme

Hereafter, a generic classification scheme is proposed to predict species distribution in montane tropical environment (Fig. 2).

2.5.1. Overstory plant species monosource classification

We first classify overstory plant species from four spectral bands (0.45–0.90  $\mu\text{m}$ ) and 0.61 m-resolution pan-sharpened Quickbird imagery. Two pan-sharpened standard scenes, acquired on 23 June 2003 at 08h06 a.m. (local time) and on 15 June 2005 at 08h25 a.m., were concatenated and orthorectified using a 5 m-resolution DEM and ground control points. A statistical computation of the region

of interest (ROI) collected in situ is performed from the four bands and informs that no spectral signature corresponds to the presence/absence of *M. calvescens* on the Quickbird scenes (data not shown, see Fig. 3).

After preprocessing, a SVM is applied on the four spectral bands and four multi-level grey level co-occurrence matrix (GLCM) texture metrics (Haralick et al., 1973) per band. Texture analysis

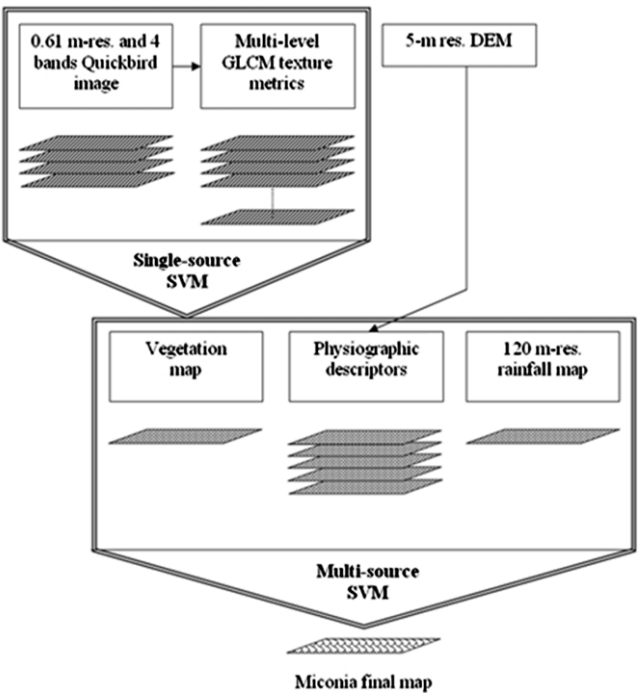


Fig. 2. Schematic diagram of the SVM-based *M. calvescens* mapping.

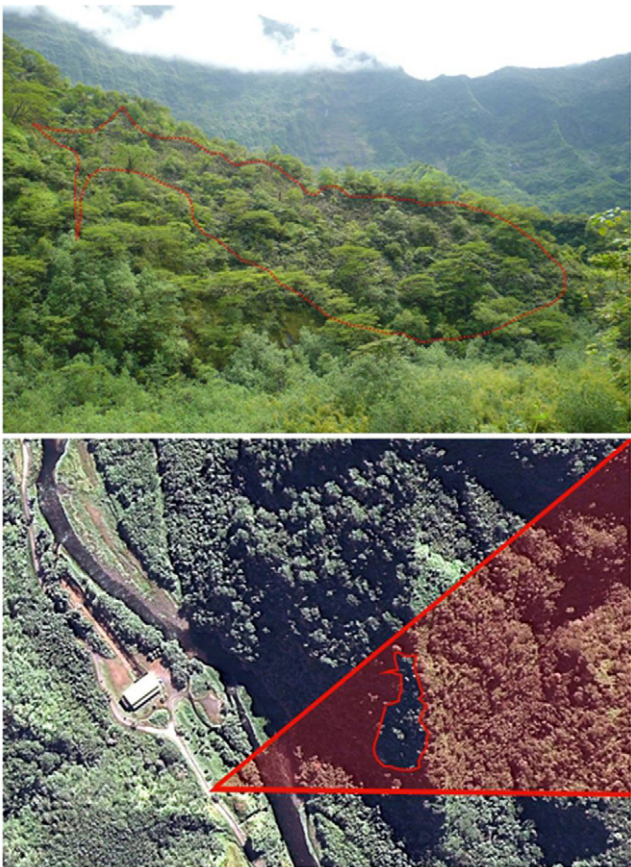
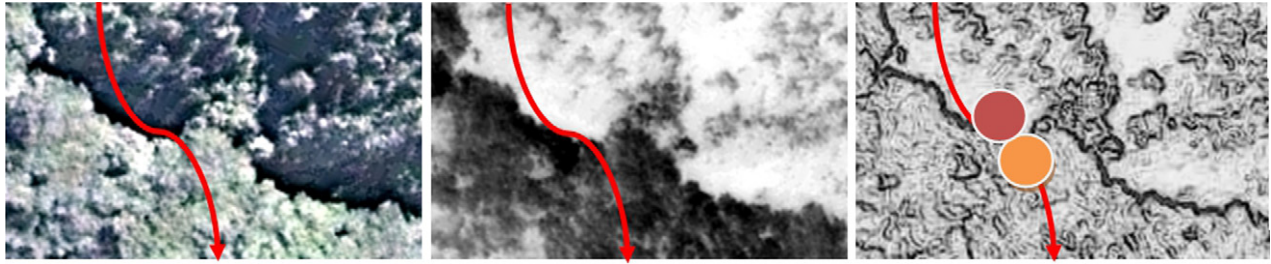


Fig. 3. Above: A stand of *M. calvescens* reaching the forest canopy in the Papenoo Valley (Tahiti) (Picture R. Pouteau). Below: The viewpoint of the above picture and localisation of the same stand invisible on the Quickbird scene because of the relief shadow.



**Fig. 4.** RGB composition of a subset of the original Quickbird image (left), principal component analysis (PCA) computed on all spectral bands and GLCM texture metrics (middle) and edge detection calculated on the PCA with two detected regions of interest (right).

assesses the spatial variability in tonal values around pixels using a panel of statistical metrics. Variance, contrast, dissimilarity and angular second moment were computed on the four bands and in the three window sizes of  $3 \times 3$ ,  $9 \times 9$  and  $15 \times 15$  pixels which visually correspond to intra-tree micro-texture, intra-tree macro-texture and inter-tree texture, respectively. Then, SVM classification is upscaled to 5 m using the nearest neighbor resampling method.

