Assessing the effectiveness of high resolution satellite imagery for vegetation mapping in small islands protected areas

A. Gil †, Q. Yu ‡, A. Lobo §, P. Lourenço †, L. Silva ‡‡ and H. Calado ‡‡

† CITA-A (Azorean Biodiversity Group), Ponta Delgada, 9501-801 Portugal; arturgil@uac.pt
‡ Department of Geosciences, University of Massachusetts, Amherst, MA 01003-9297 USA
§ ICTJA -Institut de Ciencies de la Terra Jaume I, E-08028 Spain
‡‡ Departamento de Biología Vegetal y Ecología, Centro Andaluz para la Evaluación y Seguimiento del Cambio Global, Universidad de Almería, 04120 Spain
‡‡‡ CIBIO (Azores Unit), Ponta Delgada 9501-801 Portugal

ABSTRACT


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. However, existing indigenous vascular plant taxa are particularly important because they compose a very valuable ecosystem, the Azorean Laurel Forest. One of its most significant areas is the core of Pico da Vara/Ribeira do Guilherme Special Protected Area, in the former Natural Reserve of Pico da Vará, located in the mountain complex of Serra da Tronqueira. The rapid spread of some very aggressive invasive alien species, such as Pittosporum undulatum Vent. and Clethra arborea Aiton, are causing serious damages to this ecosystem. Its direct competition with native species has resulted in a significant decline in native populations and ecosystem area. This paper assessed the effectiveness of High Spatial Resolution IKONOS satellite imagery for vegetation mapping in Pico da Vara Natural Reserve using four different supervised classification techniques: Support Vector Machine, Artificial Neural Networks (non-parametric methods), Mahalanobis Distance and Maximum Likelihood (parametric methods). The overall classification results have shown that remote sensing based vegetation mapping using IKONOS image can constitute a cost-effective approach for a continuous monitoring, characterization and assessment of these insular ecosystems. Despite the poor separability (Transformed Divergence < 1700) of some vegetation categories, Maximum Likelihood, Support Vector Machine and Artificial Neural Networks classifications have achieved good overall and per class accuracies (overall accuracy > 75% and Kappa Index Agreement > 0.6).

ADDITIONAL INDEX WORDS: Remote Sensing, Vegetation Mapping, High Resolution, Classification.

INTRODUCTION

Islands, Protected Areas and Biological Invasions

Due to their geographical framing as well as their morphological and ecological characteristics, small islands are more vulnerable to certain threats and phenomena such as climate variability and changes, proliferation of invasive exotic species, natural catastrophes and overexploitation of natural resources (Rietbergen, 2008). The effectiveness of protected areas regarding the associated goals of nature conservation and biodiversity maintainability has been widely studied and questioned. An even more central and core-based role of these spaces in the protection and preservation of biodiversity has been demanded (Hockings et al., 2005).

Biological invasions are one of the most important problems faced by island ecosystems, due to the peculiarities of the native island fauna and flora (Olesen et al. 2002). 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). However, existing indigenous vascular plant taxa are particularly important because they compose a very valuable ecosystem, the Azorean Laurel Forest. One of its most significant areas is the core of Pico da Vara / Ribeira do Guilherme Special Protected Area, in the former Natural Reserve of Pico da Vara (the case-study area), located in the mountain complex of Serra da Tronqueira. The rapid spread of some very aggressive invasive alien species, such as Pittosporum undulatum Vent. and Clethra arborea Aiton, are causing serious damages to this ecosystem (Hortal et al., 2010). Their direct competition with native species has resulted in a significant decline in native populations and ecosystem area.

Remote Sensing of Vegetation using High Spatial Resolution IKONOS satellite imagery

Consistent mapping of vegetation types requires vegetation classification being developed beforehand. Vegetation mapping is the process of delineating the geographic distribution, extent, and landscape patterns of vegetation types and/or structural characteristics. Patterns of vegetation types are best recognized after the types have been defined and described. A vegetation map unit is a collection of areas defined and named the same in terms of their component taxonomic units and/or technical groups (Brohman and Bryant, 2005). Vegetation mapping is critical to
obtain current states of vegetation cover in order to initiate 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 effective to acquire vegetation covers because they are time consuming, date lagged and often too expensive.

