Urban Vegetation Cover Extraction from Hyperspectral Remote Sensing Imagery & GIS Spatial Analysis Techniques: The Case of Athens, Greece
Petropoulos, George; Kalivas, Dionyssos; Georgopoulou, Iro A.; Srivastava, Prashant K.

Published in:
Journal of Applied Remote Sensing
DOI: 10.1117/1.JRS.9.096088
Publication date: 2015

Citation for published version (APA):
Urban vegetation cover extraction from hyperspectral imagery and geographic information system spatial analysis techniques: case of Athens, Greece

George P. Petropoulos
Dionissios P. Kalivas
Iro A. Georgopoulou
Prashant K. Srivastava
Urban vegetation cover extraction from hyperspectral imagery and geographic information system spatial analysis techniques: case of Athens, Greece

George P. Petropoulos, Dionissios P. Kalivas, Iro A. Georgopoulou, and Prashant K. Srivastava

Abstract. The present study aimed at evaluating the performance of two different pixel-based classifiers [spectral angle mapper (SAM) and support vector machines (SVMs)] in discriminating different land-cover classes in a typical urban setting, focusing particularly on urban vegetation cover by utilizing hyperspectral (EO-1 Hyperion) data. As a case study, the city of Athens, Greece, was used. Validation of urban vegetation predictions was based on the error matrix statistics. Additionally, the final urban vegetation cover maps were compared at a municipality level against reference urban vegetation cover estimates derived from the digitization of very high-resolution imagery. To ensure consistency and comparability of the results, the same training and validation points dataset were used to compare the different classifiers. The results showed that SVMs outperformed SAM in terms of both classification and urban vegetation cover mapping with an overall accuracy of 86.53% and Kappa coefficient 0.823, whereas for SAM classification, the accuracy statistics obtained were 75.13% and 0.673, respectively. Our results confirmed the ability of both techniques, when combined with Hyperion imagery, to extract urban vegetation cover for the case of a densely populated city with complex urban features, such as Athens. Our findings offer significant information at the local scale as regards to the presence of open green spaces in the urban environment of Athens. Such information is vital for successful infrastructure development, urban landscape planning, and improvement of urban environment. More widely, this study also contributes significantly toward an objective assessment of Hyperion in detecting and mapping urban vegetation cover. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JRS.9.096088]

Keywords: Earth observation; geographical information systems; Hyperion; support vector machines; spectral angle mapper; Athens; Greece.

1 Introduction

Land use and land cover (LULC) constitute a key variable of the Earth’s system that has generally shown a close relationship with human activities and the physical environment. The detection and monitoring of LULC not only has an impact on the socioeconomic status of a country, but also is a key contributor to global climate change processes. Urban areas are continuously growing, as more than half of the Earth’s population seeks greater job prospects...
In addition to this, the estimated annual urban population growth rate of 1.78% is almost twice as fast as that of the global population. An important consequence of this is evident: the change of land cover types from natural to anthropogenic impervious surfaces consisting of roofs, roads, parking lots, driveways, and sidewalks. Indeed, the fast development of human societies, particularly since the beginning of the industrial revolution, has intensified different types of human activities, resulting in a continuous and noticeable influence on LULC. To ensure the high quality of urban environments, our interest in monitoring urban vegetation has recently increased considerably. Urban vegetation is one of the central infrastructural components of any urban ecosystem, and it plays a significant role within cities. However, urbanization is arguably the most dramatic land form transformation causing an irreversible loss of vegetation in urbanized environments. Due to an unprecedented rate of growth in urbanization over the last two to three decades, monitoring the urban environment has, consequently, become a worldwide issue; currently, in the majority of countries, urban development presents itself as a perpetual challenge for city planners, who lack the tools for measuring and controlling urban sprawl processes.

As discussed above, in urban areas, vegetated spaces are facing challenges because of escalated urbanization and environmental degradation. For green or eco-city planning, in the near future, plans for the effective conservation of green spaces will be required. Such spaces have the potential to mitigate the adverse effects of urbanization, helping to clean the air, reduce noise pollution, and stabilize the microclimatic conditions within cities. In fact, urban vegetation provides ecological, social, health, and economic benefits to a city’s inhabitants. Also, accurate mapping and monitoring of vegetation distribution within an urban environment is fundamental to the understanding of urban ecosystems and could be applied in regional scale models of climatic, hydrologic, and ecologic processes. A change of urban vegetation cover impacts significantly natural resources, especially natural habitat ecosystems through their effects on soil and water quality and climatic systems, resulting in serious environmental problems. After the development of remote sensing and satellite information systems, a number of datasets are now available which can be used to assist planners and managers in green city planning. Given the importance of urban vegetation cover, quantifying and mapping its distribution and changes within urban areas has also been recognized today as a highly significant topic in land change related sciences and, in particular, for green city planning within urban areas.