SVM fits an optimal separating hyperplane by using support vector only, other vectors eq. pixels are meaningless. Pixels containing mixed spectral response are by nature in a position to be support vectors (Foody and Mathur, 2006). In our ground truth methodology, we looked for ecotones (the transition area between two adjacent but different plant communities) as area prone to overstorey vegetation mixed spectral response (Pouteau et al., 2010). In this way, we computed a principal component analysis (PCA) on all spectral bands and GLCM texture metrics. Then we detected edges on this PCA by calculating the Sobel's algorithm since it proved to have a good accuracy/confusion trade-off (Jeon et al., 2005). Finally, line transects were drawn along the paths and, when an ecotone was detected, two circular region of interest ROI are sampled in each side to train the SVM (Fig. 4).

A ROI has to be big enough to be representative and small enough to be homogenous and to limit computational times. The optimal ROI area, also called “minimal area”, can be defined as the minimal space to achieve the development corresponding to the normal species set. In order to optimize this trade-off, phyto-sociological samples are carried out in the three vegetation types following the procedure described in Gillet (2000). Species are counted in 16 successive  $5 \text{ m} \times 5 \text{ m}$  quadrats as far as obtaining a representative floristic composition of the plant community, i.e. until the curve representing the sampled surface in function of the number of species reaches a plateau. As stated in Gillet (2000), the minimal area is the surface when 80% of the total number of species is found.

### 2.5.2. Multisource SVM fusion

As reported by Richards (2005), multisource fusion emerges in the mid to late 1970s. The principle is to take advantage of the complementary content of each sensor image in order to maximize information introduced in the integration process. Applications are essentially in geosciences, defence systems, medical imaging and robotics and industrial engineering (Piella, 2003). We can separate trend aiming to produce a synthetic image with some elements from each sensor (Piella, 2003) from trend aiming to produce a classification (with discrete value) from the full set of multisource scenes, as in our case. Fusion of remotely sensed data and DEM data occurs in the literature from precursory studies (Hoffer, 1975; Strahler, 1981; Hutchinson, 1982) to more recent ones (Bruzzone et al., 1997; Treltz and Howarth, 2000; Lawrence and Wright, 2001; Halldorsson et al., 2003). These works show that the combined use of multisource remote sensing data improves accuracy for land cover classification using different classification algorithms.

SVM is a critical tool for fusion of multisource data (Waske and Benediktsson, 2007). They compare two fusion schemes: the “single SVM” and the “fused SVM”. In the single SVM method, an overall SVM is trained on all sources together whereas in the fused SVM, a single SVM is trained on each source separately, and then an additional SVM is applied as a decision rule on all monosource rule images together. The fused SVM method outperforms the single SVM and appears to be easier to implement than the classification scheme in Hill et al. (2005) or the use of composite kernel proposed in Camps-Valls et al. (2006). Consequently, we use hereinafter the fused SVM for its accuracy/ease of implementation trade-off.

In a second step, the output classification is thus inputted in an additive fused SVM as a single variable in the same way than physical descriptors in order to map *M. calvescens*. Training and validation samples are collected in a second ground truth campaign:

- 3.5 ha of *M. calvescens*-free area (called ROI1 hereafter);
- 3.5 ha of understory *M. calvescens*-invaded area (called ROI2 hereafter);
- and 3.5 ha of area invaded by *M. calvescens* reaching the canopy surface (over 50% of the area of the canopy invaded; called ROI3 hereafter). Few patches are reaching the canopy surface although not noticeable on our aerial scenes because of the systematic presence of relief shadow in these areas.

50% of samples were devoted for training and 50% for validation. To avoid possible class over-/under-representation problems (Waske et al., 2009), balanced datasets are systematically used.

### 2.6. Classification accuracy assessment

As widespread used overall accuracy (OA) can give misleading results, other metrics we are also used hereinafter to assess classification results:

- Kappa coefficient expresses whether correctly assigned pixels may have been assigned by chance or not based on the classification decision rule (Cohen, 1960). A value of 1 indicates perfect agreement and 0 indicates a pattern arising by chance;
- area under the ROC (for receiver operating characteristic) curve (AUC) was originally introduced for binary classifications. ROC metric is often represented by a curve corresponding to corrected assigned pixels rate according to the misclassification rate. The best possible prediction method would yield a point in the upper left corner or coordinate (0,1) of the ROC space, representing 100% correctly assigned pixels (Hand and Till, 2001);
- Brier metric is a scoring rule that measure the average squared deviation between predicted probabilities for a set of event and their outcomes, so lower the Brier metric, the higher is the accuracy (Brier, 1950);
- information score (IS) takes into account the classification difficulty, so higher the IS, better is the accuracy without undeserved

score obtained by uninformed classifier that classify into a single majority class for example (Kononenko and Bratko, 1991);

- producer's accuracy (PA) represents the number of correctly assigned pixels for a class divided by the actual number of ground truth pixels for that class;
- and user's accuracy (UA) is the number of correctly assigned pixels for a class divided by the total pixels assigned to that class (Congalton and Green, 2009).

### 3. Results

#### 3.1. Assessment of minimal areas

The minimal area is more than twice larger in montane cloud forest (260 m<sup>2</sup>) than in low- to mid-elevation mesic and wet forests (both achieving 125 m<sup>2</sup>). More than a natural gradient this phenomenon may be due to the invasion of alien species transforming intensely low elevation ecosystems. Indeed, the low- to mid-elevation mesic forest counts five naturalized species, namely 18% of the total number of vascular species, the low- to mid-elevation wet forest counts six naturalized species, namely 16% while we only found two naturalized species in montane cloud forest, namely 3%.

As a result, overstory plant species classification was trained on circular ROI with a 12.60 m diameter for classes occurring in mid-elevation vegetation (*H. tiliaceus* native forest, *Spathodea campanulata* forest, *F. moluccana* forest, *S. glaucifolium* bamboo grove and *D. linearis* fernland) and with a 18.20 m diameter for the montane cloud forest class. Classification is trained on 5000 pixels and validated on 5000 other pixels for each class. Balanced datasets are also used.

#### 3.2. Classification results

Concerning vegetation classes in the overstory species vegetation map, results are fairly good since PA are from 88% to 98% and UA from 86% to 100% (Table 1).

