Published by AIT
the Italian Society of Remote Sensing

Mapping invasive woody plants in Azores Protected Areas by using very high-resolution multispectral imagery

Artur Gil[1*], Agustin Lobo[2], Mohamed Abadi[3], Luis Silva[4] and Helena Calado[4]

[1]Azorean Biodiversity Group, CITA-A, Department of Biology,
University of the Azores, Rua da Mãe de Deus 13A, 9501-801 Ponta Delgada, Portugal;
[2]ICTJA -Institut de Ciències de la Terra Jaume Almera (CSIC) E-08028 Barcelona, Spain
[3]Véodis-3D, 63057 Clermont-Ferrand, France
[4]CIBIO (Azores Unit), Department of Biology, University of the Azores, 9501-801 Ponta Delgada, Portugal
*Corresponding author, e-mail address: arturgil@uac.pt

Abstract
We assessed the effectiveness of very high spatial resolution IKONOS imagery for mapping a top invasive woody plant, *Pittosporum undulatum*, in a Protected Area in S.Miguel Island. We developed a segmentation-based classification scheme. A strong separability between most important land cover classes and a high accuracy in supervised classification maps was achieved. Overall separability improved significantly after the training data depuration process. Support Vector Machine and Maximum Likelihood's supervised classifiers showed a strong agreement and a good accuracy at land-cover class level, especially with *P. undulatum*. This approach was confirmed as a cost-effective method to map woody plant invaders in Azores Protected Areas.

Keywords: Segmentation, Ikonos, invasive alien species, protected areas, Azores, *Pittosporum undulatum.*

Introduction
Climate variability and changes, the proliferation of invasive exotic species, the increasing of tourist activity, natural catastrophes, the over-exploitation of natural resources as well as the pollution and residue management are the main threats to sustainable development, to nature conservation and to small island biodiversity maintainability [CBD, 2006]. These characteristics, associated with remoteness, isolation, smallness, and particularly closed systems, make planning and management on small islands more challenging in scientific and technical terms [Calado et al., 2007; Gil et al., 2012]. Even if many oceanic islands are considered biodiversity hot spots, they are also under considerable pressure from human activities. This situation leads to habitat destruction and fragmentation and to the invasion of alien species [Caujapé et al., 2010]. The biological invasion of exotic species is considered the second leading cause for the global loss of biodiversity, being exceeded only

by habitat destruction. [Rodríguez-Echeverría et al., 2009]. Invasive plants are considered one of the major threats to biodiversity conservation in islands, including the Macaronesian Archipelagos where they invade many protected areas [Silva et al., 2008; Castro et al., 2010; Kueffer et al., 2010]. The distribution and abundance of alien plants are important to assess their impact on the invaded system [Parker et al., 1999]. Vegetation mapping is thus mandatory to obtain current states of vegetation cover in order to start vegetation protection and restoration programs [He et al., 2005]. Traditional methods (e.g. field surveys, literature reviews, map interpretation and collateral and ancillary data analysis), however, are not fully efficient to map vegetation because they are time consuming, date lagged and often too expensive [Xie et al., 2008]. Remote sensing provides repeatable and consistent assessment and monitoring of the environment. It allows independent control and its quality can be assessed. It is a tool with some very desirable characteristics for supporting environmental policy [De Leeuw, 2010]. Remote sensing allows efficient information collecting over large spatial extents at high spatial resolution. Concerning the targeted Invasive Alien Species (IAS), successful detection approaches have generally took advantage of unique phenological or biochemical properties, structural characteristics, or the spatial patterns of infestations. [Strand et al., 2007]. The most intuitive and easy approach for alien plant detection is to use high and very high spatial resolution images to map the spatial distribution of non-native species. The idea in this approach is to detect these species on the basis of their unique spatial textures/patterns or phenological characteristics. The classic source of very high resolution imagery has been aerial photography, but satellite imagery is increasingly contemplated as an alternative, in particular for those areas, such as Azores, that are distant from the bases of infrastructure that is required for aerial campaigns. The availability of Visible / Near Infrared (VNIR) spectral bands coupled to its very high spatial resolution and high temporal resolution make IKONOS satellite imagery a performing solution for species composition, land cover, phenology, habitat structure and primary productivity mapping. Therefore, several successful applications on invasive vegetation mapping have been developed using high and very high spatial resolution satellite imagery [Carson et al., 1995; Turner et al., 2003; Katoh, 2004; Casady et al., 2005; Everitt et al., 2006; Tsai and Chou, 2006; Huang and Asner, 2009]. A first attempt of Pico da Vara Nature Reserve's vegetation mapping performed by Gil et al. [2011] showed that using very high spatial resolution remote sensing data for detection and monitoring of invasive species (such as *Pittosporum undulatum* and *Clethra arborea*) and native vegetation patches (such as scrubland) can constitute a cost-effective solution to study and assess the Azorean Protected Areas. Nevertheless, the poor separability between important classes, as for instance *Pittosporum* woodland and *Cryptomeria japonica* forest stands, was identified as a major issue in order to obtain an accurate vegetation mapping of these areas.

Study Area

Pico da Vara/ Ribeira do Guilherme Special Protected Area (SPA) is located São Miguel, the largest island in the Azores. It includes one of the last main areas of native forest and scrubland in São Miguel, whose most important sub-area is located in the former Nature Reserve of Pico da Vara (815 hectares) in the mountain complex of Serra da Tronqueira (Fig. 1). It was classified as SPA in 1999 (Decree-Law 140/99 of April 24th) due to the presence and conservation status of the endemic Azores Bullfinch *Pyrrhula murina* Godman, 1866, one of the most threatened passerines in all Europe. Its estimated population of 500-800 pairs [Ceia et al., 2011] is limited to a few fragments of remaining native vegetation (Azorean Laurel Forest) seriously threatened by IAS spreading.