Earth observation (EO) technology provides a promising avenue for mapping and monitoring urban vegetation cover structures and their changes. EO has shown great potential in this direction, characterized by its ability to provide inexpensive, continuous synoptic views at a range of spatial and temporal scales. Satellite EO has the potential to provide accurate and timely geospatial information describing changes in LULC. Furthermore, its integration with geographic information systems (GIS) can provide an effective set of tools for data capture, storage, synthesis measurements, and analysis of spatial information to support reliable and consistent urban planning and decision making. The recent advances in EO radiometer technology have led to the launch of hyperspectral systems. These are sophisticated sensors able to record reflected light from land surface objects, ranging from the visible to the shortwave infrared (SWIR) parts of the electromagnetic spectrum, acquiring a vast amount of spectral information. Use of hyperspectral imagery has generally shown promising potential in terms of different land surface target identification and land cover mapping, particularly for urban vegetation cover.

Generally, a variety of techniques of different levels of complexity and assumptions have been utilized for mapping urban vegetation and its changes, with image classification being perhaps the most widely used approach. Briefly stated, the most widely used categorization of classification approaches includes three main groups, namely pixel-based, object-based, and subpixel classification techniques. Pixel-based techniques utilize the reflective characteristics of the land surface items and their associated spectral signatures in order to perform a classification by assigning pixels to land cover classes. The results of such pixel-based techniques can be achieved by either supervised or unsupervised classifiers. Unsupervised classifiers group pixels with similar spectral values into unique clusters according to statistically predefined criteria and reassign the generated spectral clusters into information classes.
classifiers use samples of a given identity for each land cover class, known as training sites, to classify image pixels of unknown identity. Supervised classifiers are also commonly divided into parametric and nonparametric, depending on whether they make assumptions about the probability of the different classes within the image scene. An example of a parametric classifier is the maximum likelihood (ML) algorithm. Among the most commonly used nonparametric classifiers are artificial neural networks (ANN), decision tree classifiers, and support vector machines (SVMs). Subpixel classification techniques work on a pixel level and aim to separate the image pixel into surface material fractions. The process involves the use of end-members or well-defined pure pixels that contribute to the varied mixtures of pixels. A number of studies have evaluated and compared the accuracy of different pixel-based classification methods for LULC mapping. For example, in a study by Petropoulos et al., the accuracy of the combined use of ANN and SAM classifiers with Landsat TM satellite imagery for burnt area mapping was examined. Results from their study indicated that overall, ANN outperformed (overall accuracy 90.29%, Kappa coefficient 0.878) the SAM classifier (overall accuracy 83.82%, Kappa coefficient 0.795). In a comparison study investigating LULC in a coastal area of northwestern Greece, Kolios and Stylios examined the accuracy of combining Landsat enhanced thematic mapper with ML, ANN, and SVMs classifying techniques. SVMs outperformed all other techniques with an overall accuracy of 96.25%, followed by ANN (96.07%) and ML (91.13%). In a similar study, Srivastava et al. focused on the comparison of three classification tools for Landsat images, ML, ANN, and SVMs, in detecting the LULC change over part of Walnut Creek, Iowa. However, in contrast to the study of Kolios and Stylios, Srivastava et al. found the ANN classification technique to be the most accurate. Therefore, more experimental and theoretical studies need to be explored to decide under what conditions ANN is better than SVMs, and vice versa. A large amount of effort has, thus, been directed toward developing and validating these sophisticated algorithms aiming to map LULC and evaluating the most accurate algorithm for use in different environments. This rapid development in image processing algorithms has also necessitated the assessment of their use with different types of EO data.

Hyperion is the first spaceborne imaging spectrometer acquiring spectral information in 242 spectral bands at the resolution of 30 m. Hyperion is onboard the Earth Observer-1 (EO-1) satellite platform, which was launched in 2000 under the NASA New Millennium Program. The sensor has two spectrometers, one in the visible and near-infrared (VNIR) (bands 8 to 57, region 427 to 925 nm) and the other in the SWIR region (bands 77 to 224, region 912 to 2395 nm). The combined use of Hyperion data with different pixel-based classification techniques has been examined in various LULC studies. Such studies have predominantly investigated the combination of Hyperion imagery with various linear and nonlinear spectral unmixing classification techniques in land classification studies (e.g.,), or object-based classification techniques for performing LULC cartography in combination with Hyperion imagery analysis. However, studies investigating the potential use of Hyperion combined with very powerful machine learning, nonparametric classifiers, such as SVMs and ANN, for LULC classification are still limited.

Understandably, being able to appreciate the extent to which different classification methods available today take advantage of the hyperspectral properties of Hyperion imagery would be of great importance. Yet, although a number of these studies have evaluated the accuracy of different classifiers for LULC studies, they are generally limited to the validation and comparison of only a distinct selection of classifying techniques. To this end, intercomparison studies of a large number of pixel-based classifying techniques combined with Hyperion hyperspectral imagery are significantly lacking. Such studies are indispensable, objectively assessing the accuracy of such techniques, their limitations for providing a measure of quality of a dataset, understanding their errors, and the potential implications of these in different applications. Performing intercomparison studies is also regarded as a necessary step toward an objective verification of the reliability, efficiency, and accuracy of such techniques. Indeed, this can provide an indication of gross dissimilarities and, potentially, insights into the causes for the differences observed. Furthermore, particularly for the case of the present study, an intercomparison would also highlight the optimum classification technique designed for implementation with hyperspectral imagery for urban vegetation mapping. Such information would be highly significant to a
number of land change related sciences and, in particular, to the use of EO technology in countering issues of increased urbanization.

