Land cover data from Landsat single-date imagery: an approach integrating pixel-based and object-based classifiers

Tommaso Ceccarelli¹, Daniela Smiraglia¹*, Sofia Bajocco¹, Simone Rinaldo¹, Antonella De Angelis¹, Luca Salvati² and Luigi Perini¹

¹Consiglio per la Ricerca e la sperimentazione in Agricoltura - Unit of Climatology and Meteorology Applied to Agriculture (CRA-CMA), Via del Caravita 7a, 00186 - Rome, Italy
²Consiglio per la Ricerca e la sperimentazione in Agricoltura - Centre for the study of Plant-Soil Interactions (CRA-RPS), Via della Navicella 2-4, 00184 - Rome, Italy

*Corresponding author, e-mail address: daniela.smiraglia@entecra.it

Abstract
This paper proposes an approach for generating land cover information from single-date Landsat images integrating pixel-based and object-based classifiers in two study areas: the province of Oristano and the region of Campania. The process consisted in: a) pre-processing; b) segmentation; c) classification based on radiometric properties and integration with textural properties and vegetation indices. A good overall classification accuracy was obtained at the first (Oristano: 87%, Campania: 88%) and at the second (Oristano: 74%, Campania: 82%) CORINE Land Cover level. Results highlight the potential of the method to be replicated in time and space in the perspective of a semi-automatic and cost-efficient transferability of the procedure.

Keywords: Pixel-Based classification, Object-Based classification, Landsat archive imagery, land cover, single-date image.

Introduction
Land cover characterization is one of the most important subjects used in projects which deal with different environmental topics, such as land cover changes [Turner et al., 2007; Soulard and Sleeter, 2012], forest monitoring [Kumar, 2011; Townshend et al., 2012], land degradation [Symeonakis et al., 2007; Bajocco et al., 2012], urbanization dynamics [Bhagyanagar et al., 2012; Fichera et al., 2012] and land quality assessment [Capotorti et al., 2012].

Aerial photographs and satellite images are the main source of information utilized to yield accurate and thorough land cover maps in space and time. Since 2008 the USGS Landsat archives provide free access to multi-spectral satellite imagery, which therefore constitute a valuable source of information for a variety of analysis concerning land cover. When applied to territorial studies at sub-national and national scales, such analysis are based on
large land cover datasets, due to their multi-temporal and spatial coverage needs [Sexton et al., 2013]. This requires automated, objective and repeatable procedures to be developed. Land cover information is generated from remote sensing data through different classification methods [Lu and Weng, 2007]. The main approaches to image analysis and classification are usually referred to as pixel-based (P-B) and object-based image analysis (OBIA) [Blaschke et al., 2008]. P-B methods are based upon the statistical classification of single pixels in a single digital image [Lillesand and Kiefer, 2000]. Recent studies indicate that P-B classification methods have a number of shortcomings; for instance they prove less effective compared to OBIA ones especially when applied to aerial photograph or to high-resolution imagery [Cleve et al., 2008; Mallinis et al., 2008]. This is due, among other reasons, to their tendency to oversample the scene [Malinverni et al., 2011], resulting in a ‘pixelized’ (salt and pepper) representation of land cover [Bock et al., 2005; Chirici et al., 2006]. In more general terms, in the classification process P-B methods are not able to represent the spatial relationships between landscape features [Schiewe et al., 2001]. To the contrary, OBIA approaches classify images based on objects and their mutual relations [Gamanya et al., 2007]. Hence, OBIA as opposed to P-B classifications may incorporate important semantic information and generate land cover objects that are uniform and meaningful from the perspective of the different application domains [Dingle Robertson and King, 2011]. For such reasons, the application of OBIA approaches as an alternative to P-B analysis has increased in the earth observation community in the last decade [Hay et al., 2005; Blaschke, 2010; Gómez et al., 2011]. Although many applications of OBIA approaches refer to very high resolution imagery, several studies tested the OBIA on high resolution satellite imagery such as ASTER [Schneevoigt et al., 2010] and Landsat [Flanders et al., 2003; Dingle Robertson and King, 2011; Gutiérrez et al., 2011]. When comparing this approach with the traditional P-B methods, many authors [e.g. Lewinski, 2006; Yan et al., 2006; Matinfar et al., 2007] documented the superiority of OBIA classifiers applied to high resolution imagery in terms of accuracy and number of classes discriminated. However, it was also shown that P-B classifications may sometimes outperform OBIA approaches in terms of accuracies, when referring to specific land cover classes [Flanders et al., 2003; Zhang et al., 2008]. Image segmentation provides the building blocks of OBIA image analysis [Blaschke, 2010]. Thanks to the recent improvements, automated image segmentation is increasingly being used in conjunction with OBIA classifiers in the identification of land cover objects [Conchedda et al., 2008; Duveiller et al., 2008]. If more than one image is available throughout the reference year, multi-seasonal analysis adds relevant information, for instance by means of vegetation indices phenology, which can be used in both the segmentation and the OBIA classification. This has been shown to be important for improving land cover classifications both in terms of accuracy and hierarchical levels. Zhang et al. [2008] and Powell and Brooks [2008] for instance, applied vegetation indices based on seasonal changes or phenological growth profiles as OBIA classifiers. In many cases, however, in the Landsat archives, only single date images (or more images of about the same period) are available for the reference year, thus limiting the power of vegetation indices in discriminating between land cover objects. This is often the case with images from the Landsat archives. In this perspective, the aims of this study are i) to generate land cover classifications from Landsat archive imagery, using single-date images, by integrating P-B and OBIA methods.
combining radiometric and textural image properties as well as vegetation indices, ii) to evaluate the accuracy of the classification and the detail of the thematic content achieved, and iii) to assess the transferability of the procedure developed in two areas which differ substantially in land cover types and heterogeneity: the province of Oristano and the region of Campania.