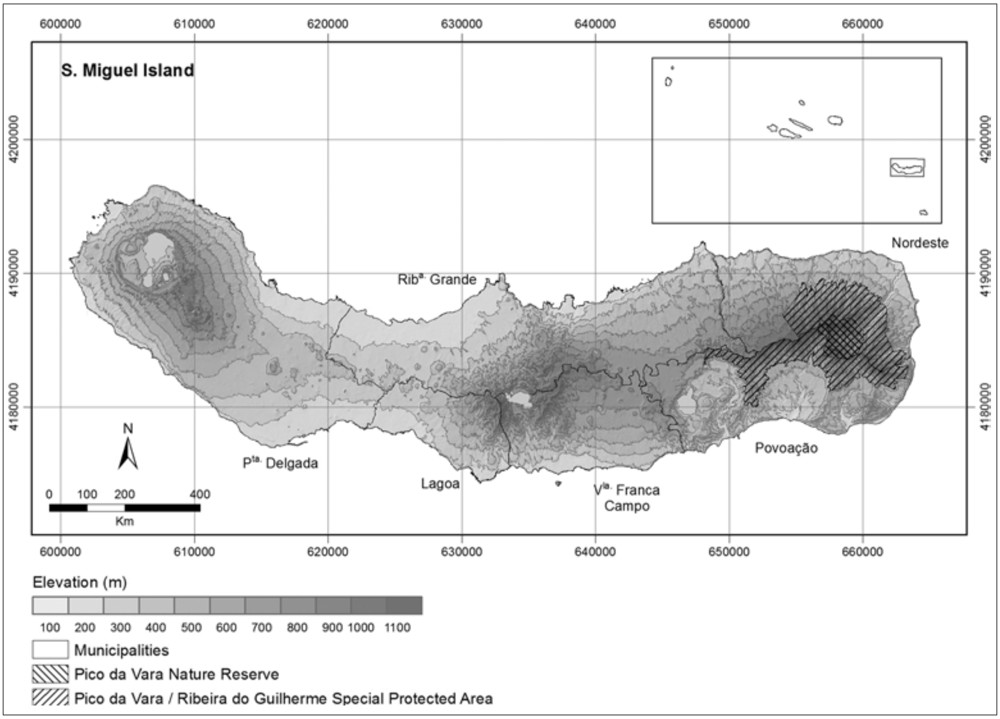

Figure 1 - Location of Pico da Vara Nature Reserve and Pico da Vara/Ribeira do Guilherme SPA (S. Miguel Island, Portugal).

Target Species: Sweet Pittosporum

S. Miguel Island's vascular plant flora (Archipelago of the Azores, Portugal) consists of approximately 1000 *taxa* and is largely dominated by non-indigenous taxa: 66% [Silva and Smith, 2004, 2006]. The native vegetation on this island is invaded by several problematic IAS, including *Pittosporum undulatum*, *Hedychium gardnerianum*, *Gunnera tinctoria* and *Clethra arborea* [Bibby and Charlton, 1991; Silva, 2001]. *Pittosporum undulatum* (Sweet *Pittosporum)* is one of the plants with the highest impact in the Azorean vegetation. According to a recent evaluation of the 100 most invasive species in Macaronesia, Sweet *Pittosporum* is considered invasive also in Madeira and the Canary islands, and ranked 8th in a total of 195 evaluated species [Silva et al., 2008]. *P. undulatum* is therefore considered one of the priority species for the implementation of control actions in the Azores, under the implementation of the Azorean Regional Program for Control and Eradication of Invasive Plants in Sensitive Areas (PRECEFIAS) [Lourenço et al., 2011]. Introduced in the Azores in the 19th century as a hedgerow species for the protection of orange tree plantations, *P. undulatum* later dispersed during the last 100 years to a wide range of habitats throughout the Azores islands, invading plant communities from 100 to 600 m elevation. *Pittosporum undulatum* is able to overgrow native vegetation, forming pure stands [Sjogren, 1973]. Its distribution might be limited by several environmental factors that vary with altitude, namely low temperatures and increased exposure to the prevailing winds at higher altitudes. [Dias, 1996; Goodland and Healey, 1996]. The Sweet *Pittosporum* in São Miguel is limited by the most important climatic gradient on the island. The steep topography of the island originates a striking altitudinal gradient, where small increases in elevation imply both strong decreases in temperature and large increases in precipitation and humidity. Some other landforms variables, such as distance from streams presented also minor significant effects [Hortal et al., 2010]. Despite its invasive behavior, as stated before, as being a woody species widely and traditionally used in some Azores' islands, an attempt for mapping *P. undulatum* woodland in this archipelago has been made on behalf of the Regional Forest Inventory (RFI) development. RFI was produced by the Azorean Forestry Regional Direction (DRRF) and was developed by delimiting forest stands through photo-interpretation of 1997 (black and white) and 2004 (real color) orthophotomaps. Field work was performed between 2003 and 2007 to identify the different types of forest stands [Lourenço et al., 2011].

This paper aims at assessing the effectiveness of very high spatial resolution satellite imagery for invasive vegetation mapping (*Pittosporum* woodland) in Pico da Vara Nature Reserve (S. Miguel Island, Azores Archipelago, Portugal), using a segmentation-based classification scheme, in order to address the most important issues identified by Gil et al. [2011], as for instance the poor separability between most relevant vegetation classes (*Pittosporum* woodland, Native scrubland and *Cryptomeria japonica*). Finally, this study aims to promote the integration of this type of satellite data and methodological approach into the Azorean Protected Area's IAS control and management decision-support process.

Data

Three different datasets were used in our study:

1) An IKONOS standard geometrically corrected image with four multispectral bands (Blue, Green, Red and Near-Infrared) acquired on Aug 18, 2005 with 11% Cloud Cover. The IKONOS System is a commercial satellite from GeoEye and offers multispectral imagery at a spatial resolution of 4 meters and panchromatic imagery at 1 meter, a short revisit time (3-5 days off-nadir and 144 days for true-nadir) and a swath of 11 km x 11 km for each single scene;

2) The Digital Terrain Model (DTM) of S. Miguel Island with a spatial resolution of 10 meters, produced by the Military Geographic Institute of Portugal (IGEOE), used for the ortho-rectification of the IKONOS image;

3) A GIS dataset of 525 points collected on the study area, by using a sub-metric GPS device. The land-cover characterization associated to each one these points was double checked by performing a photo-interpretation of the available IKONOS image (August 2005) and also by overlaying each point to the SPEA/LIFE Priolo Project's vegetation monitoring program survey (A9 Action), continuously developed between 2004 and 2008 [Teodosio et al., 2009]. Four representative land cover and vegetation classes are identified in this dataset: (1) CC - *Cryptomeria japonica* (L. fil.) D. Don ("Japanese cedar") man-planted production forest stands; (2) DD - Bare Soil and Landslide Areas; (3) LL - Native scrubland patches; (4) NN - *Pittosporum* Woodland (*Pittosporum undulatum*'s pure or largely dominated patches).

