POLITECNICO DI TORINO

Doctorate school

PhD in Environment and Territory - Environmental Protection and Management

XXV cycle

Final Dissertation

Urban land cover mapping using medium spatial resolution satellite imageries: effectiveness of Decision Tree Classifier

Steffenino Sara

Tutor: Co-tutor:

Prof. Piero Boccardo   Ing. Franca Disabato

April 2013

Table of Contents

Introduction

Disasters are crises that cause widespread damages and are beyond our ability to prevent or control. A progressive strategy for managing disasters ensures that loss of life and property is reduced in a disaster event: the primary goals of disaster management are to minimize losses from the disaster, to provide help to the victims and to facilitate a rapid recovery. Pre-disaster phases aim at identifying potential risks and the actions that can be taken before a disaster event in order to reduce its impact; post-disaster phases, instead, are aimed at providing assistance to people affected by a disaster, saving lives, protecting properties, making an affected area safe, and subsequently to regain a proper level of functioning at medium-long term.

The study, performed in collaboration within ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action) association, a research center that works in the field of emergency management by providing methods, tools and services for both pre and post-disaster phases, is inserted in the framework of the rapid mapping activities performed by the association itself.

The aim of the study is to provide reliable information about areas affected by the disaster in a few hours after it occurs; this information, properly used by civil protection operators, can help in better managing the first-aid in the field, enhancing the knowledge about extension of the affected area and occurred damages, particularly in less accessible zones. Satellite remote sensing data are a key element in rapid mapping activities (Ajmar et al. 2010b); space-based information are sources for mapping the impact of disaster, and can potentially provide the earliest information about the impact of disaster, but they can be used effectively only when the spatial data on other aspects (for example administrative boundaries, infrastructure details, settlements etc.) are integrated with the space-based information; these data layers can be used also for analysis and modelling along with scientific information to prepare risk maps related to various hazards (Ravan 2010). Where there is a lack of this kind of information, it's necessary to derive it from different data source, such as archive satellite imagery: even if automatic classification techniques (both pixel based and object-oriented) have been successfully tested in several case studies, the most common approach in an operational context is still visual interpretation (Boccardo and Giulio Tonolo 2012). Participatory mapping through media such as OpenStreetMap or Google Map Maker are very useful when data are missing on large extent: in cases such as Haiti earthquake, the city was mapped within few days thanks to these sources (Ajmar et al. 2010a). This kind of support rarely comes, and therefore the digitalization operations could become very heavy. Here's why the creation of procedures for extracting thematism in automatic or semi-automatic way is so important.

Between all necessary information, those related to the extent of urban areas are particularly relevant for supporting crisis management because it's important to know where buildings and population are placed in order to provide information about the extent of damaged areas and population potentially involved during an emergency situation: all the crisis management cycle - including damage assessment, recovery, reconstruction and planning should benefit from an improved and globally-consistent description of human settlements (Pesaresi and Halkia 2012).

The mapping of human settlements has gained more and more importance due to increasing and rapid global urbanisation and the various effects of this development (Felbier et al. 2012); as a matter of fact in 2011 the global population passed the 7 billion mark, with more than half of the population living in urban areas. Between 2011 and 2050, the world population is expected to increase by 2.3 billion, increasing to 9.3 billion and, at the same time, the urban population is projected to gain 2.6 billion, passing from 3.6 billion in 2011 to 6.3 billion in 2050. A study performed by the Population Division of the Department of Economic & Social Affairs of the United Nations (World Urbanization Prospects – the 2011 revision) in order to understand the connection between urbanization and natural hazards, using a sample of 633 cities with more than 750.000 inhabitants in 2011, demonstrated that of the more than 450 urban areas with 1 million inhabitants or more (representing 1.4 billion people), 60%, or about 890 million people, were living in areas of high risk of exposure to at least one natural hazard. In particular, cities in Latin America and the Caribbean, in Northern America, and in Asia, are the more located in areas exposed to natural hazards. Flooding is the most frequent and greatest hazard; drought, cyclones and earthquakes follow.

When looking for data about built-up areas, mostly for what mapping at small scale is concerned, the need is to have freely available data, complete over the globe, at proper resolution and in a format suitable to be used in a GIS environment: sometimes it's preferred to gain with information at lower resolution but available with global extent without any gaps rather than using costly high resolution data. Different projects made available global maps where urban areas were represented: until about fifteen years ago, global land cover datasets were based on pre-existing maps and atlases compiled from ground surveys, national mapping programs, and highly generalized biogeographic maps (Matthews 1983; Wilson M. and Henderson-Sellers A. 1984). In the 1990's, small scale land cover maps were, for the first time, based on land surface properties observed from remote sensing (Stone et al. 1994; Defries et al. 1995; Hansen et al. 2000; Loveland et al. 2000). Nowadays, nearly all global urban maps exploit remotely sensed imagery, either directly as input data or indirectly using remote sensing-based products (Schneider et al. 2010).

Actually, best resolution to map Urban areas at global level is 300 m (GlobCover, Arino et al. 2009); best resolution is available when mapping is performed at continental or country level, but this lead to gap of information: not all areas are covered in the same way and often data are missing where more necessary, like in developing countries.

Although high resolution (HR, < 10 meter spatial resolution) and even very high resolution (VHR, < 1 meter) imageries with an almost global coverage are available, no consistent global coverage of settlements derived from those datasets exists. This is due to the fact that HR and VHR satellite data are operated on a commercial basis and consequently complete global coverage is costly; moreover, for some applications like change detection, urban sprawl characterization, urban vegetation monitoring and (partially) disaster management, the availability of low-cost images as well as expensive fine resolution data sets is equally important, and the lower cost of coarser-resolution images makes them more appealing for these applications. In the public sector a need emerges to find out possible uses of available sensors for new, interesting characterizations, and this is the case of urban studies using Landsat multispectral data (Lisini et al. 2005). Multi-spectral data, ranging from 20 meter to 100 meter ground pixel size, have proven to be very effective in

environmental analysis at local to regional scale, particularly in areas where human influence on the natural environment is heavy, as in the case of urban and sub-urban areas, due to the good compromise between the spatial and spectral details provided (Villa 2012).

The main aim is therefore to develop algorithms for the classification of medium-resolution data that can be applied at global level.

Aim of the research

This research work is performed in the above-mentioned framework with the aim of developing algorithms for the extraction of urban layer by means of medium spatial resolution Landsat data processing; decision tree is investigated as classification technique, due to its ability in establishing which are the most relevant information to be used for the classification process and its capability of extracting rules that can be later applied to any desired input.

The classifier, that is constituted by rules that, once applied to the image, allow to generate the classification, is created on the basis of a sample that contains classified elements: the algorithm evaluates which are the most significant attributes in order to separate classes, together with thresholds that have to be used in order to create rules for the land cover extraction.

The aim of this work is to evaluate which steps to perform in order to obtain a good classification procedure, mainly focusing on processing that can be applied to images and on training set features.

Therefore, the training set is evaluated on the basis of:

1. Number of classes with whom it is created;
2. Temporal extension;
3. Input attributes;

While images are subject to:

4. Different kind or radiometric pre-processing;
5. Different post-processing

For what the different variables examined for the training set generation is concerned, the following elements are considered:

1. The aim of the classification process is to generate a Urban/Non urban binary mask; from a theoretical point of view, it could be enough to use a training set with only Urban and Non Urban elements in order to learn the classifier, taking into account land cover classes such as residential areas, industrial areas and streets for what Urban is concerned, and vegetation, cultivated fields, bare soil and everything else for what Non urban is concerned. Since very different classes would be joined together, mainly for what Non urban class is concerned, it wonder if this wouldn't cause any problems in detecting rules. The first test is therefore aimed at analyzing this task: taking into consideration different land cover classes, from one side classifiers will be learned on the basis of two macro-classes of Urban/Non urban, from the other they will be learned on the basis of all considered land cover classes and than rules for the desired land cover class extracted.

2. Since the aim is to create rules that can later be applied to scenes acquired at different times and in different locations, it is necessary to consider what are the

4

most suitable scenes for the training sets extraction. The concept of time is considered in this second part. From one side there is the need to extract information about the different land cover classes that is representative of more than one acquisition, from the other side the aim is to evaluate if the change that necessarily occours in some land cover class according to the season (e.g. vegetation) and less in other (e.g. urban) could be a useful element to separate classes and detect built-up areas. In order to analyze this task, information from a single image will be extracted in one case, in the other information will be extracted as a mean over a multitemporal set.

3. The third test concern input attributes. Spectral information constitutes the first element to be used in order to characterize the different land cover classes. Other elements can be considered: indexes were evaluated in this framework. They allow enhancing some kind of behaviour and are derived from pixel spectral information. The aim is, in this case, to evaluate how to combine indexes and spectral information in order to create the training set (that is to determine if it's better to use only spectral information, only indexes, or both) and to determine which are the most suitable indexes to better separate classes (that is indexes that are, more frequently, chosen by the algorithm to be part of the classifier).

The second macro-analysis is related to the detection of best processing for input data (satellite images) and for the classification process output (urban layer).

4. Images need, before being used, to be corrected from some errors that are inherent to the acquisition process; particularly, the focus is on radiometric errors that depends upon sensor, system geometry and atmosphere. Effects due to sensor and system geometry are corrected through a calibration into reflectance (thus allowing a better comparison between different scenes) while effects due to the presence of atmosphere have to be removed in order to better compare images acquired in different moments. Atmospheric correction can be performed with simplified methods or with rigorous models: the aim is here to evaluate if the application of more complex pre-elaborations (thus more time-demanding) is needed in this specific context in order to enhance classification accuracies.

5. Classification results can be further submitted to post-processing in order to reduce the noise typical induced by pixel-based classifiers. This happens when single pixels are classified differently than the surrounding area and homogenous regions cannot be generated: post-processing can help in reducing this effect. When applying post-processing with a clustering effect, the Minimum Mapping Unit is increased. In this framework different post-processing were tested on small areas, then the best one applied to all images in order to compare classification results with or without post-processing.

The work is performed on medium spatial resolution Landsat ETM+ data; above-mentioned variables are compared and results evaluated on the basis of reached accuracies. Data used for the validation is derived from the Digital Regional Technical Map. The validation mask, generated from a more refined data (one order of magnitude larger), is a very accurate data to be used for

validation purpose: the mask itself or the procedure used to obtain it, can be used for further applications.

Therefore, the thesis is subdivided in the following sections:

Chapter 1 – it aims at providing an overview of built-up areas dataset freely and globally available to be used for mapping purposes; the review is extended also to dataset covering only Europe, in order to provide examples on how spatial resolution can be improved focusing on a smaller coverage. A second part is aimed at describing the main phases of a classification process, as well as providing a review of the available classification methods; lastly, the final part is devoted at describing which methods were used for the production of the above mentioned datasets, focusing on the ones derived from remotely-sensed imagery.

Chapter 2 analyzes data, methods and software used in the framework of this research. The classifier, the classification process, the algorithm used for the study together with its main features are analyzed in the first part; data used for the study, both in input and for the validation phase, are described later. As far as input data is concerned, the main technical features, information about the mission, and reasons that led to the choice of this data are analyzed; data used for the validation are instead examined also on the basis of elements that are part of the dataset itself. In the end, main software used for data analysis and management are briefly described.

Chapter 3 describes methodology used to examine the above mentioned tasks; all phases are analyzed in the detail, and a summary of created procedures is also provided. A section is also devoted to provide detailed information about the study area and specific used scenes.

Chapter 4, at last, provides results of the performed analysis. They are evaluated both in qualitative way, in order to evaluate the generated output and to obtain information about most-common errors of the classification process, both in quantitative way, using the Confusion Matrix and accuracy measures derived from it. An overview about rules generated by classifiers that provided better results is also given.

1   Literary review

1.1   Introduction

Satellite imageries are considered a key element in the framework of emergency mapping services, but a great importance is also given to reference data that are essential in order to provide a base upon which add information about damages caused by the event. Updated and reliable data inventory are desirable in this framework: freely available data, provided at global level and in a format suitable to be used in a GIS environment is the three main requirements.

This work focuses on a particular thematism that is built-up areas. Built-up areas can be intended in two different ways: as a cartographic base layer, where elements are grouped in different categories and coded, or as a thematism, where the information is linked to the concept of land cover: elements are not coded in this case, just built-up area boundaries are provided. Cartographic base layers, that contain a huge amount of information, are usually provided with a limited extension; on the contrary, "built-up areas" thematism can be produced on a wider extent. Since our interest is to have information about the extension of urban areas, and not to single elements that are in the urban area, and since our aim is to deal with data available at global level, we will refer to urban with the meaning of urban thematism.

The aim of this chapter is to provide the state of the art of urban areas datasets available at global level (chapter 1.2.1); a section is also devoted to provide an inventory of dataset available at European level (chapter 1.2.2), in order to point out how spatial resolution can be improved focusing on smaller extent. A general overview of the main phases of a classification process and taxonomy of image classification methods and techniques is instead provided in chapter 1.3.1 and 1.3.2; a wide description of how the described methods and techniques were used for the production of the global and continental urban maps is provided in chapter 1.3.3.

1.2   Urban areas: inventory of existing dataset

Until about fifteen years ago, global land cover datasets were based on pre-existing maps and atlases compiled from ground surveys, national mapping programs, and highly generalized biogeographic maps (Matthews 1983; Wilson M. and Henderson-Sellers A. 1984). In the 1990's, small scale land cover maps were, for the first time, based on land surface properties observed from remote sensing (Stone et al. 1994; Defries et al. 1995; Hansen et al. 2000; Loveland et al. 2000). Nowadays, nearly all global urban maps exploit remotely sensed imagery, either directly as input data or indirectly using remote sensing-based products (Schneider et al. 2010).

Satellite remote sensing, in fact, made it possible to gather information about not accessible areas (like mountainous area, foreign or dangerous country, etc.), offers wide regional coverage and good spectral resolution, enables continuous acquisition of data and serves as a large archive of historical data.

Information about urban areas derived from remote sensing data can be extracted at different scales and with different coverage; usually, datasets available at global level suffer from lack of detail (ESA provides, with the Globcover dataset at 300 m resolution, the best dataset for built-up areas), while dataset at higher resolution are provided at country level or less. A wide percentage of landscape could not be covered with data at proper scale.

This section provides an overview about datasets of built up areas; only datasets with the following features are considered:

- freely available;
- globally consistent;
- in a format suitable to be used in a GIS environment

Why are these features so important?

The importance of having freely available data is easily explained: organizations not always have to possibility to sustain the cost necessary for the purchase of commercial data.

The consistency of the data is necessary to be certain not having any gaps, and having all areas provided in the same way and with almost the same level of accuracy. Let's think on data such as the one provided by OpenStreetMap: a huge amount of data is freely available, and information such as buildings and land use are provided. This kind of information could be very useful for emergency mapping purpose, after a preliminary quality check, but the coverage is not uniform: example of different coverage is provided in Figure 1 and Figure 2.
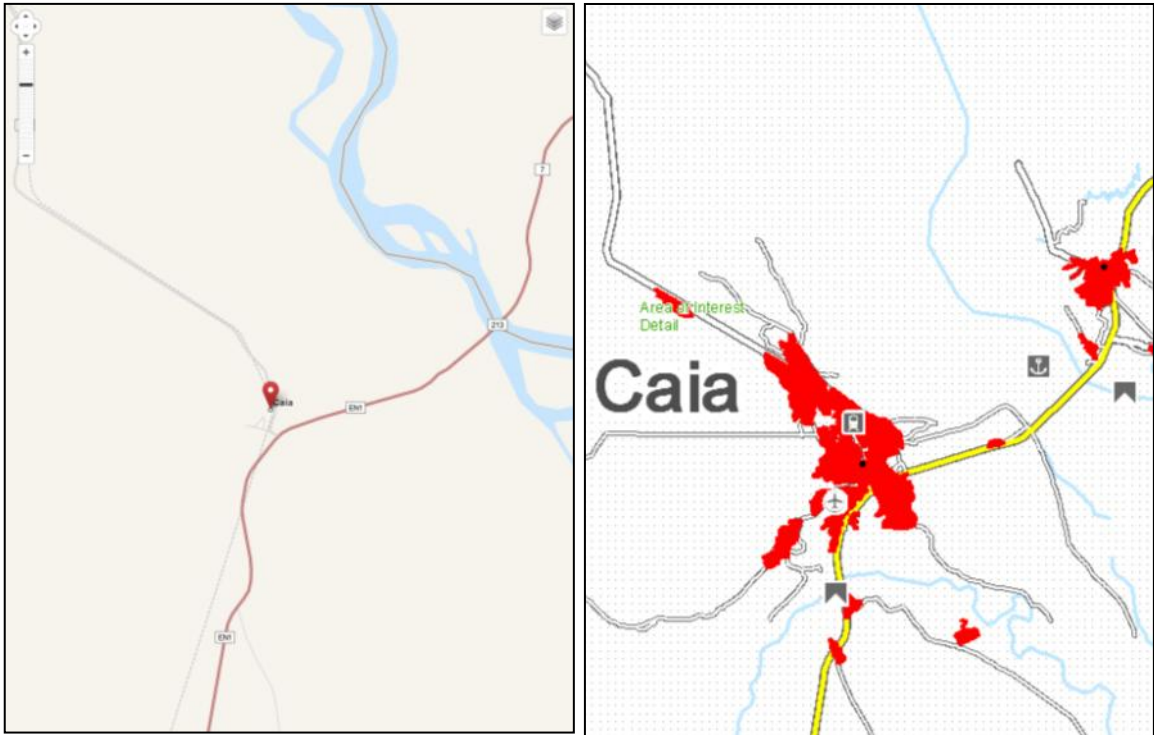
Figure 1 – Example of reference data necessary for the hypothetic production of cartography at medium-small scale (1:350.000) for emergency management: built up area boundaries and major roads are needed (on the right). Only two major roads and no information about built-up area is provided by OpenStreetMap (on the left)

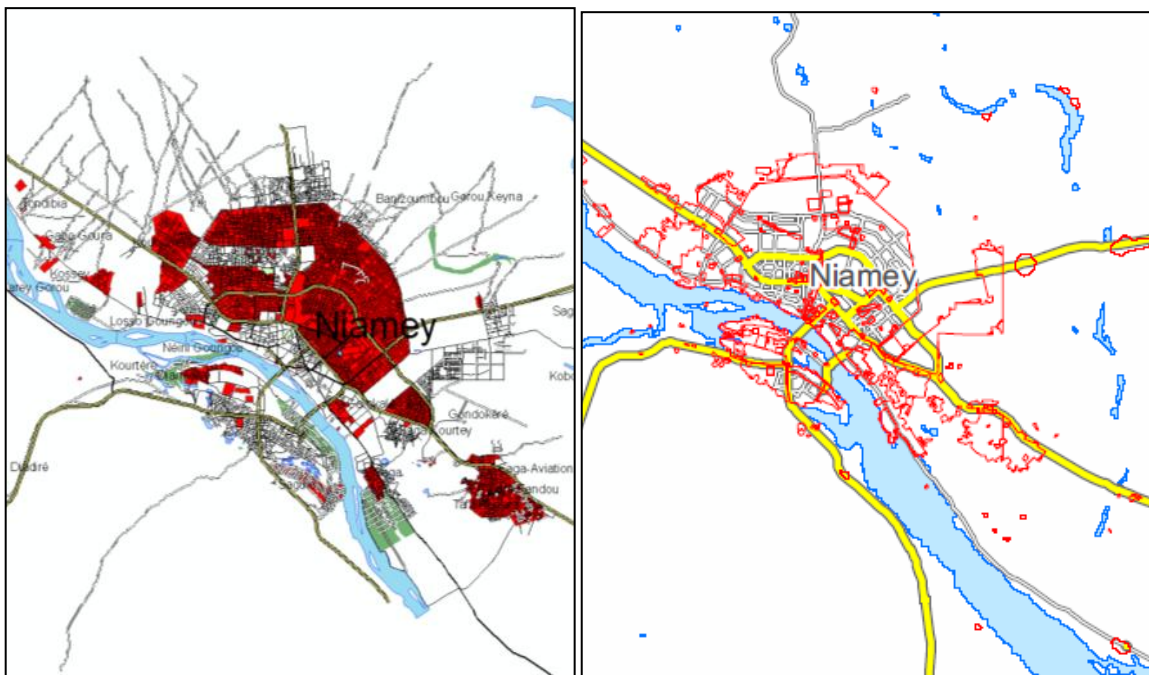**Figure 2 – Example of area where lot of information, provided by OpenStreetMap (on the left), can be used for the production of cartography at medium-small scale (1:250.000) for emergency management (on the right). Built-up areas appear in red colour**

For what the third requirement is concerned, the way information is provided is very important: it's necessary to deal with data in vector format, or in raster format with unique and coded values that allows easily extracting the thematism. Land cover datasets are a suitable type of information, while cartography, in a raster format, could be difficult to manage, also if georeferenced.