Study areas

The Oristano province

The Oristano province is located in the coastal zone of the central-western part of the Sardinia island (Italy) (Fig. 1), covering a surface of around 850 km². It is mainly occupied by agricultural areas (72% of the total area), most of which are arable lands (53%). These are either placed in the lowlands, in association with dairy production, or at the foothills, in association with rain-fed cultivation and goat and sheep farming systems. A peculiarity of this region is the presence of rice paddy fields (4%) as well as different types of wetlands (salt marshes, salines, inland marshes). The northern and eastern parts of the study area are characterized by a hilly topography dominated by sclerophyllous vegetation (5%). Artificial areas (6%) are principally located in the lowlands and consist mainly of the urban areas of Oristano and Cabras.

The Campania region

The Campania region is located in southern Italy (Fig. 1), covering a surface of around 13,600 km². The landscape is characterized by the mountain range of the central Apennines in the inland, volcanic areas in the coastal zone, rocky promontories and wide plains. Land cover is characterized by agricultural areas (55%), above all arable lands (24%) prevailing in the wide plains, and natural and semi-natural areas (38%), most of which are forests covering hills and mountains (28%). Artificial areas (7%) consist almost entirely of urban areas located above all along the coast and around the city of Naples.

Figure 1 - The study areas (Oristano, Sardinia, and the Region of Campania, Italy).
Reference data

LandSat images
To derive the land cover classification, we used Landsat 7 (ETM+) images freely available from the Landsat archives (http://edcsns17.cr.usgs.gov/NewEarthExplorer/). The spatial resolution of the Landsat ETM+ sensors is 30m for the bands 1–5 and 7, and 15m for the panchromatic band.
As for the province of Oristano, the original scene (path 193, row 32, acquisition date: 2003-05-19), covers a large part of the Sardinia region and we extracted only the portion of the study area. All archive scenes with such spatial extent refer to the same period of the year. The region of Campania instead is within the path 183, row 31 and the path 189, row 32 (acquisition date: 2001-05-01).

