Classification of agricultural fields by using Landsat TM and QuickBird sensors. The case study of olive trees in Lesvos island.

Christos Vasilakos, University of the Aegean, Department of Environmental Studies, University Hill, 81100, Mytilene, Greece, chvas@env.aegean.gr
John Hatzopoulos, University of the Aegean, Department of Environmental Studies, University Hill, 81100, Mytilene, Greece, ihatz@aegean.gr
Kostas Kalabokidis, University of the Aegean, Department of Geography, University Hill, 81100, Mytilene, Greece, kalabokidis@aegean.gr
Kostas Koutsovilis, Extended University Program – Environmental Cartography, University of the Aegean, University Hill, 81100, Mytilene, Greece, koutsobilis.k@envcart.aegean.gr
Argiroula Thomaidou, Extended University Program – Environmental Cartography, University of the Aegean, University Hill, 81100, Mytilene, Greece, thomaidou.a@envcart.aegean.gr

Remote sensing techniques are the main methods used during the last two decades for information extraction and digital mapping of spatial features. Traditional methods include supervised classification of data with 30-m spatial resolution (i.e. Landsat TM) by collecting Ground Truth Data with Global Positioning Systems (GPS) in the field. These methods encounter some problems especially in complex landscapes like agricultural fields in the Mediterranean region. Spatial and spectral uncertainties lower the accuracy of the classification results.
Nowadays, however, high-resolution sensors with spatial resolution of 2.5-m can be used for direct mapping by visual interpretation or automatic object recognition algorithms. The main disadvantages of these methods are the specialized technological know-how that is required, and the whole coverage of the study area with expensive high-resolution data. The purpose of this study is to evaluate the improvement of Landsat TM data supervised classification by using QuickBird high-resolution sensor for Ground Truth Data collection, in order to retrieve the spatial distribution of agricultural fields of olive tree orchards in Lesvos island. The resulted classification scheme includes different qualitative classes of the fields as well as an effort to estimate the tree densities.

Keywords: Agriculture, Olive Trees, Supervised Classification, Landsat TM, QuickBird, Remote Sensing.

1. Introduction

Olive fields are the major cultivation type in Mediterranean region. The importance of olive tree in Greece appears in ancient years, while at nowadays, its farming consists the major or complementary income activity for the population of Lesvos island. European Commission’s Regulation (EC) No 2366/98 of 30 October 1998 laid down detailed rules for the application of the system of production aid for olive oil. According to the latter, farmers receive production aid only if they submit a registration of their fields and locate them on the appropriate cartographic background. The area of these orchards is estimated, traditional, only by farmer’s registration, thus, various research projects were carried out the last two decades aiming to map, sufficiently, agricultural fields by using satellite remote sensing data. The processing of these data in order to retrieve land-use and land-cover types is quite difficult to Mediterranean region due to complex landscape. Steep slopes and unclear field boundaries among with medium-resolution sensors leads supervised classification to considerable errors regarding this kind of application. This research attempts to study the usefulness of QuickBird data in retrieving ground truth data for Landsat TM classification. Opposed to previous methods where the sample data were collected in field by using low resolution GPS measurements, the user can identify objects and patterns by taking advantage the 2.5 m pixel resolution.

2. Description of Study Area

The island of Lesvos covers an area of 1672 Km^2 with a variety of geological formations, climatic conditions and vegetation types (i.e., forests, shrublands, grasslands and agricultural lands). The climate in these areas is typically Mediterranean, with warm and dry summers and mild and moderately rainy winters. Annual precipitation averages around 670 mm, but there is a rainfall gradient of more than 45% from East (subhumid) to West (semiarid). The average annual air temperature is 18 ^{o}C with high oscillations between maximum and minimum daily temperatures. The terrain in the study areas is rather hilly and rough, with its highest peak at 800 m ASL. Slopes greater than 20% are dominant covering almost 2/3's of the area. Narrow alluvial plains are present in the lower parts along the coast, having a shallow groundwater table. Prominent arid lands are mainly found in the western part of the island, in which acid volcanic rocks dominate.