The resulting 5 m-resolution classification of *M. calvescens* is very relevant with an OA of 0.97, a Kappa coefficient of 0.95, an AUC of 0.99, a Brier score of 0.03 and an IS of 1.88 (Table 2). Fused SVM outperforms the niche-distribution model GARP used in LaRosa et al. (2007) for mapping *M. calvescens*. GARP is more pessimistic since it notably overestimates the *M. calvescens* distribution (Fig. 5),

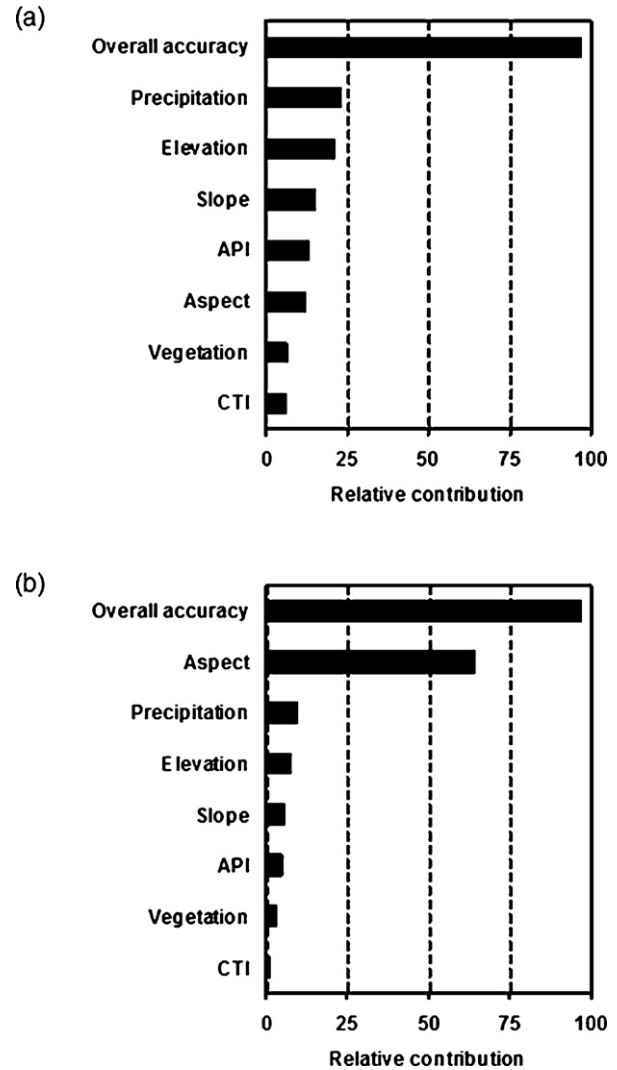
**Table 1**  
Vegetation classes status and class-specific accuracies (%) of the overstory species SVM classification.

Class	Status	PA	UA
Cloud forest (CF)	–	90	97
<i>Dicranopteris linearis</i> fernland (Dic.)	Native	98	100
<i>Schizostachyum glaucifolium</i> grove (Sch.)	Polynesian introduction	98	87
<i>Falcataria moluccana</i> forest (Fal.)	Alien invasive	88	100
<i>Spathodea campanulata</i> forest (Spa.)	Alien invasive	94	96
<i>Hibiscus tiliaceus</i> forest (Hib.)	Native	95	86

PA, producer's accuracy; UA, user's accuracy.

**Table 2**  
Accuracy metrics of *M. calvescens* classifications outputted by the two algorithms GARP and SVM. OA, overall accuracy; kappa, kappa coefficient; AUC, area under the receiver operating characteristic curve; Brier, Brier score; and IS, information score.

Accuracy	GARP	SVM
OA	0.63	0.97
Kappa	0.26	0.95
AUC	0.86	0.99
Brier	0.36	0.03
IS	1.21	1.88



**Fig. 5.** Relative contribution of each biophysical variable in the SVM fusion trained with ROI1 and ROI2 (a) and with ROI1, ROI2 and ROI3 (b).

giving an OA of 0.62, a Kappa coefficient of 0.26, an AUC of 0.86, a Brier score of 0.36 and an IS of 1.21 (Table 2).