Methods

Our methodological approach is based on producing vegetation maps through supervised object-oriented IKONOS image, instead of the pixel-based classification approach. In very high spatial resolution imagery, such as IKONOS, a group of pixels can represent the characteristics of land-cover targets better than single pixels, so groups of adjacent and similar pixels are organized into objects (or segments) and each of the objects is treated as a minimum classification unit. The image segmentation approach consists in the partition of the image into homogeneous elements (segments) that are thereafter classified. It has several advantages over conventional per-pixel methods and the "textural channels" approach. Discriminating segments instead of individual pixels greatly reduces the number of elements to be classified because there are fewer segments than pixels. It facilitates the application of more complex methods [Lobo, 1997; Yu et al., 2006]. The supervised image classification process is generally guided by expert to give the desired land-cover/vegetation classes. First, training samples which are representative and typical for that information class are defined, and secondly all input pixels (or segments) are labeled according to their class [Lenka and Milan, 2005; Xie et al., 2008]. In this study we apply three different classifiers: K-Nearest Neighbor (KNN), Maximum Likelihood classifier (MLC) and Support Vector Machine (SVM). This part of the process was run in R [R Development Core Team, 2011] with packages MASS [Venables and Ripley, 2002], e1071 [Dimitriadou et al., 2010] and raster [Hijmans and Van Etten, 2011]. K-Nearest Neighbor (KNN) is a non-parametric method commonly used in remote sensing, pattern recognition and statistics to classify objects into a predefined number of categories based on a given set of predictors [Franco-Lopez et al., 2001]. An object is classified by a majority vote of its neighbors in the feature or statistical space, with the object being assigned to the most common class amongst its k-nearest neighbors (k is a positive integer). It is widely employed in national forest inventory applications using remote sensing data [Mäkelä and Pekkarinen, 2001; Chirici et al., 2012]. Maximum-Likelihood Classifier (MLC) is usually regarded as the classic and most widely used supervised classification for satellite images resting on the statistical distribution pattern [Sohn and Rebello, 2002; Xu et al., 2005]. Support vector machines (SVM) have considerable potential as classifiers of remotely sensed data. This approach seeks to find the optimal separating hyperplane between classes by focusing on the training cases that lie at the edge of the class distributions, the support vectors, with the other training cases effectively discarded. Thus, yielding high accuracy with small training sets may be expected, which could be very advantageous, given the costs of training data acquisition in remote sensing. A constraint on their application in remote sensing is their binary nature, requiring multiclass classifications to be based upon a large number of binary analyses [Brown et al., 2000; Foody and Mathur, 2004].

Another innovative methodological aspect of our approach is the previous depuration of the dataset of sites providing information for training and validation. To simulate the interactive outlining of training and validation fields, we selected the segments that included photo-interpreted sites. In accordance with the segmentation-based classification concept, training and validation segments were preferred to the original single point sites. The training and validation dataset was first submitted to an interactive process to eliminate outliers in the statistical space, in order to increment its quality and to maximize the separability between land-cover and vegetation classes (to achieve higher classification accuracy).

Processing chain

The methodological approach is schematically described in Figure 2. The first phase consisted of pre-processing the IKONOS Visible and Near Infrared (VNIR) multi-spectral image, by performing three tasks. The first one was the ortho-rectification of the multispectral image by using the Rational Polynomial Coefficients (RPC) IKONOS sensor model (RPC files were supplied with the IKONOS image), 25 Ground Control Points (GCP) and the Digital Terrain Model (DTM) of S. Miguel Island. The root-mean-square error (RMSE) which was calculated for the whole image was of 10 meters [Ganas et al., 2002; Hale and Rock, 2003]. Afterwards, an atmospheric correction was performed by using Quick Atmospheric Correction (QUAC) method. This method determines the atmospheric compensation parameters directly from the information contained within the scene using the observed pixels spectra. The approach is based on the empirical finding that the spectral standard deviation of a collection of diverse material spectra, such as the endmember spectra in a scene, is essentially spectrally flat [Bernstein et al., 2005]. Finally, a cloud cover mask was built by performing a GIS-based photo-interpretation of a true-color (blue: band 1; green: band 2; red: band 3) composition. The vectorization of all clouded areas within study site was undertaken and the resulting mask was applied to the four IKONOS image multispectral bands.

The second phase of our methodological approach was the image segmentation procedure. This process was performed by applying the Mean Shift algorithm [Commaniciu, 2002] which was used as implemented in Monteverdi and Orfeo Toolbox [Christophe et al., 2011].

The third phase consisted of the field data depuration process, which started with the selection of most suitable data: segments that included photo-interpreted sites were selected. The second step was the calculation of segmentation-based and pixel-based image statistics to create a dataset of inspected test segments. Then, the whole dataset was displayed in reduced space in order to allow an interactive selection of outliers. Field data was tested for within-class consistency in terms of its average spectral properties according to the IKONOS image. After calculating the average Digital Number (DN) values in each spectral band, a Principal Component Analysis (PCA) was performed and the projected values were displayed in bi-plots of Principal Component (PC1, PC2 and PC3). PC1, PC2 and PC3 accounted for 99.4% of the total variance. The PC plots were used to interactively identify all the polygons which PC values were actually overlapping a different class considering all 3 PCs. Afterwards, an inspection and diagnostic of outlier segments was performed by using the image and aerial photography. An action of “confirmation”, “correction” and/or “elimination” of outliers was developed. The selected outliers were geographically located within a GIS environment and the reason for their peculiar values was identified. All remaining cloudy and shadowed segments were removed from the training dataset. Those who had been assigned to a wrong class were corrected. These actions were performed repeatedly until no more outliers were detected. From this point, the statistics of the remaining training and validation segments were arranged in two separated datasets, differing for their diagnostic variable: Dataset A with 4 classes (CC - *Cryptomeria japonica*; DD - Bare Soil and Landslide Areas; LL - Native scrubland; NN -*Pittosporum* woodland); and Dataset B with 5 classes (CC; DD; LL; NN and NN2 - resulting from the splitting of NN in two classes). In order to reduce the redundancy and to maximize the effectiveness of classes, the statistical separability among the land cover/vegetation classes was examined before and after the segmentation process (Tab. 1) using the Jeffries-Matsushita (J-M) separability measure ($0 \leq$ J-M ≤ 2) [Richards and Jia, 1999]. As a general rule, if the J-M value is greater than 1.9, then separation is good. If the J-M is between 1.7 and 1.9 then separation is fairly good; below 1.7 the separation is considered to be poor. Finally, Datasets A (with 4 classes: CC; DD; LL; NN) and B (with 5 classes: CC; DD; LL; NN and NN2) were randomly split into training (TA and TB) and validation (VA and VB) datasets. Both TA and TB datasets will be used for classification, whilst VA and VB will be used for assessing classification accuracy.