3. Data and Methodology

Satellite images of medium-high (Landsat TM) and very high resolution (QuickBird) were used as original data. The Landsat TM image has been acquired in 28/5/1999. It is presumed that the last 3 years, there are no major modifications in land cover types able to be mapped in 30 by 30 m. Geometric corrections for Landsat TM were applied by using high-accuracy GPS (3-5 m) measurements and geo-referenced into the Hellenic Geodetic Reference System 87. Selected sceneries of QuickBird were rectified by using same Ground Control Points (GCPs) with Landsat TM.

QuickBird is the world's highest resolution commercial satellite with ground pixel resolution 61cm for panchromatic band and 2.44 m for multispectral (Eurimage, 2003). Imagery is collected in 11-bit format (2048 grey levels) while four-band multi-spectral imagery consists of blue, green, red and near-infrared bands are available (Table 1). QuickBird images were acquired during of summer 2002 and the whole study area was covered by more than 15 sceneries. Disadvantages of acquiring new high-resolution data is the high price and the risk of having maximum 20% of area of interest covered by clouds without including haze and shadows.

Table 1: Band's wavelength for Landsat TM and Quickbird multispectral sensors

Band \ Sensor	Landsat TM	QuickBird
Band 1	450 - 520 nm	450 - 520 nm

Band 2	520 - 600 nm	520 - 600 nm
Band 3	630 - 690 nm	630 - 690 nm
Band 4	760 - 900 nm	760 - 900 nm
Band 5	1550-1750 nm	-
Band 6	10400-12500 nm	-
Band 7	2080-2350 nm	-

In supervised training, user relies on its own pattern recognition skills and *a priori* knowledge of the data to help the system determine the statistical criteria (signatures) for data classification. Usually, this knowledge is acquired by field work in terms of interpreting landscape into classes and assigning the corresponded X and Y values acquired by GPS. This procedure has a lot of man-hours cost, travel cost and requires the ability of user to travel in the field. Within this research, field work replaced by QuickBird data, so the classifier interprets the different classes from images and collects the corresponding Areas of Interest (AOIs) that have been used for the supervised classification of Landsat TM.

The classes that have been recognized in order to classify Landsat TM are:

- Salt mines, coastal areas, water
- Sea
- Urban area
- Barren soil
- Pine forest
- Olive tree orchards
- Oak forest
- Chestnut tree forest
- Evergreen shrubs
- Deciduous tree forest
- Agriculture land
- Grazing land

Olive tree orchards separated into 5 different classes according to their type and trees' density. Figure 2 shows AOIs for these classes overlaid on QuickBird images assigning the Near Infrared channel to Red, Blue channel to Blue and Green channel to Green.

- E1: 90-120 trees per hectar
- E2: 130-150 trees per hectar
- E3: 160-190 trees per hectar
- E4F: 200-230 trees per hectar at flat area
- E4M: 200-230 trees per hectar at mountainous area

After collecting AOIs, Landast TM classified into the above classes with supervised classification by using the feature space decision rule as a non-parametric rule and maximum likelihood as a parametric rule. The feature space decision rule determines whether or not a candidate pixel lies within the nonparametric signature in the feature space image. When a pixel's data file values are in the feature space signature, then the pixel is assigned to that signature's class. The maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions. If there is *a priori* knowledge that the probabilities are not equal for all classes, weight factors can be specified for particular classes (Hord, 1982). The equation for the maximum likelihood/Bayesian classifier is as follows where the pixel is assigned to the class, *c*, for which *D* is the lowest:

$$D = \ln(ac) - \left[0.5\ln(|Covc|)\right] - \left[0.5(X - Mc)T(Covc - 1)(X - Mc)\right] \tag{1}$$

Where:
D = weighted distance (likelihood)
c = a particular class
X = the measurement vector of the candidate pixel
Mc = the mean vector of the sample of class *c*
ac = percent probability that any candidate pixel is a member of class *c*
Covc = the covariance matrix of the pixels in the sample of class *c*
|*Covc*| = determinant of *Covc* (matrix algebra)
Covc-1 = inverse of *Covc* (matrix algebra)
T= transposition function (matrix algebra)

The accuracy assessment was evaluated with the contingency matrix of the classification. The contingency matrix is a quick classification, with the same rules, of the pixels in a set of training samples to see what percentage of the sample pixels are actually classified as expected. The pixels of each training sample are not always so homogeneous that every pixel in a sample is actually classified to its corresponding class. Each sample pixel only weights the statistics that determine the classes. However, if the signature statistics for each sample are distinct from those of the other samples, then a high percentage of each sample's pixels is classified as expected (ERDAS, 1994).