In the purview of the above, the purpose of this research is to test and compare the applicability of both the SAM and SVMs classification algorithms for deriving information on the regional distribution of urban land cover types within a Mediterranean environment, and to compare the statistical accuracy of each approach utilizing hyperspectral (EO-1 Hyperion) data. As a case study, the city of Athens has been selected due to its high heterogeneity in terms of urban vegetation features and structures.

2 Study Area and Datasets

2.1 Study Area

Our study area includes the wider area of Athens, the capital of Greece, situated in the prefecture of Attiki, extending from 23°39’ to 23°42’ East and from 38°4’ to 37°55’ North (Fig. 1). The area represents a densely populated city with complex urban infrastructure, occupied primarily by a high number of buildings and a dense network of roads with lesser vegetation cover. The climate is Mediterranean type, characterized by relatively cold winters and often quite warm summers. This is a very suitable site for this type of research from both a scientific and practical aspect. Information on the urban vegetation structure plays a very important role in the microclimatic

Fig. 1 Geographical location of the study area.
conditions observed. Second, due to its rich biodiversity, a part of the wider area has been designated as a national park as well as a biodiversity conservation site. In addition, all data required were already available from previous works conducted in the region.38

2.2 Datasets

The Hyperion imagery used in the present study was acquired on August 27, 2009, obtained at no cost from the United States Geological Survey archive. The imagery was received as a full long scene (185-km strip) and at a level 1 [L1GST (Radiometrically corrected and resampled for geometric correction and registration to a geographic map projection)] processing level in GeoTIFF format, written as band-interleaved-by-line (BIL) files stored in 16-bit signed integer radiance values. The L1G product is radiometrically corrected, geometrically resampled, and registered to a geographic map projection with elevation correction applied. In addition, near co-orbital Google Earth imagery was used to support different aspects of methodological implementation and results interpretation. A shapefile was also created by digitizing all the green areas of Ilion municipality using Google Earth high-resolution imagery near concurrent to the Hyperion imagery acquisition date. This additional dataset was used as reference to assess the accuracy of each classifier in mapping the urban areas from the Hyperion imagery, as explained later (Sec. 3.3).

3 Methods

A supervised classification was implemented to the Hyperion imagery by applying the spectral angle mapper (SAM) and SVMs techniques to extract the urban vegetation cover within our study area. An overview of the methodology implemented is depicted in Fig. 2. All image processing was conducted in ENVI 5.0 and ArcMap 10.1 software platforms.

3.1 Data Preprocessing

Data preprocessing included the following; the acquired Hyperion imagery was converted to ENVI format files that contain wavelength, full width half maximum, and bad band information

---

**Fig. 2** Overview of the methodology implemented to extract urban vegetation cover.
using the Hyperion_tools.sav toolkit acquired from ITT Visual Information Solutions. The sensor non-calibrated bands (i.e., bands 1 to 7, 58 to 76, 225 to 242) and the water absorption bands (namely bands 120 to 132, 165 to 182, 185 to 187, 221 to 224) were then eliminated. This was done to reduce the influence of atmospheric scatter, water vapor absorption caused by well-mixed gasses to the data. Bands 77 and 78 were also removed, as they had a low signal-to-noise ratio value and overlapped with bands 56 and 57, respectively. Bands with vertical striping were then visually identified and manually removed from further analysis (namely bands 8 and 9, 56 and 57, 79 and 80, 218 to 220). Then the at-sensor radiance was computed from the raw digital number (DN) values for all remaining spectral bands. This was derived by dividing the pixel’s DN by a constant value, which was 40 for the VNIR (bands 8 to 57) and 80 for the SWIR (bands 79 to 224).

Subsequently, a minimum noise fraction (MNF) was applied to the Hyperion image. This aimed to separate noise from data and minimize the influence of systematic sensor noise during image analysis. MNF transformation was performed in ENVI as a linear transformation, essentially a two principal component transformation. The first transformation, based on an estimated noise covariance matrix, decorrelates and rescales the noise in the data. Then a standard principal component transformation is implemented, which creates several new bands containing the majority of the information. MNF was applied by handling the VNIR and SWIR data separately. The resulting MNF bands were analyzed for their spectral information content using eigenvalue plots and individual MNF gray-scale bands. For the inverse MNF transformation, nine components in the VNIR and seven in the SWIR were used. After the inverse MNF implementation, final Hyperion imagery consisted of 136 bands in total (46 in the VNIR and 90 in the SWIR). The resulting image was reduced to a subset of the studied region. These final 136 bands were used next to perform classifications using the SAM and SVMs to map the urban vegetation cover over the study area.