The fourth phase of our methodological approach consisted of the segmentation-based supervised classification of the IKONOS multispectral image (ortho-rectified and atmospherically corrected) by using TA and TB as training datasets, and by applying three different classifiers (SVM; MLC and KNN). In total, 6 classification maps were produced (three for each training dataset).

Finally, the accuracy assessment of six outputted classification maps was performed by computing overall and "per class user's accuracy, producer's accuracy and overall Kappa coefficient of agreement, using both validation datasets [Congalton and Green, 1999; Foody, 2002].

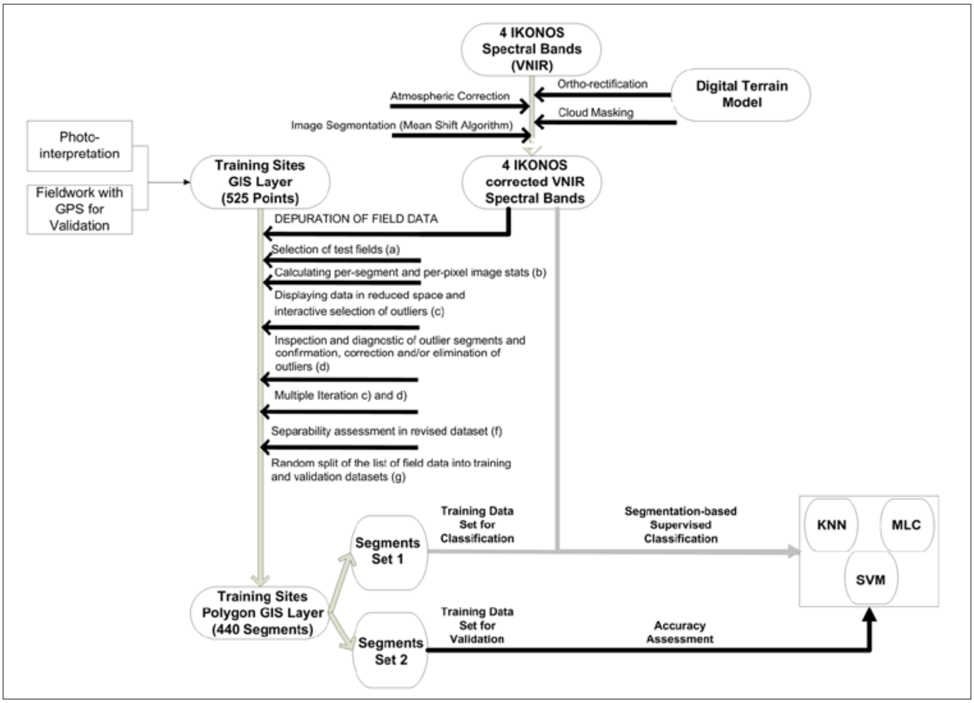

Figure 2 - Methodological flowchart.

Results and Discussion

Depuration of field data by interactive outlier detection and deletion reduced the number of segments from 525 to 440. It also improved overall separability (Tab. 1) in both datasets A (4 classes) and B (5 classes).

Table 1 - Separability Assessment comparison between initial and depurated training data. (CCa, DDa, LLa, NNa - Separability Assessment before the training data segmentation process; CCb, DDb, LLb, NNb - Separability Assessment after the training data segmentation process using 4 classes (CC, DD, LL, NN); CCc, DDc, LLc, NNc, NN2c - Separability Assessment after the training data segmentation process using 5 classes (CC, DD, LL, NN, NN2)).

	CCa	CCb	CCc	DDa	DDb	DDc	LLa	LLb	LLc	NNa	NNb	NNc	NN2c
CC	-	-	-	1.6	1.8	1.4	0.9	1.4	1.8	2.0	2.0	2.0	2.0
DD	1.6	1.8	1.4	-	-	-	1.7	1.9	1.9	2.0	2.0	2.0	1.9
LL	0.9	1.4	1.8	1.7	1.9	1.9	-	-	-	2.0	2.0	2.0	1.8
NN	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	-	-	-	2.0
NN2	-	-	2.0	-	-	1.9	-	-	1.8	-	-	2.0	-

Our main target vegetation class, *Pittosporum* woodland (highly aggressive IAS) showed a maximum separability value (2.0) when using 4 classes (unique NN), and only a fairly good separability with LL when splitting NN into NN and NN2. This slight decrease of separability between NN2 and LL classes can be explained by the higher similarity between less developed *Pittosporum* woodland and Native scrubland's structures. All the other classes benefited from the depuration process by reaching a good or fairly good separability from each other. The only exceptions are:

- *Cryptomeria japonica* (CC) vs. Native scrubland (LL) in Dataset A (4 classes), which separability improvement (0.9 to 1.4) is insufficient (<1.7). This low separability could be explained by a significant portion of training sites located in CC's forest stands marginal and transitional areas to LL patches.
- *Cryptomeria japonica* (CC) vs. Bare Soil/Landslide areas (DD) in Dataset B (5 classes), which separability decreased (1.6 to 1.4). This lower separation between these two classes could be explained by the presence of some misclassified training sites.

The main separability issues involving our main target vegetation class *Pittosporum* woodland patches that were identified by Gil et al. [2011] have been addressed and solved by increasing the quality of the training data, by applying the depuration process. Nevertheless, the problems involving *Cryptomeria japonica* require increasing the quantity and quality of CC's training data, and by including less sites located in marginal and transitional areas to others land cover/vegetation categories.

Segmentation-based Supervised Classification

At overall level, the best overall classifications when using Training Dataset TA (4 classes) were obtained by applying classifiers MLC (Overall KIA = 0.97) and SVM (Overall KIA = 0.93), showing a strong agreement and a good accuracy (Overall KIA ≥ 0.8). KNN classification was slightly less accurate (Overall KIA = 0.74).

When using Training Dataset TB (5 classes), the overall accuracy results were slightly less accurate. The best overall classifications with this dataset were also obtained by applying MLC (Overall KIA = 0.93) and SVM (Overall KIA = 0.90). KNN classification didn't reach a strong agreement (Overall KIA = 0.72) (Tab. 2). At overall level, the segmentation-based approach used in this study has shown a significantly higher accuracy when compared to standard pixel-based approach performed by Gil et al. [2011] for the same area. MLC's overall accuracy (0.78) and MLC's overall Kappa (0.74) showed a good agreement, as it also occurred with SVM's overall accuracy (0.78) and SVM's overall Kappa (0.73).