E1

E1

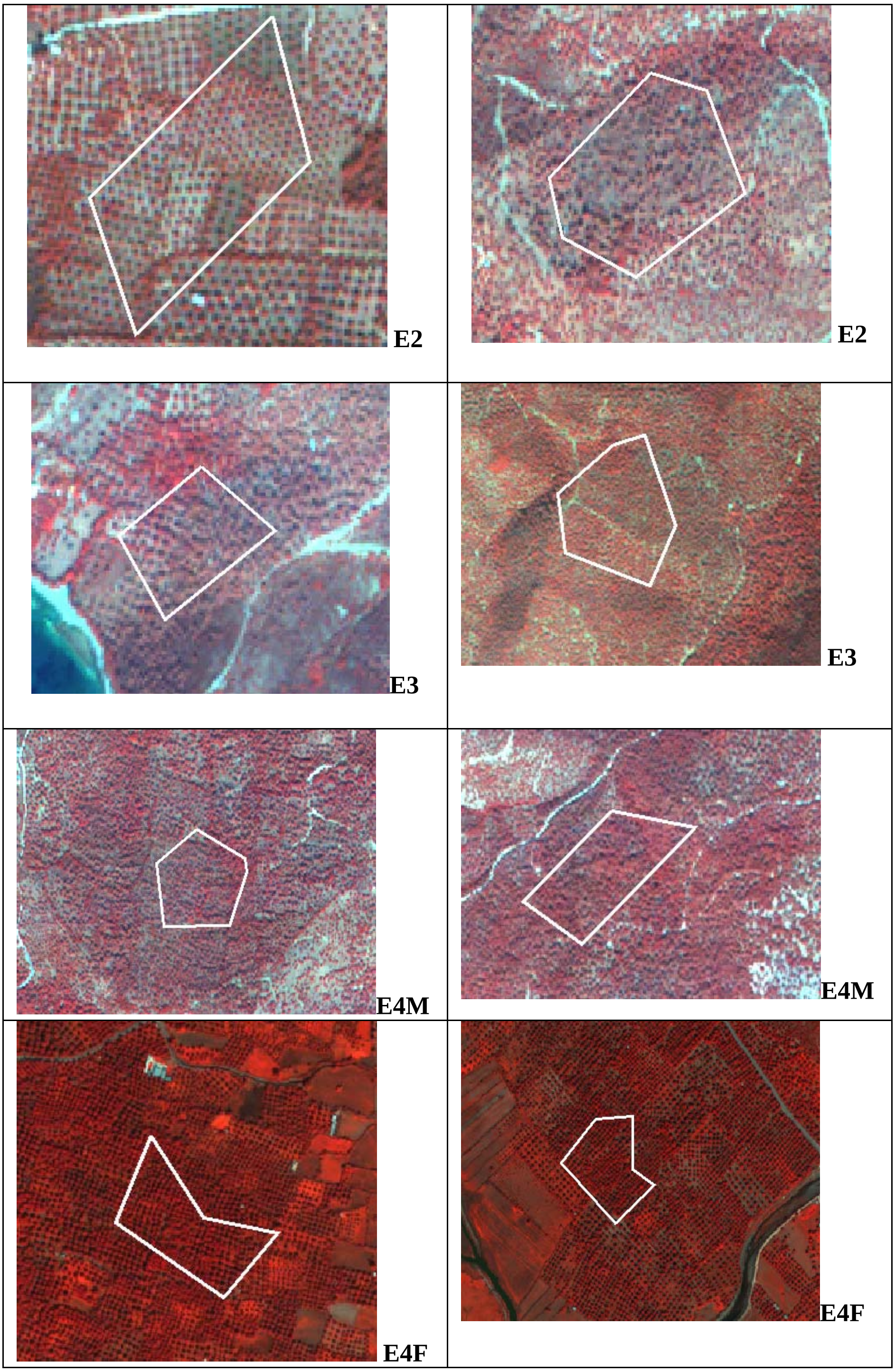

Figure 2**:** Sample plots of olive tree orchards, over QuickBird images.