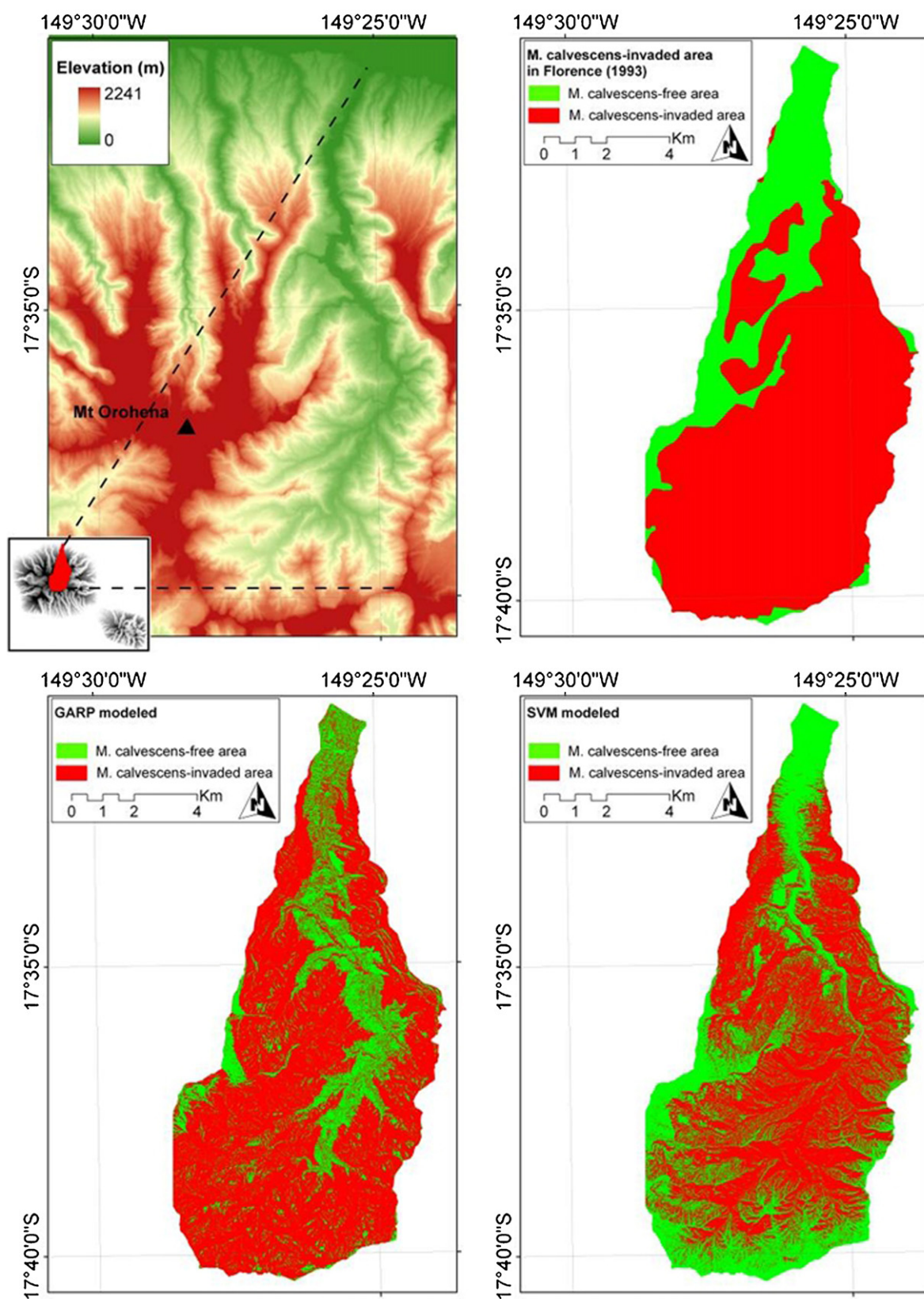
#### 3.3. Synoptical results

Fused SVM classification performance suggests that inputted variables are necessary and sufficient to explain the distribution of *M. calvescens*, confirming that its potential area is currently saturated.

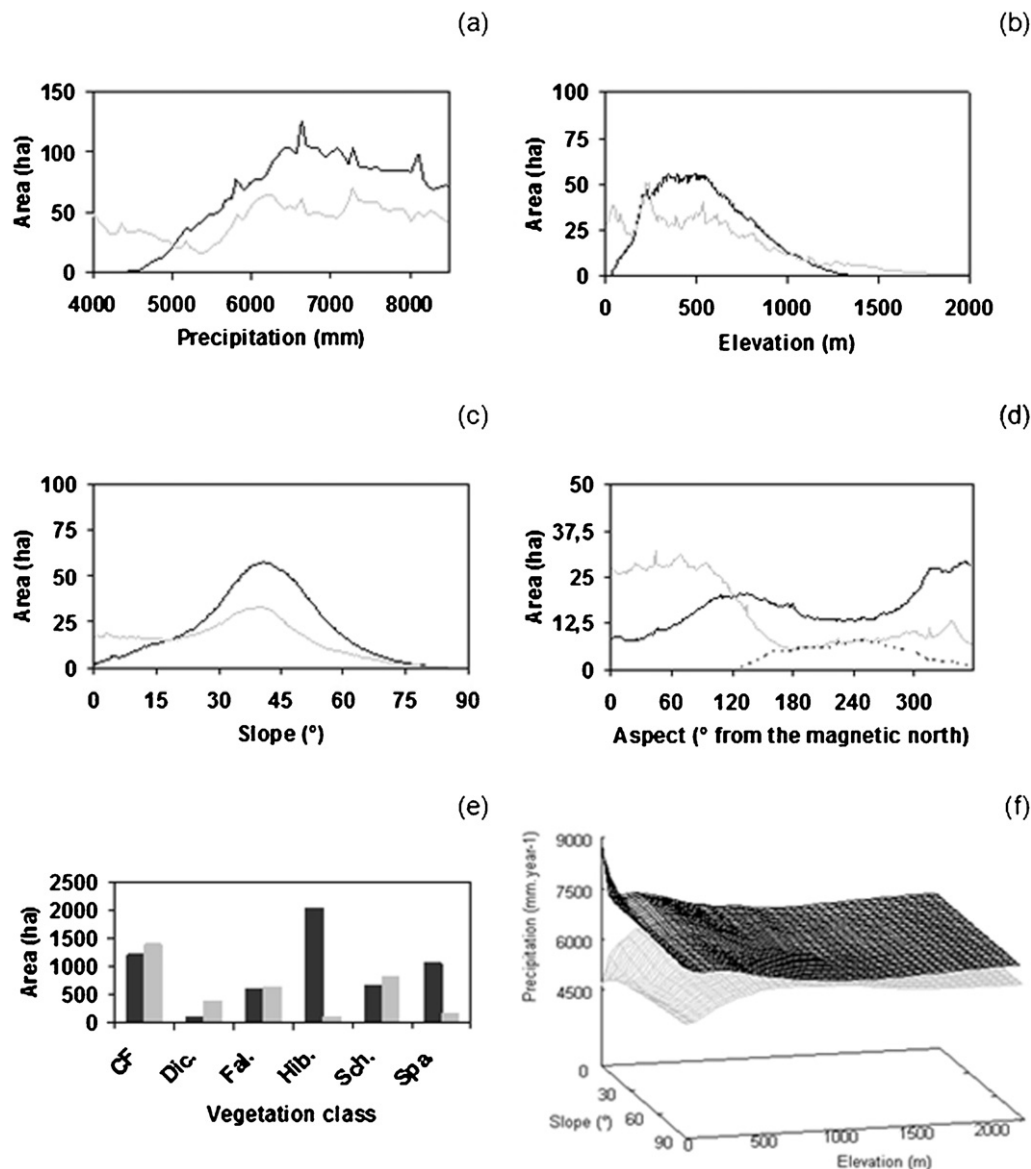
Chosen variables are not significantly correlated since the most correlated ones are both elevation and slope steepness with a correlation coefficient  $R=40\%$  (data not shown). The three most contributing variables are interestingly the three ones used in LaRosa et al. (2007), namely precipitation, elevation and slope steepness, representing two thirds of the OA (Fig. 5a). So, the new biophysical descriptors in the present study (API, slope exposure, CTI and overstory plant species) may explain the rest of the spatial variability of *M. calvescens* in high volcanic islands in the Pacific.

The stretch of the invasion of *M. calvescens* presented in Fig. 6 concerns 5284 ha i.e. 59% of the entire study site. The newly constructed map of *M. calvescens* is coarsely corroborating the pioneering island-scale map of Florence (1993). However a finer analysis highlights the fundamental need to work at the landscape scale dealing with ecosystem monitoring: area mapped as fully





**Fig. 6.** Digital elevation model (above on the left) of the Papenoo valley and *M. calvenscens* distribution according to Florence (1993) (above on the right), to the GARP model (below on the left) and to fused SVM (below on the right).



**Fig. 7.** Surface of *M. calvenscens*-invaded (black continuous line) and *M. calvenscens*-free (grey continuous line) areas in function of biophysical variables. For (d) only, the black dashed line represents *M. calvenscens* reaching the canopy surface and was plotted with the three classes ROI1, ROI2 and ROI3. (f) shows range of precipitation, elevation and slope steepness where *M. calvenscens* is invasive (black surface) and where *M. calvenscens* is absent (grey surface). Surfaces were fitted from 2,113,773 and 1,470,137 pixels respectively. Precipitation range of (a) may be over-estimated.

invaded is actually comprised of a mosaic of *M. calvenscens*-free areas and invaded ones. Species distribution varies according to the considered scale. In the present study, by means of fine resolution imagery and associated emerging statistical tools we access the target species distribution at a lower integration level – hence the differences between the two maps.