Table 2 - Accuracy assessment of classification maps. (P: Producer Accuracy (0-1); U: User Accuracy (0-1); K: Kappa Index of Agreement).

Set		**A - With unique NN**			**B - Whith NN + NN2**	
	MLCa	**SVMa**	**KNNa**	**MLCb**	**SVMb**	**KNNb**
CC (P|U|K)	0.96 0.92 0.91	0.92 0.89 0.87	0.88 0.54 0.47	0.96 0.96 0.96	0.92 0.96 0.95	0.88 0.52 0.46
DD (P|U|K)	1.00 1.00 1.00	0.89 1.00 1.00	1.00 1.00 1.00	1.00 1.00 1.00	0.89 1.00 1.00	1.00 1.00 1.00
LL (P|U|K)	1.00 0.98 0.96	0.99 0.94 0.89	0.92 0.92 0.85	1.00 0.96 0.93	1.00 0.91 0.83	0.92 0.93 0.87
NN (P|U|K)	0.95 1.00 1.00	0.93 0.98 0.98	0.53 0.74 0.64	0.88 0.81 0.78	0.88 0.88 0.86	0.33 0.89 0.88
NN2 (P|U|K)	-	-	-	0.79 0.96 0.96	0.79 0.96 0.96	0.59 0.61 0.53
Overall Kappa	0.97	0.93	0.74	0.93	0.90	0.72
Overall Accuracy	0.98	0.95	0.82	0.95	0.93	0.80

At the land cover/vegetation class level, *Cryptomeria* forest (CC) mapping was more accurate when training MLC (K = 0.96) and SVM (K = 0.95) with Dataset TB (5 classes). When using Dataset TA (4 classes), MLC was the most accurate classifier (K = 0.91). KNN was the least effective classifier using both training datasets (K = 0.47 for A and K = 0.46 for B). The accuracy of both SVM and MLC results for CC in this study exceeded the best result (K = 0.88 with SVM) reached by Gil et al. [2011] under a standard pixel-based approach. Therefore, a segmentation-based approach should be considered in this archipelago for operational forestry decision-support regarding *Cryptomeria japonica* stands mapping, monitoring and management.

Landslide and Bare Soil areas (DD) mapping has proven to be highly accurate with all three classifiers and both datasets (K = 1.0). Those results exceeded the best one achieved by applying Mahalanobis Distance classifier (K = 0.71) when using a standard pixel-based approach for the same area [Gil et al., 2011]. Therefore this segmentation-based approach can be highly effective for landslides detection and monitoring when using high spatial resolution satellite imagery and its integration within Regional Natural Hazard Monitoring System should be seriously considered.

Native scrubland patches (LL) mapping was highly accurate when applying each one of the classifiers to both datasets ($0.83 \leq K \leq 0.96$), especially when using MLC (K for Dataset A = 0.96; K for Dataset B = 0.93).

Regarding the main goal of this study, which was to map accurately the spatial distribution of *Pittosporum* woodland in the Pico da Vara Nature Reserve, the results were very positive in two cases. First, when using Dataset A (single NN class), *Pittosporum* woodland mapping has been extremely accurate when applying MLC (K = 1.0) and SVM (K = 0.98) classifiers.

These results exceeded significantly the best one achieved by applying Artificial Neural Networks classifier (K = 0.66) under a standard pixel-based approach [Gil et al., 2011]. Second, when using Dataset B (NN split into NN and NN2), both sub-classes reached quite different accuracy results. KNN (K = 0.88) and SVM (K = 0.86) have been the most effective methods to map the first sub-class (NN), while MLC and SVM (K = 0.96) have proven to be highly accurate when mapping the second sub-class (NN2). KNN was ineffective to map accurately this sub-class (K = 0.53). These results showed that creating the Dataset B (with 5 classes) to split the NN initial class in two sub-classes (NN and NN2) had no positive effects on the accuracy of mapping at overall and land cover/vegetation class levels, excepting for the case of *Cryptomeria* forests (K = 0.96), which was slightly more accurate than by using Dataset A with 4 classes (K = 0.91).
To perform a cost-benefit analysis to support a more realistic and cost-effective *Pittosporum* woodland control and management in Pico da Vara Nature Reserve under PRECEFIAS development, the approximated total area occupied by this invasive woody species has been estimated based on the best classification performed using MLC and dataset TA (K for *Pittosporum* woodland = 1.0). Being the IKONOS multispectral band's spatial resolution equal to 4m (16 m² of minimal area unit), according to MLC results (Fig. 3), these IAS patches occupy 126 hectares of Pico da Vara protected area (815 hectares in total). According to the Regional Forest Inventory (RFI) the estimated area of *Pittosporum* Woodland in Pico da Vara Nature Reserve in 2007 is 35 hectares only.

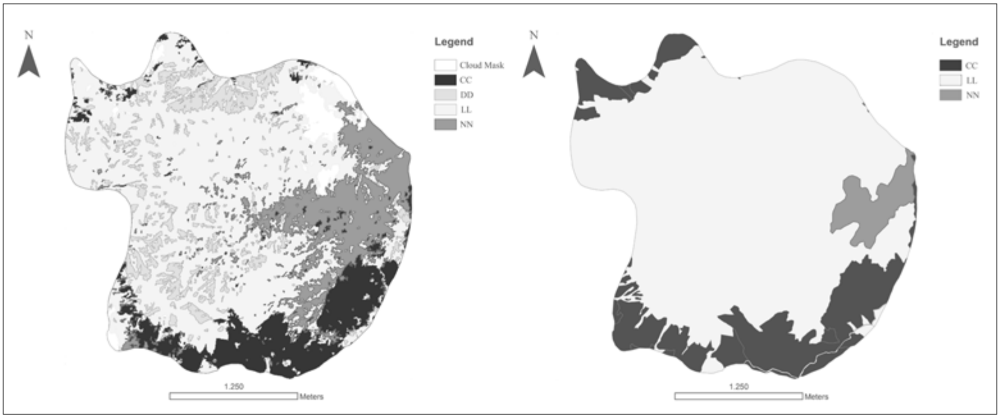

Figure 3 - Comparison between classification map obtained by applying MLC to dataset A (left) and Regional Forest Inventory mapping for Pico da Vara Nature Reserve (right).