4. Results and Discussion

The accurate area calculation of olive trees orchards and the number of trees is quite difficult in the Greek insular environment due to complex landscape, the abandonment of land and the lack of statistical and historical data. Olive trees mapping by satellite data classification has promising results and new techniques are explored in order to improve them. By using high-resolution imagery as a tool of ground truth data collection, the total area coverage and the number of olive trees amplitude was estimated (Table 2). The total area was estimated 47561,5 hectares while within a previous research where the ground truth data were collected by using a 30-m accuracy GPS, the total area was estimated 43324,1 hectares (Hatzopoulos et al., 1992, Koukoulas, 1994). The overestimation is mainly explained by the time-difference between the two satellite datasets which is 9 years and the different approaches in ground truth data collection. Figure 3 shows the output maps of Landsat TM classification regarding land cover types and olive tree orchard types.

Table2: Classification results of Landsat TM

Olive Trees class	Area (hectares*)	Number of trees per area unit	Number of trees (min)	Number of trees (max)
E1	20064	90-120	1805760	2407680
E2	6230	130-150	809900	934500
E3	19171,2	160-190	3067392	3642528
E4M	1648,7	200-230	329740	379201
E4F	447,6	200-230	89520	102942
Total	**47561,5**	**-**	**6102312**	**7466851**

* 1 hectare =10000 m^2

Table 3 shows the contingency matrix for the plot samples of the land cover types that has been selected as the final classification scheme. The overall accuracy is 99.38%, meaning that this percentage of the sample pixels is classified correct. Regarding olive trees, only 1.7% of the pixels that have been identified as olive trees by the user over QuickBird image, have been classified in 5 other classes. That is explained because of the larger training dataset selected for olive trees. Furthermore, there is an overestimation of 5.5% of the pixels that belong to other classes and they have been classified as olive trees. Comparing to former procedure of supervised classification (Hatzopoulos et al., 1992), where the user has to collect ground truth data with GPS, the classification results seems to be much better.