Land cover maps
The reference land cover maps used in this study was made available by the Cartographic and GIS Service of the Autonomous Region of Sardinia (Carta dell’Uso del Suolo, UDS, 2003) and by the Agricultural Department of the Region of Campania (Carta dell’Utilizzazione Agricola dei Suoli della Campania, CUAS, 2001), in accordance with the CORINE Land Cover (CLC) nomenclature. The CLC project (http://www.eea.europa.eu/publications/COR0-landcover) has been carried out by the European Environment Agency (EEA) to provide pan-European land cover maps. The CLC classification scheme is composed by several land cover categories grouped into a hierarchical nomenclature at three levels. Instead of the three levels originally envisaged, in both maps the nomenclature was extended to five levels in order to encompass the specificity of local landscapes and land uses. The UDS map has a nominal scale of 1:25,000 and a minimum mapping unit of 1 ha in urban areas and 1.5 ha in non-urban areas. The CUAS map has a nominal scale of 1:50,000 and a minimum mapping unit of 1 ha.
For the purpose of this study different levels of the nomenclature were used depending on the classes as shown in Table 1: level II was used for classes 1 (Artificial surfaces), 2 (Agricultural areas) and 3 (Forest and semi natural areas), while for the remaining classes (4, Wetlands and 5, Water bodies), the level I was retained.

<table>
<thead>
<tr>
<th>CORINE Land Cover I/II level classes</th>
<th>I level</th>
<th>II level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Artificial Surfaces</td>
<td>11 Urban fabric</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12 Industrial, commercial and transport units</td>
<td></td>
</tr>
<tr>
<td>2 Agricultural Areas</td>
<td>21 Arable land</td>
<td></td>
</tr>
<tr>
<td></td>
<td>22 Permanent crops</td>
<td></td>
</tr>
<tr>
<td>3 Forest And Semi Natural Areas</td>
<td>31 Forests</td>
<td></td>
</tr>
<tr>
<td></td>
<td>32 Scrub and/or herbaceous vegetation associations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33 Open spaces with little or no vegetation</td>
<td></td>
</tr>
<tr>
<td>4 Wetlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Water bodies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Methods

The proposed methodology integrates textural analysis, vegetation index extraction, multi-resolution segmentation, P-B supervised classification and OBIA classification (Fig. 2). The procedure can be divided into the following steps:

a) pre-processing of the Landsat images, including textural analysis and the derivation of a vegetation index;

b) segmentation of the pre-processed images;

c) classification of the segmented images based on radiometric properties and integration with textural properties and vegetation indices by means of an object oriented classification;

d) evaluation of the classification results against reference land cover maps.

Pre-processing of satellite data

Textural analysis

As discussed for instance by Møller-Jensen [1997], Pesaresi and Bianchin [2001], Yan [2006] and Matinfar [2007], a more traditional classification strategy based entirely on the spectral properties of individual pixels is not enough for adequately classifying land cover from Landsat, especially in the case of urban fabric [Møller-Jensen et al., 2005].

Many texture measures have been developed [He and Wang, 1990; Unser, 1995] that are mainly used for land cover classification [Herold et al., 2003; Yu et al., 2006]. For instance, texture analysis applied to the identification of the urban footprint is grounded on the consideration that urban areas can be defined on the basis of urban elements. Accordingly, a definition of urban texture can be given as the geometrical structure formed by the spatial distribution of urban elements as buildings, roads and green areas [Ober et al., 1997]. Based on these consideration, in order to improve the classification results, textural analysis was also applied in this research. A number of derivative bands, using a 3x3 moving window and measuring different textural properties, were generated from the original data. The texture bands were derived from the application of the convolution filter on the panchromatic band. Three different texture measures were computed: Mean, Range and Variance [Anyes et al., 1994; ENVI, 2000] and the texture value was assigned to the central pixel of each window location of the 9 pixels involved. As a result, three images were obtained, one for each texture measure.

Vegetation indexing

Satellite-based spectral vegetation indices are widely used to characterize the type, amount and condition of the vegetation within the scene [Jackson and Huete, 1991; Pettorelli et al., 2005]. For this purpose, greenness indices such as the NDVI (Normalized Difference Vegetation Index) are commonly used [Wylie et al., 2002; Ahl et al., 2006; Beck et al. 2006]. The NDVI is computed as the normalized difference between the reflectance in the Red and Near InfraRed (NIR) spectral bands and ranges from -1 to 1, where the negative values refer to non-vegetated areas, while positive low values indicate sparsely vegetated zones and values close to 1 indicate densely vegetated zones. NDVI is typically used to characterize the phenological state and the seasonal dynamics of vegetation [Zhang et al., 2003; Ahl et al., 2006] because of its demonstrated direct relation with plant cover activity [Reed et al., 2003] and biophysical parameters like the fraction of absorbed photosynthetically active radiation [Grace et al., 2007].
Figure 2 - Workflow of the integrated classification procedure.