The technology of remote sensing offers a practical and economical means to study vegetation cover changes, especially over large areas (Langley et al. 2001; Nordberg and Everston 2003). Remote sensing provides repeated and consistent assessment and monitoring of the environment; it allows independent control and its quality can be assessed. As such, it is a tool with some very desirable characteristics for supporting environmental policy (De Leeuw, 2010). Using satellite data to classify and map various forest and/or land-use types has historically been the most frequent use of satellite data (Iverson et al., 1989). Scale, image resolution, and the user’s need are the most important factors affecting the selection of remotely sensed data. The user’s need determines the nature of classification and the scale of the study area, thus affecting the selection of suitable spatial resolution of remotely sensed data. The idea in this approach is to pinpoint these species based on their unique spatial textures/patterns or phenological characteristics. The IKONOS System is a commercial satellite from GeoEye and offers multispectral imagery at a spatial resolution of 4 meter and panchromatic imagery at 1 meter. Furthermore, IKONOS spectral resolution (Visible/Near Infrared), revisit time (3-5 days off-nadir and 144 days for true-nadir) and swath (11 km x 11 km each single scene) make it an interesting solution for species composition, land cover, phenology, habitat structure and primary productivity mapping (Huang and Asner, 2009; Xie et al., 2008; Tsai and Chou, 2006; Katoh, 2004; Turner et al., 2003).

**Satellite Imagery Classification Techniques**

Image classification, in a broad sense, is defined as the process of extracting differentiated classes or themes (e.g. land use categories, vegetation species) from raw remotely sensed satellite data (Xie et al., 2008). A supervised classification method is learning an established classification from a training dataset, which contains the predictor variables measured in each sampling unit and assigns prior classes to the sampling units (Lenka and Milan, 2005).

Maximum-Likelihood Classifier (ML) and Mahalanobis Distance (MD) classifiers are 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). Both classifiers can show less satisfactory successes because their assumption that the data follow Gaussian distribution may not always be held in complex areas (Kapoor et al. 2010; Xie et al., 2008).

Artificial Neural Network (ANN) is appropriate for the analysis of nearly any kind of data regardless their statistical properties. ANN is very useful in extracting vegetation-type information in complex vegetation mapping problems (Filippi and Jensen 2006), although it is at the expense of the interpretability of the results. ANN deploys a black-box approach which hides the underlying prediction process (Cerna and Chytry, 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 a very advantageous feature, 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 (Foody and Mathur, 2004; Brown et al., 2000).

**MATERIALS AND METHODS**

**The Case-Study Area: Pico da Vara Natural Reserve**

The Pico da Vara / Ribeira do Guilherme Special Protected Area was classified in 1999 (Decree-Law 140/99 of April 24th) due to the presence and conservation status of the Azores Bullfinch Pyrrhula murina Godman, 1866; an endemic bird to São Miguel Island in the Azores. Its estimated population is of 500-800 pairs (Teodosio et al., 2009), limited to only a few fragments of remaining native vegetation. It is one of the most threatened passerine birds in all Europe, and is currently listed in the Red List Data Book as Endangered species.

This Special Protected Area (SPA) is located in São Miguel Island, the largest island on the Archipelago of the Azores, and it comprises the municipalities of Nordeste at north and Povoação at south. In the present the SPA covers an area of 6,067.27 hectares that represents 28.3% of the total area of the two municipalities. This SPA comprises one of the last main areas of native Laurel Forest in São Miguel, whose most important sub-area is located in the former Natural Reserve of Pico da Vara in the mountain complex of Serra da Tronqueira (Figure 1).