The provided inventory focus on features such us the used source data (name and date of the resource), the dataset content (information that has been extracted from the satellite data), the dataset type (raster or vector), the scale or spatial resolution of the final output and the way to access the resource (link ftp or download from the website).
The analysis was aimed at collecting information about dataset at global level; it order to provide examples of how much can the spatial resolution be improved focusing on smaller extent, a section was devoted to analyze data available in Europe.

A first data inventory of global land cover datasets was performed by the author in the framework of a preparatory activity, promoted by the European Commission, in support to the GMES (Global Monitoring for Environment) initiative. The initiative, headed by the European Commission (EC) in partnership with the European Space Agency (ESA) and the European Environment Agency (EEA), is an Earth observation programme aimed at providing accurate, timely and easily accessible information to improve the management of the environment, understand and mitigate the effects of climate change and ensure civil security (http://www.esa.int/Our_Activities/Observing_the_Earth/GMES). The preparatory activity performed in this framework, "Implementation of an initial GMES service for geospatial reference data access covering areas outside Europe – Lot 2", was aimed at realizing a geospatial reference data access through the following phases: data gathering activities, definition of a protocol for data quality checking, transformation of data into formats suitable for storage and their distribution. A part of the performed inventory, available in the Task 10 of the project delivery (Boccardo et al. 2012), was extracted and extended in this thesis in order to include also dataset specific for Urban thematism and to include datasets at higher spatial resolution: Europe was chosen as test site.

Therefore, the description of existing works is subdivided the following sub-sections:

- Global urban maps (chapter 1.2.1);
- Datasets of urban areas available in Europe at continental level (chapter 1.2.2).

1.2.1 Global Urban Maps

In this section, the inventory performed in the framework of the GMES project (paragraph 1.2), adapted and extended according to our needs, is provided; information about urban maps available at global level is summarized in Table 1; Figure 3 represents the correlation between spatial resolution and reference data of global urban maps.

Datasets of urban areas are usually provided at a global level as part of more complex databases that include not only this feature, but, generically, all information concerning land cover; only a few dataset are specific for the urban thematism. In land cover datasets, urban can be defined differently according to the adopted Classification System: for example, in the Land Cover Classification Scheme (LCCS, Di Gregorio and Jansen 2005) proposed by FAO, urban is defined as "Artificial Surfaces and associated areas", according to the 17-class International Geosphere–Biosphere Programme Classification (IGBP, Loveland and Belward 1997) is mentioned as "Urban and Built-up", while in the 14-class University of Maryland classification (UMD; Hansen et al. 2000), in the 10-class system used by the MODIS LAI/FPAR algorithm (Myneni et al. 2002; Lotsch et al. 2003), in the 8-Biome classification proposed by Running et al. (1995) and in the 12-Class plant functional type classification described by (Bonan et al. 2002), it is simply mentioned as "urban".

Between selected global urban maps, one is a multi-thematic database that includes information about population and built-up areas (VMap0), six are general multi-class land cover maps that include an urban class (AVHHR Global Land Cover Classification, GLCC, GLC2000, GLCNMO, MCD12Q1, GlobCover 2009) while two are binary (presence/absence) maps devoted entirely to urban land: MOD500 and the Dynamics of Global Urban Expansion. Among these, The Dynamics of Global Urban Expansion contains maps only for 120 cities in the globe with more than 100.000 inhabitants.

It's important to mention other two projects devoted to the creation of global maps of urban settlements: the Global Human Settlement Layer (GHSL), performed by the Joint Research Centre (JRC) of the European Commission, and the Global Urban Footprint (GUF) of the German Aerospace Center (DLR). Since they are not actually complete or are only in a preliminary phase, they have been described but not mentioned in the following tables.
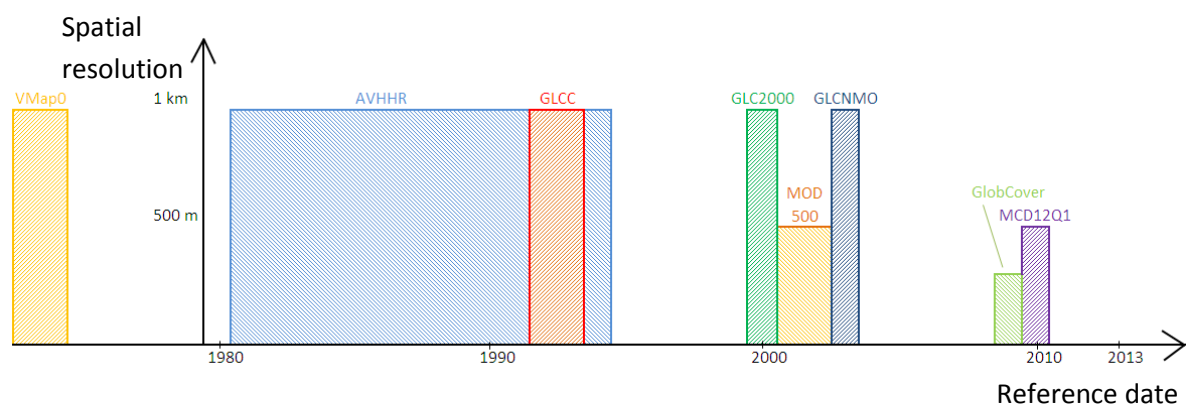
Figure 3 - Representation of existing global urban maps with reference to scale (or spatial resolution) and reference date

Product name	Producer	Source data (name and date of the resource)	Dataset content	Dataset type	Scale or spatial resolution	Access instruction
Vector Map Level 0 (VMap0)	NGA	Version 5 of VMap0 is based on the newest editions of the 1:1,000,000 Operational Navigation Charts, published by the *United States National Imagery and Mapping Agency* on September 30, 2000	Urban areas	vector	1:1.000.000	http://www.mapability.com/info/vmap0_index.html
AVHHR Global Land Cover Classification	University of Maryland Department of Geography (UMD)	AVHRR imageries acquired between 1981 and 1994	land cover	raster	1° 8 km 1 km	ftp.glcf.umiacs.umd.edu/data/landcover/data.shtml http://glcf.umiacs.umd.edu/data/landcover/data.shtml
Global Land Cover Characterizatio n database (GLCC)	USGS, UNL, EC JRC	AVHRR imageries acquired between April 1992 through March 1993	land cover	raster	1 km	http://edc2.usgs.gov/glcc/glcc.php

Product name	Producer	Source data (name and date of the resource)	Dataset content	Dataset type	Scale or spatial resolution	Access instruction
Global Land Cover 2000 DB (GLC 2000)	GVM Unit – JRC (in collaboration with over 30 research teams around the world including FAO, UNEP/GRID and CIESIN)	GLC 2000 makes use of the VEGA 2000 dataset: a dataset of 14 months of pre-processed daily global data acquired by the VEGETATION instrument on board the SPOT 4 satellite	land cover	raster	1 km	http://bioval.jrc.ec.europa.eu/products/glc2000/products.php
Global Land Cover by National Mapping Organizations (GLCNMO) - part of Global Map -	Japan GSI (project coordination); National Mapping Organizations participating in Global Mapping project (data production)	MODIS data observed in 2003	land cover	raster	30 arc-sec (about 1 km at equator)	http://www.iscgm.org/gmd/
MODIS Land Cover Type Product (MCD12Q1)	University of Boston	A full year of composites 8-day MODIS observation (from 2001 to 2010, one product every year)	land cover	raster	500 m	REVERB: http://reverb.echo.nasa.gov/reverb/ DATA POOL: https://lpdaac.usgs.gov/get_data/data_pool

Product name	Producer	Source data (name and date of the resource)	Dataset content	Dataset type	Scale or spatial resolution	Access instruction
MODIS Urban Land Cover 500m (MOD500)	University of Wisconsin and Boston University (US-NASA)	MODIS 500 m satellite data (2001-2002)	urban areas	raster	500 m	http://sage.wisc.edu/people/14chneider/research/data.html
Globcover 2009	GlobCover is an ESA initiative which began in 2005 in partnership with JRC, EEA, FAO, UNEP, GOFC-GOLD and IGBP	MERIS fine resolution mode data acquired between January 2009 and December 2009	land cover	raster	300 m	http://ionia1.esrin.esa.int/index.asp
The Dynamics of Global Urban Expansion	This research was supported by the World Bank's Urban Development Division and by the National Science Foundation	Landsat TM (circa 1990) Landsat ETM+ (circa 2000)	urban areas (only for 120 cities with more than 100000 inhabitants)	raster	30 m	http://clear.uconn.edu/projects/Urban_Growth/data.htm

Table 1 – Freely available urban maps actually distributed at global level

VMap0

The Vector Map Level 0 (VMap0) database is an up-to-date and improved version of the NIMA's (National Imagery and Mapping Agency) Digital Chart of the World that provides worldwide coverage of vector geospatial data which can be viewed at 1:1.000.000 scale. The primary source for the database is the 1:1.000.000 scale Operational Navigation Chart (ONC) series co-produced by the military mapping authorities of Australia, Canada, United Kingdom, and the United States. The database contains data organized into 10 themes (Boundaries, Data quality, Elevation, Hydrography, Industry, Physiography, Population, Transportation, Utilities and Vegetation): information about urban areas is included in the class Population. Vmap0 is the earliest global map that includes an urban class.

Urban areas are detected as built-up areas in the Population theme.

AVHHR Global Land Cover Classification

The first global land cover map compiled from remote sensing, AVHHR Global Land Cover Classification, was generated by The University of Maryland, Department of Geography, in 1998; it made use of the best remote sensing data available at that time for global land cover applications. It was generated at 1° spatial resolution (DeFries and Townshend 1994) firstly, then at 8 km resolution (Defries et al. 1998), and lastly at 1 km resolution (Hansen et al. 2000). Data were acquired from the AVHRR satellites between 1981 and 1994.

Urban areas are detected only in the 1km version and identified as "Urban and Built-up" (value: 13).

Global Land Cover Characteristics Data Base (GLCC)

Developed by the U.S. Geological Survey (USGS), the University of Nebraska-Lincoln (UNL), and the European Commission's Joint Research Centre (JRC), GLCC was developed on a continent-by-continent basis using AVHHR NDVI composites covering 1992-1993 (Loveland et al. 2000). All continental databases share the same map projections (Interrupted Goode Homolosine and Lambert Azimuthal Equal Area), have 1 km nominal spatial resolution, and are based on 1 km AVHRR data spanning April 1992 through March 1993. It was released in 1997.

The database is provided with different classification legends (Global Ecosystems, IGBP, Biosphere Atmosphere Transfer Scheme, Simple Biosphere Model Scheme, Simple Biosphere 2 Model, USGS Land Use/Land Cover Scheme); urban areas are identified as "Urban" (value: 1) in Global Ecosystems, "Urban and Built-up" (value: 13) in IGBP, "Urban and Built-up land" (Code: 100) in USGS Land Use/Land Cover Scheme.

Global Land Cover Database 2000 (GLC 2000)

Produced by the GVM (Global Vegetation Monitoring) Unit of the JRC in collaboration with over 30 research teams from around the world with the aim to provide for the year 2000 a harmonized land cover database over the whole globe, GLC2000 (Bartholomé and Belward 2005) adopts a so called bottom-up approach: 19 Regional Land Cover datasets have been produced by regional experts and then harmonized into a full resolution global product. In order to guarantee consistency all partners used the global Land Cover Classification System (LCCS) provided by FAO. LCCS describes land cover according to a hierarchical series of classifiers and attributes. Coding each class with LCCS allows the map producer to create a regional legend, composed of individual classifiers, which hierarchically map into the more general global legend. The global scale GLC2000 legend documents 22 land cover types whereas the more detailed regional legends vary between 5 and 44 classes.

Urban areas are detected as "Artificial surfaces and associated areas" (code 22) and provided with a resolution of 1km. The last release of the global version is dated January 2004, while regional versions can be more updated.

Global Land Cover by National Mapping Organizations (GLCNMO), part of Global Map

Proposed by the Ministry of Construction of Japan in 1992, coordinated by Geographical Survey Institute of Japan (GSI) and produced by the National Mapping Organizations (NMO) participating in the Global Mapping project, Global Map is a 1:1.000.000 scale framework dataset of the world.
It consists of vector and raster layers of transport, administrative boundaries, drainage, elevation, vegetation, land use and land cover data.
In Global Map, land cover layer is provided by the Global Land Cover by National Mapping Organizations (GLCNMO) dataset, a data of 1km grid with 20 land cover items. Data were created by using MODIS data observed in 2003.

The classification is based on LCCS developed by FAO. Therefore, it is possible to compare and integrate GMLNMO and other land cover data products based on LCCS. Urban areas are identified with value: 18.

MODIS Land Cover Type Product (MCD12Q1)

With the launch, on December 1999, of Terra satellite that carried on the MODerate resolution Imaging Spectroradiometer (MODIS) sensor it was possible to have an improved basis for monitoring and mapping global land cover. MODIS land cover products were thereby generated firstly at 1 km resolution (Friedl et al. 2002), and later improved; the last version (Version 5), developed from the MODIS Land Cover group, Department of Geography and Environment at Boston University, has an increased spatial resolution of 500 m. A lower spatial resolution climate modeling grid (MCD12C1) is produced at 0.05° resolution for user who do not require the spatial

detail afforded by the 500-m land cover product. MCD12Q1 is provided at annual time steps for 2001 – 2007 in Version 5, and for 2001 – 2010 in Version 5.1 (Friedl et al. 2010).

It contains multiple classification schemes : IGBP (Loveland and Belward, 1997), UMD from the University of Maryland (Hansen et al., 2000), MODIS LAI/FPAR scheme (Myneni et al., 1997; Lotsch et al, 2001), the MODIS-derived Net Primary Production (NPP) scheme proposed by Running et al. (1995) and the Plant Functional Type (PFT) scheme described by Bonan et al. (2002).

MODIS Urban Land Cover 500 m (MOD500)

The MODIS 500-m global map of urban extent (Schneider et al. 2009; Schneider et al. 2010) was produced by Annemarie Schneider at the University of Wisconsin-Madison, in partnership with Mark Friedl at Boston Univeristy and the MODIS Land Group. The goal of this project was to generate a current, consistent, and seamless circa 2001-2002 map of urban, built-up and settled areas for the Earth's land surface. This work builds on previous mapping efforts using Moderate Resolution Imaging Spectroradiometer (MODIS) data at 1-km spatial resolution (Schneider et al. 2003), which was included as part of the MODIS Collection 4 Global Land Cover Product (Friedl et al. 2002; Friedl et al. 2010).

The classes are defined according to the IGBP program; urban areas are detected as "Urban areas" with code 13.

Globcover 2009

Began in 2005 as an ESA initiative in partnership with JRC, EEA, FAO, UNEP, GOFC-GOLD and IGBP, the GlobCover project (Arino et al. 2007) has developed a service capable of delivering global composites and land cover maps using as input observations from the 300 m MERIS sensor on board the ENVISAT satellite. Delivered for the first time in 2008 (Globcover 2005) as the very first 300 m global land cover map, it has been updated in 2010 with the time series of MERIS FR 2009 mosaics (Arino et al. 2009).

The land cover map counts 22 land cover classes defined with the United Nations LCCS. Artificial surfaces and associated areas (urban areas >50%) are coded with the value: 190.

The Dynamics of Global Urban Expansion

The Dynamics of Global Urban Expansion is a study, commissioned by the World Bank, with the aim of examining how cities expand: developing country cities should be making realistic yet minimal plans for urban expansion, designating adequate areas for accommodating the projected expansion, investing wisely in basic trunk infrastructure to serve this expansion, and protecting sensitive land from incursion by new urban development. In the framework of this study, an universe of 3943 cities with population in excess of 100.000 was defined and stratified global

sample of 120 cities from this universe (Angel et al. 2005) was drawn. The 120 cities were identified from remote sensing data acquired in two different periods: circa 1990 and circa 2000. Data were acquired from Landsat TM in the first case and Landsat ETM+ in the second case.

Also if this work doesn't provide a consistent layer of built-up areas, it is equally mentioned since it can be useful for some applications and also because it deals with imageries at higher resolution than the previous mentioned works.

Global Human Settlement Layer (GHSL)

Performed by the Joint Research System (JRC) of the European Commission and released for the first time in November 2012, the Global Human Settlement Layer is the first ever compiled dataset of human settlement containing globally-consistent physical measurements describing human settlements at different scales, generated systematically and automatically. It's provided with three nominal reference scales: 1:10.000 (local), 1:50.000 (regional) and 1:500.000 (global). It makes use of multi-resolution, multi-platform and multi-sensor image data: 75m ENVISAT-ASAR data, 30m Landsat TM, 2.5 m SPOT till to 0.5-0.6 GeoEye and Worldview data. The first version of the public GHSL platform is only experimental and is delivering a small subset of the information layers currently under test in the internal GHSL platform prototype. The public GHSL can be accessed using a web interface (http://ghslsys.jrc.ec.europa.eu/concept.php) but it can be requested access to the WMS service. Future versions will allow a 'process on demand', where interested users may process their satellite or airborne data to derive the GHSL information layers of their data.

Global Urban Footprint (GUF)

The German Aerospace Center (DLR) is developing the Urban Footprint Processor (UFP) in order to generate a global binary map of human settlements of the year 2011 and 2012 - the Global Urban Footprint (GUF). It will exploit the TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurements) mission that acquired two complete coverages of the earth's surface of the years 2011 and 2012 (Felbier et al. 2012).