This relevant difference between our and RFI results (Fig. 3) can be explained by some different factors. The remoteness and lack of access to the whole Pico da Vara Nature Reserve area make intensive, effective and accurate fieldwork difficult. At the cartographic level, the difference between the minimum spatial unit of the RFI (1 ha) and the spatial resolution of the IKONOS image (16 m²) could have implied the omission of many small *P. undulatum* patches in the FRI that were included in the satellite image classification process. Furthermore, the low quality of 1997's black and white orthophotomaps that were the first cartographic basis of RFI makes the photo-interpretation task very difficult and

resulted in low accuracy. Also important is the high frequency of extremely shadowed areas within the study area in real-color orthophotomaps of 2004. Finally, the eventual spread of *Pittosporum* woodland from 1997 (year of production of the first collection of orthophotomaps) to 2005 (year of production of the IKONOS image) cannot be ignored.
Therefore, the use of very high spatial resolution multispectral imagery (as IKONOS) and the application of this segmentation-based classification scheme should be integrated as core components of PRECEFIAS decision-support system. This methodological proposal constitutes a cost-effective solution for woody IAS patches (as Pittosporum woodland) detection, assessment and monitoring in the Azorean Protected Areas. Furthermore, this segmentation-based approach has proven to be particularly effective and accurate for forest areas mapping, exceeding significantly the results obtained in the same area by using the standard pixel-based approach [Gil et al., 2011], addressing and solving the most relevant separability issues between land-cover/vegetation classes. To address and map most vegetation classes in the Azores Islands, more categories (and more geographically dispersed training data) should be added to the classification scheme when mapping larger Protected Areas than Pico da Vara Natural Reserve (815 hectares). A good example of a relevant vegetation class in the Azores which should be added to a broader vegetation mapping scheme is “Exotic Woodland”, constituted by a mixture of woody species dominated by the non-indigenous *Acacia longifolia* and *Eucalyptus globulus*.

Conclusions

The results have shown that applying a segmentation-based approach to very high spatial resolution multispectral data (as IKONOS) can constitute a cost-effective approach for a continuous assessment of woody IAS spatial distribution and spread within Small Islands Protected Areas. The segmentation-based approach combined to the depuration of the initial field training dataset by interactive deletion of outliers allowed a significant increase in the separability between most relevant land-cover/vegetation classes, addressing and solving the most important problems that were identified in previous studies in the same area. Our main target vegetation class, *Pittosporum* woodland (highly aggressive IAS) showed a maximum separability value (2.0) in all pairwise comparisons when using a 4 class’s legend.
At overall level, SVM and MLC classifications showed a strong agreement and a good accuracy (Overall KIA $\geq$ 0.90). KNN classifications showed a good agreement and a lower accuracy (Overall KIA $\geq$ 0.70). At land-cover/vegetation class level, Native scrubland patches (LL) mapping was highly accurate when applying all classifiers ($0.83 \leq K \leq 0.96$). SVM, MLC and KNN have also proven to be highly accurate (K = 1.0) when mapping Bare Soil / Landslide areas (DD). Therefore this segmentation-based approach could be integrated into the Regional Natural Hazards Monitoring System to detect and map landslides. *Cryptomeria* forest (CC) mapping has been highly accurate when applying MLC and SVM (K > 0.90). Therefore, a segmentation-based approach should be considered for operational forestry decision-support regarding *Cryptomeria japonica* stands mapping, monitoring and management.
Pittosporum woodland (NN) mapping has been extremely accurate when applying MLC (K = 1.0) and SVM (K = 0.98) classifiers. Therefore, this methodological approach has proven to be extremely effective to map this aggressive woody IAS. These results will allow regional authorities to perform a more realistic, adequate and cost-effective *Pittosporum* woodland management in Azorean Protected Areas under PRECEFIAS development, by clearly identifying

the location, the dimension and the logistics constraints associated to the sites to be intervened and managed. Additionally, this new and more detailed data on *Pittosporum undulatum* spatial distribution will allow more accurate ecological modeling studies of this IAS in the Azores.
The comparison between our MLC's *Pittosporum* woodland mapping (126 hectares) and the *Pittosporum* woodland mapping performed on behalf of the photo-interpretation and fieldwork-based Regional Forest Inventory (35 hectares) for the same area (Pico da Vara Nature Reserve) has shown a very significant difference (91 hectares). This can be explained by some factors that interfered with an effective and accurate fieldwork and photo-interpretation tasks performed during the RFI production: (1) remoteness and lack of access to some relevant forest areas; (2) Low quality of orthophotomaps; (3) the continuous spread of *Pittosporum* Woodland; (4) the difference between the minimum spatial unit of the RFI (1 ha) and the spatial resolution of IKONOS image (16 m²).
Finally, this segmentation-based vegetation mapping approach could be easily integrated into other Protected Areas monitoring schemes by addressing their specific decision-support requirements and by mapping their most representative land-cover/vegetation classes. Protected Areas located in Macaronesian archipelagos as Madeira (Portugal) and Canary Islands (Spain) could especially benefit from the adaptation and application of this methodological approach because of their similarities with Azores in terms of existing vegetation classes and structure.

Acknowledgements

This research has been developed on the behalf of a Doctoral (M3.1.2/F/025/2007) and a Post-Doctoral (M3.1.7/F/005/2011) Research Project, both supported by the Regional Fund for Science (FRC) of the Azorean Regional Government. This particular case-study research has been developed with the support of project Woody Biomass, funded by the Green Islands program (MIT Portugal and Azorean Regional Government). We would like to thank António Moniz and Rita Godinho (University of the Azores) for their cartographic support. We would like to thank SPEA/LIFE Priolo Project's staff for sharing their vegetation monitoring program survey data. We are also thankful to the R Development Core Team, and to T.H. Keitt, R. Bivand, E. Pebesma and B. Rowlingson for the development of *rgdal;* R.J. Hijmans and Jacob van Etten, for the development of package *raster,* Evgenia Dimitriadou, Kurt Hornik, Friedrich Leisch, David Meyer and Andreas Weingessel for package *e1071* and to the CNES of France for the Orfeo Toolbox.

References

Bernstein L.S., Adler-Golden S.M., Sundberg R.L., Levine R.Y., Perkins T.C., Berk A. (2005) - V*alidation of the QUick Atmospheric Correction (QUAC) algorithm for VNIR-SWIR multi - and hyperspectral imagery*. In: Shen S.S., Lewis P.E. (Eds.), SPIE Proceeding Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XI, 5806: 668-678.

Bibby C.J., Charlton T.D. (1991) - *Observations on the São Miguel Bullfinch*. Acoreana, 7: 297-304.