### 3.2 Hyperion Classification

A pixel-based supervised classification using both SAM and SVMs classifiers was implemented on the Hyperion image through three main steps. First, the classification key was decided, consisting of the classes shown in Table 1. Second, training sites representative of each class were collected from the Hyperion image following a stratified random sampling strategy. Selection of the training sites was guided by Google’s very high-resolution imagery photo interpretation, supported by selected field visits conducted in the study area. The training sites were selected from homogeneous regions. Approximately 200 to 300 pixels per class (in total 1007 pixels) were identified as training data. Subsequently, the two classification algorithms were implemented using the same training sites collected. The separability of the selected training points for all classes was examined in ENVI, which allows for the computing of the spectral separability between selected regions of interest for a given input file. These values range from 0 to 2.0 and indicate how well the selected regions of interest pairs are statistically separate. Values >1.9 indicate that the compared pairs have good separability, whereas very low separability values (<1) indicate that it might be appropriate to combine the compared spectra into a single one. In this study, mean pixel values of the selected training sites for performing the supervised

<table>
<thead>
<tr>
<th>Class name</th>
<th>ID</th>
<th>Class description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asphalt</td>
<td>1</td>
<td>Surfaces covered principally with asphalt, roads</td>
</tr>
<tr>
<td>Buildings</td>
<td>2</td>
<td>Urban fabric, urban area</td>
</tr>
<tr>
<td>Bare soil</td>
<td>3</td>
<td>Open areas with no vegetation, rocks or previously burnt</td>
</tr>
<tr>
<td>Trees</td>
<td>4</td>
<td>Surfaces principally covered with trees</td>
</tr>
<tr>
<td>Low vegetation</td>
<td>5</td>
<td>Open areas with little or vegetation of low height</td>
</tr>
</tbody>
</table>
classifications were reported to always have a separability index > 1.78 for all the class pairs. An illustration of the mean spectra of the training sites used in this study is provided in Fig. 3. Third, the SAM and SVMs classifiers were developed and implemented in ENVI image processing environment, using the collected training sites. The classification key used in this study is indicated in Table 1.

3.2.1 Spectral angle mapper

SAM is a spectral classification method that performs classification based on the spectral similarity between image spectra and reference spectra. The algorithm determines the spectral similarity between two spectra using an n-D angle approach to match pixels to reference spectra. The angles between the spectra are represented as vectors in a space with dimensionality equal to the number of bands. The n-D angle multispectral space is measured with respect to the axes that define the coordinate system of the space. Figure 4(a) provides a simplified illustration of the angle between a reference spectrum and a test spectrum from two-band data in a two-dimensional plot obtained from two points within an image. The lines connecting each spectrum point (a and b) and the origin contain all possible positions for the sample, possibly corresponding to intensity changes due to the illumination variability. A threshold angle value is required for the separation between the end-member spectrum vector and the pixel vector in n-dimensional space defined by the number of bands of each Hyperion image. However, the angle (α) between two vectors is independent of their lengths. Small angles between two spectra indicate a high similarity and high angles indicate a low similarity, whereas pixels with an angle larger than the tolerance level of the specified maximum angle threshold are not classified. The angle between the spectrum vector to be classified, \( \vec{a} \), and the reference spectrum vector, \( \vec{b} \), is computed by applying the following equation:

\[
\alpha = \cos^{-1}\left( \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| ||\vec{b}||} \right).
\]

The main advantage of the SAM classifier is that it is based only on the angular information and, hence, not affected by solar illumination factors. Also, since this method uses only the direction of a vector, it is also insensitive to albedo effects. Herein, for SAM classification, a single value of 0.3 radians was used as the maximum classification threshold value for all classes, which was selected after careful consideration of different threshold angle values on the basis of the classification accuracy assessment results. Reference spectra in our study were extracted directly from the remote sensing imagery. The n-dimensional space for the SAM implementation
was defined by using the total number of each Hyperion image after the implementation of the MNF transformation as input.

3.2.2 Support vector machines

SVMs are a nonlinear and nonparametric large margin classifier implementing Vapnik’s structural risk minimization principle. SVMs fit a decision surface [called hyperplane, Fig. 4(b)] that provides the best separation between the two classes in a multidimensional feature space. This hyperplane maximizes the distance between the hyperplane and the nearest positive and negative training examples, called the margin. Generally, the larger the margin, the lower is the generalization error of the classifier. From a given set of training data, the optimization problem is solved to find the hyperplane that leads to a sparse solution. Hence, only a subset of the training samples, those that lie on the margin (called support vectors), are used to define the hyperplane. Such solutions guarantee a minimal generalization error. By using nonlinear kernel functions (e.g., Gaussian radial based function (RBF), polynomial), the SVMs implicitly work linearly in a higher-dimensional space, corresponding to a nonlinear solution in the input space where the data naturally live. Thus, often not all of the obtainable training examples are used in the design of the separating hyperplane, and this is a key generalization feature of SVMs in comparison to other nonparametric classifiers. In order to represent more complex hyperplane shapes than the linear examples, the techniques can be extended by using kernel functions. In this case, the problem transforms into an equivalent linear hyperplane problem of higher dimensionality. The use of the kernel function essentially allows the data points to be classified to spread in a way that allows the fitting of a linear hyperplane. SVMs also introduce a cost parameter C to quantify the penalty of misclassification errors in order to handle nonseparable classification problems.