**Segmentation**
Image segmentation is the partitioning of raster images into spatially continuous, disjointed and homogeneous regions, i.e. segments, based on pixel values and locations [Jensen, 2004]. Pixels having similar feature (textural and spectral) values that are spatially connected are grouped in single segments or objects, minimizing their heterogeneity. In this research, the image segmentation was used in two different ways. In the first place for generating the segments which were then used as training and test samples in the P-B classification. In the second place as a preliminary step in the integrated classification. The segmentations were performed based on the software eCognition®.

In order to make the resulting segments as comparable as possible in terms of resolution and to test the repeatability of the procedure to different images, the same method and parameters (scale, color, smoothness) were applied in the generation of samples and in the classification of image objects to both study areas. The multiresolution segmentation was selected as the most appropriate method. Only the feature variables used differed in the two segmentations.

**Pixel-Based classification**
The P-B supervised classification was performed after the selection of training samples which are representative and typical of all information classes. In order to derive training...
samples for the P-B classification, the Landsat images were first segmented based on all spectral bands available with the following parameters: scale 30, color 0.8 and smoothness 0.5. The parameters were applied to the Oristano test area as well as the Campania region. After that, a certain amount of segments were randomly selected in each of the five classes at I level, using the Geostatistical Analyst extension of ArcGIS®. The training samples were identified, out of those segments, using two steps: the first consisted in an automatic selection of the segments which had a minimum of 70% area overlap with the corresponding polygons of the land cover maps (UDS and CUAS); the second was a manual photo-interpretation of the resulting segments checking both the correspondence with land cover maps classes and the degree of homogeneity in each of them. Due to the relatively high heterogeneity of the class arable land (2.1), this was further divided into two classes: vegetated and not-vegetated arable land. Subsequently, polygons having a high degree of heterogeneity were discarded altogether. At the end of the process, all remaining polygons were finally labelled as the corresponding land cover class. About 50% of the polygons were retained as training samples and the remaining 50% as the test samples. The training samples were used in the P-B classification while the test samples were used to construct the confusion matrix and verify the accuracy of the classification obtained.

The P-B classification was based on the Maximum Likelihood algorithm (ML) according to the Bayes theorem. A detailed description of ML classifiers can be found in various reference texts [Richards and Jia, 1999; Jensen, 2004]. The classification was performed using the IDRISI® SELVA software and a same prior probability of 5% for the signatures of all classes. Pixels with probabilities lower than this value were not classified.

Integrated classification

The integrated approach combines radiometric and textural properties as well as vegetation indices, and is based on image objects rather than on single pixels. The classification model was developed using the eCognition® object-oriented algorithms. The image segmentation, used as a preliminary step to the integrated approach, was based on the following feature variables:

a) Landsat image layers (for all bands);

b) P-B classified image;

c) NDVI.

The same segmentation method (multiresolution) and parameters used for the selection of the training and test samples were applied in this case. The weight associated to each feature variable was 1, apart from the panchromatic band where a weight of 5 was used.

In this classification, objects originated from the segmentation are evaluated instead of individual pixels. The classification algorithm analyses features in each image object against a list of selected classes and determines its membership value of the basis of class descriptions. The classification adopted was based on fuzzy membership functions acting as class descriptors. Fuzzy membership functions describe intervals of feature values wherein the objects belong to a certain class or not by a certain degree [Pakhale and Gupta, 2010]. This is especially relevant when the same value cannot be assigned unambiguously to a class, as it was found to be the case for all the classifiers in the present study. Each class was described combining class descriptors for the different feature variables by means of fuzzy-logic operators and the objects were then grouped and located in the corresponding classes.
For this classification, in addition to the feature variables used in the segmentation, the variance, mean and range images were used, applying a 3x3 moving window over the panchromatic band. Also the Rectangular Fit, a feature based on shape, was used. This shape feature describes how well an image object fits into a rectangle of similar size and proportion. Specific object features from the above-mentioned variables were used in the classification rule-set of each land cover class, as summarized in Table 2.