The final product will be made available as a public domain product and will therefore have a resolution of 3 arc second, which corresponds to a resolution of about 75 m.

1.2.2   Urban maps available at continental level: the case of Europe

The outcomes of the performed inventory state that the best resolution available for dataset of urban areas currently completed and released at global level is 300 m (Globcover, Arino et al. 2007); layers of urban areas at higher resolution are provided with smaller geographic extent.

In this section an overview about dataset available at European level is provided; information is summarized in Table 2. Focusing on smaller extent, it's possible to find datasets with improved spatial resolution.

Between selected European urban maps, three are general multi-class land cover maps that include an urban class (GlobCorine, CLC, and Urban Atlas), while one, the High Resolution Soil Sealing Layer, characterize urban land as a continuous variable: the fraction of impervious surface. One of the three multi-class land cover maps, Urban Atlas, provides information only on cities with more than 100.000 inhabitants.

GlobCorine

The GlobCorine project builds on the GlobCover findings and aims to make the full use of the MERIS time series for frequent land cover monitoring at the pan-European scale. This dataset will not be as precisely as the Corine Land Cover (CLC), described later, but will significantly shorten the time response and expand the coverage. The main source is MERIS 300m Full Resolution Full Swath (FRS) composites as produced and delivered by the GlobCover processing chain (acquired between January and December 2009).
The legend associated with the GlobCorine 2009, identical to the GlobCorine 2005 one, was defined to be as compatible as possible with the CLC aggregated typology and with the LCCS system. "Urban and associated areas" areas are identified with code 10.

Corine Land Cover and High Resolution Soil Sealing layer

Born at European level for "for gathering, coordinating and ensuring the consistency of information on the state of the environment and natural resources in the Community" (Official Journal L 176, 6.7.1985), the first realization of the project date back at 1990 (CLC90); it has been further updated in 2000 through the project "Image & Corine Land Cover 2000" and again in 2006 within the initiative "Fast Track Service on Land Monitoring" (FTSP) of Global Monitoring for Environment and Security (GMES) Program.

Among services that have been realized through CLC2006, there are:

- Corine land cover mapping 2006;

- Corine land cover change mapping 2000–2006;

- Built-up areas including degree of soil sealing, 2006

Imageries provided by SPOT-4 and IRS P6 satellites acquired in 2006 +/- 1 year have been employed for the production of this dataset; they have, respectively, pixel size of 20 (multi-spectral) and 10 meters (panchromatic) in case of SPOT and pixel size of 23 meters in case of IRS P6 satellite. They have been resampled to 20 m and the output provided in vector format with a nominal scale of 1:100.000. Output is also provided in raster format resampled to 100 m or 250 m.

The standard CLC nomenclature includes 44 land cover classes. These are grouped in a three-level hierarchy, with 44 classes at third level, 15 at second and 5 at first level. The five main (level-one) categories are: 1) artificial surfaces, 2) agricultural areas, 3) forests and semi-natural areas, 4) wetlands, and 5) water bodies.

Information about urban areas is therefore available within this layer with increasing level of detail.

In the framework of the same program it has been commissioned the production of a *High resolution soil sealing layer*; obtained from the same imageries acquired for CLC2006, it provides information about the sealing degree of built-up areas (ranging from 0 to 100%).

Urban Atlas

The European Urban Atlas is part of the local component of the GMES land monitoring services, one of the services planned in the framework of GMES to manage and protect the environment and natural resource, and ensure civil security. It provides high-resolution land use maps for 305 Large Urban Zones and their surroundings (more than 100.000 inhabitants as defined by the Urban Audit) for the reference year 2006 at scale 1:10.000.
Urban is detected and subdivided in 5 levels according the density of buildings, other classes have been created for certain categories, such as industrial, commercial or sports and leisure facilities.

Product name	Producer	Source data (name and date of the resource)	Dataset content	Dataset type	Scale or spatial resolution	Access instruction
GlobCorine	Launched by ESA in a joint initiative with EEA and implemented by Université Catholique de Louvain.	MERIS 300m Full Resolution Full Swath (FRS) composites as produced and delivered by the GlobCover processing chain (acquired between January 2009 and December 2009)	land cover	raster	300 m	http://ionia1.esrin.esa.int/globcorine/
Corine Land Cover (CLC)	The project is coordinated, at European level, by the European Commission and the European Environment Agency	For the last version (2006): SPOT-4 HRVIR at 10/20 m resolution and/or IRS P6 LISS III at 23 m resolution acquired in 2006 +/- 1 year	land cover	vector or raster	1:100K (100 m or 250 m for rater format)	Data are distributed from EEA at: http://www.eea.europa.eu/data-and-maps/data in vector or raster format (at 100 m or 250 m resolution) in country files or as a unique layer

Product name	Producer	Source data (name and date of the resource)	Dataset content	Dataset type	Scale or spatial resolution	Access instruction
High resolution Soil Sealing Layer	Commissioned in the framework of GMES Fast Track Service on Land Monitoring Programme from EEA, realized from an International Team composed by Infoterra GmbH (Germany), Planetek Italia, Metria (Sweden), Geoville (Austria), GISAT (Czech republic) e Tragsatec (Spain) and validated from single Member States	SPOT-4 HRVIR at 10/20 m resolution and/or IRS P6 LISS III at 23 m resolution acquired in 2006 +/- 1 year	degree of imperviousness	raster	20 m - 100 m	Raster version at 100 m resolution is distributed from EEA at: http://www.eea.europa.eu/data-and-maps/data Version at 20 m resolution is visible at: http://www.eea.europa.eu/data-and-maps/explore-interactive-maps/european-soil-sealing-v2
Urban Atlas	It is a joint initiative of the EC DG for Regional Policy and the DG for Enterprise and Industry with the support of ESA and the EEA. The Urban Atlas was executed by the French company Systèmes d'Information à Référence Spatiale (SIRS).	Earth Observation (EO) data acquired between 2005 and 2007 with 2.5 m spatial resolution multispectral or pan-sharpened (multispectral merged with panchromatic) data. Multispectral data includes near-infrared band.	land use/ land cover (only for cities with more than 100000 inhabitants)	vector	1:10K	http://www.eea.europa.eu/data-and-maps/data/urban-atlas

Table 2 – Freely available urban maps distributed at European level

1.3   Image classification methods and techniques

Basically, classification can be defined as the ordering of entities into groups or classes on the basis of their similarity. This means that a set of entities is arranged into groups, so that each group is as different as possible from all other group, but internally as homogeneous as possible (Bailey 1994). In remote sensing, the classification process brings to the creation of a thematic map, where each pixel is associated with semantic information related to a specific theme, class or category.

Classifying remotely sensed data into a thematic map remains a challenge because of many factors, such as the complexity of the landscape in a study area, the selection of remotely sensed data and image-processing and classification approaches, may affect the success of a classification (Lu and Weng 2007).

The major steps that may be involved in image classification are described in section 1.3.1; an overview about major classification methods is provided in section 1.3.2, while section 1.3.3 describes the classification approaches used by the urban maps described in sections 1.2.1 and 1.2.2.

1.3.1   Remote sensing classification process

Remote-sensing classification process starts from the analysis of the observed ground in order to produce a thematic map. This process can be performed from the interpreter through visual interpretation or by means of automatic classification that reduce the subjectivity of the interpreter.
In this case, the interpretation process may occur in different ways, but the major steps are the following (Lu and Weng 2007):
- Determination of a suitable classification system;
- Selection of training samples;
- Image pre-processing;
- Feature extraction;
- Selection of suitable classification approaches;
- Post-classification processing;
- Accuracy assessment.

A suitable *classification system* and a sufficient number of training samples are prerequisites for a successful classification. A classification system determine the system used in order to classify things; when planning a project involving remotely sensed data, it is very important that sufficient effort be given to the classification scheme to be used. In many instances, this scheme is an existing one; in other cases, the classification scheme is dictated by the objectives of the project. A few simple guidelines should be followed. First of all, any classification scheme should be mutually exclusive and totally exhaustive. In other words, any area to be classified should fall into one and only one category or class. In addition, every area should be included in the classification.

Finally, if possible, it is very advantageous to use a classification scheme that is hierarchical in nature. If such a scheme is used, certain categories within the classification scheme can be collapsed to form more general categories (Congalton 1991).

A sufficient number of training samples and their representativeness are critical for image classifications (Hubert-Moy et al. 2001; Chen and Stow 2002; Landgrebe 2003; Mather 2004). Training samples are usually collected from fieldwork, or from fine spatial resolution aerial photographs and satellite images. When the landscape of a study area is complex and heterogeneous, selecting sufficient training samples becomes difficult. This problem would be complicated if medium or coarse spatial resolution data are used for classification, because a large volume of mixed pixels may occur. Therefore, selection of training samples must consider the spatial resolution of the remote-sensing data being used, availability of ground reference data, and the complexity of landscapes in the study area.

Image *pre-processing* may include the detection and restoration of bad lines, geometric rectification or image registration, radiometric calibration and atmospheric correction, and topographic correction. Accurate geometric rectification or image registration of remotely sensed data is a prerequisite for a combination of different source data in a classification process, while atmospheric correction is mandatory when multitemporal or multisensor data are used.

Selecting suitable variables is a critical step for successfully implementing an image classification. Many potential variables may be used in image classification, including spectral signatures, vegetation indices, transformed images, textural or contextual information, multitemporal images, multisensor images, and ancillary data, but the use of too many variables in a classification procedure may decrease classification accuracy (Hughes 1968; Price et al. 2002), and is therefore important to select only the variables that are most useful for separating land-cover or vegetation classes, especially when hyperspectral or multisource data are employed.

The *classification method* should be selected taking into account many factors, such as spatial resolution of the remotely sensed data, different sources of data, a classification system and availability of classification software. Different classification results may be obtained depending on the classifier chosen; a detail summarization of major classification methods is provided in section 1.3.2.

Post-classification processing is an important step in improving the quality of classifications (Harris & Ventura 1995; Murai & Omatu 1997; Stefanov et al. 2001; Lu & Weng 2004); usually, a majority filter is applied to reduce the noise that come out from traditional per-pixel classifier ("salt and pepper" effect) or ancillary data are used to modify the classification image based on established expert rules (For example, data describing terrain characteristics can be used to modify classification results of forest distribution, or population density data to correct some classification confusions between commercial and high-intensity residential areas in classification of urban areas).

Evaluation of classification results is an important process in the classification procedure. Different approaches may be employed, ranging from a qualitative evaluation based on expert knowledge to a quantitative accuracy assessment based on sampling strategies. Three basic components are included in a classification accuracy assessment: sampling design, response design and analysis procedures.

The error matrix is one of the most widely used approach in accuracy assessment; after generation of an error matrix, other important accuracy assessment elements, such as overall accuracy, omission error, commission error, and kappa coefficient, can be derived. All these components are widely descript in section 3.2.8.

1.3.2   Image classification approaches

Classification methods can be grouped differently according to the considered criteria; this section aims at providing a first overview of image classification methods subdivided according to the selected criteria. For more information about each method, reference bibliography should be consulted (Brivio et al. 2006; Gomarasca 2009). As mentioned in Lu & Weng 2007, classifiers can be grouped as:

- supervised or unsupervised;
- parametric or non parametric;
- hard and soft (fuzzy);
- per-pixel or sub pixel;
- objected-oriented or per- field;
- spectral, contextual or both.

Supervised/Unsupervised

Classification can be supervised or unsupervised whether training samples are used or not. In a supervised approach classes are previously defined and a certain number of pixels is selected as representative for each class and then used to train the classifier in order to classify the spectral data into a thematic map: the classification method depend on the knowledge of the ground truth. Maximum Likelihood, Minimum Distance, Decision Tree (Hansen et al. 1996; Friedl & C. Brodley 1997; Defries et al. 1998; Friedl et al. 1999; Defries & Chan 2000; Pal & Mather 2003; Lawrence et al. 2004) are example of supervised classifiers.

Unsupervised classification approach makes use of clustering-based algorithms that are used to partition the spectral image into a number of spectral classes based on the statistical information inherent in the image. Spectral classes are then merged and labelled by the analyst. ISODATA and K-means clustering algorithms are unsupervised classifiers.

Parametric/Non parametric

Classifiers can be either parametric or non-parametric whether parameters (such as mean, standard deviation, etc.) are used or not. Parametric classifiers are based on distribution model (e.g. Gaussian) defined by geometric-statistical parameters that are generated from training samples. Maximum Likelihood, Minimum Distance, Parallelepiped are example of parametric

classifiers. In non-parametric classifier no assumption about data is required. They don't employ statistical parameters to calculate class separation. Artificial neural network, Decision Tree classifier and Support Vector machine are good example.

Per pixel/Sub pixel

Classifiers are divided in per-pixel and sub-pixel according to which kind of pixel information is used. Traditional per-pixel classifiers typically develop a signature by combining the spectra of all training-set pixels for a given feature. The resulting signature contains the contributions of all materials present in the training pixels, but ignores the impact of the mixed pixels. Most of the classifiers own at this category. To overcome this problem, sub pixel classification approaches have been developed to provide a more appropriate representation and accurate area estimation of land covers than per-pixel approaches, especially when coarse spatial resolution data are used (Foody & Cox 1994; Binaghi et al. 1999; Ricotta & Avena 1999; Woodcock & Gopal 2000); the spectral value of each pixel is assumed to be a linear or non-linear combination of defined pure materials (or endmembers), providing proportional membership of each pixel to each endmember. The fuzzy-set technique (Foody 1996; Maselli et al. 1996; Mannan et al. 1998; Zhang & Kirby 1999; Shalan et al. 2003) and spectral mixture analysis (SMA) classification (Adams et al. 1995; Roberts et al. 1998; Rashed et al. 2001; Lu et al. 2003) are the most popular approaches used to overcome the mixed pixel problem.

Hard/soft

Classifier can be either hard or soft (fuzzy) whether the output is a definitive decision about land cover class or not. Hard classifier are based on the hypothesis that all pixel in the image have pure spectral characteristics, that every pixel belong to just one class, and that all land cover classes have constant spectral signatures (Brivio et al. 2006). Soft classifiers, instead, provide for each pixel a measure of the degree if similarity for every class. They provide more information and, potentially, a more accurate result especially for coarse spatial resolution data. Classifiers that provide this information are Fuzzy set classifiers and Spectral Mixture Analysis.

Object oriented/Per-field

Classifier can be either Object-oriented or Per-field according to which kind of pixel information is used. In object-oriented the classification is conducted based on object, instead of an individual pixel, and object are obtained from processing such as Image Segmentation. No GIS vector data are used. In Per-field classifiers, vector or raster data are integrated in the classification; for example, vector data can be used to subdivide an image into parcels, and classification is based on parcels.

Spectral/Contextual

Depending on whether spatial information is used or not, classifier can be defined as spectral, contextual, or spectral-contextual; in spectral classifiers pure spectral information is used, while contextual classification exploits spatial information among neighbouring pixels to improve classification results (Flygare 1997; Stuckens et al. 2000; Hubert-Moy et al. 2001; Magnussen et al. 2004). A contextual classifier may use smoothing techniques, Markov random fields, spatial statistics, fuzzy logic, segmentation, or neural networks (Binaghi et al. 1997; Cortijo and De La Blanca 1998; Kartikeyan et al. 1998; Keuchel et al. 2003; Magnussen et al. 2004). In general, pre-smoothing classifiers incorporate contextual information as additional bands, and a classification is then conducted using normal spectral classifiers, while post-smoothing classification is conducted on classified images previously developed using spectral-based classifiers. The Markov random field-based contextual classifiers, such as iterated conditional modes, are the most frequently used approaches in contextual classification (Cortijo and De La Blanca 1998; Magnussen et al. 2004), and have proven to be effective in improving classification results.

In order to improve classification accuracy, different characteristics of remote sensing data have been taken into consideration and employed in image classification. In particular, use of spatial information, integration of different sensor data, use of multi-temporal data, data transformation techniques and use of GIS have been taken into consideration and will be described.

Use of spatial information

Spatial resolution determines the level of spatial detail that can be observed on the Earth's surface. Fine spatial resolution data greatly reduce the mixed-pixel problem, providing a greater potential to extract much more detailed information on land cover structures than medium or coarse spatial resolution data. However, some new problems associated with fine spatial resolution image data emerge: the high spectral variation within the same land-cover class is one of the major disadvantages. In order to make full use of the rich spatial information inherent in fine spatial resolution data, it is necessary to minimize the negative impact of high intraspectral variation. Exploiting the spatial information, information important for the understanding of an image is not represented in single pixels but in meaningful image objects and their mutual relations (Blaschke et al. 2000). Spatial information may be used in different ways, such as in contextual-based or object-oriented classification approaches, or classifications with textures. Many texture feature extraction methods exist. Tuceryan and Jain (1993) identify four major categories of texture feature analysis methods: 1) statistical (such as those based on the computation of the Gray-Level Co-Occurrence matrix, GLCM), 2) geometrical (including structural), 3) model-based, such as Markov random fields (MRF), and 4) signal processing (such as Gabor filters). GLCM is one of the most popular methods for extracting textural feature from images

Integration of different sensor data

Images from different sensors contain distinctive features. Data fusion or integration of multi-sensor or multi-resolution data takes advantage of the strengths of distinct image data for improvement of visual interpretation and quantitative analysis. Data fusion involves two major procedures: (1) geometrical co-registration of two datasets and (2) mixture of spectral and spatial information contents to generate a new dataset that contains the enhanced information from both datasets. Accurate registration between the two datasets is extremely important for precisely extracting information contents from both datasets, especially for line features, such as roads and rivers. Radiometric and atmospheric calibrations are also needed before multi-sensor data are merged.

Fused images may provide increased interpretation capabilities and more reliable results since data with different characteristics are combined. The images vary in spectral, spatial and temporal resolution and therefore give a more complete view of the observed objects.

Many methods have been developed to integrate spectral and spatial information in previous literature. In general, the techniques can be grouped into two classes: (1) Colour related techniques, and (2) Statistical/numerical methods (Pohl and Van Genderen 1998).
The first comprises the colour composition of three image channels in the RGB colour space as well as more sophisticated colour transformations, e.g., *IHS* and *HSV*.
Statistical approaches are developed on the basis of channel statistics including correlation and filters. Techniques like *Principal Component Analysis* (PCA) and *Regression* belong to this group.
The numerical methods follow arithmetic operations such as image differencing and ratios but also adding of a channel to other image bands. A sophisticated numerical approach uses *Wavelets-merging techniques*.

Use of multi-temporal data

The use of different seasons of remotely sensed data has proven useful for improving classification accuracy; multitemporal data can be useful in classification in order to correct cloud cover, to enhance different behaviour of objects according to the season, to assess change detection and for monitoring issues.