Herein, SVMs were implemented on the final preprocessed Hyperion image. The RBF kernel function was used for performing the pair-wise SVMs classification. This kernel requires the definition of only a small number of parameters to run and has already shown generally good results. RBF kernel is defined from the following equation:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0.$$  \hspace{1cm} (2)

The $\gamma$ parameter is a user-defined parameter and was set to a value equal to the inverse of the number of the spectral bands of the imagery used each time (i.e., in our study, 0.007). In order to ensure accuracy, the penalty parameter, which controls the trade-off between allowing training errors and forcing rigid margins, was set to its maximum value (i.e., 100). The pyramid parameter was set to a value of zero, meaning that each image should be processed at full resolution. A classification probability threshold of zero was applied, forcing all image pixels to be classified into one class label.

Fig. 4 (a) A graphical illustration of the spectral angle mapper (SAM) operation. Vector representation of a reference spectrum and test spectrum for a two-band image; $a$ is the spectral angle. (b) An illustration of the support vector machines (SVMs) operation, where black examples denote support vectors.
3.2.3 Accuracy assessment

The computation of error matrix statistics, namely the overall accuracy (OA), user’s accuracy (UA), and producer’s accuracy (PA), and the Cohen’s Kappa ($K_c$) statistics, was used to assess the different classification methods’ performances. $K_c$ measures the actual agreement between reference data and the classifier used to perform the classification, versus the chance of agreement between the reference data and a random classifier. PA, expressed as a percentage, represents the correctly classified ground classes by the analyst (omission error). Likewise, UA expresses the percentage of pixels that do not truly belong to the reference class, but are committed to other ground truth classes (commission error). In mathematical terms, these parameters are expressed by the equations shown in Table 2.

In computing the above statistical measures, ~40 to 60 sampling points per class were selected directly from the Hyperion imagery following a random sampling strategy, and these points formed the validation dataset. The selection of these points was performed following exactly the same criteria used for the selection of training points, and their spectral separability was also assessed using the same procedures implemented for the training sites selection to ensure consistency. They were selected from Hyperion imagery and the interpretation was completed with the support of an image of the same date in Google Earth. To ensure consistency in our comparisons, the same set of validation points was used in evaluating the accuracy of all thematic maps produced from the implementation of the different classifiers.

In addition, the urban vegetation results predicted by the SAM and SVMs were compared against the reference estimate of the municipality of Ilion digitized previously from the high-resolution Google image covering the same area. This method was originally proposed by Kontoes et al. for evaluating burnt area estimates from different algorithms; yet this rationale has also been implemented in other land cover classification studies. For accuracy assessment of classified image, all the trees and low vegetation were considered together as green areas. All processing and computations concerning this validation step are described below. Briefly, the accuracy of the urban vegetation area detection (the trees and low vegetation classes were considered as one class) was expressed in terms of detected green area efficiency, skipped green area rate (omission error), and false green area rate (commission error). These accuracy figures were calculated on the basis of the equations presented in Table 2. To enable overlay and facilitate efficiency in the urban vegetation comparisons, the estimates of the urban vegetation area from the two classifiers were extracted from the corresponding classification maps and were subsequently transformed into vector (i.e., shapefile) GIS products. Finally, since the Hyperion

<table>
<thead>
<tr>
<th>Name</th>
<th>Mathematical definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy</td>
<td>$OA = \frac{1}{N} \sum n_{ij}$</td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>$PA = \frac{n_{ij}}{n_{icol}}$</td>
</tr>
<tr>
<td>User’s accuracy</td>
<td>$UA = \frac{n_{ij}}{n_{irow}}$</td>
</tr>
<tr>
<td>Kappa coefficient</td>
<td>$K_c = N \sum n_{ij} - \sum n_{icol}n_{irow}/N^2 - \sum n_{icol}n_{irow}$</td>
</tr>
<tr>
<td>Algorithm detection efficiency</td>
<td>Detected area efficiency $=</td>
</tr>
<tr>
<td>Skipped area rate</td>
<td>Skipped area rate $=</td>
</tr>
<tr>
<td>False area rate</td>
<td>False area rate $=</td>
</tr>
</tbody>
</table>

Table 2 Definition of the quantitative measures used to assess the agreement between the two classifiers in terms of deriving the urban green areas. $n_{ij}$ is the number of pixels correctly classified in a category; $N$ is the total number of pixels in the confusion matrix; $r$ is the number of rows; and $n_{icol}$ and $n_{irow}$ are the column (reference data) and row (predicted classes) total, respectively. DGE is the detection efficiency rate, SGA is the skipped green area rate, detected green area (DGA) and FGA is the false green area rate.
spatial resolution is 30 m, an additional analysis was performed by separating the digitized green areas into two groups: areas with surface ≤900 m² and >900 m². This allowed examining the detection efficiency more objectively, accounting as well for the land fragmentation in respect to the Hyperion pixel size (30 m).