Fig. 6 clearly shows the dependence of *M. calvenscens* on topography: steepest slopes, highest and driest areas always result in *M. calvenscens*-free area.

#### 4. Discussion

##### 4.1. 4.1. Biophysical descriptors as proxies of *M. calvenscens* ecology

##### 4.1.1. Biophysical descriptors as proxies of *M. calvenscens* distribution

Concerning the three most contributing variables, *M. calvenscens*-free areas are in the majority where habitat is less than 5,000 mm

per year (Fig. 7a), neither too low (<200 m) nor too high (>1100 m) (Fig. 7b) and with low slope steepness (<15°) (Fig. 7c).

The reason of the abundance of *M. calvenscens* in slopes between 15° and 60° could be the perturbation due to severe soil erosion favorable to *M. calvenscens* and/or could be linked to the species biological traits: (i) endozoochory (bird-dispersal) plus barochory (gravity) is an ideal dispersal strategy to colonize complete slopes; (ii) slopes release micro-sites favorable to germination; (iii) the tentacular root system might be a good adaptation for thin soils on steep slopes.

Observed upper elevation limit of *M. calvenscens* reaches 1353 m which is consistent with data obtained during extensive field surveys in Tahiti (between 1300 and 1400 m, Meyer, 1998; Meyer et al., 2008). The cause of this limit is still unknown but we can hypothesize that rather than low temperatures (assuming a 0.6 °C decrease per 100 m, there is a 8.4 °C decrease from sea level to 1400 elevation) the limiting factor could be the change of vegetation structure at higher elevation (from tall canopy forest to shrubby vegetation) not suitable to *M. calvenscens* growth.



In comparison, upper limit of *M. calvescens* is lower in the Hawaiian archipelago (between 300 and 820 m, Maui Invasive Species Committee, personal communication). Such difference between the Hawaiian archipelago and Tahiti island is not likely to be due to genetic or ecological differences but more probably to the later introduction of *M. calvescens* in Hawaii (1960s) than in Tahiti (1937) and nearly two decades of containment/eradication efforts (Le Roux et al., 2008). A more comparable variable is slope steepness reaching 73° in the Maui Island (Maui Invasive Species Committee, personal communication) which is interesting for its close correspondence with our observations in Tahiti (76°) and stable within invaded sites.

Meyer (1996, 1998) stated that the lower rainfall limit for *M. calvescens* distribution is about 2000 mm per year for the island of Tahiti. In the Papenoo valley, our results show that this limit is 4500 mm per year. Three hypotheses might explain this difference: (i) absolute values of the precipitation map of Tahiti can be biased because it was interpolated from 43 meteorological stations with only seven above 500 m elevation and three located in the Papenoo valley at 4, 116 and 350 m elevation. However, relative values seems to be reliable since it is the first explicative descriptor; (ii) affinity of *M. calvescens* with steep slopes compared to flat areas can explain its absence in the lower part of the valley; (iii) some areas in the lower part of the valley are dominated by stands of the large alien tree *F. moluccana* (Fal.) with few *M. calvescens* in the understorey, and the Polynesian bamboo *S. glaucifolium* (Sch.), a species forming impenetrable layer of rhizomes and almost monospecific groves characterized by low light incidence.

Site discrepancy in biophysical limits suggests that they should not be considered alone but as a system of interacting factors with synergistic and antagonistic effects. Such complex interactions are illustrated in Fig. 7f. Precipitation i.e. humidity at the landscape-scale is intensely impacting *M. calvescens* distribution whereas CTI i.e. local effect of microtopography on humidity not as much.

In regard to the overstorey plant species influence, the invasion is generally extended and all vegetation classes are concerned. Nevertheless, *M. calvescens* is predominantly present in mixed forests (Hib. and Spa. classes) and less frequent, with more or less 50% occurrence, in *F. moluccana* (Fal.) mesic forests, in cloud forests (CF), in open and relatively dry *D. linearis* fernlands (Dic.) and in *S. glaucifolium* bamboo groves (Sch.) (Fig. 7e).

#### 4.1.2. Biophysical descriptors as proxies of *M. calvescens* height

Although slope exposure does not noticeably contribute to explain *M. calvescens* distribution, it is the best proxy of *M. calvescens* height (Figs. 6b and 7d). Now, considering *M. calvescens* reaching the canopy surface as a full class, the two thirds of the classification OA (achieving the same score than with two classes only: 97%; data not shown) are obtained with the slope exposure variable lonely.

Slope exposure groups wind and insolation effects together. Wind may not have much influence since the valley is generally protected from it by deeply steep-sided slopes. Insolation varies at annual and diurnal temporal scales. Slope exposure is predominantly contributing in the face of API, showing no prevailing contribution in Fig. 5, and *M. calvescens* reaching the canopy surface is faced from 120° to 360° i.e. in relief shadow during the morning, which suggests the critical role of diurnal insolation on *M. calvescens* height. We hypothesize that it is due to light resource and/or to the fact that *M. calvescens* favors nocturnal cool temperature and high humidity maintained longer during the day in the shade. Thus, while it is usually benefiting from shadow under the upper vegetation stratum, *M. calvescens* would be also capable of morphological plasticity, forming monospecific cover by using relief shadow. That could be why *M. calvescens* reaching the canopy

surface is not visible in our Quickbird scenes taken early in the morning.

#### 4.2. GARP model vs fused SVM

Machine learning algorithms are more adequate to classify multisource data than conventional multivariate statistical classifiers constrained by the distribution of input data. Machine learning algorithms are non-parametric and enable the weighting of the different sources (Waske and Benediktsson, 2007) which is critical when using redundant data. Even so, regarding the literature, disparities between machine learning algorithms performances are existing. SVM, one of the machine learning algorithms, is successfully used for classifying monosource and multisource remotely sensed data in many application studies and outperforms other machine learning algorithms in many comparative studies. Therefore we assessed the ability of SVM to integrate a set of biophysical variables for predicting distribution of *M. calvescens*. Fused SVM has been compared with GARP model, another machine learning algorithm classically used in species predictive mapping. According to all the comparison metrics we used, GARP model performs worse than fused SVM in our study case which lead us to the following observations.