Data transformation techniques

A large number of spectral bands provide the potential to derive detailed information on the nature and properties of different surface materials on the ground, but the bands also create difficulty in image processing and high data redundancy due to high correlation in the adjacent bands. An increase in spectral bands may improve classification accuracy, but only when those bands are useful in discriminating the classes (Thenkabail et al. 2004). Several techniques have been developed to transform the data from highly correlated bands into a dataset. Vegetation indices, principal component analysis, tasselled cap, and minimum noise fraction, are among the most commonly used ones (Oetter et al. 2000; Wu and Linders 2000). Wavelet transform and

spectral mixture analysis have also been used in recent years (Roberts et al. 1998; Rashed et al. 2001; Lu and Weng 2004).

Use of ancillary data

Ancillary data, such as topography, soil, road, and census data, may be combined with remotely sensed data to improve classification performance. For example, topographic data are valuable for improving land-cover classification accuracy, especially in mountainous regions (Janssen et al. 1990, Meyer et al. 1993, Franklin et al. 1994), and DEM-derived variables may be used in the image-pre-processing stage for topographic correction or normalization so the impact of terrain on land-cover reflectance can be removed (Teillet et al. 1982; S. 1996; R. 1997; Gu and Gillespie 1998; Dymond and Shepherd 1999; Tokola et al. 2001).

Moreover, data related to human systems such as population distribution and road density are frequently incorporated in urban classifications (V. 1998; Epstein et al. 2002; Zhang et al. 2002).

GIS plays a critical role in handling multisource data, because it helps to manage multisource data, it's able to convert different data formats into a uniform format and evaluating the data quality and it further develop suitable models for classification.

1.3.3 Classification approaches used in global and continental urban maps

Maps of land cover are usually produced with one of these three methods: clustering, maximum likelihood classification and decision tree classifiers. Maps of urban areas at higher spatial resolution (layer of land cover at European level) still exploit, sometime in consistent way, visual interpretation techniques (Corine Land Cover, High Resolution Soil Sealing Layer), while almost all makes use of different methodology or ancillary data derived from other sources for the extraction of urban areas: AVHHR and GLCC rasterized urban polygons from DCW, GLCNMO used population data and other satellite data, MCD12Q1, GlobCover and GlobCorine used a specific supervised approach for urban areas.

Methods used for the production of above-mentioned urban datasets were analyzed and described below, focusing on the ones derived from remote sensing; Table 3 and Table 4 summarize collected information (only for datasets currently completed).

Product name	Methods	Supervisioned	Unsupervisioned	Parametric	Non parametric	Per - pixel	sub - pixel	Object-oriented	Per - field	Hard	Soft	Spectral	Contextual	Spatial information	Different sensor data	Multitemporal data	Data transformation techniques	Use of ancillary data	Other
AVHHR Global Land Cover Classification	Maximum Likelihood Classification (1° version)	o		o		o				o		o		no	no	o	o	no	
	Decision tree (8 km version)	o			o	o				o		o		no	no	o	o	no	
	Decision tree (1 km version)	o			o	o			-	o		o		no	no	o	o	o	
Global Land Cover Characterizatio n database (GLCC)	Unsupervised clustering		o			o			-			o		no	no	o	o	o	
Global Land Cover 2000 DB (GLC 2000)	Different according to region; mainly unsupervised clustering	-	-	-	-	-				-		-		-					

Product name	Methods	Supervisioned	Unsupervisioned	Parametric	Non parametric	Per - pixel	sub - pixel	Object-oriented	Per - field	Hard	Soft	Spectral	Contextual	Spatial information	Different sensor data	Multitemporal data	Data transformation techniques	Use of ancillary data	Other
Global Land Cover by National Mapping Organizations (GLCNMO) - part of Global Map -	Maximum Likelihood Classification	o		o		o			-	o		o		no	o	o	o	o	
MODIS Land Cover Type product (MCD12Q1)	Decision tree and boosting	o			o	o			-	o		o		No	no	o	o	o	
MODIS Urban Land Cover 500m (MOD500)	Decision tree and boosting	o			o	o				o					no	o	o	o	filter
Globcover 2009	Supervised for land cover that are not well represented (urban, wetland) and unsupervised for other	o	o	-		o			-	o		o		No	no	o	o	o	

Product name	Methods	Supervisioned	Unsupervisioned	Parametric	Non parametric	Per - pixel	sub - pixel	Object-oriented	Per - field	Hard	Soft	Spectral	Contextual	Spatial information	Different sensor data	Multitemporal data	Data transformation techniques	Use of ancillary data	Other
The Dynamics of global urban expansion	ISODATA in 2 steps		o			o				o				No	no	no	no	o	

Table 3 – A survey of classification methods used in order to produce global urban maps derived from remote sensing described in Table 1

Product name	Methods	Supervisioned	Unsupervisioned	Parametric	Non parametric	Per - pixel	sub - pixel	Object-oriented	Per - field	Hard	Soft	Spectral	Contextual	Spatial information	Different sensor data	Multitemporal data	Data transformation techniques	Use of ancillary data	Other
GlobCorine	Supervised and unsupervised	o	o			o						o		no	no	o	o	o	
Corine Land Cover (CLC)	Photo interpretation	o				o				o			-	no	no	no	no	no	
High resolution Soil Sealing Layer	Maximum Likelihood classification and visual interpretation	o		o		o				o		o		no	o	o	o	no	
Urban Atlas	Segmentation and clustering		o			o		o		o		o		no	o	no	no	o	

Table 4 – A survey of classification methods used in order to produce urban maps available at European level described in Table 2

AVHHR Global Land Cover Classification was produced with different methodologies according to the input used dataset; the version at 1° spatial resolution was produced using maximum likelihood classification of monthly composited AVHRR normalized difference vegetation index (NDVI) data (DeFries and Townshend 1994), while the 8km spatial resolution version was made using a decision tree classification technique (Defries et al. 1998). The 1 km version was obtained, more recently, with a supervised classification strategy using decision trees (Hansen et al. 2000). A set of 41 metrics, derived from bands 1-5 and NDVI was created for input into the decision tree, and pruning based on visual interpretation. The urban and built-up class wasn't, however, derived from the AVHHR imageries but developed from the populated places layer that is part of the Digital Chart of the World (Danko 1992).

The methods used for *Global Land Cover Characteristics Data Base* can be described as multitemporal unsupervised classification of NDVI data with post-classification refinement using multi-source earth science data (http://edc2.usgs.gov/glcc/globdoc2_0.php). Data are initially processed by recomposing the 10 day composites into monthly data sets. Masks representing water bodies, snow and ice, and barren or sparsely vegetated areas are then developed to eliminate NDVI data from the composites for those areas where the meaning of the NDVI values is ambiguous. An unsupervised clustering is then applied in order to segment the 12-month NDVI composites into seasonal greenness classes and a preliminary labeling performed in order to provide a general understanding of the characteristics of each cluster and to determine which classes have two or more disparate land cover classes represented within their spatial distribution. A post-classification stratification is used at this step in order to separate classes containing two or more disparate land cover types and initial criteria therefore refined and finally used to permanently modify the original class.
As AVHHR Global Land Cover Classification, also GLCC rasterized the urban polygons from the Digital Chart of the World.

As mentioned before, *GLC2000* was created with a bottom up approach, by harmonizing 19 regional land cover mapping performed by regional experts. Each partner had to use data based on SPOT-4 VEGETATION VEGA2000 dataset, with the opportunity to choose the preferred method of mapping and the regional legend. A less thematically detailed global legend was also created to harmonize regional legends into one consistent product. It derives a large amount of techniques that are summarized in Annex I: most of them exploit unsupervised classification techniques (clustering) with an a-posteriori labeling using available regional dataset as reference (Fritz et al. 2003).

In *GLCNMO* the source data were processed to remove cloud contamination, reprojected into latitude/longitude coordinate system, and mosaicked to make five continental data (Eurasia, North America, South America, Africa, and Oceania) and two island regions in the Ocean. Training data for supervised classification were collected using Landsat image, MODIS NDVI seasonal change pattern, and comments from National Mapping Organizations with the reference

of Google Earth, Virtual Earth and existing regional maps; a maximum likelihood method was applied.

Six land cover classes, which are Water, Urban, Tree open, permanent snow/ice, Mangrove, and Wetland, were extracted independently. Particularly, urban areas were extracted using population density data (Landscan dataset, http://www.ornl.gov/sci/landscan/index.html), MODIS/NDVI data to exclude large green area such as a park in populated area and DMSP/OLS data (http://www.ngdc.noaa.gov/dmsp/maps.html) to exclude villages in developing countries with large population (Tateishi et al. 2003).

The *MODIS Land Cover Type Product (MCD12Q1)* is produced using a supervised classification algorithm that is estimated using a database of high quality land cover training sites (Friedl et al. 2010). The training site database was developed using high resolution imagery in conjunction with ancillary data. The base algorithm is a decision tree (C4.5; Quinlan 1993) in conjunction with a technique for improving classification accuracies known as boosting. Boosting improves classification accuracies by iteratively estimating a decision tree while systematically varying the training sample.

Input features used in the algorithm include spectral and temporal information from MODIS BRDF/albedo (Schaaf et al. 2002) product (that provides surface reflectance measurements normalized to a consistent nadir view geometry based on BRDF- models of surface anisotropy, thereby minimizing the effect of variable view geometry in surface reflectance data) bands 1–7, supplemented by the enhanced vegetation index (EVI; Huete et al. 2002) and Collection 5 MODIS Land Surface Temperature (LST; Wan et al. 2002), which was not used in previous Collections.

Since some classes are particularly problematic and difficult to map (urban land use, wetlands, and deciduous needleleaf forests), they have been classified with different methods. Urban areas, in particular, were mapped using an ecoregion-based stratification with eighteen strata, where training data and supervised classifications were developed and tuned to each stratum.

MOD500, the first map developed at global level specifically for urban areas, exploit a classification methodology based on a supervised decision tree algorithm (C4.5), in conjunction with boosting (Schneider et al. 2009; Schneider et al. 2010).

To facilitate processing and classification of the MODIS data, a global stratification of the Earth's land surface was developed in order to define 16 quasi-homogeneous strata (urban ecoregions); regions were then processed on a case-by-case basis. The classification approach employs a one-year time series of MODIS data to exploit spectral and temporal properties in land cover types: differences in temporal signatures for urban and rural areas (that result from phenological differences between vegetation inside and outside the city) were used. MODIS 8-day NBAR values for the seven bands, aggregated to 32 day averages, and EVI index were used, together with monthly and yearly minima, maxima and means for each band in order to increase classification accuracy.

The classification algorithm was run twice: it first utilizes the full set of land cover exemplars that includes urban sites and then excludes the urban training sites. The first classification classifies both the urban core and mixed urban spaces correctly, with the caveat that some non-urban areas are erroneously labeled urban land (e.g. expanses of semi-arid shrubland). The urban class

probabilities are then extracted from the first classification and areas of confusion determined on the basis of low membership to the urban class. These pixels take advantage of the information in the second classification (without urban sites) to modify the urban class probabilities using Bayes rule. This step relies on the urban ecoregion stratification previously described. Regional post-processing includes application of the MODIS 500-m water mask, use of a spatial filter to remove single, stand-alone urban pixels, and hand editing.

GlobCover 2009 has been developed considering two major modules: a pre-processing module, leading to global mosaics of land surface reflectance at 300 m, and a classification module leading to the final land cover map at 300 m resolution (Bontemps et al. 2011). The pre-processing includes the following steps: geometrics correction of the input data to achieve at least 150 m geo-location accuracy, atmospheric correction, including aerosol correction, cloud screening and shadow detection, land/water reclassification and a correction of the smile effect, Plate-carrée projection, Bidirectional Reflectance Distribution Function (BRDF) effect reduction and temporal compositing. The classification module starts with the stratification of the world in equal-reasoning areas, in order to reduce the land cover variability in the dataset and to improve the discrimination efficiency of the classification algorithm. The classification process runs independently for each delineated equal-reasoning area. The spectral classification consists of a supervised and an unsupervised classification. The supervised classification aims at identifying land cover classes that are not well represented, i.e. urban and wetland areas. The pixels classified through this process are masked and an unsupervised classification is then applied on the remaining pixels to create clusters of spectrally similar pixels. Clusters produced by the unsupervised classification are then temporally characterized and then grouped together into a manageable number of spectro-temporal classes according to their similarity in the temporal space. The labeling procedure is automated and based on a global reference land cover database which is, in the GlobCover 2009 project, the GlobCover 2005 (V2.2) land cover map. Finally, gaps due to uneven acquisitions of MERIS data are filled out using the GlobCover 2005.

In the *Dynamics of Global Urban Expansion*, input data (Landsat imageries) were first subset on the basis of an administrative district boundary map files, the Socioeconomic Data and Applications Center48 (SEDAC), part of Columbia University's Center for International Earth Science Information Network (CIESIN); only those parts of the image that were in the subset of districts containing the metropolitan area were selected for classification.

An unsupervised classification approach was then chosen for the classification; the ISODATA clustering algorithm was used to partition the subset scenes into 50 spectrally separable classes. Each of the 50 clusters was then placed, using the Landsat data themselves and independent reference data when available, into one of seven pre-defined cover classes: water, urban, vegetation, barren (including bare soil agriculture), clouds/ shadow, snow/ice, and "undetermined" (usually pixels confused between urban and barren). Those pixels labeled as such were extracted and submitted to a second clustering in an attempt to maximize the separability among those spectrally similar classes. The clusters from this second iteration were labeled into one of the six informational classes. The classification was further refined through a careful, section-by-section examination of the Landsat imagery (Angel et al. 2005).

Global Human Settlement layer is produced by automatic image information extraction techniques, using in input several different digital images coming from different satellite platforms. A reference set, that includes a global mosaic of Landsat image data, the Open Street Map (OSM) vector, the Landsacan population data, MODIS urban areas and a so called BURef data layer containing the best estimation of presence of built-up areas at the GHSL global scale, has the function to support the optimization of the spatial and thematic consistency during the GHSL production. In the pre-processing, Landsat data and the OSM are used to control the spatial consistency of the data; in the processing, three main image-derived characteristics are used: radiometric, textural and morphological/shape criteria. The textural features are derived from the GLCM matrix, while the morphological information is a product of a multi-scale morphological analysis protocol referred to as the "mtDAP" (http://ghslsys.jrc.ec.europa.eu/tecnology.php). The derived features are subsequently classified after they underwent a learning process, where 3 different stages can be observed. In the 'adaptive learning' modality, the system optimizes the decision thresholds in the input features using a given reference layer. The 'meta-learning' modality is used to study the behavior of these decision thresholds in the set of scenes processed and to detect regularities: for example typical thresholds for a given sensor in specific regions. The output of the meta-learning is then exploited during the 'discovery' modality that can be activated in order to have the chance of recovering image information lost because of errors (incompleteness, inconsistencies) in the reference data, or different scale generalization of the image-derived information and in the available reference data. In practice, the adaptive learning optimizes consistencies between the image information under processing and the reference data, while the meta-learning and discovery modes take the risk of the image information recognition also in cases where reference data is not available with the necessary thematic, spatial precision. Prior to the final mosaicking and output data preparation all results are validated by combining a global reference data set with limited visual interpretation (Pesaresi and Halkia 2012).

The *Global Urban Footprint* is generated by the Urban Footprint Processor. Two main processing stages are performed: the Basic Product Generation and the Final Product Generation.

Within the Basic Product Generation, four processing steps are accomplished: An Amplitude calculation, a texture analysis to derive a Speckle Divergence image, a classification step to extract vertical urban structures and a multilooking step to reduce the immense amount of data. The second processing stage (Final Product Generation) is a post processing step, in which the final result – the Global Urban Footprint (GUF) – is generated. This processing stage consists of four processing steps again: a generalization is performed, which takes all image components of the first stage into account, to delineate a binary map of human settlements. Additionally, the Final Product Generation includes the geocoding of the generalized binary map and applies a slope correction based on a global DEM to eliminate false classifications in the map induced by highly textured mountainous regions. A mosaicking step merges various GUF maps of predefined spatial extent together (Felbier et al. 2012).

The methodology used for the creation of *GlobCorine 2009* is based, as the one used for GlobCorine 2005, on both supervised and unsupervised techniques. The pre-processed daily images are composited into MERIS FRS seasonal syntheses, through a Mean Compositing (MC) algorithm and then merged to produce MERIS FRS seasonal mosaics. The pan-European continent is then stratified in 5 equal-reasoning areas from an ecological and a remote sensing point of view: each delineated equal-reasoning area is then classified independently. A supervised classification is performed on the land cover classes that are not well represented in the strata (urban areas and wetlands). An unsupervised spectral classification is then performed on the pixels that were not classified in the first step. The labeling procedure is automated. Pixels classified as crop or mosaic classes after the step 2 are then further processed with an unsupervised classification on MERIS 10-day NDVI profiles created in order to have clusters of pixels similar in the temporal space; new smaller and more homogeneous spectro-temporal classes are created and a second labeling procedure performed. An external dataset, the SRTM Water Body dataset (SWBD), was finally used to improve the "water bodies" delineation in the GlobCover classification (Defourny et al. 2010).

A *Corine Land Cover* map for the year 2006 (*CLC2006*) has been produced by integrating the data of land cover changes 2000–2006 with the land cover map from the year 2000 (CLC2000). The map of changes in land cover between the years 2000 and 2006 has been based on visual image comparison in a dual-window environment. The CLC2006 database has then been generated in a, mostly, automatic way with some human interaction by combining CLC2000 and the photointerpreted CLC-Changes. Only land cover changes that are larger than 5 ha and wider than 100 m are mapped (EEA 2007).

Multi-sensor and bi-temporal, orthorectified satellite imagery (IMAGE2006) was used to derive the first soil sealing database for Europe, produced as part of the GMES Fast Track Service on Land Monitoring (Land FTS LM) in 2006-2008: the *High Resolution Soil Sealing Layer*. The database was implemented in three phases: Supervised classification of built-up areas followed by visual improvement of classification result performed by Member States and then derivation of degree of soil sealing on the basis of calibrated NDVI (Maucha et al. 2010).

Urban Atlas makes use of Earth Observation data (SPOT, ALOS, QUICK BIRD, RAPIDEYE) with 2.5 m spatial resolution, topographic maps at a scale of 1: 50.000 or larger and COTS (commercial off the shelf) navigation data for the road network as input data; pre-processing on data mainly concern the geometry: all dataset should be congruent with the EO data.
Automatic classification routines, such as segmentation and clustering, are applied whenever appropriate to achieve an initial differentiation between basic land cover classes (urban vs. forest vs. water vs. other land cover); then COTS navigation are used to generate the street and railroad network of the mapping product and finally objects are interpreted using the EO data, topographic maps, COTS navigation data and auxiliary information including local expertise (Pesaresi and Halkia 2012).

2 Data, methods and software

2.1 Methods

2.1.1 Urban classification using decision tree technology

The main aim that of this work is the automation of procedures for features extraction. The approach that was chosen to deal with is related to the extraction of rules that allow identifying the desired land cover type. Rules are based on the definition of the attributes to include inside together with the choice of decision thresholds; if, on one side, thresholds could be obtained using the knowledge provided by experts, nevertheless it's advisable to explore data mining approaches for the identification of suitable bands for classification as well as determining the decision thresholds. In fact, experts may disagree on the decision boundaries, while with the use of data mining techniques, reliable, transferable and reproducible decision thresholds can be obtained (Otukei and Blaschke 2010).