4 Results and Discussion

The urban vegetation cover maps produced after the implementation of the SVMs and SAM classifiers to the Hyperion image of the study region are illustrated in Fig. 5. The statistical results obtained after the classification accuracy assessment conducted based on the error matrix results are summarized in Table 3.

Both classification methods produced comparable results in terms of describing not only the spatial distribution but also the cover density of each land cover category in the test site. It appears that SVMs generally outperformed the SAM classifier in both OA and individual classes’ accuracies. SVMs’ OA and Kappa coefficient were 86.53% and 0.823, respectively, while SAM’s classification OA and Kappa coefficient were 75.13% and 0.672, respectively. Among the two classifiers, as evidenced from the statistics of the individual urban vegetation cover classes’ results for both OA and Kappa statistics, SVMs was more accurate than SAM in describing the spatial distribution of urban vegetation for our study site. In terms of the individual classes’ accuracy, PA varied from 57.6 to 96.9% and UA varied from 73.3 to 90.5% for SVMs, whereas PA varied from 27.3 to 98.5% and UA ranged from 52.9 to 86.1% for SAM. Less accurate results for the case of classifying urban areas from both techniques can be ascribed to a wide range of spectrally complicated properties across the electromagnetic region range. This is because of the presence of numerous spectrally unique and ambiguous materials, such as dark-shingles and asphalt roads, makes the classification process slightly inaccurate. Other factors that further complicate the analysis of urban areas, leading to high within-class spectral variability, include the three-dimensional heterogeneity of urban areas and urban

Fig. 5 The Hyperion classification using SVMs RBF classifier (a) and SAM classifier (b).
vegetation cover material aging.\textsuperscript{49} In addition, some buildings and open spaces are covered by spectrally similar urban surface materials, which further hamper a clear discrimination between them.

Overall classification results obtained herein from both classifiers are also in agreement with or are better than those reported in previous studies exploring the use of these specific classifiers using hyperspectral data. For example, Zulhaidi et al.\textsuperscript{50,51} compared different classifiers using airborne hyperspectral image data and reported classification accuracies of 50\% for ANN and SAM, respectively. Another study\textsuperscript{52} examined the combined use of SVMs with hyperspectral imagery from the DAIS airborne sensor for a test region in Spain and reported an overall classification accuracy for SVMs >91\%. More recently, Koetz et al.\textsuperscript{53} combined hyperspectral and LiDAR data with SVMs for fuel types mapping for a region in France and reported an OA and $K_c$ of 71\% and 0.755, respectively. Notably, previous works by the authors have also evaluated the implementation of these advanced classification algorithms with both hyperspectral and multispectral data. For example, Petropoulos et al.\textsuperscript{25,34,38} compared different classifiers using both Landsat TM and EO-1 advanced land imager data for land cover mapping with emphasis on burnt area delineation in a heterogeneous Mediterranean study site. Although a direct comparison of both techniques was unavailable, reported classification accuracies of both SVMs and SAM classifications were notably better than the results presented herein; however, in agreement with the present study, the implementation of SVMs (OA 91 to 95.87\%, $K_c$ 0.885 to 0.948) outperformed SAM (OA 83.82\%, $K_c$ 0.795). Authors suggested that the lower overall classification accuracy, as well as the lower total burnt area estimate of SAM may be attributed to the relative insensitivity of this method to brightness variations. SAM is not responsive to the effects of solar illumination since it uses only the angle between the two compared spectra and not the vector length.

A number of previous works by the author have also specifically evaluated the use of these two classification techniques implemented with hyperspectral Hyperion data, although again a direct comparison was not examined. For example, Petropoulos et al.\textsuperscript{27} compared different classifiers using Hyperion hyperspectral imagery for discriminating land-cover classes in a typical Mediterranean setting. They reported close classification accuracy between both SVMs (OA 89.26\% and $K_c$ 0.880) and ANNs (OA 85.95\% and $K_c$ 0.842), with SVMs outperforming the ANNs by 3.31\% (OA) and by 0.038 ($K_c$), respectively. Although both classifiers produced close results, SVMs generally appeared more useful in describing the spatial distribution and the cover density of each land cover category. In a subsequent study, Petropoulos et al.\textsuperscript{28} compared the potential of SVMs and the object-based approach using Hyperion imagery in a Mediterranean setting. For the SVMs classifier, OA and $K_c$ were 76.30\% and 0.719, and for the object-based classification, OA and $K_c$ were 81.30\% and 0.779, respectively. It should be noted that results from the previous studies, which utilized the SVMs classifying algorithm with Hyperion data, returned accuracies comparable to the results presented herein. However, in