An example of rule-set for the class “Urban Fabric” at the 2nd level classification for Oristano, is shown below:

And (min)

\[
\begin{align*}
&\text{NDVI} \\
&\text{MeanBand4 (NearInfrared)} \\
&\text{MeanBand7 (FarInfrared)} \\
&\text{MeanPAN_TEX_DataRange} \\
&\text{Or (max)} \\
&\text{BestClassificationPO_Class11} \\
&\text{BestClassificationPO_Class21}
\end{align*}
\]

In order to be assigned to the class, all the conditions have to be applied (logical operator “and”(min)). Moreover, the image object had to be classified at the same time as “urban fabric” or (logical operator “or”(max)) “arable land” in the pixel oriented classification. Class descriptions were defined based on their capacity of discriminating among the different classes. This was achieved by means of the eCognition® “sample editor”, inspecting image values of the samples especially in relation to classes with similar properties. From this analysis it emerged that the NDVI as well as the Mean NearInfrared and Mean Far Infrared bands were the most suited feature variables contributing to the identification of image objects without vegetation. Otherwise texture features (DataRange) especially helped in discriminating between non-vegetated arable land and “urban fabric”.

The same rule-set applied in the Oristano area, was also used for the classification of the whole Region of Campania, with minor changes in the object features used and adjustments in the class descriptions, based on samples from the available reference map. For example, still concerning the class “Urban fabric”, only the variable “Brightness” is added in rule-set classification.

**Table 2 - Summary of the key object features utilized in the integrated classification for each class (2nd level CLC ).**

<table>
<thead>
<tr>
<th>Class</th>
<th>P-B classes</th>
<th>NDVI</th>
<th>Band 4</th>
<th>Band 7</th>
<th>Texture</th>
<th>Rectangular Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>21</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Accuracy assessment of the classifications

The quality of the classifications obtained through the integrated approach was assessed taking into account the hierarchical levels of the land cover classification scheme and the accuracy in the same classifications. The accuracy of the different classification schemes was evaluated by means of confusion matrices according to the land cover maps (UDS and CUAS). Therefore, the observed values were obtained from the land cover maps, while the predicted values are the result of the integrated classifications. All confusion matrices were constructed accordingly, firstly at the CLC level I and secondly at a II level for specific land covers, as early discussed. The test samples in the P-B classification were used for evaluating the accuracy of the integrated classifications.

Results

The integrated approach to the classification was performed initially at the first level of the CLC nomenclature, i.e. generating five, more generalized land cover classes. Figure 3 shows the classified image together with a true color composite of the satellite image for the Oristano area and Table 3 provides the results of the accuracies attained by means of a confusion matrix.

An overall accuracy of 86.6 % was attained with a Kappa Index of Agreement (KIA) of 75%. A number of classes was more accurately identified than others as class 5 (Water bodies: user accuracy 92.5%, producer accuracy 93.8%) and class 2 (Agricultural areas: user accuracy 87.7%, producer accuracy 94.4%). For class 1 (Artificial surfaces), the user accuracy achieved was 81% and the producer accuracy 51.6%. For class 3 (Forest and semi natural areas), the user accuracy was 81% and the producer accuracy 73.2%. Finally class 4 (Wetlands) had a producer accuracy of only 53.7% and an user accuracy of 11.5%.

Table 3 - Confusion matrix of the classification of the integrated approach at the CLC level I, Oristano area. Occurrences for each land cover class are in hectares. Producer (PA), user’s (UA), and overall (OA) accuracies in percentage as well as the KIA are reported.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference map</th>
<th>Total</th>
<th>PA %</th>
<th>UA %</th>
<th>OA %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>695</td>
<td>1,347</td>
<td>51.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13,817</td>
<td>14,629</td>
<td>94.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3,117</td>
<td>4,259</td>
<td>73.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>348</td>
<td>11.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2,706</td>
<td>2,884</td>
<td>93.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>858</td>
<td>23,468</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA %</td>
<td>81.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OA %</td>
<td>86.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIA</td>
<td>75%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A second classification illustrates the potential of the approach in identifying a number of classes up to the II level of the CLC nomenclature. This is illustrated by the confusion matrix in Table 4.
In this case the overall accuracy was found to be around 74% and the KIA 57%. Apart for class 5 again, the best results were achieved for class 11 (Urban fabric: user accuracy 80%, producer accuracy 62.7%) and class 2.1 (Arable land: user accuracy 77.4%, producer accuracy 94.3%); in all remaining classes user and producer accuracies were below 60%, except for the producer accuracy in class 3.1 (Forests: 81.1%).