Decision Tree have been widely used for remote sensing application, particularly for land cover mapping from coarse resolution data (Hansen et al., 2000; Friedl et al., 2002). Between global urban maps previously described, AVHHR, MCD12Q1 and MOD500 makes use of Decision Tree technology (Table 3). Decision threshold are also used in GHSL.
This classification method is able to handle large, non-parametric datasets with noisy or missing data, complex, non-linear relationships between features and classes, and problems that require a many-to-one mapping approach (Fayyad and Irani 1992; Friedl and Brodley 1997b).
This classifier was chosen for this study for the following reasons: first, it's able to extract rules in an automatic way, that can be later applied to different images; second, the model built by the algorithm is easily interpreted, and can therefore been modified according to the user need; third, the choice of which attributes include in the classifier is performed by the algorithm, on the basis of the criteria used by each one. Finally, decision tree algorithms don't require additional information besides that already contained in the training data (e.g. domain knowledge or prior knowledge of distributions on the data or classes) and generally display good classification accuracy compared to other techniques.

The classifier, the classification process, the algorithm used for this study together with his main features, the parameters that can be set by the user according to his needs, the way used by the algorithm to perform a first accuracy assessment, are described in the following section.

2.1.2 Decision Tree classifier: algorithm description

Decision Tree is a tree structure, which is composed of a root node and a series of internal nodes and leaf nodes, each node with only one father node and two or more sub-nodes. Each leaf node in Decision Tree corresponds to one category attribute value, different leaf node being able to correspond to the same category attribute value (Figure 4). Decision Tree can be expressed as a group of rules with IF-THEN form in addition to the tree form (Shen et al. 2011).
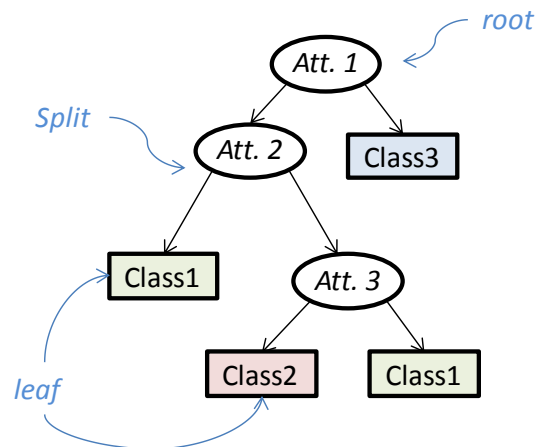
Figure 4 – Representation of a Decision Tree classifier

In classification, a set of example records is given, called the training data set, with each record consisting of several attributes. One of the categorical attributes, called the class label, indicates the class to which each record belongs. The objective of classification is to use the training data set to built a model of the class label such that it can be used to classify new data whose class labels are unknown (Figure 5).

ID	Attr 1	Attr 2	Attr 3	Attr 4	Attr 5	Attr 6	Class
1	0.12	0.25	0.32	0.44	0.76	0.02	Urban
2	0.25	0.34	0.2	0.54	0.12	0.07	Soil
3	0.34	0.15	0.32	0.65	0.81	0.4	water
…	…	…	…	…	…	…	…

TRAINING SET

ID	Attr 1	Attr 2	Attr 3	Attr 4	Attr 5	Attr 6	Class
25	0.44	0.02	0.12	0.25	0.12	0.44	?
26	0.54	0.07	0.25	0.34	0.25	0.54	?
27	0.65	0.4	0.34	0.15	0.34	0.65	?
…	…	…	…	…	…	…	…

TEST SET

Figure 5 – Creation of a model for classification

The tree classifier used to perform the analysis is J48, an open source Java implementation of C4.5 data mining algorithm developed by Quinlan (1993), one of the most popular decision tree algorithms. Also used in the MODIS Land Cover product (Schneider et al. 2009), it's an extension of Quinlan's earlier ID3 algorithm that is able to:

- Convert decision tree to equivalent production rules;
- Handling both continuous and discrete attributes;
- Handling training data with missing attribute values;
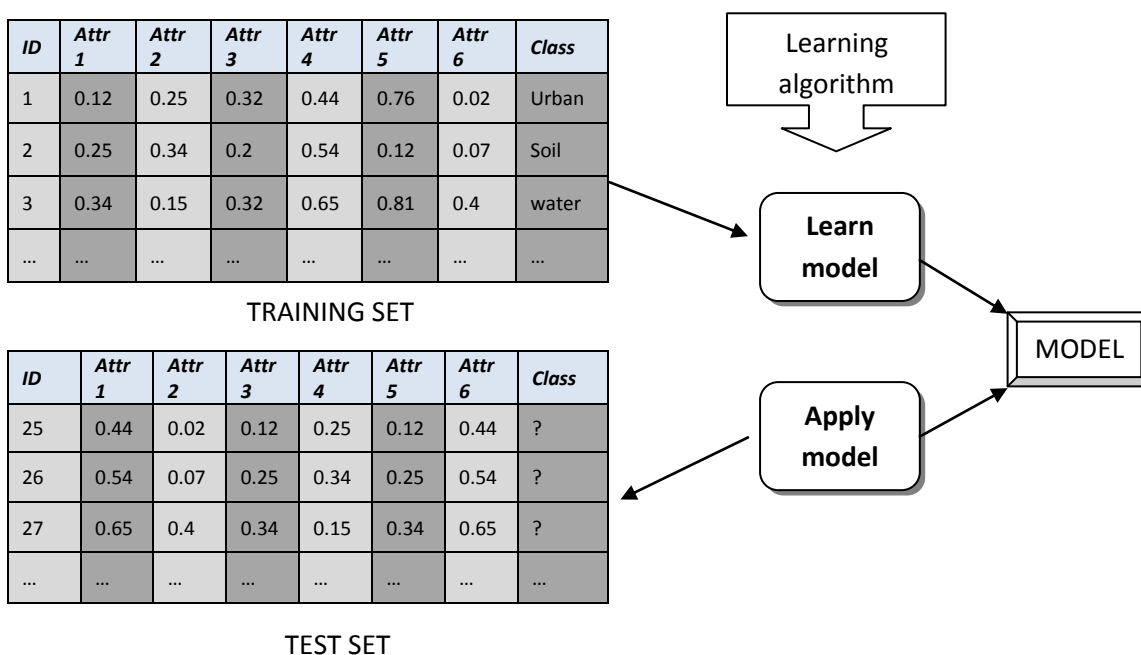- Pruning trees after creation;

C4.5 performs classification in two phases: *Tree Building* and *Tree Pruning*. In tree building, the decision tree model is built by recursively splitting the training data set based on a locally optimal criterion until all or most of the records belonging to each of the partitions bear the same class label.

The building phase constructs a perfect tree that accurately classifies every record from the training set. However, one often achieves greater accuracy in the classification of new objects by using an imperfect, smaller decision tree rather than one which perfectly classifies all known records (Quinlan and Rivest 1989). The reason is that a decision tree which is perfect for the known records may be overly sensitive to statistical irregularities and idiosyncrasies of the training set. Thus, most algorithms perform a pruning phase after the building phase in which nodes are iteratively pruned to prevent "overfitting" and to obtain a tree with higher accuracy (Rastogi and Shim 2000).

There are two main operations during tree building: the first is related to the evaluation of splits for each attribute and selection of the best split, while the second concern the creation of partitions using the best split. The determination of the best split for each attribute is performed through the application of the splitting criterion to the data. Several splitting schemes have been proposed in the past (Rastogi and Shim 2000); the one used by C4.5 is named *Gain RATIO* (Kohavi and Quinlan 1999).

It starts from the measures of the *Entropy* at each node that is the measure of homogeneity of a node. The Entropy, calculated as:

$$\textbf{Entropy (t)} = -\sum_j \textbf{p(j|t)} \textbf{log(j|t)} \qquad \textbf{[2.1]}$$

Where p(j|t), that is the relative frequency of class j at node t, is minimum (0) when all records belong to one class, while is maximum (tends to 1) when records are equally distributed.

Based on the entropy, the *Information Gain,* that measures reduction in entropy achieved because of the split, is computed. The measure of Entropy at each node is connected to determination of the best split in this way:

$$\textbf{GAIN}_{\textbf{split}} = \textbf{Entropy(p)} - \left(\sum_{i=1}^{k} \frac{\textbf{n}_i}{\textbf{n}} \textbf{Entropy(i)}\right) \qquad \textbf{[2.2]}$$

Where p is the parent node, n_i is the number of records in partition, k is the number of partition of the parent node.

The disadvantage, that is that it prefer splits that results in large number of partition (each being small but pure), is overtaken with the use of the following index:

$$\text{GainRATIO} = \frac{GAIN_{split}}{-\sum_{i=1}^{k} \frac{n_i}{n} log \frac{n_i}{n}} \qquad \textbf{[2.3]}$$

That penalizes the use of a large number of small partitions.

In order to perform a first evaluation about the error rate of the learning technique, the so called "*10-fold cross-validation*" approach can be adopted.

The data is divided randomly into 10 parts in which the class is represented in, approximately, the same proportions as in the full dataset. Each part is held out in turn and the learning scheme trained on the remaining nine-tenths; then its error rate is calculated on the holdout set. Thus the learning procedure is executed a total of 10 times on different training sets (each of which have a lot in common). Finally, the 10 error estimates are averaged to yield an overall error estimate (Witten and Frank 2005). With this method it's possible to have first information about the performance of the classifier, also if it's obviously essential to perform other kind of accuracy assessment.

In order to obtain the final classifier, different parameters can be set by the user according to his needs. Particularly, the user can decide to use the *Reduce Error Pruning* technique instead of the one used by the C4.5 algorithm for pruning, to consider or not the *Subtree raising operation* when pruning, and, finally, the Minimum Number of instances that are demanded in each leaf. This last parameter allows performing a pruning acting on values that depend on the size of the initial training set.

2.2   Data used for the analysis

The aim of this work is the automation of procedures for urban land cover extraction; in spite this work just provides a first assessment about features to be considered for the classifier learning, the main project is conceived to be effective at a global level. Therefore, data to be used in the analysis should be provided with a global coverage.

The use of medium resolution multispectral data for urban characterization has not been largely considered. It is customary to think that high and very high spatial resolution sensors have bypassed this kind of data, providing to the final users more accurate data sets. Although the availability in time of images from privately owned HR and VHR sensors is almost the same as for publicly owned medium resolution satellites, costs are different, and for some applications, like change detection, urban sprawl characterization, urban vegetation monitoring and (partially) disaster management, the availability of low-cost images as well as expensive fine resolution datasets is equally important, and the lower cost of coarser-resolution images makes them more appealing for these applications. Moreover, in the public sector a need emerges to find out possible uses of available sensors for new, interesting characterizations, and this is the case for urban studies using Landsat multispectral data (Lisini et al. 2005). Multi-spectral data, ranging from 20 meter to 100 meter ground pixel size, have proven to be very effective in environmental analysis at a local to regional scale. This is particularly true in areas where human influence on the natural environment is heavy, as in the case of urban and sub-urban areas, due to the good compromise between the spatial and spectral details provided (Villa 2012).

Therefore Landsat data have been chosen as input data; both Landsat 5 and Landsat 7 satellite are still in orbit today, but the following restrictions need to be considered:

- The maximum spatial resolution of Landsat 5 data is limited to 30 meters;
- Landsat 7, whose spatial resolution can be improved up to 15 meters resolution, had a failure in the Scan Line Corrector (SLC) in 2003 (as described in 2.2.1.2); therefore most update images need a gap-filling processing to remove black lines before their use.

Landsat 7 data, however, were included in a collection, GLS2005, created in order to provide a global dataset with Landsat sensor in use at that time (2005); to be included in this collection, data were pre-processed in order to remove black line (as described in 2.2.2): data available in GLS2005 are therefore very suitable for this study and were used for the classification process, while Landsat 7 data in modality SLC-on (prior to 2003) were used for the training set on multitemporal data.

The analysis was conducted on Piedmont region, thanks to the availability of an open source Regional Cartography (CTRN, section 2.2.3) in vector format for validation purpose, with an update compatible with data to be used in the classification phase.

Data used for the analysis are described below.

2.2.1 Landsat Archive Collection

2.2.1.1 Landsat Mission

Landsat represents the world's longest continuously acquired collection of space-based moderate-resolution land remote sensing data. Nearly four decades of imagery provides a unique resource for those who work in agriculture, geology, forestry, regional planning, education, mapping, and global change research. Landsat images are also invaluable for emergency response and disaster relief.

The U.S. entered into a partnership with NASA in the early 1970's to assume responsibility for the archive management and distribution of Landsat data products.

In July 1972, NASA launched the first in a series of satellites designed to provide repetitive global observations of the Earth's land masses. Part of NASA's Earth Resources Survey Program, the Earth Resources Technology Satellite (ERTS) Program was later renamed Landsat to better represent the emphasis of the prime civil satellite program on remote sensing of land resources.

Landsat satellites can be subdivided into 2 groups:

- First generation satellite (Landsat 1, 2, 3), launched at an altitude of about 920 km, with an orbit period of 103 minutes and a revisit time of 18 days; Landsat 1, 2, and 3 carried the Multispectral Scanner (MSS) sensor and also included the experimental Return Beam Vidicon (RBV) cameras.

- Second generation satellite (Landsat 4,5,7), launched at an altitude of about 705 km, with an orbit period of 98.9 minutes and a revisit time of 16 days; the Landsat-4 satellite carried the MSS and Thematic Mapper (TM) sensors, as does the still currently orbiting Landsat-5 satellite. Landsat 6 failed at launch in 1993, but his twin, Landsat 7, was successfully launched on 15 April 1999 carrying on-board the Enhanced Thematic Mapper Plus sensor (ETM+).

Table 5 provides a summary of Landsat satellites series.

Satellite	Launched	Decommissioned	Sensors
Landsat 1	1972	1978	MSS/RBV
Landsat 2	1975	1982	MSS/RBV
Landsat 3	1978	1983	MSS/RBV
Landsat 4	1982	2001	MSS/TM
Landsat 5	1984	-	MSS/TM
Landsat 6	1983	Did not achieve orbit	ETM
Landsat 7	1999	-	ETM+

Table 5 – Mission dates and sensors achieved by Landsat satellites series (source: http://landsat.usgs.gov)

The Landsat Data Continuity Mission (LDCM) is the next-generation of Landsat satellites and is scheduled no earlier than February 11, 2013. This mission will ensure the continued acquisition and availability of Landsat-like data well beyond the duration of the current Landsat 5 and Landsat 7 missions.

2.2.1.2   Landsat 7

After the failure in the launch of Landsat 6 on 5 October 1993, Landsat 7 (Figure 6) was successfully launched on 15 April 1999, positioned on a repetitive, circular, sun-synchronous, near-polar orbit at 705 km altitude. This generation of Landsat improved technical devices maintaining the same base characteristics (swath width of 185 km and 16 day temporal resolution), which ensure the continuity of Landsat missions since 1982. Configuration of Landsat 7 and Landsat 5 orbits allows acquisition of an image on a certain area every 8 days, thus doubling the acquisition possibility.
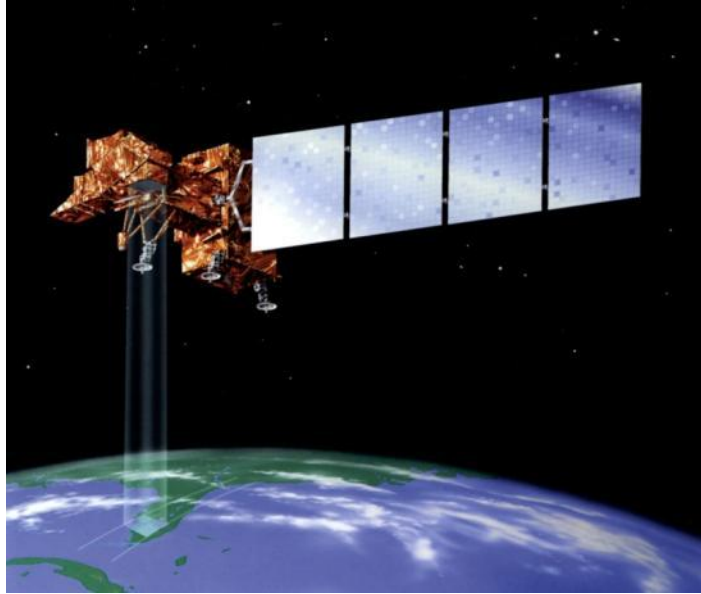
Figure 6 – Landsat 7 satellite (credits: NASA)

Landsat 7 is equipped by 1 panchromatic (Pan), 6 multispectral (MS) reflective bands and 2 thermal infrared (TIR) bands. Compared to his precursor Landsat 4 and 5, Landsat 7 makes use of an updated version of the TM, the Enhanced Thematic Mapper Plus (ETM+), a single nadir-pointing instruments.

The sensor is built in a "whisk broom" configuration (Figure 7). It uses rotating mirrors to scan from side to side perpendicular to the direction of the sensor platform. The rotating mirrors redirect the reflected light to a single or a few detectors (Gomarasca 2009). Data are collected one pixel at a time.

Figure 7 – Whiskbroom scanner (used by Landsat satellites)

ETM+ sensor provides:

- A panchromatic band at 15 meters geometric resolution with spectral resolution ranging from visible to near infrared, acquired simultaneously with the other 7 bands and thus being perfectly co-registered with them;

45

- A thermal band at 60 meters resolution, available in high and low restitution versions; this band is resampled at 30 meters resolution in imageries acquired after February 2010, in order to align the thermal band with the multispectral bands.

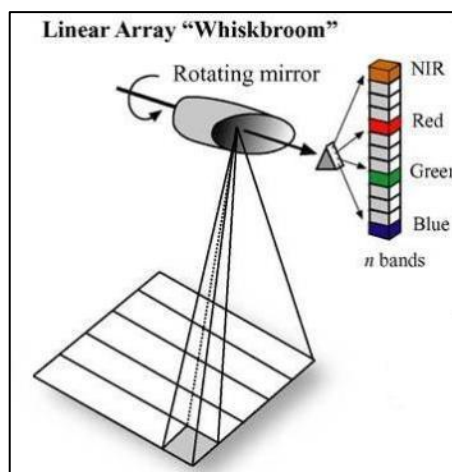The introduction of the panchromatic band is the most important innovation; covering the visible range (0.5 – 0.7 µm) as well as part of the near infrared (VIR: 0.7 – 0.9 µm), a better signal-to-noise ratio is obtained, improving the image readability. Data fusion techniques such as pan-sharpening make possible the integration of multispectral bands with the panchromatic one in order to create a single 15-meters resolution image. The simultaneous acquisition of panchromatic and multispectral images with homogeneous characteristics from the same satellite platform allows co-registration, time saving and processing cost reduction. The introduction of a thermal band with improved characteristics, both geometric and radiometric, allows enlarging the potential applications, such as pollution, thermal emissions, volcanic phenomena, etc. The improved geometric resolution and the confirmed temporal acquisition can give contribution to global change assessment, both providing elements for verifying land cover changes and defining local process that can induce these modifications (Gomarasca 2009).Bandwidth specifications related to Landsat 7 are provided in Table 6.