**Satellite Imagery and Ground-truth Data**

Three different data sets were used in our study:

- An IKONOS-2 standard geometrically corrected image with four multispectral bands (Blue, Green, Red and Near-Infrared) acquired on Aug 18, 2005 with 11% Cloud Cover (Figure 2);
- The Digital Terrain Model (DTM) of S.Miguel Island with an equidistance of 10m between contour isolines, produced in GIS format by the Military Geographic Institute of Portugal (IGEOE) at a 1/25000 scale, used for the ortho-rectification of the IKONOS-2 image;
- A main GIS data set of 1292 polygonal test sites covering the Pico da Vara Natural Reserve, divided in an aleatory way in two smaller shapefiles: the larger one containing 861 features (2/3 of total features) for training and the other smaller one containing 431 for test (1/3 of total features). All those 1292 polygonal test-sites are the result of a simultaneous interpretation and cross-validation of the available IKONOS-2 image (August 2005) with the SPEA/LIFE Priolo Project’s vegetation monitoring program survey (A9 Action), continuously developed between 2004 and 2008 (Teodosio et al., 2009). Eight representative classes are included in both data sets:
1) Ortho-rectification of the four IKONOS-2 image multispectral bands using the Digital Terrain Model (DTM) of S.Miguel Island;
2) Atmospheric correction of the four IKONOS-2 image multispectral bands using the QAC (Quick Atmospheric Correction) module within ENVI 4.7® software;
3) Training and test site generation by interpretation and cross-validation of the IKONOS-2 image with the SPEA/LIFE Priolo Project’s vegetation monitoring program survey (A9 Action), continuously developed between 2004 and 2008 (Teodosio et al., 2009);
4) Separability analysis between pairwise endmembers (8 categories of the Pico da Vara Natural Reserve vegetation map) using “Transformed Divergence” index (Lillesand and Kiefer, 2000);
5) Supervised classification of the IKONOS-2 Ortho-Rectified and Atmospherically Corrected Image Multispectral bands (VIS/NIR) using 4 different classifiers within:
   i. Maximum Likelihood (parametric method)
   ii. Mahalanobis Distance (parametric method)
   iii. Artificial Neural Network (non parametric method): logistic activation, training threshold contribution and training momentum = 0.9, training rate = 0.2, training RMS exit criteria = 0.1, number of hidden layers = 1, number of training iterations = 1000
   iv. Support Vector Machine (non parametric method): Kernel Type = Radial Basis Function, Gamma in Kernel function = 0.25, penalty parameter = 200, pyramid levels= 1
6) Accuracy assessment of classification maps by computing user's accuracy, producer's accuracy and overall Kappa coefficient.

RESULTS AND DISCUSSION

It is important to examine the separability before classification. Transformed Divergence values are calculated with training data. As a general rule, if the TD value is greater than 1900, then classes can be separated; between 1700 and 1900, the separation is fairly good; below 1700, the separation is poor (Lillesand and Kiefer, 2000). Regarding the 8 classification categories described above (CC, CL, DD, LL, NN, NU, PN, SO), the separability is fairly good or very good in all pairwise combinations except for the following ones:

Methodology

This paper intends to assess the effectiveness of High Spatial Resolution IKONOS satellite imagery VIS-NIR spectral bands for vegetation mapping of Pico da Vara Natural Reserve, using, comparing and assessing the accuracy of four different supervised classification techniques: Support Vector Machine (SVM), Neural Networks (non-parametric methods), Mahalanobis Distance and Maximum Likelihood techniques (parametric methods). In order to achieve this goal, a standard land cover classification approach (Lu et al., 2004) has been applied to the case-study IKONOS image (Figure 3):
Four supervised classification techniques have been applied to the ortho-rectified and atmospherically corrected IKONOS-2 image (Figure 4). In order to assess effectively the accuracy of these 4 classifications maps, User Accuracy (0-100%), Producer Accuracy (0-100%), Kappa Index of Agreement (0-1) have been calculated for each map category (Table 1). Overall Accuracy (0-100%, with 95% Confidence Interval) and “Overall Kappa Index of Agreement” (0-1) have also been calculated for each classifier (Lillesand and Kiefer, 2000):