### Table 3 Summary of the results from the classification accuracy assessment conducted.

<table>
<thead>
<tr>
<th>Ground truth (percent)</th>
<th>SVMs PA (%)</th>
<th>SVMs UA (%)</th>
<th>SAM PA (%)</th>
<th>SAM UA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asphalt</td>
<td>96.92</td>
<td>88.73</td>
<td>98.46</td>
<td>79.01</td>
</tr>
<tr>
<td>Buildings</td>
<td>95.00</td>
<td>86.36</td>
<td>77.50</td>
<td>86.11</td>
</tr>
<tr>
<td>Bare soil</td>
<td>64.71</td>
<td>73.33</td>
<td>76.47</td>
<td>56.52</td>
</tr>
<tr>
<td>Trees</td>
<td>94.74</td>
<td>85.71</td>
<td>73.68</td>
<td>80.00</td>
</tr>
<tr>
<td>Low vegetation</td>
<td>57.58</td>
<td>90.48</td>
<td>27.27</td>
<td>52.94</td>
</tr>
<tr>
<td><strong>OA (%)</strong></td>
<td><strong>86.53</strong></td>
<td><strong>75.13</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Kappa coefficient</strong></td>
<td><strong>0.823</strong></td>
<td><strong>0.672</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SVM, support vector machine; PA, producer’s accuracy; UA, user’s accuracy; SAM, spectral angle mapper.
a similar study by Elatawneh et al.24 object-based, SAM, and ANN classifiers were implemented to Hyperion imagery for a Mediterranean site, with object-based classification outperforming the other techniques with an OA of 83%. Subpixel classification based on ANN showed an OA of 52%, very close to that of SAM (48%). However, SAM applied using the training sites selected directly from the Hyperion imagery supported by QuickBird imagery and field visits returned an increased accuracy of 64%, underlining the effect of the relatively low spatial resolution of the Hyperion imagery on the spectral separation among the land use/cover classes. Similar issues have been identified in previous studies as well as in the present work. It should be noted that in contrast to all other hyperspectral-based studies, results reported in the study of Elatawneh et al.24 returned significantly less accurate results compared to the present study. Similar results were reported in Petropoulos et al.39 with object-based classification clearly outperforming SAM by 7.91% (OA); however, the overall classification accuracy of SAM was significantly higher comparable to results recorded for the present study (OA 74.10%, \( K_c \) 0.695). However, a direct comparison of both SAM and SVMs classification algorithms implemented to hyperspectral data for LULC change is yet to be explored. Thus, results from this study provide another component to the comprehensive evaluation of advanced classification algorithms implemented with Hyperion hyperspectral data for LULC, in particular for their use in urban vegetation mapping.

On the basis of the PA statistical measure, it can be observed that the class of low vegetation produced the lowest percentage in both techniques. In addition, it can be observed that the low vegetation class yielded the lowest percentage in the SAM technique and the class of bare soil performed with the lowest percentage in the SVMs. This can be attributed, in part, to the similar spectral characteristics between the two classes, which were most probably affected by the mixed pixels combined with the low spatial resolution of the Hyperion sensor.

The validation results obtained after the classification accuracy conducted using all the digitized reference dataset for the Ilion municipality for both classifiers are presented in Table 4. The surface validation of approaches in the case of urban green area is depicted in Fig. 6, where detected, skipped, and false areas are shown. The false green areas were found to be higher in the case of the SVMs' classified images. Similarly, the common green areas between the classification and GIS digitization were also higher for SVMs. The skipped green area indicated higher values in the case of the SVMs classifier compared to SAM. Results from both classifiers demonstrated a comparatively similar performance in terms of detection efficiency rate and commission error with slightly higher values for SVMs in comparison to SAM. By contrast, the false alarm rate values showed the opposite trend with a higher estimate in the case of SAM (0.38) than the SVMs' (0.32) classified image. Most of the detected green areas are located along the borders of the extreme northwestern and northeastern side of Athens municipality.

<table>
<thead>
<tr>
<th>Common GA (m²)</th>
<th>False GA (m²)</th>
<th>Skipped GA (m²)</th>
<th>Detection efficiency rate</th>
<th>Skipped area rate</th>
<th>False area rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparisons for all reference dataset</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAM</td>
<td>1,016,506</td>
<td>385,838</td>
<td>632,398</td>
<td>0.62</td>
<td>0.28</td>
</tr>
<tr>
<td>SVMs</td>
<td>1,127,339</td>
<td>609,199</td>
<td>521,566</td>
<td>0.68</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Comparisons for green areas ≤900 m²</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAM</td>
<td>11,978</td>
<td>1,390,368</td>
<td>61,618</td>
<td>0.16</td>
<td>0.84</td>
</tr>
<tr>
<td>SVMs</td>
<td>18,074</td>
<td>1,718,465</td>
<td>55,521</td>
<td>0.25</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>Comparisons for green areas &gt;900 m²</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAM</td>
<td>1,004,529</td>
<td>397,816</td>
<td>570,781</td>
<td>0.64</td>
<td>0.36</td>
</tr>
<tr>
<td>SVMs</td>
<td>1,109,265</td>
<td>627,274</td>
<td>466,045</td>
<td>0.70</td>
<td>0.30</td>
</tr>
</tbody>
</table>
The best result in terms of absolute difference from the reference polygon is shown by the SVMs classifier, which displayed a lower skipped rate than the SAM classifier, concentrated mostly toward the northeastern side of Athens.