The classified image of the region of Campania is shown in Figure 4 and the confusion matrix at the I level of the CLC nomenclature in Table 5.

The overall accuracy attained was 87.7% with a KIA of 78%. In terms of individual classes, the classes which were better classified were class 1 (Artificial surfaces: user accuracy 96.2%, producer accuracy 82.7%), class 3 (Forest and semi natural areas: user accuracy 96.5%, producer accuracy 87.7%), and class 2 (Agricultural areas: user accuracy 75%, producer accuracy 91.2%). For class 5 (Water bodies) a lower accuracy was achieved (user accuracy 61.1%, producer accuracy 89.7%) compared to the province of Oristano. In the region, there are virtually no areas falling under the class 4 (Wetlands). A limited amount of unclassified objects correspond to clouds, mostly concentrated in the central part of the study area.

As reported in Table 6, at the II level of the CLC nomenclature, the overall accuracy was 82.1% with a KIA of 71%. When considering specific land cover classes, the better accuracies achieved were for class 33 (Open spaces with little or no vegetation: user accuracy 93.1%,
producer accuracy 75.4%), class 31 (Forests: user accuracy 96.5%, producer accuracy 84.2%), class 11 (Urban fabric: user accuracy 78.5% producer accuracy 79.4%). It was not possible to further discriminate class 2 (Agricultural Areas) between Arable land and Permanent Crops. Class 32 (Scrub and/or herbaceous vegetation associations) was also poorly discriminated (user accuracy 26% and producer accuracy 39.1%).

Table 4 - Confusion matrix of the integrated classification at the CLC level I/II, Oristano area. Occurrences for each land cover class are in hectares. Producer (PA), user’s (UA), and overall (OA) accuracies in percentage as well as the KIA are reported.

<table>
<thead>
<tr>
<th>Reference Map</th>
<th>Classification</th>
<th>PA %</th>
<th>OA %</th>
<th>KIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>646</td>
<td>62.7</td>
<td>80.0</td>
<td>57%</td>
</tr>
<tr>
<td>12</td>
<td>30</td>
<td>4.8</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>11,988</td>
<td>94.3</td>
<td>77.4</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>1,616</td>
<td>5.1</td>
<td>36.2</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>1,554</td>
<td>81.1</td>
<td>59.4</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>2,186</td>
<td>13.4</td>
<td>47.3</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>1,915</td>
<td>7.9</td>
<td>53.7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>140</td>
<td>11.5</td>
<td>92.5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>141</td>
<td>93.8</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2,924</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

The results of the classifications showed high levels of accuracy for both the Oristano and Campania areas. When looking at inter-class performances at the I level of the CLC nomenclature, in most cases user and producer accuracies are above 75%. Exceptions are in the Oristano area where a high omission error was observed for class 1 (Artificial surfaces) given the importance of small rural residential settlements. Also class 4 (Wetlands) was poorly discriminated (Tab. 3) due to the considerable differences existing between the different types of wetlands in the area (salt marshes, salines, inland marshes). In Campania, a high commission error was achieved (around 39%) only for class 5 (Water bodies), probably due to the limited representativeness of the test samples. For the mentioned classes, ancillary information (e.g. distance to the coastline to discriminate between salt and inland marshes) as well as a different sampling design could improve the classification while introducing only minor modifications in the general classification model.

At the II level of the CLC nomenclature, inter-class accuracies showed satisfactory results for a number of classes. In the Oristano area the best accuracies were achieved for the class 21 (Arable land) and the class 11 (Urban fabric). In Campania good results were similarly achieved for the class 11. This shows the potentiality of this method in identifying and monitoring urban dynamics, as also obtained by Shaban and Dikshit [2001] who
investigated an Indian urban area, and found that a combination of texture and spectral features improved significantly the classification accuracy. However, a lower accuracy regards class 12 (Industrial, commercial and transport units), highlighting the difficulty to well identify this land cover class in the integrated classification method. Also in this case, ancillary information such as features extracted from topographic maps, could help improving the accuracy classification, with limited adjustments in the segmentation and classification scheme.