Brown M., Lewis H.G., Gunn S.R. (2000) - *Linear spectral mixture models and support vector machines for remote sensing*. IEEE Transactions on Geoscience & Remote Sensing, 38: 2346-2360. doi: http://dx.doi.org/10.1109/36.868891.

Calado H., Quintela A, Porteiro J. (2007) - *Integrated Coastal Zone Management Strategies* *on Small Islands*. Journal of Coastal Research, SI 50: 125 - 129. Gold Coast, Australia.

Carson H.W., Lass L.W., Callihan R.H. (1995) - *Detection of yellow hawkweed (Hieracium pratense) with high resolution multispectral imagery*. Weed Technology, 9 (3): 477-483.

Casady G.M., Hanley R.S., Seelan S.K. (2005) - *Detection of leafy spurge (Euphorbia esula) using multidate high-resolution satellite imagery*. Weed Technology, 19: 462-467. doi: http://dx.doi.org/10.1614/WT-03-182R1.

Castro S.A., Daehler C.C., Silva L., Torres-Santana C.W., Reyes-Betancort J.A., Atkinson R., Jaramillo P., Guezou A., Jaksic F.M. (2010) - *Floristic homogenization as a teleconnected trend in oceanic islands*. Diversity and Distributions. doi: http://dx.doi.org/10.1111/j.1472-4642.2010.00695.x.

Caujapé J., Tye A.T., Crawford D.J., Santos-Guerra A., Sakai A., Beaver K., Lobin W., Florens F.B.V., Moura M., Jardim R., Gomes I., Kueffer C. (2010) - *Conservation of oceanic island floras: present and future global challenges*. Perspectives in Plant Ecology, Evolution and Systematics, 12: 107-129. doi: http://dx.doi.org/10.1016/j.ppees.2009.10.001.

CBD (2006) - *Eighth Ordinary Meeting of the Conference of the Parties to the Convention on Biological Diversity,* Decision VIII/1, appendix - Goal 6, activity: 6.1.1.3.

Ceia R., Sampaio H., Parejo S., Heleno R., Arosa M.L., Ramos J.A., Hilton G. (2011) - *Throwing the baby out with the bathwater: does laurel forest restoration remove a critical winter food supply for the critically endangered Azores bullfinch?* Biological Invasions, 13: 93-104. doi: http://dx.doi.org/10.1007/s10530-010-9792-x.

Chirici G., Corona P., Marchetti M., Mastronardi A., Maselli F., Bottai L., Travaglini D. (2012) - *K-NN FOREST: a software for the non-parametric prediction and mapping of environmental variables by the k-Nearest Neighbors algorithm.* European Journal of Remote Sensing, 45: 433-442. doi: http://dx.doi.org/10.5721/EuJRS20124536.

Christophe M.E., Michel J., Inglada J. (2011) - *Remote Sensing Processing: From Multicore to GPU*. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4 (3): 643-652. doi: http://dx.doi.org/10.1109/JSTARS.2010.2102340.

Commaniciu C. (2002) - *Mean shift: A robust approach toward feature space analysis.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 24 (5): 603-619. doi: http://dx.doi.org/10.1109/34.1000236.

Congalton R., Green K. (1999) - *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, FL. CRC/Lewis Press, pp. 137.

De Leeuw J., Georgiadou Y., Kerle N., de Gier A., Inoue Y., Ferwerda J., Smies M., Narantuya D. (2010) - *The function of remote sensing in support of environmental policy*. Remote Sensing, 2: 1731-1750. doi: http://dx.doi.org/10.3390/rs2071731.

Dias E. (1996) - V*egetação Natural dos Açores*. PhD Thesis. Universidade dos Açores, Angra do Heroísmo, pp. 302.

Dimitriadou E., Hornik K., Leisch F., Meyer D., Weingessel A. (2010) - *e1071: Misc Functions of the Department of Statistics (e1071), TU Wien*. R package version 1.5-24.

Everitt J.H., Yang C., Deloach C.J. (2006) - *Remote sensing of giant reed with QuickBird satellite imagery.* Journal of Aquatic Plant Management, 43: 81-85.

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: 1336-1343. doi: http://dx.doi.org/10.1109/TGRS.2004.827257.

Foody G.M. (2002) - *Status of land cover classification accuracy assessment*. Remote Sensing

of Environment, 80: 185-201. doi: http://dx.doi.org/10.1016/S0034-4257(01)00295-4.
Franco-Lopez H., Ek A.R., Bauer M.E. (2001) - *Estimation and mapping of forest stand density, volume, and cover type using the knearest neighbors method*. Remote Sensing of Environment, 77: 251-74. doi: http://dx.doi.org/10.1016/S0034-4257(01)00209-7.
Ganas A., Lagios E., Tzannetos N. (2002) - *An investigation on the spatial accuracy of the IKONOS 2 orthoimagery within an urban environment*. International Journal of Remote Sensing, 23: 3513-3519. doi: http://dx.doi.org/10.1080/01431160210127047.
Gil A., Fonseca C., Lobo A., Calado H. (2012) - *Linking GMES Space Component to the development of land policies in Outermost Regions - the Azores (Portugal) case-study.* European Journal of Remote Sensing, 45: 263-281. doi: http://dx.doi.org/10.5721/EuJRS20124524.
Gil A., Yu Q., Lobo A., Lourenço P., Silva L., Calado H. (2011) - *Assessing the effectiveness of high resolution satellite imagery for vegetation mapping in small islands protected areas*. Journal of Coastal Research, 64: 1663-1667.
Goodland T., Healey J.R. (1996) - *The invasion of Jamaican montane Rainforests by the Australian Tree Pittosporum undulatum*. School of Agricultural and Forest Sciences, University of Wales, Bangor, pp. 54.
Hale S.R., Rock B.N. (2003) - I*mpacts of topographic normalization on land-cover classification accuracy*. Photogrammetric Engineering & Remote Sensing, 69: 785-792.
He C., Zhang Q., Li Y., Li X., Shi P. (2005) - *Zoning grassland protection area using remote sensing and cellular automata modeling - A case study in Xilingol steppe grassland in northern China.* J Arid Environ, 63 (4): 814-826. doi: http://dx.doi.org/10.1016/j.jaridenv.2005.03.028.
Hijmans R.J., Van Etten J. (2011) - *Raster: Geographic analysis and modeling with raster data*. R package version 1.1.2/r948.
Hortal J., Borges P.A.V., Jiménez-Valverde A., Azevedo E.B., Silva L. (2010) - *Assessing the areas under risk of invasion within islands through potential distribution modeling: the case of Pittosporum undulatum in São Miguel, Azores*. Journal for Nature Conservation, 18: 247-257. doi: http://dx.doi.org/10.1016/j.jnc.2009.11.002.
Huang C., Asner G.P. (2009) - *Applications of Remote Sensing to Alien Invasive Plant Studies*. Sensors, 9: 4869-4889. doi: http://dx.doi.org/10.3390/s90604869.
Katoh M. (2004) - *Classifying tree species in a northern mixed forest using high-resolution IKONOS data*. Journal of Forestry Research, 9 (1): 7-14. doi: http://dx.doi.org/10.1007/s10310-003-0045-z.
Kueffer C., Daehler C.C., Torres-Santana C.W., Lavergne C., Meyer J.Y., Rüdiger O., Silva L. (2010) - *Magnitude and form of invasive plant impacts on oceanic islands: a global comparison*. Perspectives in Plant Ecology, Evolution and Systematics, 12: 145-161. doi: http://dx.doi.org/10.1016/j.ppees.2009.06.002.
Lenka C., Milan C. (2005) - *Supervised classification of plant communities with artificial neural networks*. Journal of Vegetation Science, 16: 407-414. doi: http://dx.doi.org/10.1111/j.1654-1103.2005.tb02380.x.
Lobo A. (1997) - I*mage segmentation and discriminant analysis for the identification of land cover units in ecology*. IEEE Transactions on Geoscience and Remote Sensing, 35 (5): 1136-1145. doi: http://dx.doi.org/10.1109/36.628781.
Lourenço P., Medeiros V., Gil A., Silva L. (2011) - *Distribution habitat and biomass of Pittosporum undulatum, the most important woody plant invader in the Azores Archipelago*. Forest Ecology and Management, 262 (2): 178-187. Environmental Stress