- MD: Overall Accuracy = 66.04 – 68.09%; Overall Kappa = 0.61 (good agreement);
- ML: Overall Accuracy = 76.93 – 78.75%; Overall Kappa = 0.74 (good agreement);
- ANN: Overall Accuracy = 76.25 – 78.09%; Overall Kappa = 0.73 (good agreement);
- SVM: Overall Accuracy = 76.95 – 78.76%; Overall Kappa = 0.73 (good agreement).

The separability analysis using Transformed Divergence examines whether IKONOS multispectral imagery is effective for vegetation mapping in Azores. The poor separation (TD<1700) between both Pittosporum undulatum Vent. and Cryptomeria japonica (L. fil.) D. Don woody forests could be a serious obstacle to produce accurate vegetation maps, where both species are largely represented. The poor separability between Pittosporum undulatum Vent. and Clethra arborea can be explained by the arbustive structure that both species present in higher-altitude patches, where Pittosporum undulatum Vent. has much more difficulties in growing and spreading.

Finally, the poor separability between Natural pasture areas and Landslides/Bare Soil Areas is due to the low soil coverage where some patches of “Natural Pastures” can present because of their slope and consequent superficial erosion. An increased proportional number of test-sites for classification and the integration of ancillary GIS Data (elevation, slope, aspect, distance to water streams, soil type and lithology) could mitigate this separability issues between categories with quite similar greenness and texture (CC and NN; NN and CL; DD and PN). The overall classifications accuracies and Kappa Index of Agreement results showed that ML, SVM and ANN methods have all produced good and quite similar overall results (accuracy > 76% and good agreement), while MD performance was the worst of all, at both single class and overall levels. As highlighted values show in Table 1, at single class level, regarding user accuracy assessment, ANN performed a better classification of CC, CL, DD and SO, while SVM performed a better classification of NN and PN. LL was better classified using ML and NU using MD. Regarding Kappa Index assessment, SVM performed a better classification of CC, LL and NU, while ANN was better classifying NN and PN; MD for DD and SO; and, finally, ML for CL and NU. Therefore, non-parametric methods SVM and ANN performed better than parametric methods (ML and MD), mostly for the less separable (CC, NN, PN, DD and CL) and heterogeneous (PN and LL) classes. When classifying complex data sets, SVM and NN appear to be better options because they don’t assume any data distribution, while ML should be used with good results when data distribution is Gaussian.

CONCLUSIONS

The classification results have shown that using IKONOS multispectral imagery for vegetation mapping in Azores terrestrial Protected Areas can constitute a cost-effective approach for a continuous monitoring, characterization, and assessment of these insular ecosystems.
Due to the poor separability and high heterogeneity of some representative vegetation categories, classifications using SVM, ANN or ML techniques could be improved by increasing the quantity and quality of training sites, by integrating other features such as texture, or by using ancillary information, such as elevation, slope, aspect, distance to water streams, soil type and lithology.

LITERATURE CITED


Hortal, J., Borges, P.A.V., Jiménez-Valverde, A., Azevedo, E.B. and Silva, L., 2010. Assessing the areas under risk of invasion within islands through potential distribution modelling: the 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 SPEALIFE Priolo Project’s staff, especially Azucena Cruz, Carlos Silva, Rita Bastos, Rui Botelho and Sérgio Timóteo for sharing their vegetation monitoring program survey data and for being always available to discuss and comment on it.

Acknowledgement

This research has been developed on the behalf of a Doctoral Research Project led by the first author, whose main goal is to assess the effectiveness of processing and integrating remote sensing data in the framework of Small Islands’ Natural Resources Planning and Management instruments. This Ph.D. Research Project (M3.1.2/F/025/2007) is supported by the Azorean Regional Secretary of Science, Technology and Equipments. 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).