Figure 4 - True color composite of the Landsat 7 image and the integrated classification at CLC level I for the Region of Campania.

In Campania good results were also attained for specific classes in the Forest and semi natural areas (31, Forest, and 33, Open spaces with little or no vegetation) indicating the potential of the proposed approach in discriminating between natural spaces and cultivated areas, also providing information about their fragmentation state. On the contrary, it was not possible to further discriminate between arable land and permanent crops, and the accuracy obtained for class 32 (Scrub and/or herbaceous vegetation associations) was also low (Tab. 6). This clearly indicates the limits of relying on single-date information. When multi-temporal profiles and trends based on vegetation indices can be related to crop phenology or crop cycles, it is expected that these classes can be better classified even on the same Landsat images, and with limited adjustments in the segmentation as well as in the classification schemes.
Table 5 - Confusion matrix of the integrated classification at the CLC level I, Campania Region. Occurrences for each land cover class are in hectares. Producer (PA), user’s (UA), and overall (OA) accuracies in percentage as well as the KIA are reported.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference map</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Unclassified</th>
<th>Total</th>
<th>PA %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4,221</td>
<td>771</td>
<td>101</td>
<td>0</td>
<td>9</td>
<td>101</td>
<td>5,101</td>
<td>82.7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>167</td>
<td>17,060</td>
<td>1,475</td>
<td>0</td>
<td>11</td>
<td>185</td>
<td>18,713</td>
<td>91.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>4,610</td>
<td>34,007</td>
<td>0</td>
<td>118</td>
<td>35</td>
<td>38,766</td>
<td>87.7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>57</td>
<td>18</td>
<td>0</td>
<td>17</td>
<td>96</td>
<td>106</td>
<td>96.5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>20</td>
<td>2</td>
<td>0</td>
<td>218</td>
<td>39</td>
<td>243</td>
<td>89.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4,387</td>
<td>22,732</td>
<td>35,245</td>
<td>0</td>
<td>357</td>
<td>557</td>
<td>63,279</td>
<td>OA %</td>
<td></td>
</tr>
<tr>
<td>UA %</td>
<td>96.2</td>
<td>75.0</td>
<td>96.5</td>
<td>61.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIA</td>
<td>78%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6 - Confusion matrix of the integrated classification at the CLC level I/II, Campania Region. Occurrences for each land cover class are in hectares. Producer (PA), user’s (UA), and overall (OA) accuracies in percentage as well as the KIA are reported.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference map</th>
<th>11</th>
<th>12</th>
<th>2</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>4</th>
<th>5</th>
<th>unclass.</th>
<th>Total</th>
<th>PA %</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>3,005</td>
<td>120</td>
<td>572</td>
<td>61</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>18</td>
<td>3,785</td>
<td>79.4</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>727</td>
<td>369</td>
<td>199</td>
<td>0</td>
<td>19</td>
<td>13</td>
<td>0</td>
<td>6</td>
<td>83</td>
<td>1,417</td>
<td>26.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td>102</td>
<td>17,060</td>
<td>655</td>
<td>819</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>185</td>
<td>18,897</td>
<td>90.3</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>0</td>
<td>3,152</td>
<td>29,858</td>
<td>2,338</td>
<td>0</td>
<td>0</td>
<td>92</td>
<td>34</td>
<td>35,474</td>
<td>84.2</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>2</td>
<td>1,400</td>
<td>336</td>
<td>1,118</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2,856</td>
<td>39.1</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>30</td>
<td>1</td>
<td>58</td>
<td>2</td>
<td>0</td>
<td>354</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>470</td>
<td>75.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0</td>
<td>57</td>
<td>15</td>
<td>0</td>
<td>37</td>
<td>0</td>
<td>17</td>
<td>96</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>3</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>218</td>
<td>39</td>
<td>282</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3,830</td>
<td>596</td>
<td>22,518</td>
<td>30,928</td>
<td>4,294</td>
<td>380</td>
<td>0</td>
<td>374</td>
<td>359</td>
<td>63,279</td>
<td>OA %</td>
<td></td>
</tr>
<tr>
<td>UA %</td>
<td>78.5</td>
<td>61.9</td>
<td>75.8</td>
<td>96.5</td>
<td>61.0</td>
<td>58.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIA</td>
<td>71%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the Campania region were attained retaining the same feature variables as well as almost the same rule sets in the classification. Only minor modifications were required in the definition of the class descriptions constructed, as usual, on the available reference maps.