Enhanced Thematic Mapper Plus (ETM+)			
Landsat 7	Wavelength [µm]	Resolution [m]	Comparison with Landsat 5
Band 1 (VIS)	0.45-0.52	30	
Band 2 (VIS)	0.52-0.60	30	
Band 3 (VIS)	0.63-0.69	30	
Band 4 (VNIR)	0.77-0.90	30	
Band 5 (SWIR)	1.55-1.75	30	
Band 6 (TIR)	10.40-12.50	60	120 meters
Band 7 (SWIR)	2.09-2.35	30	
Band 8 (PAN)	0.52-0.90	15	Not available

Table 6 – Landsat 7 ETM+ band designation (and comparison with the ones provided by Landsat 5).

On 31/05/2003, the Scan Line Corrector (SLC), which compensates for the forward motion of Landsat 7, failed. Subsequent efforts to recover the SLC were not successful, and the failure appeared to be permanent. Without an operating SLC, the Enhanced Thematic Mapper Plus (ETM+) line of sight now traces a zig-zag pattern along the satellite ground track (Figure 8) with resulting data gaps that form alternating wedges that increase in width from the center of the image to the edge; an estimated 22% of any given scene is lost because of the SLC failure.
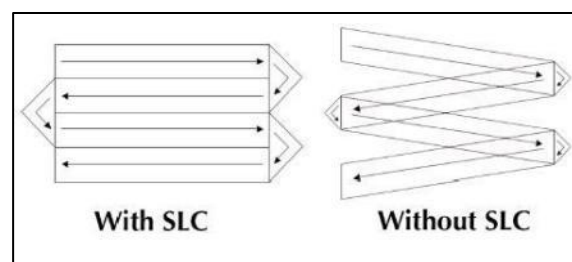

Figure 8 – SLC failure occurred on Landsat ETM+ on 31/05/2003

However, Landsat 7 is still capable of acquiring useful imagery worldwide with the SLC turned off (SLC-off mode), particularly within the central part of any given scene, and these data are of the same high radiometric and geometric quality as data collected prior to the SLC failure.

The USGS has provided the user community with methods to fill gaps, both for visualization or for scientific analysis; they makes use of other Landsat scenes acquired on the same site shortly before or after the gap. Band-specific gap mask files are included with every SLC-off standard data product. These ancillary data allow users identify the location of all pixels affected by the original data gaps in the primary SLC-off scene. Please refer to http://landsat.usgs.gov/using_Landsat_7_data.php for further information about available methods.

A summary of Landsat 7 collection main features is represented in Table 7. Landsat scenes held in the USGS archive are available for download using USGS Global Visualization Viewer (GloVis) or EarthExplorer.

Landsat Archive Collection	Data availability	Spatial resolution	Access instruction
Landsat 7 ETM+ SLC-off	2003 - present	Until 15 m	http://glovis.usgs.gov/ http://earthexplorer.usgs.gov/
Landsat 7 ETM+ SLC-on	1999 - May 2003	Until 15 m	

Table 7 – Main features of Landsat 7 collections

2.2.2   Global Land Survey (GLS)

The USGS and the NASA collaborated on the creation of global land datasets using Landsat data from 1972 through 2008; the dataset is named Global Land Survey, and each collection that constitute is was created from the primary Landsat sensor in use at the time (Table 8): the Multispectral Scanner (MSS) in the 1970s, the Thematic Mapper (TM) in 1990, the Enhanced Thematic Mapper Plus (ETM+) in 2000, and a combination of TM and ETM+, as well as EO-1 ALI data, in 2005.
Recently, NASA and the USGS have again partnered to develop the Global Land Survey 2010 (GLS2010), a new global land data set with core acquisition dates of 2008-2011. This dataset consists of both Landsat TM and ETM+ images that meet quality and cloud cover standards established by the earlier GLS collections. Data acquired in 2011 are used to fill areas of low image quality or excessive cloud cover.

GLS Collections	Images acquisition date	Satellite/ Sensor
GLS 1975	1972 – 1983 1982 - 1987	Landsat 1-3 MSS Landsat 4-5 MSS (to fill gaps in Landsat 1 – 3 data)
GLS 1990	1987 - 1997	Landsat 4 – 5 TM
GLS 2000	1999 - 2003	Lansdat 7 ETM+
GLS 2005	2003 - 2008	Landsat 5 TM; Landsat 7 ETM+; EO-1 ALI
GLS 2010	2008 - 2011	Landsat 5; Landsat 7 ETM+

Table 8 – Global Land Survey collections (source: http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/GLS)

GLS2005 was used in this thesis: the goal of this collection is to provide one clear image during leaf-on conditions for every location of the globe between 2004 and 2007, with priority on 2005. In this case, the choice of "primary" sensor was not obvious. Since the May 2003 failure of the scan-line corrector mechanism aboard Landsat-7, ETM+ images have been afflicted with cross-track, wedge- shaped gaps that eliminate 24% of the image area. Although USGS EROS has implemented techniques to merge ("gap fill") several ETM+ acquisitions from the same season, it was not clear if these products would be suitable for rigorous land cover change analyses. Conversely, although Landsat-5 TM is still in service, the radiometric and geometric properties of Landsat-5 TM images are worse than Landsat-7 ETM+ (Masek 2005). Thus, the priority assigned to the choice of the reference scene was the following:

- One Landsat 5 image OR one gap-filled Landsat 7 pair;

Where Landsat-5 was not available/suitable and no cloud free L7 pair was available:

- Two Landsat 7 SLC-off images, without gap filling;

For humid tropics where no cloud free L7 pair was available:

- Three Landsat 7 SLC-off images, without gap filling;

Where no Landsat data were available:

- EO-1 ALI or ASTER image

Images were gap-filled through an algorithm, created by EROS, which used one SLC-off image as a "base" image, and designated one or more images as "fill" images. Ideally base and fill images should have been acquired within 1-2 orbital cycles, or at least within the same season. In addition, since the position of the gaps varies randomly between images, merging any two images may not completely fill the missing area of the base image. Since the radiometry changed between the base and fill image (due to atmospheric effects, BRDF changes, phenology, etc) the fill images were radiometrically adjusted to match the base image before compositing. A small moving window (7-11 pixels) was applied to a pair of base and fill images, and a per-band linear regression applied. The regression was used to calculate a gain and bias, which were then applied to the fill image to radiometric match it to the base image. After radiometric adjustment, pixels

from the fill image were substituted for the missing "gaps" in the base image, in order to create a gap-filled image; a "gap mask" was also created, in order to allow users to determine which pixels in the final scene were derived from which input scene (Figure 9).

In GLS2005 a large majority of the scenes have less than 5% residual gap area after filling.
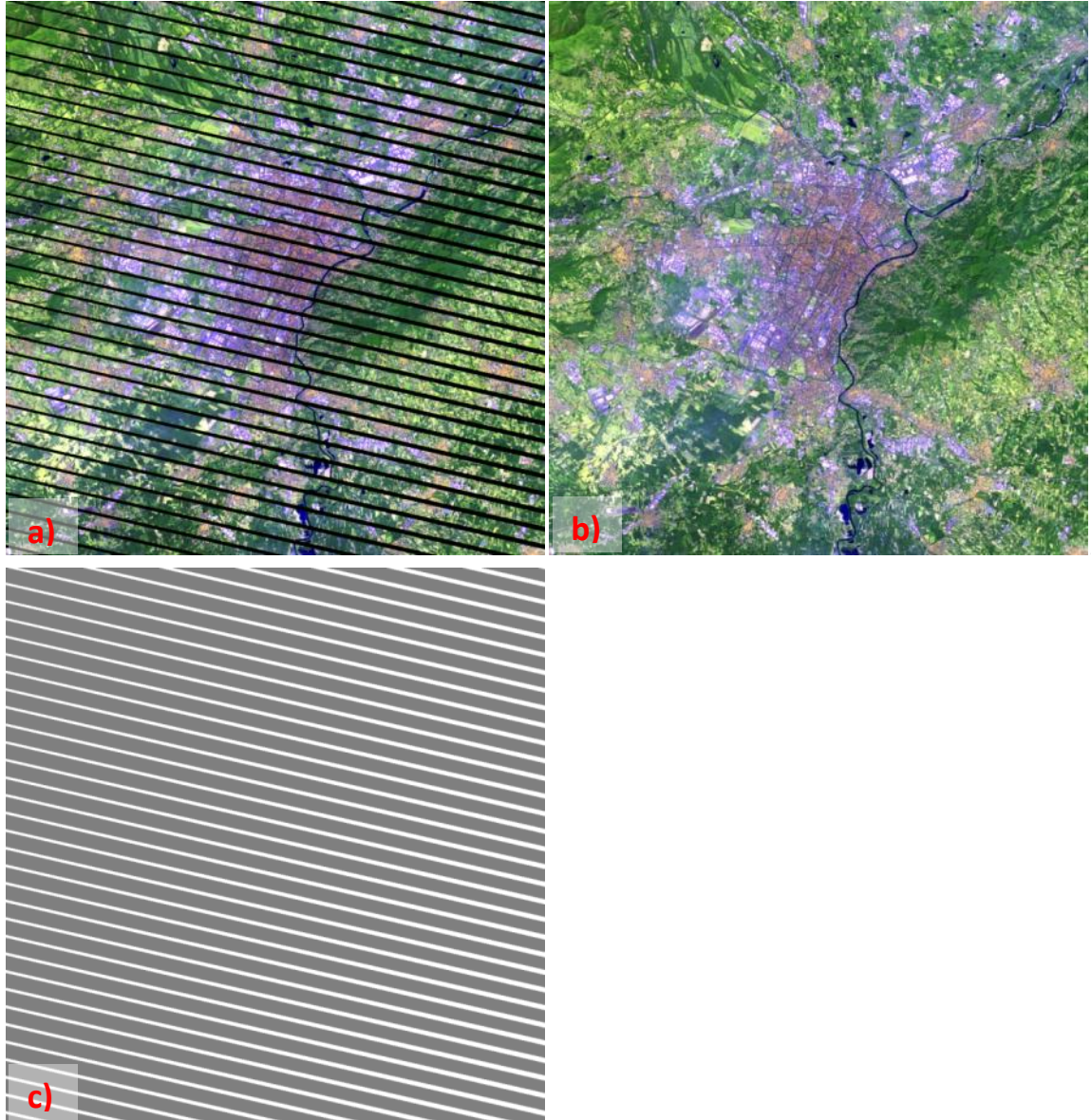
Figure 9 – a) Subset of gap-filled product (GLS 2010) acquired at Path/Row 195/29. Base image: 5/08/2009; Fill image: 6/09/2009. Gap fill percent: 98.8%; b) Subset of the Base image as acquired by Landsat 7 ETM+ on 05/08/2009 without gap-filling; c) Subset of gap mask

2.2.3 Digital Regional Technical Map (CTRN)

The Digital Regional Technical Maps (Carta Tecnica Regionale Numerica - CTRN) of Piedmont Region is composed, approximately, by 800 sections at scale 1:10.000.

CTRN was initially derived from aerial images acquired in 1991 for all the region except for the Susa valley whose images where acquired in 1995; later updates were derived from further acquisitions carried out between 1999 and 2005. Different places refer, therefore, to different acquisition dates; Figure 10 provides a summary of the actual frame.

1991	primo impianto
1995	valle di Susa
1999	aggiornamento provincia di Torino (parte)
2001	prototipo Piemonte Est (parte province di Novara e Alessandria)
2004	provincia di Asti
2005	area metropolitana (fogli 155-156)

Figure 10 – Update of CTRN in Piedmont

Dataset is provided in a vector format and organized into a hierarchic way; the first level is composed by 11 groups:

1) Streets and Railways;
2) Buildings and street forniture;
3) Water and handiworks related to water;
4) Lines and pipes for energy, liquid and soil material, people transport (lines);
5) Ground divisions;

6) Ground elements;
7) Vegetation;
8) Altimetry (lines, points);
9) Point of interest;
10) Administrative boundaries (lines);
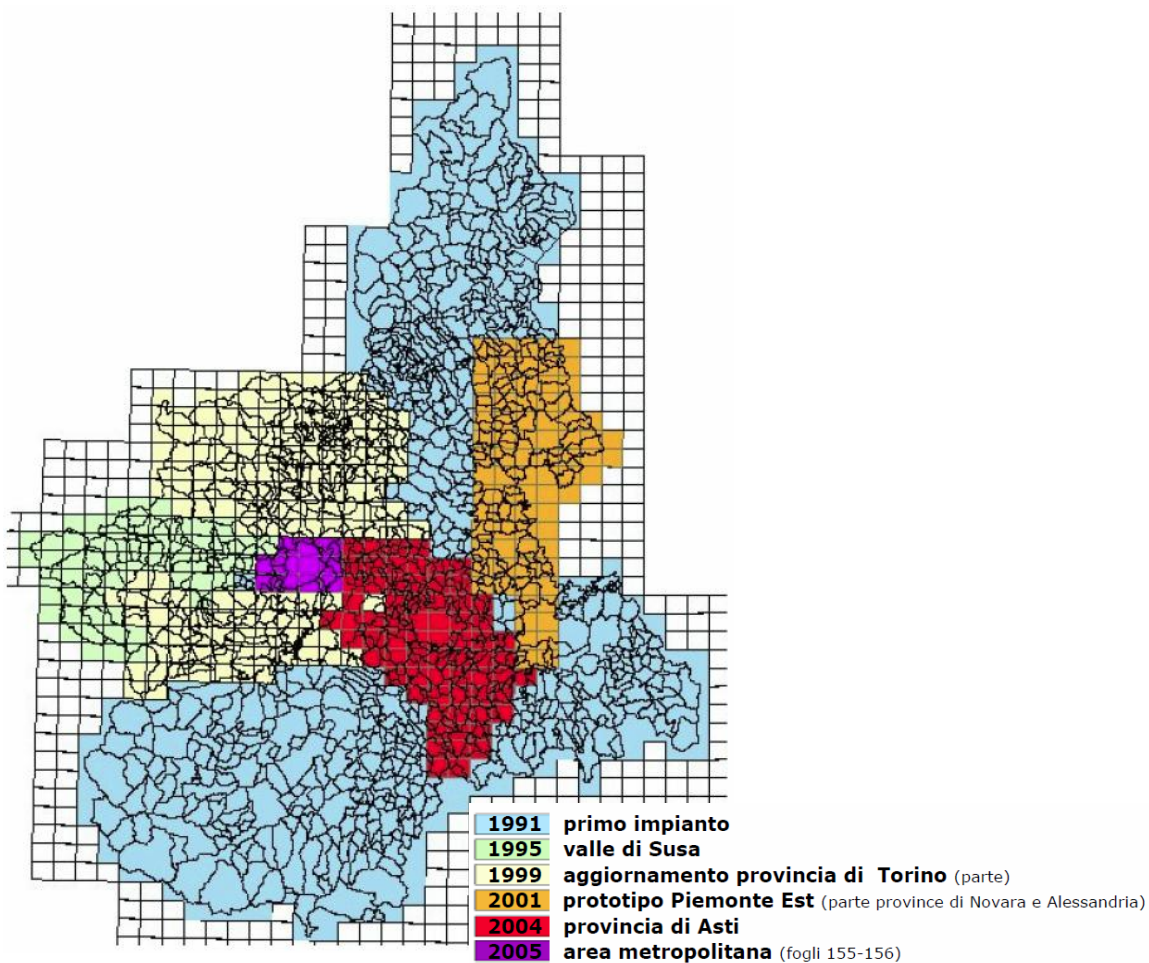11) Positioning network (points).

The elements that constitute these groups can be punctual, linear or areal; since the cartography was converted into raster format for validation purposes, only areal elements have been taken into consideration. Groups 4, 8, 10, 11 are the only one not to contain any areal element.

Data are georeferenced into WGS84/ UTM 32 Nord Reference system and accessible at: http://www.regione.piemonte.it/repcarj/welcome.do?ric=3&ric=3

A detailed description of elements available in each group is provided in Annex II.

2.3 Software

A description of the software used for the study is provided below:

- *ENVI[1] (ENvironment for Visualizing Images) 4.7*, a commercial software for processing and analyzing geospatial imagery, commonly used for remote sensing application; it has been used for the pre and post processing, classification and validation phase. Procedures have also been implemented in IDL (Interactive Data Language), ENVI scientific programming language, in order to automatize the different steps of the workflow.
- Weka (Waikato Environment for Knowledge Analysis) 3.6, a free data mining software available under the GNU General Public License, developed at the University of Waikato, New Zealand, by the Machine Learning Group (Hall et al. 2009); it provides a collection of machine learning algorithms for data mining tasks and contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It has been used for the learning process: algorithms provided by the software have been used to learn the classifier, later applied to images through ENVI software.
- ArcGis 10.0[2], a commercial software property of Environmental Systems Research Institute, Inc. (Esri) for working with maps and geographic information, usable for creating and using maps, compiling geographic data, analyzing mapped information, using maps and geographic information in a range of applications and managing geographic information in a database. It has been used for the creation of the validation mask.

---

[1] © 2012 Exelis Visual Information Solutions
[2] © 1995–2012 Esri.

3 Methodology

3.1 Introduction

This research work was developed with the aim of supporting the production of cartography for emergency response, thus constituting a first step towards the definition of a methodology for automatic urban extraction. The step that was investigated through this work is related to the definition of a procedure for the classification of urban areas, that is the set of step to follow in order to obtain a good classification procedure. As previously mentioned in section 1.3.1, the classification method selection, together with the choice of variables and attributes to use as input to the classifier, are considered key elements in the classification process.

The chosen classification method, Classification Tree, is based on the extraction of rules that allow identifying the desired land cover type; rules are, in turn, based on the definition of the attributes to include inside together with the choice of decision thresholds.

This work examined the potentiality of different variables for the classification process: features for the learning process, as well as the use of different pre and post-processing. Considered features are: attributes in input to the classifier, use of multitemporal data and the selection of 2 or more classes to train the classifier. A summary of the considered variables is provided in Figure 11.
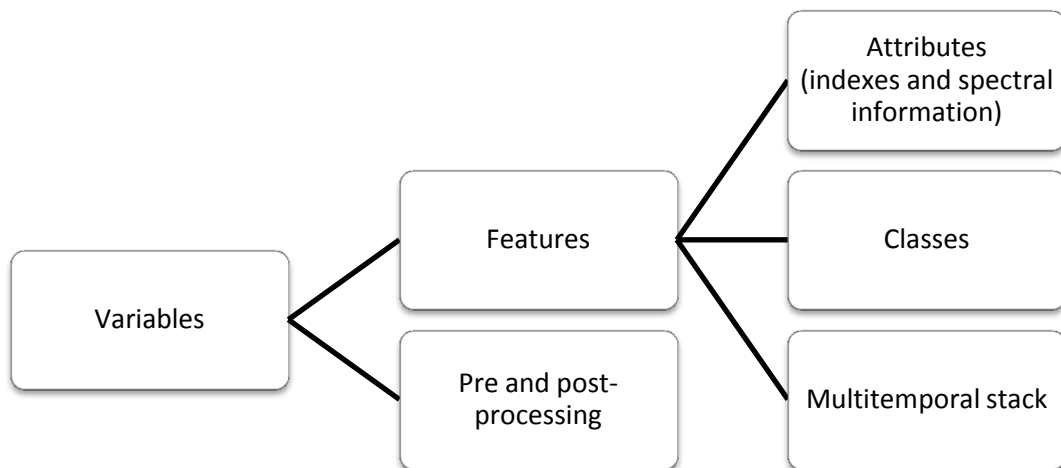

Figure 11 – Summary of analyzed variables

The first section of this chapter describes the methodology used in this work, while the second part describes how this methodology was applied to different case studies.