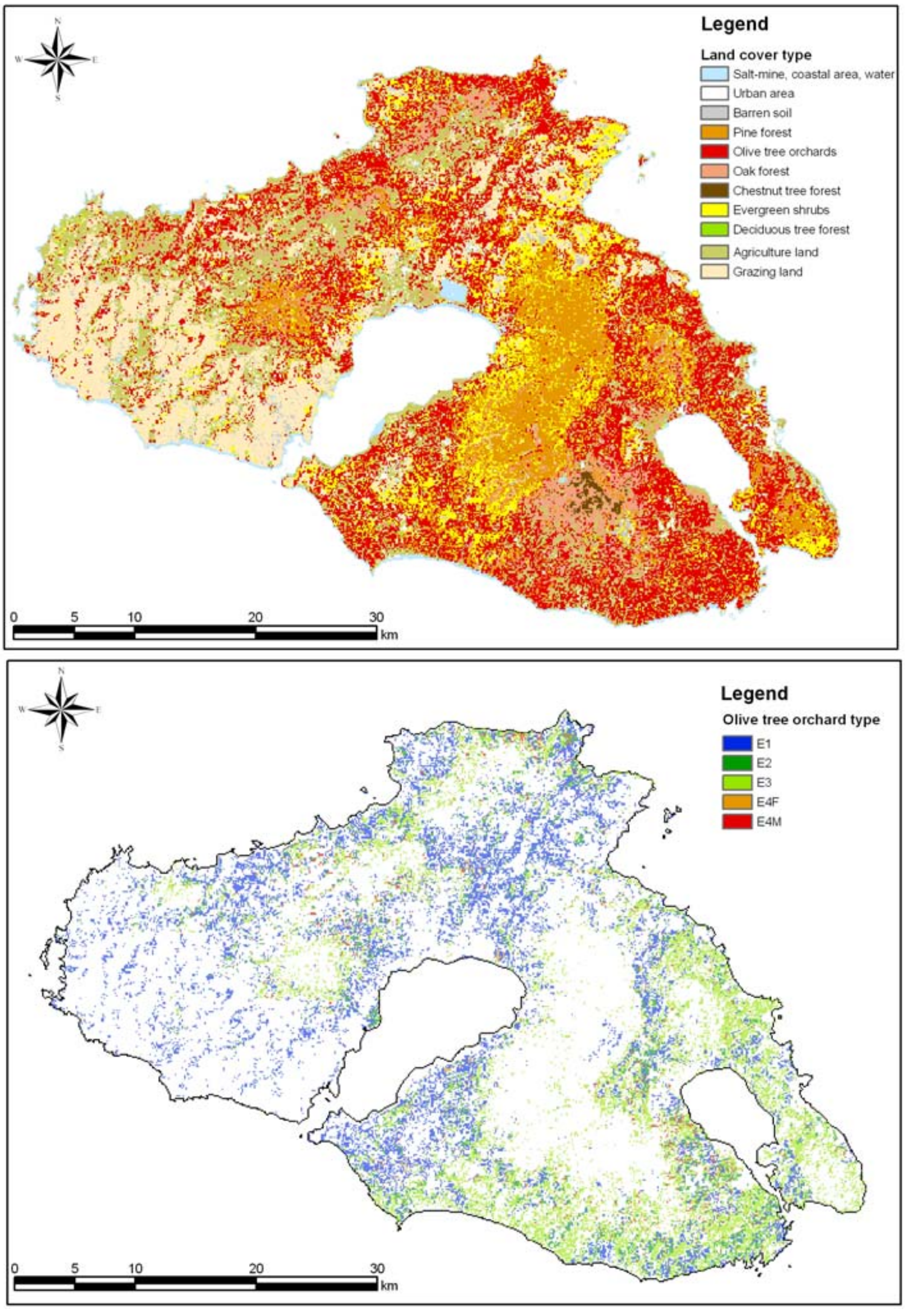

Figure 3: Land cover types (above) and Olive tree orchard types (below) maps from Landsat TM classification

Table 3: Contingency matrix of the classification (Numbers in percentage)

Classification Data	Pine Forest	Salt-mine etc	Olive trees	Urban area	Deciduous trees	Barren Soil	Evergreen shrubs	Agriculture land	Grazing land	Chestnut trees	Oak forest
Pine Forest	99.6	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Salt-mine etc	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Olive trees	0.4	0.0	98.3	0.0	0.0	0.0	2.9	0.0	2.2	0.0	0.0
Urban area	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Deciduous trees	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.2
Barren Soil	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
Evergreen shrubs	0.0	0.0	0.5	0.0	0.0	0.0	97.1	0.0	0.0	0.0	0.0
Agriculture land	0.0	0.0	0.3	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0
Grazing land	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	97.8	0.0	0.0
Chestnut trees	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
Oak forest	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.8

5. References

ERDAS Inc, (1994), ERDAS Field Guide, Atlanta, GA.

Eurimage (2003), QuickBird products at http://www.eurimage.com/products/quickbird.html

Hatzopoulos I.N., Giourga C., Koukoulas S., Margaris N., (1992). Land Cover Classification of Olive trees in the Greek Islands Using Landsat TM Images. Symposium ASP&RS, USA.

Hord, R. M., (1982). Digital Image Processing of Remotely Sensed Data. New York. Academic Press.

Koukoulas S., (1994). Development of a methodology framework for mapping the productivity status of olive trees in Lesvos island by using Landsat TM satellite data. Bachelor Thesis. Mytilene. Greece. (In Greek)