This indicates the solidity of the approach when transferring the classification procedure to a very different geographical context. It also shows the potential for covering large areas with very limited adjustments required from the analyst.
Conclusion

The recent free access to the Landsat archives, makes it possible the use of satellite imagery for the classification of land cover in the past two decades and provides a unique and invaluable data source for tracking change in land covers and landscapes [Woodcock et al., 2001], relating it to a number of environmental dynamics and problems.

Several studies utilized Landsat imagery to test the performances (i.e. in terms of classification accuracy and classes identified) of the two major image classification methods: pixel-based and object-based [Desclée et al., 2006; Lewinski, 2006; Blaschke et al., 2008; Gutiérrez et al., 2011]. The approach proposed in this paper focused on the integration of the two methods to improve the results of land cover classification, based on archive Landsat ETM+ and when only single-date imagery is available.

Also with reference to previous work in this domain [Wang et al., 2004; Yuan and Bauer, 2006; Malinverni et al., 2011], this study attempted to demonstrates the potential of an integrated approach based on features derived from spectral, textural as well as vegetation indices properties, in improving the classification performance. The results suggest the possibility of applying the integrated methodology to discriminate not only among very generalized land cover categories, but also, in some cases at least, between specific vegetation typologies. The integrated classification allowed insights at fairly detailed levels of the thematic classification, suggesting the applicability of this procedure to complex landscapes. More specifically the proposed approach proved to achieve satisfactory results both at the I and at the II level of the CLC nomenclature identifying, in the last case, a number of classes with high levels of accuracy.

The approach also showed the limitations encountered in discriminating a number of specific classes, and suggested that in this case specific ancillary information as well as multi-temporal series of images could improve on the results, with limited additions in terms of feature variables and adjustments in the segmentation and classification procedure.

The transferability of the procedure in different spatial contexts was also investigated. Results are encouraging showing that only minor adaptations (and hence inputs from the analyst) are needed for achieving satisfactory results over large and potentially heterogeneous areas in terms of land covers and landscapes.

In perspective, this appears to be a promising achievement towards a semi-automatic and, thus, cost-efficient replication of the procedure in space (for covering different and larger areas, i.e. different scenes) and potentially also in time (for constructing comparable land cover time series).

Indeed, the proposed approach could help to derive a classification which can meet the requirements of the land cover change analysis, even when the availability of images in the archives does not allow a multi-temporal evaluation. Noteworthy is that the accuracy values obtained in the integrated classification were based only on one satellite images for a reference year that, notwithstanding their intrinsic limits, allowed reaching acceptable results. An additional drawback of the single-date image analysis is in the fact that vegetation indices cannot be fully exploited, e.g. drawing profiles or trends related to crop phenology or crop cycles. In this context the proposed procedure guarantees in any case good results in the classification.

While this study aims above all at classifying archive imagery, and is based on Landsat 7, the approach could be extended to future satellite data acquisition, providing a valuable
monitoring tool. The condition is of course that earth observation programmes such as the Landsat 8 - continuity mission and the European Sentinel Missions (especially the Sentinel 2 constellation), will provide in the future similar satellite data, and in the context of similar (free or low cost) pricing policies.

Acknowledgement
The research of this paper was financed by ‘Agroscenari’ project (research unit 6a) funded by the Italian Ministry of Agricultural and Forestry Policies.

References


Cleve C., Kelly M., Kearns F.R., Moritz M. (2008) - Classification of the wildland–urban


Soulard C.E., Sleeter B.M. (2012) - *Late twentieth century land-cover change in the basin and range ecoregions of the United States*. Regional Environmental Change, 12: 813-823. doi: http://dx.doi.org/10.1007/s10113-012-0296-3.