A major limit of GARP model is its unstable behavior which leads to various solutions. In contrast, fused SVM has shown to be a stable algorithm that integrates effectively heterogeneous data and that can be trained on relatively small sets of meaningful pixels (Foody and Mathur, 2006). In this way, fused SVM should be used to map rare endemic species for which, in general, few presence ground control pixels are known. As stated by Anderson et al. (2003), one of the major limitations of the GARP model (in its Desktop version) is that only presence data are inputted rather than both presence and absence as for fused SVM. Another significant improvement with the fused SVM is its ability to integrate the Quickbird image directly whereas the GARP model needs a vegetation map built a priori as input. Finally, GARP model has been designed for species predictive mapping. With regard to SVM, originally introduced as a binary classifier, it has been extended to deal with multi-class (Q) problem and become both a classification and a regression classifier. In consequence, GARP model can only produce presence area of the target species while SVM can also be used for mapping quantitative demographic traits of the target species such as height, density or biomass for example.

#### 5. Conclusion

In this distribution model study, we proposed to integrate biophysical descriptors from multisource remotely sensed data by means of SVM fusion. The latter has led to very accurate analytical results and coherent interpretations. Our integrated synoptical procedure has permitted to emphasize additional ecological traits of *M. calvescens* and factors that determine its distribution have been quantified and organized into an hierarchy. The methods and results presented in this study can contribute to a better understanding of the dynamics of plant invasions in rainforests, so that decision-makers can develop efficient strategies to improve the ecological sustainability of vulnerable landscapes in general, as well as in the Papenoo valley.

The proposed method can be adapted and should be applied to plant or animal, invasive or native species other than *M. calvescens* in forest ecosystems, particularly in complex terrain. These species will be of two types: (i) species with a saturated distribution for which the produced map will be an actual distribution map; (ii) species with an unsaturated distribution for which the produced map will be a potential distribution map or a target species niche

map. In the latter case, the accuracy is not adapted to evaluate the classification but provides a useful figure: the percentage of the potential habitat which is actually inhabited (provided that used descriptors are sufficient to explain the distribution of the target species).

If the classification accuracy could not be calculated to assess the model for species which distribution is unsaturated, the present study results underline several successful traits of the model: (i) selected biophysical descriptors are sufficient to explain the distribution of *M. calvescens*; (ii) SVM is an ideal statistical tool to perform image fusion and to integrate such a complex set of descriptors; and (iii) in a generic point of view, the model may be robust for the reasons mentioned above and because SVM are non-parametric classifiers and adapted to mixture of qualitative and quantitative variables.

Regarding the use of the model for other species, several data are required to train the SVM and can constitute a limitation: (i) sufficient field data is needed which can be restricting in the case of rare species known in few stations; and (ii) free areas have to be identified in situ and differentiated from non-inhabited potential areas which can be hard in the case of species with unsaturated distribution.

In a future work, it will be interesting to compare the valley scale-model presented here with an entire island-scale model (in Tahiti and/or Moorea, where the invasion seems to be also saturated). The weight of the biophysical descriptors may vary according to the scale of observation as observed for a range of ecological processes such as soil composition or air temperature and we can suppose that other variables will have to be taken into consideration at an island-scale such as windwardness.

The present research is a preliminary but essential step to be able to model the potential distribution of *M. calvescens* in area where it has been recently introduced and is spreading rapidly such as the Hawaiian islands, New Caledonia, the tropical region of Queensland in Australia, or in other French Polynesian islands such as Raiatea and Tahaa (Society archipelago), Fatu Hiva and Nuku Hiva (Marquesas).

## Acknowledgements

The authors are grateful to the Government of French Polynesia, its Service de l'Urbanisme for providing the Quickbird and the DEM data and its Délégation à la Recherche for financial support. We deeply thank the "Maui Invasive Species Committee" for sharing data on the ecological range of *M. calvescens* on the Hawaiian islands, Lloyd L. Loope (Pacific Island Ecosystems Research Center, USGS) for his relevant comments, Ravahere Taputuarai and Vetea Liao (UC Berkeley R. Gump South Pacific Research Station) for revising the English on an early draft of this paper, and the anonymous reviewers for improving it.