3.2   Applied methodology

The process of image classification may include, as previously described in section 1.3.1, seven main steps: determination of a suitable classification system, selection of training samples, image pre-processing, feature extraction, selection of suitable classification approaches, post-classification processing and accuracy assessment;

Taking into account the above mentioned steps, and adapting it to our needs, the workflow can be subdivided in eight steps, as reported in Figure 12; it starts from the definition of the classes to be mapped together with some information about the possible output of the procedure, than goes on with the selection of sample areas, that serve to learn the classifier, analyzes pre-processing to apply to images and features to be used in the training set, and arrives to the creation of different training sets and corresponding models (classifiers) that are than applied to different images in order to provide thematic maps; finally, different kind of post processing are evaluated.

In this work some of these steps are considered variables; which pre-processing allows obtaining higher accuracy? Which features can I use in my classification? How can I improve my results with post-classification? All these variables are described in the following sections.

More specifically, the procedure applied in this work is the following: first, sample areas, to be used to learn the classifier, are detected on satellite images with the help of higher resolution images. Training sets for the detected sample areas are then built taking into account different variables: the radiometric pre-processing applied to images, the attributes used as input, the number of classes use to learn the classifier and, finally, the temporal extension of the dataset. The combination of all these variables led to the creation of 36 different training set; in this work the term "sample area" is used to indicate the geographic location of elements that are in the training set, while the term "training set" is used to indicate the type of information associated to each pixel detected in the sample area. For example, if 3 different kind of radiometric correction are performed on the same image, three different training sets result: the radiometric information associated to the same pixel in different images is, indeed, different. Each training set is then used as input to the classifier; the application of the same algorithm to different training set led to the definition of different classifiers, each with given attributes and decision thresholds. Each classifier is then applied to one common image; it derives as many classifications as classifiers are.

Finally, different kinds of post processing are applied, thus resulting in a further increase of thematic map produced. At the end, each thematic map is evaluated trough the accuracy assessment described in section 3.2.8; a schema of the above mentioned procedure is provided in Figure 13. The numeric label associated to the narrow serves as link between the main outputs of the process, as described in Figure 13, and the different steps of the workflow, as represented in Figure 12.

1. Determination of the classification system and associated information

2. Selection of the sample areas

3. Image pre-processing

4. Feature extraction

5. Training set generation

6. Classification

7. Image post-processing

8. Validation
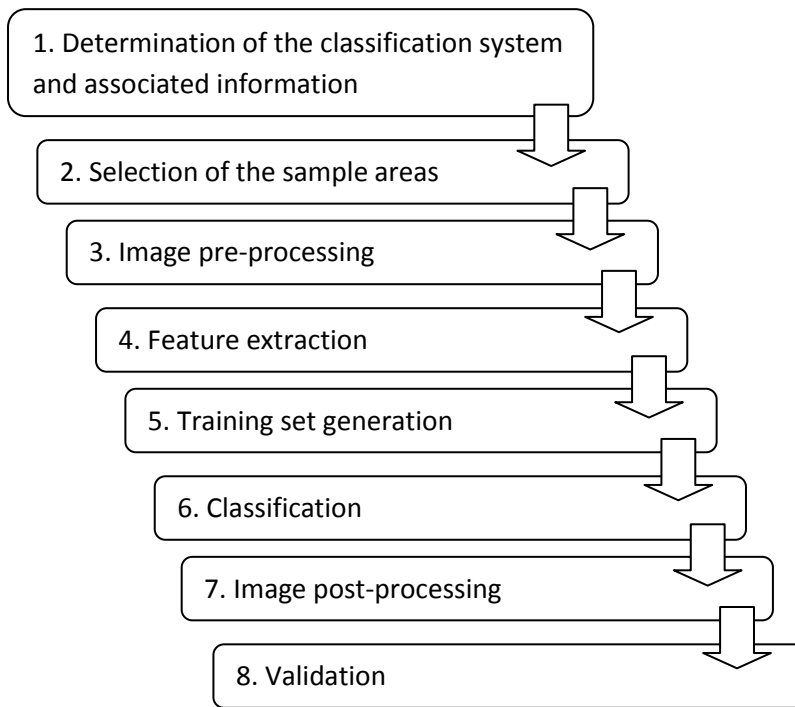
Figure 12 – Workflow of the applied methodology (processing)
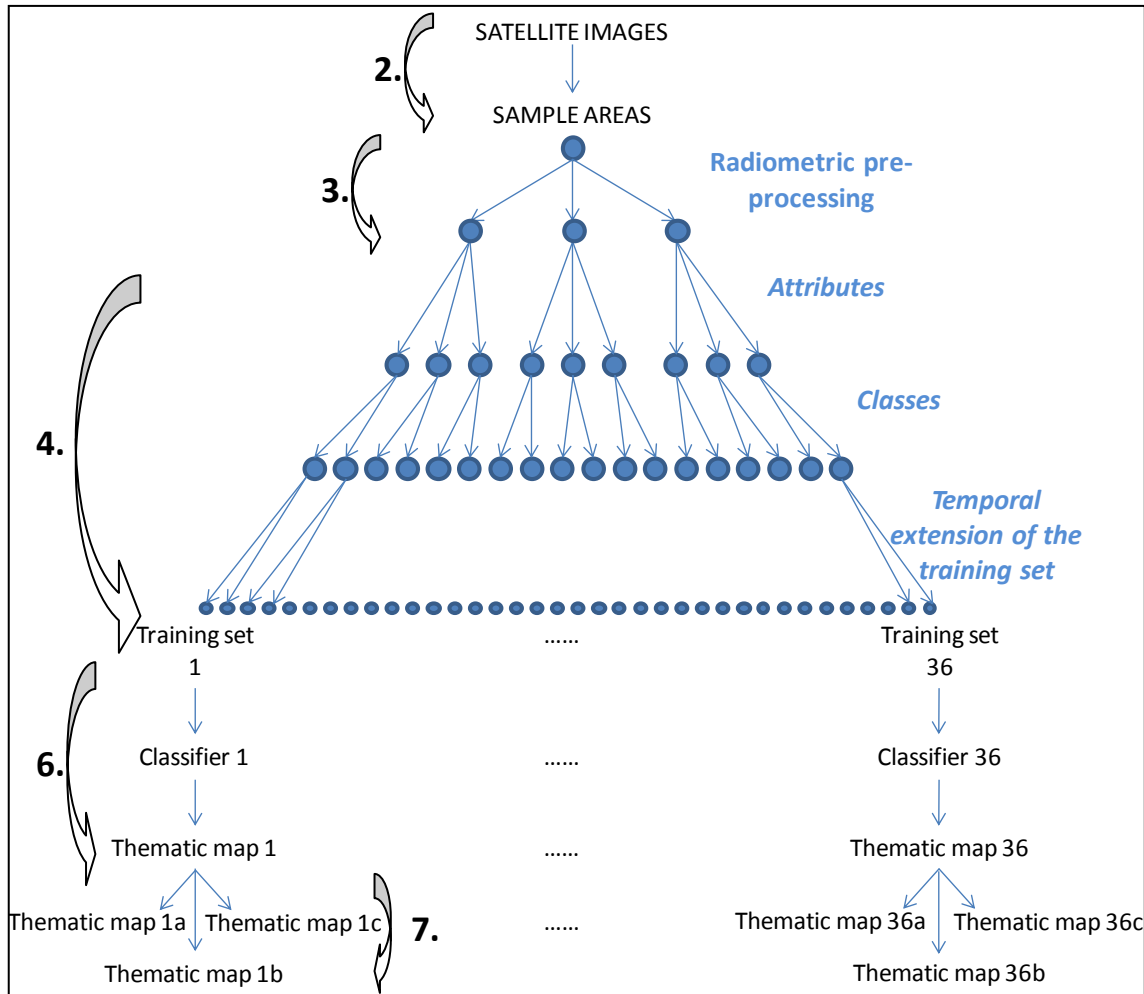
Figure 13 – Output achieved at each step of the applied methodology (in black). Numbers refer to processing shown in Figure 12. In blu: intermediate steps.

3.2.1 Preliminary observation

It's very important, when planning a project involving remotely sensed data, that sufficient effort is given to the classification scheme to be used.

In this case, the aim is to distinguish between urban and non urban surfaces, but a definition of "urban" and "non urban" must be given.

Within the context of this work, the term "urban" refers to all man-made features; the concept is linked to the definition of surface materials and properties, rather than to the actual land use. All impervious surfaces (residential buildings, industrial buildings, streets, pavement, etc.) are considered "urban areas".

Some consideration should be performed also on the scale suitable for the output: which is the ideal scale, or the accettable range of map scales, in funcion of the resolution of the input data?

An image's geometric resolution limit is given by the matrix cells size; enlargement over 2 pixels/mm generate the perception of a single cell as a discrete element (Gomarasca 2009). For

same applications where the spectral information is more important, it is possible to extend this resolution, producing out-of-focus, pixelated images. A relationship between the most frequently used satellite imagery, their applications and the most appropriate scales of cartographic restitution is provided in Table 9.

Satellite/sensor	Scale								
	2.5 Mil	1 mil	500k	250k	100k	50k	25k	10k	5k
AVHHR SPOT veg	GC LC G								
Resurs – 01		GC LC G							
Landsat MSS			GC LC G						
Landsat TM				GC E	LC G				
Landsat ETM+					GC E	LC G			
SPOT XS, XI/IRS, LISS					GC E	LC G			
SPOT Pan							T LC G		
IRS Pan							T LC G		
Ikonos, Quickbird, Orbview, DK 1-2								T LC G	

GC: General cartography   LC: Land cover   G: Geological cartography
E: Environmental thematic cartography   T: Topographic cartography

Table 9 – Relationship between the most frequently used satellite imageries, their applications and the most appropriate scales of cartographic restitution (Gomarasca 2009)

It gather that Landsat ETM+ imageries could be used to produce land cover maps at scales ranging from 1:25.000 to 1:100.000.

Taking into consideration the guideline of the American Society for Photogrammetry and Remote Sensing (ASPRS Specification Standards Committee, 1990) based on the cartographic convention to fix the minimum readable tickness for graphic elements of a map to 0.2 – 0.3 mm, the ideal scale could be set to 1:120.000 $\left(\frac{30\,m}{0.00025\,m}\right)$ and, considering the maximum accettable scale corresponding to the dimension of 0.5 mm for each information pixel, the maximum scale could be set to 1:60.000.

Finally, also CLC2006 dataset, developed from SPOT-4 HRVIR at 10/20 m resolution and/or IRS P6 LISS III at 23 m resolution and his predecessor, CLC2000, developed from Landsat 7 ETM+ data at 30 m resolution, are provided for maps at scale 1:100.000.

It stands to reason to fix the output scale of the urban maps at 1:100.000.

3.2.2   Selection of the sample areas

Sample areas are detected on the reference image in order to belong to these classes:

1. Urban (residential buildings)
2. Industrial areas (industrial buildings)
3. Vegetation
4. Water
5. Bare soil

In order to correctly identify sample areas, higher resolution maps (Google Earth) and the CTRN are used to perform a quality check on identified areas.

3.2.3   Radiometric pre processing

The data collected by sensors on board different platforms, before being used for the interpretation, need to be processed to correct errors due to the noise and distortions generated during acquisition and transmission. In this study, the attention was mainly focused on radiometric pre-processing: the aim is to obtain comparable units in order to analyze and measure the earth surface and its changes over the time. Radiometric errors depend on:

- the sensor;
- the system geometry;
- the atmosphere.

Effects due to the sensor acquisition can be removed through a radiometric calibration into radiance; this process transforms the image Digital Number (DN) into radiance (L) value measured by the optical systems' detectors. This transformation, that is called r*adiometric calibration into radiance* (Figure 14), is generally expressed by parametric linear functions related to calibration coefficients specific to each sensor and functional to the wavelength or the spectral band. Radiance (L) can be obtained as:

$$L(\lambda) = DN(\lambda) * gain(\lambda) + offset(\lambda)$$

(3.1)

Where:

L(λ): spectral radiance $\left[\frac{W}{m^2*sr*\mu m}\right]$

57

DN(λ): Digital Number derived from the input spectral radiance

gain(λ): gain coefficient or system's amplification $\left[\frac{W}{m^2*sr*\mu m*DN}\right]$

offset(λ): movement coefficient $\left[\frac{W}{m^2*sr*\mu m}\right]$

λ: wavelength or spectral range [μm]

Obtained radiance values are still affected by other factor, such as terrain characteristics, illumination and observation conditions, and the presence of atmosphere.

A first reduction of the variability among the scenes can be obtained through the normalization with respect to the incident radiation, going from the spectral radiance (L) to the apparent reflectance (ρ_{app}), with the process called *radiometric calibration into reflectance* (Figure 14)*. This correction is based on a flat surface patter.

The relationship is the following:

$$\rho_{app}(\lambda) = \frac{\pi \; d^2 L(\lambda)}{E_0(\lambda)\cos\theta_z}$$

(3.2)

Where:

$\rho_{app}(\lambda)$: apparent reflectance (as perceived by the sensor)

L(λ): spectral radiance measured by the sensor above the atmosphere $\left[\frac{W}{m^2*sr*\mu m}\right]$

d: Earth-sun distance during the scenes' acquisition

E_0: solar irradiance above the atmosphere incident on a surface perpendicular to the radiation $\left[\frac{W}{m^2*\mu m}\right]$

θ_z: incident angle of the solar flux directed towards the Earth Surface, defined as Sun zenith angle.

This is just a first step in order to compare satellite data acquired at different times; atmospheric corrections are needed (Figure 14) in order to depurate the signal from the atmospheric effects and to obtain the *albedo* that is the ground reflectance (ρ_g). This correction, besides clearing the signal from atmospheric absorption and scattering effects, also includes the radiometric calibration and the normalization of the effects due to the system geometry. In practice, through the atmospheric correction, the main radiometric effects in remotely sensed images are normalized. There is no unique method for the atmospheric effect correction that is simple, accurate, and widely used (Gomarasca 2009). As a consequence, numerous methods have been developed, for specific kinds of problems and different levels of accuracy. Based on the quantity

and on the accuracy of atmospheric parameters needed to apply them, the methods for studying atmospheric effects can be distinguished into two groups:

- Models based on the physics of the radiative transfer;
- Methods based on the images (simplified methods).

The first category includes the models solving the equation of electromagnetic energy radiative transfer through the atmosphere (Radiative Transfer Code). To accurately describe the radiation propagation, these models need in situ collection measures of the optical properties of the atmosphere acquired at the same time as the scene acquisition. In general radiative transfer codes combined with in situ atmospheric measures produce the most accurate assessment. Nevertheless, all these procedures are often too expensive and complicated to be commonly applied. The biggest disadvantage of this kind of correction is that it requires measures of atmospheric parameters simultaneously with each passing of the sensor. This is often impossible in monitoring programmes, or when historical series have to be analysed (Gomarasca 2009). In order to overcome this inconvenience, other correction methods have been proposed, where the information about the atmospheric properties can be retrieved from the image. Despite these methods produce less accurate results compared to radiative transfer models, however they are widely used and are a good alternative when the atmospheric properties during the image acquisition are not known.

Among the most known image based methods, the Dark Subtraction method is based on the hypothesis that in the scene at least one pixel has reflectance equal to zero. This pixel's radiance contribution hence depends only on the atmospheric component which is subtracted from all the image's pixels. The operation is applied band by band. Among the surfaces that better reply to the dark pixel hypothesis, there are deep and oligotrophic water bodies or sharp shadowed areas.
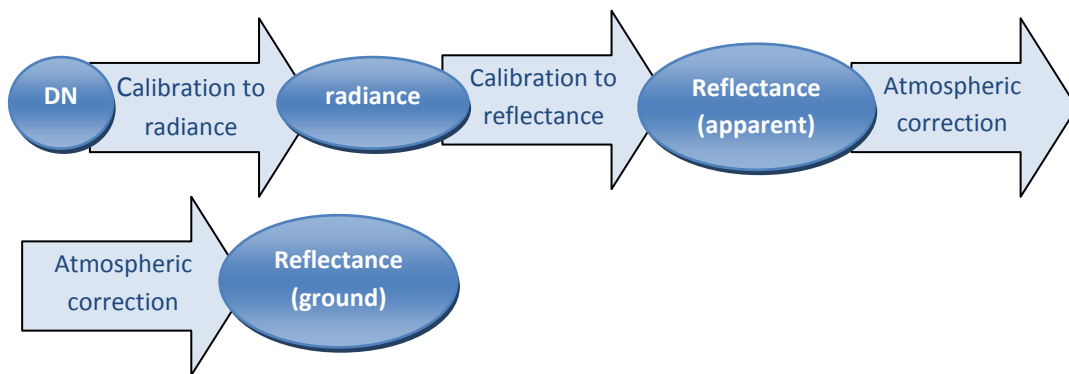

Figure 14 – Radiometric pre-processing needed to obtain comparable measures

In this case study classification obtained with different level of radiometric pre-processing (calibration into reflectance, atmospheric correction with a simplified method - Dark Subtraction and correction with a rigorous method - FLAASH) were compared (Figure 15).
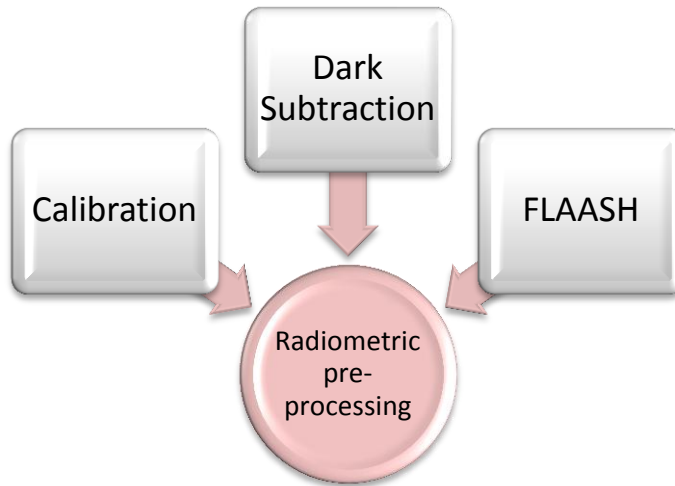
Figure 15 – Radiometric pre-processing compared in the study

Calibration is performed with the Landsat Calibration tool provided by ENVI© that automatically reads calibration parameters from the metadata associated to the image (Data acquisition, sun elevation, satellite sensor) and performs the pre-processing.