and Forest Ecosystems: Case studies from Estonia, 15 July 2011. doi: http://dx.doi.org/10.1016/j.foreco.2011.03.021.

Mäkelä H., Pekkarinen A. (2001) - *Estimation of timber volume at the sample plot level by means of image segmentation and Landsat TM imagery*. Remote Sensing of Environment, 77: 66-75. doi: http://dx.doi.org/10.1016/S0034-4257(01)00194-8.

Parker I.M., Simberloff D., Lonsdale W.M., Goodell K., Wonham M., Kareiva P.M., Williamson M.H., Holle V., Moyle P.B., Byers J.E., Goldwasser L. (1999) - *Impact: toward a framework for understanding the ecological effects of invaders*. Biological Invasions, 1: 3-19. doi: http://dx.doi.org/10.1023/A:1010034312781.

R Development Core Team (2010) - *R: A Language and Environment for Statistical Computing*.

Richards J.A., Jia X. (1999) - *Remote Sensing Digital Image Analysis*. Springer-Verlag, Berlin, pp. 363. doi: http://dx.doi.org/10.1007/978-3-662-03978-6.

Rodriguez-Echeverría S., Crisóstomo J.A., Nabais C., Freitas H. (2009) - *Belowground mutualists and the invasive ability of Acacia longifolia in coastal dunes of Portugal*. Biological Invasions, 11 (3): 651-661. doi: http://dx.doi.org/10.1007/s10530-008-9280-8.

Silva L., Ojeda-Land E., Rodríguez-Luengo J.L. (2008) - *Invasive terrestrial flora and fauna of Macaronesia*. Top 100 in Azores, Madeira and Canaries. ARENA, Ponta Delgada, pp. 546.

Silva L., Smith C.W. (2006) - A *quantitative approach to the study of non-indigenous plants: an example from the Azores Archipelago*. Biodiversity and Conservation 15: 1661-1679. doi: http://dx.doi.org/10.1007/s10531-004-5015-z.

Silva L., Smith C. (2004) - A *characterization of the non-indigenous flora of the Azores Archipelago*. Biological Invasions, 6: 193-204. doi: http://dx.doi.org/10.1023/B:BINV.0000022138.75673.8c.

Silva L. (2001) - *Plantas vasculares invasoras no Arquipélago dos Açores. Caracterização geral e estudo de um caso: Clethra arborea Aiton (Cletheraceae)*. PhD Thesis. Universidade dos Açores. Ponta Delgada, Portugal, pp. 541.

Sjogren E. (1973) - R*ecent changes in the vascular flora and vegetation of the Azores Islands*. Memórias da Sociedade Broteriana, 22: 1-113.

Sohn Y., Rebello N.S. (2002) - *Supervised and unsupervised spectral angle classifiers*. Photogrammetric Engineering & Remote Sensing, 68: 1271-1280.

Strand H., Höft R., Strittholt J., Miles L., Horning N., Fosnight E., Turner W. (2007) - *Sourcebook on Remote Sensing and Biodiversity Indicators*. Secretariat of the Convention on Biological Diversity, Montreal, Technical Series 32, pp. 203.

Teodosio J., Ceia R., Costa L.T. (2009) - S*pecies action plan for the Azores Bulfinch Pyrrhula murina in the European Union,* pp. 19.

Tsai F., Chou M.J. (2006) - *Texture Augmented Analysis of High Resolution Satellite Imagery in Detecting Invasive Plant Species*. Journal of the Chinese Institute of Engineers, 29 (4): 581-592. doi: http://dx.doi.org/10.1080/02533839.2006.9671155.

Turner W., Spector S., Gardiner N., Fladeland M., Sterling E., Steininger M. (2003) - *Remote sensing for biodiversity science and conservation*. Trends in Ecology and Evolution, 18: 306-314. doi: http://dx.doi.org/10.1016/S0169-5347(03)00070-3.

Venables W.N., Ripley B.D. (2002) - *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. doi: http://dx.doi.org/10.1007/978-0-387-21706-2.

Xie Y., Sha Z., Yu M. (2008) - *Remote sensing imagery in vegetation mapping: a review.* Journal of Plant Ecology, 1 (1): 9-23. doi: http://dx.doi.org/10.1093/jpe/rtm005.

Xu M., Watanachaturaporn P., Varshney P.K., Arora M.K. (2005) - *Decision tree regression for soft classification of remote sensing data*. Remote Sensing of Environment, 97: 322-336. doi: http://dx.doi.org/10.1016/j.rse.2005.05.008.

Yu Q., Gong P., Clinton N., Biging G., Kelly M., Schirokauer D. (2006) - *Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery*. Photogrammetric Engineering & Remote Sensing, 72 (7): 799-811.

Received 11/12/2011, accepted 09/03/2013