The digitized green areas in Ilion municipality with a surface area $\leq 900$ m$^2$ included 210 polygons with a total area of 73,595 m$^2$, while surface areas $> 900$ m$^2$ included 155 polygons with a total area of 1,573,310 m$^2$. The SVMs classifier identified common parts of green areas only in 64 polygons out of the 210 small polygons, while in the 155 large polygons, common parts were found in 129 polygons. Similar results were found in the SAM classifier. In small green areas (Table 4), the common green areas between classification and GIS digitization (DGA) are 1,109,265 m$^2$ in SVMs, a greater area in comparison to that of the SAM image (1,004,529 m$^2$). One advantage of using the SVMs classifier is that it has fewer skipped green areas (466,045 m$^2$) than SAM (570,781 m$^2$). The detection efficiency rates in small green areas ($\leq 900$ m$^2$) was 0.16 and 0.25 for SAM and SVMs respectively, while those for areas with a surface $\geq 900$ m$^2$ were 0.64 and 0.70, respectively. The skipped green areas were higher in the case of the $> 900$ m$^2$ green areas as compared to the $\leq 900$ m$^2$ category.

The higher performance of SVMs can be attributed to the high capability to identify an optimal separating hyperplane for separation of classes, which other machine learning classifiers may not be able to locate. The analysis indicates that SVMs are also able to optimally generalize this, separating hyperplane unseen samples with the least errors among all the separated hyperplanes, thus producing the best class separation at the end of classification as compared to SAM. SVMs are successful in addressing ill-posed problems, providing high classification accuracy results in comparison to SAM, even in cases when small training sets are used. In comparison to SVMs, the main advantage of SAM is that it is an easy and rapid method for mapping the spectral similarity of image spectra to reference spectra. It is also a powerful classification method because it represses the influence of shading effects to accentuate the target reflectance characteristics.

\section{5 Conclusions}

The combined use of Hyperion hyperspectral imagery with the SVMs and SAM classifiers for discriminating different urban classes at a typical urban setting was the main objective of this
study. Our results confirmed the ability of both classifiers when combined with Hyperion imagery to extract urban vegetation cover for the case of a densely populated city with complex urban features, such as Athens. SVMs outperformed SAM in terms of both overall and individual class accuracy, at least in our experimental setting. The higher classification accuracy reported by SVMs was mainly attributed to some key characteristics accompanying the implementation of this classifier. However, misclassifications were also found by both classifiers. Those were attributed largely to their inability to account for the subpixel heterogeneity, which is very high at the Hyperion spatial resolution of 30 m in urban areas, particularly in sites like that of our study site. Spectral confusion in pixels may potentially lead to classification errors of a spectral class, which can provide an important barrier to both the classification approaches examined, especially when applied to coarser-resolution EO imagery. Consequently, there is a demand for further evaluation of classification algorithms and further research to be conducted in techniques that perform at the subpixel level to deal with the heterogeneity, perhaps with the eventual development of a synergistic approach.

The findings of this study could provide a significant advantage to planners and managers to formulate effective policies for conserving this important environmental attribute. Such information is important in successful infrastructure development, urban landscape planning, and improvement of the urban environment in general. More widely, these results also provide a significant contribution toward an objective assessment of satellite hyperspectral data, such as those from Hyperion, in mapping urban vegetation cover from space, given that, to our knowledge, very limited research has been explored in this direction.

Acknowledgments

The authors wish to thank the United States Geological Survey (USGS) for the free provision of the Hyperion imagery used in performing the present study. The authors wish to also thank Antony Smith from Aberystwyth University for his assistance in the preparation of some of the figures included in the paper and Gareth Ireland for his help in the manuscript preparation.

References


George P. Petropoulos is a senior lecturer in remote sensing and GIS in the Department of Geography & Earth Sciences at Aberystwyth University, UK. His research work focuses on exploiting EO data alone or synergistically with models for deriving key parameters of the Earth’s energy and water budget and on land cover mapping/changes from either anthropogenic activities or natural hazards. He is author/co-author of +38 peer-reviewed journal articles, +75 international conferences, +12 book chapters, and a book editor.

Dionissios P. Kalivas is an associate professor teaching GIS and spatial analysis. He is also responsible of the GIS Research Unit of the Agricultural University of Athens. He started his academic career as visiting assistant professor in the Department of Environmental Engineering, Democritus University of Thrace (2000 to 2002). From 2002 to 2005, he was a lecturer at Panteion University. He is author and reviewer of many research papers in scientific journals and congress proceedings.

Iro A. Georgopoulou graduated from the Department of Natural Resources Management and Agricultural Engineering of Agricultural University of Athens (2003 to 2009). She holds an MSc in applications of geoinformatics in natural resources (AUA, 2011-2013). She participates in the research project NEA GNOSI in collaboration with the Technological Educational Institute of Epirus, where she works on the spatial distribution and observation of insect Metcalfa pruinosa in the area of Epirus using GIS-based spatial analysis techniques.

Prashant K. Srivastava obtained a PhD from the Department of Civil Engineering, University of Bristol, United Kingdom. He has also worked as an assistant professor in India. Currently, he is working as a research scientist with ESSIC/NASA GSFC, Hydrological Sciences Laboratory on SMAP satellite soil moisture retrieval algorithm development and its applications. He has published several peer-reviewed papers, book chapters, and two books.