## References

- Anderson, R.P., Lew, D., Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162, 211–232.
- Asner, G.P., Jones, M.O., Martin, R.E., Knapp, D.E., Hughes, R.F., 2008a. Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment* 112, 1912–1926.
- Asner, G.P., Hughes, R.F., Vitousek, P.M., Knapp, D.E., Kennedy-Bowdoin, T., Boardman, J., Martin, R.E., Eastwood, M., Green, R.O., 2008b. Invasive plants transform the three-dimensional structure of rain forests. *Proceedings of the National Academy of Sciences of the USA* 105, 4519–4523.
- Atherton, J., et al. (Eds.), 2007. Polynesia-Micronesia biodiversity hotspot. Critical ecosystem profile. Ecosystem Partnership Fund, International-Melanesia Center for Biodiversity and Conservation, Arlington, Virginia.
- Baruch, Z., Goldstein, G., 1999. Leaf construction cost, nutrient concentration, and net CO<sub>2</sub> assimilation of native and invasive species in Hawaii. *Oecologia* 121, 183–192.
- Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78, 1–3.
- Bruzzone, L., Conese, C., Maselli, F., Roli, F., 1997. Multisource classification of complex rural areas by statistical and neural-network approaches. *Photogrammetric Engineering and Remote Sensing* 63, 523–533.
- Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2, 121–167.
- Camps-Valls, G., Gomez-Chova, L., Muñoz-Marí, J., Vila-Francés, J., Calpe-Maravilla, J., 2006. Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters* 3, 93–97.
- Chen, J., Saunders, S.C., Crow, T.R., Naiman, R.J., Brosfokske, K.D., Mroz, G.D., Brookshire, B.L., Franklin, J.F., 1999. Microclimate in forest ecosystem and landscape ecology. *Bioscience* 49, 288–297.
- Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational Psychological Measurement* 20, 37–46.
- Congalton, R.G., Green, K., 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC Press, Boca Raton, FL.
- Drake, J.M., Randin, C., Guisan, A., 2006. Modelling ecological niches with support vector machines. *Journal of Applied Ecology* 43, 424–432.
- Duncan, R.A., Fisk, M.R., White, W.M., Nielsen, R.L., 1994. Tahiti: geochemical evolution of a French Polynesian volcano. *Journal of Geophysical Research* 99, 24,341–24,357.
- Ebeling, S.K., Welk, E., Auge, H., Bruehlheide, H., 2008. Predicting the spread of an invasive plant: combining experiments and ecological niche model. *Ecography* 31, 709–719.
- Florence, J., 1993. La végétation de quelques îles de Polynésie. Planches 54–55. In: Dupon, F. (Coord. Ed.), *Atlas de la Polynésie française*. ORSTO, Paris, France.
- Florence, J., Lorence, D.H., 1997. Introduction to the flora and vegetation of the Marquesas Islands. *Allertonia* 7, 226–237.
- Franklin, J., 1995. Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography* 19, 474–499.
- Foody, G.M., Mathur, A., 2004. A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* 42, 1335–1343.
- Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment* 103, 179–189.
- Fourdrigniez, M., Meyer, J.-Y., 2008. Liste et caractéristiques des plantes introduites naturalisées en Polynésie française. Contribution à la Biodiversité de Polynésie française no.17, Délégation à la Recherche, Papeete, French Polynesia.
- Fu, P., Rich, P.M., 2002. A geometric solar radiation model with applications in agriculture and forestry. *Computers and Electronics in Agriculture* 37, 25–35.
- Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L., Holmes, K., 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. *Soil Science Society of America Journal* 64, 2046–2056.
- Gillet, F., 2000. La phytosociologie synusiale intégrée – Guide méthodologique. Technical report, Laboratoire d'écologie végétale et de phytosociologie, Neuchâtel University, France.
- Guo, Q., Kelly, M., Graham, C.H., 2005. Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecological Modelling* 182, 75–90.
- Halldorsson, J.H., Benediktsson, J.A., Sveinsson, J.R., 2003. Support vector machines in multisource classification. *IEEE International Geoscience and Remote Sensing Symposium* 3, 2054–2056.
- Hand, D.J., Till, R.J., 2001. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning* 45, 171–186.
- Haralick, R., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems Man and Cybernetics* 3, 610–621.
- Hill, M.J., Ticehurst, C.J., Lee, J.-S., Grunes, M.R., Donald, G.E., Henry, D., 2005. Integration of optical and radar classifications for mapping pasture type in Western Australia. *IEEE Transactions on Geoscience and Remote Sensing* 45, 1665–1681.
- Hoffer, R.M., 1975. Natural resource mapping in mountainous terrain by computer analysis of ERTS-1 satellite data. Technical Note, Laboratory for Applications of Remote Sensing, Purdue University, USA.
- Hsu, C.W., Chang, C.C., Lin, C.J., 2009. A practical guide to support vector classification. Technical Note, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan.
- Hutchinson, C.F., 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing* 48, 123–130.
- Invasive Species Specialist Group (ISSG), 2010. Available on line at [www.issg.org/database/](http://www.issg.org/database/).
- Jeon, M., Alexander, M., Pedrycz, W., Pizzi, N., 2005. Unsupervised hierarchical image segmentation with level set and additive operator splitting. *Pattern Recognition Letters* 26, 1461–1469.
- Joshi, C., De Leeuw, J., van Duren, I.C., 2004. Remote sensing and GIS applications for mapping and spatial modelling of invasive species. In: *Proceedings of the International Society for Photogrammetry and Remote Sensing*, Istanbul, Turkey (CD-ROM, unpaginated).
- Joshi, C., De Leeuw, J., van Andel, J., Skidmore, K.A., Lekhak, H.D., van Duren, I.C., Norbu, N., 2006. Indirect remote sensing of a cryptic forest understorey invasive species. *Forest Ecology and Management* 225, 245–256.
- Kononenko, I., Bratko, I., 1991. Information-based evaluation criterion for classifier's performance. *Machine Learning* 6, 67–80.
- Krantz, T.P., 1997. Results of the *Miconia calvescens* color infrared aerial photography detection project, Moorea, French Polynesia. With imagery by Mr T.

- Norris, with assistance from Dr. J.-Y. Meyer and the Office of the Ministère de l'Environnement, and the Richard B. Gump Research Station, University of California at Berkeley.
- Kueffer, C., Daehler, C.C., Torres-Santana, C.W., Laverigne, C., Meyer, J.-Y., Otto, R., Siolva, L., 2010. A global comparison of plant invasions on oceanic islands. *Perspectives in Plant Ecology, Evolution and Systematics* 12, 145–161.
- Lardeux, C., Frison, P.-L., Tison, C., Souyris, J.-C., Stoll, B., Fruneau, B., Rudant, J.-P., 2009. Support vector machine for multifrequency SAR polarimetric data classification. *IEEE Transactions on Geoscience and Remote Sensing* 47, 4143–4152.
- LaRosa, A.M., Purrell, M., Franklin, J., Denslow, J., 2007. Designing a control strategy for *Miconia calvenscens* in Hawaii using spatial modelling. In: 9th International Conference on the Ecology and Management of Alien Plant Invasions (poster), Perth, Australia.
- Lawrence, R.L., Wright, A., 2001. Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering and Remote Sensing* 67, 1137–1142.
- Laurent, V., Maamaatuaiahutapu, K., Maiau, J., Varney, P., 2004. Atlas climatologique de la Polynésie française. Météo-France, Direction interrégionale de Polynésie française, Papeete, French Polynesia.
- Lennon, M., Mercier, G., Hubert-Moy, L., 2002. Classification of hyperspectral images with nonlinear filtering and support vector machines. *IEEE International Geoscience and Remote Sensing Symposium*, 1670–1672.
- Le Roux, J.J., Wicczorek, A.M., Meyer, J.-Y., 2008. Genetic diversity and structure of the invasive tree *Miconia calvenscens* in Pacific islands. *Diversity and Distribution* 14, 935–948.
- Linderman, M.A., Liu, J., Qi, J., An, L., Ouyang, Z., Yang, J., Tan, T., 2004. Using artificial neural networks to map the spatial distribution of understory bamboo from remotely sensed data. *International Journal of Remote Sensing* 25, 1685–1700.
- Loo, S.E., Mac Nally, R., Lake, P.S., 2007. Forecasting New Zealand mudsnail invasion range: model comparisons using native and invaded ranges. *Ecological Applications* 17, 181–189.
- Loope, L.L., Mueller-Dombois, D., 1989. Characteristics of invaded islands, with special references to Hawaii. In: Drake, et, al. (Eds.), *Biological Invasions: a Global Perspective*. John Wiley and Sons, Chichester, UK.
- Meyer, J.-Y., 1996. Status of *Miconia calvenscens* (Melastomataceae), a dominant invasive tree in the Society Islands (French Polynesia). *Pacific Science* 50, 66–76.
- Meyer, J.-Y., Florence, J., 1996. Tahiti's native flora endangered by the invasion of *Miconia calvenscens* DC. (Melastomataceae). *Journal of Biogeography* 23, 775–781.
- Meyer, J.-Y., 1998. Observations on the reproductive biology of *Miconia calvenscens* (Melastomataceae), an alien invasive tree on the island of Tahiti (South Pacific Ocean). *Biotropica* 30, 609–624.
- Meyer, J.-Y., 2004. Threat of invasive alien plants to native flora and forest vegetation of Eastern Polynesia. *Pacific Science* 58, 357–375.
- Meyer, J.-Y., 2007. Conservation des forêts naturelles et gestion des aires protégées en Polynésie française. *Bois et Forêts des Tropiques* 291, 25–40.
- Meyer, J.-Y., Taputuarai, R., Killgore, E.M., 2008. Dissemination and impact of the fungal pathogen *Colletotrichum gloeosporioides* f. sp. *miconiae* on the invasive alien tree *Miconia calvenscens* (Melastomataceae) in the rainforests of Tahiti (French Polynesia, South Pacific). In: Julian, M.H., Sforza, R., Bon, M.C., Evan, H.C., Hatcher, P.E., Hinz, H.L., Rector, B.G. (Eds.), *Proceedings of the XII International Symposium on Biological Control of Weeds*, La Grande Motte, France. CAB International, Wallingford.
- Meyer, J.-Y., 2009. Flore et végétation du plateau marécageux de Anaor'i'i (vallée de la Papeno'o, Tahiti): une zone humide unique en Polynésie française. Technical report. French Polynesian Department of Research, Papeete, French Polynesia.
- Meyer, J.-Y., 2010. Montane cloud forests on remote islands of Oceania: the example of French Polynesia (South Pacific Ocean). In: Bruijnzeel, L.A., Scatena, F.N., Hamilton, L.S. (Eds.), *Tropical Montane Cloud Forests*. Cambridge University Press.
- Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well Being: Current State and Trends*. Island Press.
- Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: a review of hydrological, geomorphological and biological applications. *Hydrological Processes* 5, 3–30.
- Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* 57, 443–452.
- Nagy, L., Grabherr, G., Körner, C., Thompson, D.B.A. (Eds.), 2003. *Alpine Biodiversity in Europe*. Springer Verlag, Berlin, Germany.
- Nemmour, H., Chibani, Y., 2006. Multiple support vector machines for land cover change detection: an application for mapping urban extensions. *ISPRS Journal of Photogrammetry and Remote Sensing* 61, 125–133.
- Pal, M., Mather, P.M., 2004. Assessment of the effectiveness of support vector machines for hyperspectral data. *Future Generation Computer Systems* 20, 1215–1225.
- Paulay, G., 1994. Biodiversity on oceanic islands its origin and extinction. *American Zoologist* 34, 134–144.
- Papy, R.H., 1951–1954. Tahiti et les îles voisines. La végétation des îles de la Société et de Makatea (Océanie française), 2<sup>e</sup> partie. Travaux du laboratoire forestier de Toulouse, tome V, 2<sup>e</sup> sect. vol. I. art. III, Toulouse, France.
- Peterson, A.T., Scachetti-Pereira, R., Kluza, D.A., 2003. Assessment of invasive potential of *Homalodisca coagulata* in Western North America and South America. *Biota Neotropica* 3, 2–7.
- Piella, G., 2003. A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion* 4, 259–280.
- Pouteau, R., Stoll, B., Chabrier, S., 2010. Ground truth method assessment for SVM based landscape classification. *IEEE International Geoscience and Remote Sensing Symposium*, 2715–2718.
- Richards, J.A., 2005. Analysis of remotely sensed data: the formative decades and the future. *IEEE Transactions on Geoscience and Remote Sensing* 43, 422–432.
- Riegel, G.M., Miller, R.F., Krueger, W.C., 1992. Competition for resources between understory vegetation and overstory *Pinus ponderosa* in Northeastern Oregon. *Ecological Applications* 2, 71–85.
- Roura-Pascual, N., Suarez, A.V., Gómez, C., Pons, P., Touyama, Y., Wild, A.L., Peterson, A.T., 2004. Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society London B* 271, 2527–2535.
- Roura-Pascual, N., Brotons, L., Peterson, A.T., Thuiller, W., 2008. Consensual predictions of potential distributional areas for invasive species: a case study of Argentine ants in the Iberian Peninsula. *Biological Invasions* 11, 1017–1031.
- Sánchez-Flores, E., 2007. GARP modelling of natural and human factors affecting the potential distribution of the invasives *Schismus arabicus* and *Brassica tournefortii* in “El Pinacate y Gran Desierto de Altar” Biosphere Reserve. *Ecological Modelling* 204, 457–474.
- Schölkopf, B., Smola, A., 2002. *Learning with Kernels*. Cambridge, MA, MIT Press.
- Stockwell, D., Peters, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13, 143–158.
- Strahler, A.H., 1981. Stratification of natural vegetation for forest and rangeland inventory using Landsat digital imagery and collateral data. *International Journal of Remote Sensing* 2, 15–41.
- Treitz, P., Howarth, P., 2000. Integrating spectral, spatial, and terrain variables for forest ecosystem classification. *Photogrammetric Engineering and Remote Sensing* 66, 305–317.
- Vapnik, V., Chervonenkis, A., 1998. *Statistical learning theory*. In: *Support Vector Machines for Pattern Recognition*. John Wiley & Sons, New York.
- Walsh, S.J., McCleary, A.L., Mena, C.F., Shao, Y., Tuttle, J.P., Gonzalez, A., Atkinson, R., 2008. Quickbird and Hyperion data analysis of an invasive plant species in the Galapagos Islands of Ecuador: implications for control and land use management. *Remote Sensing of Environment* 112, 1927–1941.
- Waske, B., Benediktsson, J.A., 2007. Fusion of support vector machines for classification of multisensor data. *IEEE Transactions on Geoscience and Remote Sensing* 45, 3858–3866.
- Waske, B., Benediktsson, J.A., Sveinsson, J.R., 2009. Classifying remote sensing data with support vector machines and imbalanced training data. In: *Multiple Classifiers Systems*. Springer, Berlin/Heidelberg.
- Wibaux, B., 1990. Étude sur les potentialités de l'outil télédétection à contrôler les états d'évolution de la prolifération du genre *miconia* à Tahiti. Délégation à l'environnement/Station Polynésienne de Télédétection. Convention 89/1653, Unpublished report.
- Wilson, J.P., Gallant, J.C., 2000. *Terrain Analysis: Principles and Applications*. John Wiley & Sons, New York.
- Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. *Plant Ecology* 1, 9–23.