Dark subtraction is performed with the Dark Subtraction tool provided by ENVI© that applies scattering correction to image data by subtracting the band minimum DN to every pixel in the selected band (alternatively, other values can be subtracted, depending on the user needs).

Atmospheric correction using a model based on the physics of the radiative transfer code is performed using FLAASH, a first-principles atmospheric correction tool that corrects wavelengths in the visible through near-infrared and shortwave infrared regions, up to 3 µm.
Unlike many other atmospheric correction programs that interpolate radiation transfer properties from a pre-calculated database of modeling results, FLAASH (Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes) incorporates the MODTRAN4 radiation transfer code.
It starts from a standard equation for spectral radiance at a sensor pixel, L, which applies to the solar wavelength range and flat, Lambertian materials or their equivalents. The equation is as follows:

$$L = \left(\frac{A * \rho_g}{1 - \rho_e * S} \right) + \left(\frac{B * \rho_e}{1 - \rho_e * S} \right) + L_a \qquad \text{(3.3)}$$

where:

ρ_g is the pixel surface reflectance

ρ_e is an average surface reflectance for the pixel and a surrounding region

S is the spherical albedo of the atmosphere

L_a is the radiance back scattered by the atmosphere

A and B are coefficients that depend on atmospheric and geometric conditions but not on the surface.

All these variables depend on the spectral channel; the first term in Equation (3.3 corresponds to radiance that is reflected from the surface and travels directly into the sensor, while the second

term corresponds to radiance from the surface that is scattered by the atmosphere into the sensor. The distinction between ρ and ρ_e accounts for the adjacency effect (spatial mixing of radiance among nearby pixels) caused by atmospheric scattering. The values of A, B, S and L_a are determined from MODTRAN4 calculations that use the viewing and solar angles and the mean surface elevation of the measurement, and they assume a certain model atmosphere, aerosol type, and visible range. The values of A, B, S and L_a are strongly dependent on the water vapor column amount, which is generally not well known and may vary across the scene. If no water information is available, it's possible to select the known or expected surface air temperature which tends to correlate with water vapour. If the temperature is unknown, it's possible to select a Model Atmosphere based on Latitudinal/seasonal dependence of surface temperature. This is the schema adopted to select the column water vapour amount. It starts from the detection of the latitude scene in order to select the correct Model Atmosphere (Table 11), and then connect the estimated water amount (Table 10).

Since input images were acquired at about 45° latitude and in the period June-November, SAS and MLS models have to be used.

Model Atmosphere	Water vapour [g/cm^2]	Surface air temperature [°C]
Sub-artic winter (SAW)	0.42	- 16
Mid-Latitude Winter (MLW)	0.85	-1
US Standard (US)	1.42	15
Sub-artic summer (SAS)	2.08	14
Mid-latitude Summer (MLS)	2.92	21
Tropical (T)	4.11	27

Table 10– Column water vapour amounts and surface temperatures for the MODTRAN Model Atmospheres

Latitude [°N]	Jan	March	May	July	Sept	Nov
80	SAW	SAW	SAW	MLW	MLW	SAW
70	SAW	SAW	MLW	MLW	MLW	SAW
60	MLW	MLW	MLW	SAS	SAS	MLW
50	MLW	MLW	SAS	SAS	SAS	SAS
40	SAS	SAS	SAS	MLS	MLS	SAS
30	MLS	MLS	MLS	T	T	MLS
20	T	T	T	T	T	T
10	T	T	T	T	T	T

Table 11– Selection of MODTRAN Model Atmosphere Based on Latitudinal/Seasonal Dependence of Surface Temperature

FLAASH includes a method for retrieving the aerosol amount and estimating a scene average visibility using a dark pixel reflectance ratio method based on work by Kaufman et al. (1997). The dark-land pixel-retrieval method requires the presence of sensor channels around 660 nm and 2100 nm. A dark-land pixel is defined to be one with a 2100 nm reflectance of 0.1 or less and a

660:2100 reflectance ratio of approximately 0.45. If the input image contains bands near 800 nm and 420 nm, an additional check is performed, requiring the 800:420 radiance ratio to be 1 or less, which eliminates pixels likely to be shadows and water bodies. In Landsat images, band 3 and band 7 corresponds to wavelength of 660 nm and 2100 nm. If no dark pixels are found, then an Initial Visibility (estimate of the scene visibility in km) is used.

Finally, an aerosol model should be chosen; the choice of the model is not critical if the visibility is high. The model choices are the following:

- *Rural*: Represents aerosols in areas not strongly affected by urban or industrial sources. The particle sizes are a blend of two distributions, one large and one small;
- *Urban*: A mixture of 80% rural aerosol with 20% soot-like aerosols, appropriate for high-density urban/industrial areas;
- *Maritime*: Represents the boundary layer over oceans, or continents under a prevailing wind from the ocean. It is composed of two components, one from sea spray and another from rural continental aerosol (that omits the largest particles);
- *Tropospheric*: Applies to calm, clear (visibility greater than 40 km) conditions over land and consists of the small-particle component of the rural model. For more details on MODTRAN aerosol models, see Abreu and Anderson (1996).

To sum up, the following parameters were set to correct images from atmospheric effects using FLAASH:

Subset	Input image	Acquisition time	Model Atmosphere	Aerosol model	Ground elevation	Scene center location
Torino	30/11/1999	10:10:04	SAS	rural	0.4	45°4'31.18'' N 7°38'4.94'' E
	28/09/2002	09:58:52	SAS			
	06/10/1999	10:04:09	SAS			
	24/08/2001	09:59:54	MLS			
	05/03/2000	10:09:54	SAS			
	30/07/2001	10:06:28	MLS			
	02/07/2005	10:00:00	MLS			
Asti	01/05/2000	10:03:24	SAS	rural	0.27	45°49'40.57'' N 8°12'13.47'' E
	06/10/1999	10:04:09	SAS			
	21/06/2001	10:00:35	SAS			
	23/11/1999	10:03:53	SAS			
	24/08/2001	09:59:54	MLS			
	28/09/2002	09:58:52	SAS			
	02/07/2005	10:00:00	MLS			

Table 12 – Set parameters for the atmospheric correction with FLAASH

3.2.4   Feature extraction

3.2.4.1   Attributes: indexes and spectral information

Spectral indexes are often used in remote sensing applications in order to predict, model or infer surface process. They can be used to assess in the monitoring of several different land change processes, such as vegetation health and status, burned area, fire severity, or to help in the identification of a certain thematism.

An optimal spectral index is very sensitive to the desired information (e.g., the amount of vegetation), and as insensitive as possible to perturbing factors (such as soil color changes or atmospheric effects). Since both the desired signal and the perturbing factors vary spectrally, and since the instruments themselves only provide data for particular spectral bands, optimal indexes should be designed for specific applications and particular instruments (Verstraete and Pinty 1996).

A spectral index should:

- Maximize the sensitivity of certain surface properties. Ideally, such responses should change linearly to allow both ease of scaling and use over a wide range of surface conditions;
- Normalize or reduce effects due to sun angle, viewing angle, the atmosphere, topography, instrument noise, etc, to allow consistent spatial and temporal comparisons;
- Be linked to specific and measurable surface processes (e.g. biophysical parameter such as LAI, biomass, APAR, etc) – i.e. be related to a measurable parameter or process.

In this study classifications obtained with different kind of information (only spectral information, only indexes, spectral information and indexes) were compared (Figure 16).

The aim is dual: on one side, the purpose is to evaluate if indexes are selected as suitable variables from the algorithm, on the other it's intended to know which indexes more easily allow to separate classes.
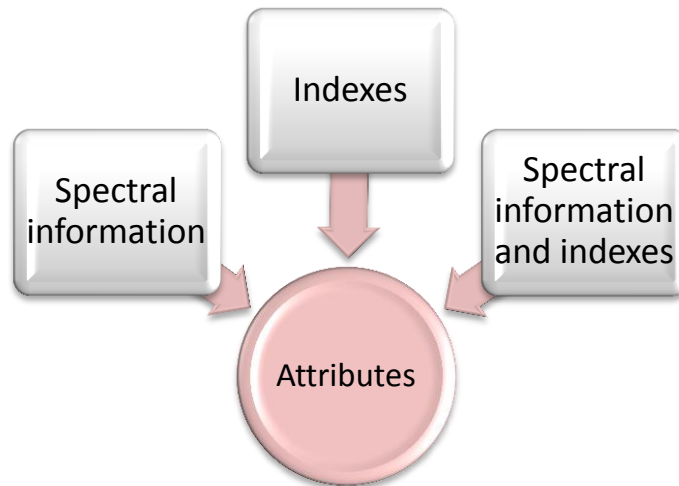
Figure 16 – Set of attributes compared in the study

After a literary review, and according to our needs, indexes that were taken into account in the study are the following:

- Normalized Difference Built-up Index (NDBI)

Proposed by Zha et al. in order to automate the process of mapping built-up areas; it takes advantage of the unique spectral response of built-up areas and other land covers (Zha et al. 2003).

$$\text{NDBI} = \frac{\text{TM5} - \text{TM4}}{\text{TM5} + \text{TM4}}$$

(3.4)

Used in conjunction with a traditional Normalized Difference Vegetation Index (NDVI), the NDBI was reported to be an effective technique to map urban built-up areas with the total accuracy 92% in the city of Nanjing, eastern China. In this area, a binary image was created with the hypothesis that the positive value of NDBI should be built-up areas and the positive value of NDVI should be vegetation. Such recoding process made the approach unable to separate urban areas from barren and bare land and suggested the approach's universality need to be tested in other geographic areas due to the complicated spectral response pattern of vegetation" (He and Xie 2007)

- Normalized Difference Blue Band Built-up index (NDBBBI)

Suitable for detecting built-up areas, a normalized difference index accounting for the blue-band component in built-up areas and barren land (NDBBBI) was defined by Baraldi et al in 2006 as:

$$\text{NDBBBI} = \frac{\text{TM1} - \text{TM5}}{\text{TM1} + \text{TM5} + 0.001}$$

(3.5)

- Normalized Difference Vegetation Index (NDVI)

One of the most frequently used vegetation index, NDVI relates the spectral absorption of chlorophyll in the red with a reflection phenomenon in the near infrared, influenced by the leaf structure type. It is defined as:

$$\text{NDVI} = \frac{\text{TM4} - \text{TM3}}{\text{TM4} + \text{TM3}}$$

(3.6)

Negative values corresponds to water, positive values near to zero corresponds to soils while higher values, between 0.2 and 0.6, are related to the presence of vegetation, till to 0.8 for very dense vegetation (Brivio et al. 2006).

- Built-up index (BUI)

Used in Lee et al. and defined as:

$$\text{BUI} = \text{NDBI} - \text{NDVI}$$

(3.7)

BUI takes the form of high values in the urban areas and low values in the non-urban areas

- Test_index

$$\text{test}_{\text{index}} = \text{TM5} + \frac{\text{TM4}}{\text{TM7}}$$

(3.8)

The proposed index takes into account that in urban areas (such as in bare soil), reflectance in band 4 is similar to reflectance in band 7 but urban areas have lower values than bare soil in band 5.

Employed attributes	Attributes	Spectral range/formula
Spectral information	Band 1	0.45 – 0.52 μm
	Band 2	0.52 – 0.60 μm
	Band 3	0.63 – 0.69 μm
	Band 4	0.76 – 0.90 μm
	Band 5	1.55 – 1.75 μm
	Band 7	2.08 – 2.35 μm
Indexes	NDBI	$\dfrac{TM5 - TM4}{TM5 + TM4}$
	NDBBBI	$\dfrac{TM1 - TM5}{TM1 + TM5 + 0.001}$
	NDVI	$\dfrac{TM4 - TM3}{TM4 + TM3}$
	BUI	NDBI - NDVI
	Test_index	$TM5 + \dfrac{TM4}{TM7}$

Table 13 – Attributes used for the classification process

A summary of tested attributes is provided in Table 13. Spatial information and indexes ere used both individually, both together.

3.2.4.2   Training of the classifier on multiple land cover types

Sample areas were detected on the reference image in order to belong to these classes:

- Urban (residential buildings)
- Industrial areas (industrial buildings)
- Vegetation
- Water
- Bare soil

Two options were tested during the creation of the training set: in one case input related to residential and industrial buildings were merged together in order to create a unique "urban" class, and all the other merged together in order to create a unique "non urban" class, in the second case all classes were considered separately. We refer to a 2-class classifier or to a 5-class classifier (Figure 17), thus resulting in classifiers with two or five leaf nodes.
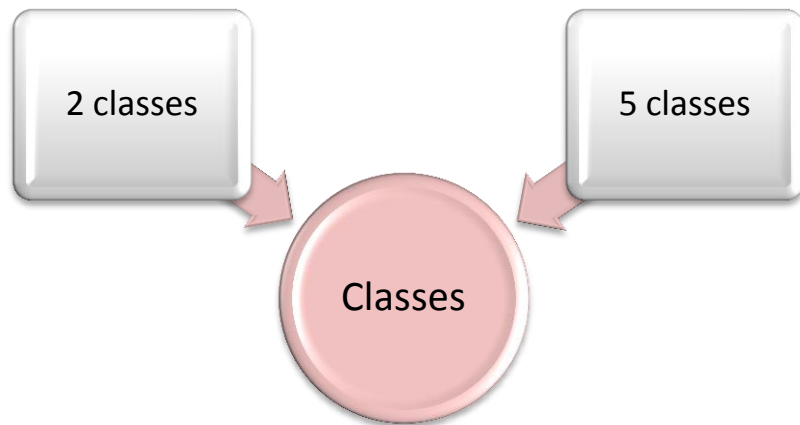
Figure 17 – Number of classes used to learn the classifier

This step allows assessing whether the use of a higher number of classes improves the classification accuracy compared to the use of two larger classes.

3.2.4.3   Training of the classifier on multi-temporal data

Sample areas detected on the reference image led to the creation of two training set: information were derived from a single scene in on one case, on the other information related to every single pixel were obtained as a mean over the chosen multitemporal stack (Figure 18).

The multitemporal stack is a set of images acquired on the same area in different periods of the year: a spectral signature obtained as a mean over the chosen multitemporal stack is more representative of possible modifications as the season changes, but can also create confusion in the learning phase: the aim was to evaluate if the use of a multitemporal stack could be a pro or a con for the classification process.

The conception was that certain kinds of land cover change more according to the season (e.g. vegetation, crops, etc), while other are more stable (e.g. urban, bare soil): the employment of a multitemporal stack could be able to detect this property.

Cloud presence or possible variations of land cover in the multitemporal set were evaluated case by case.

Figure 18 – Temporal extension of the training set

3.2.5 Training set generation

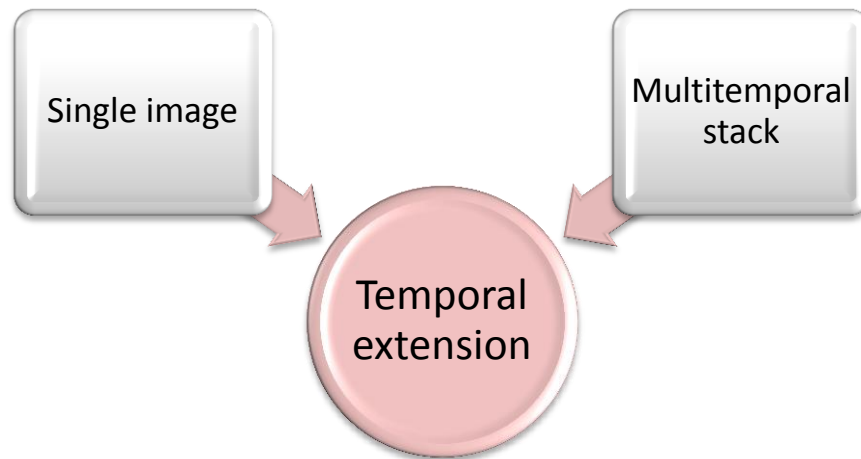Sample areas detected on the reference images have to be intended as geographic location of pixels where information (spectral or not) are extracted; in order to obtain training sets built with different kind of features, the step "Training set generation" was performed. Obtained training sets were later used as input to train the classifier. Since the software used to process the images (ENVI©, section 2.3) is different from the one used for the learning process (Weka, section 2.3), the procedure also takes into consideration the conversion in a format and structure usable in the following steps.

Figure 21 shows procedures used to obtain the desired training sets; every image (X_1 ... X_n) is first pre-processed in order to obtain a calibrated image, an atmospherically corrected with Dark Subtraction image and an atmospherically corrected with FLAASH image; then, indexes are computed from every image with the procedure *Compute_index.pro*; the procedures, *Single_image_stat.pro* and *Multitemp_image_stat.pro* were then used in order to obtain the 36 training sets as represented in Figure 13. These procedures combine image pre-processing (Section 3.2.3) with the different chosen variables, that is the use of different attributes (Section 3.2.4.1), the training of the classifier on two or more land cover types (Section 3.2.4.2) and the use of different temporal extension of the training set (Section 3.2.4.3).

Indexes computation

Indexes (NDBI, NDVI, NDBBBI, BUI, test_index, section 3.2.4.1) were computed from each pre-processed image with the procedure *Compute_index.pro* (Figure 21). The procedure gives in output an image with 5 bands, one for each index. In case of atmospherically corrected with FLAASH images, the input is first converted in a radiance value ranging from 0 to 1 (FLAASH

automatically scale the output with a factor 10000), then a check is performed on DN values in order to set value out of the range (0;1) equal to null (-999). Infact, if very low radiance values are present in the image, then FLAASH may not be able to model the reflectance accurately and some negative reflectance values can be found on the image; the same thing can also happen to very high radiance value that are converted in values greater than 1.

Statistics computation

In this step, spectral information and indexes values are extracted from the regions of interest, the sample areas. Information must be provided in a format compatible with Weka, whose native format is an ARFF (Attribute-Relation file format) file. The bulk of an ARFF file consists of a list of independent, unordered instances (no relationship among instances is involved), and the attribute values for each instance, separated by commas. The header of the file (example in Figure 19c) is composed by the name of the relation (TO_2005_spectral_DS_5class) and by a block defining the attributes (B1, B2, B3, B4, B5, B7, target). Nominal attributes are followed by the set of values they can take on, enclosed in curly braces (for example, in target attribute: {'urban ','industrial ','water ','vegetation ','bare soil '}). Numeric values are followed by the keyword numeric.
After the attribute definitions there is an @data line that signals the start of the instances in the dataset. Instances are written one per line, with values for each attribute in turn, separated by commas.

Nevertheless, CSV (comma-separated value) format is also accepted from Weka, thanks to the file format converters available in the software. The appropriate converter is used based on the extension and the .csv file is automatically converted in ARFF format.

Figure 19 provides an example of the export of a spreadsheet (a) in CSV format (b); it consists in a list of records with commas between items. This file, readable from Weka, can be used as it is and eventually also exported in ARFF format (c).

Single_image_stat.pro and *Multitemp_image_stat.pro* aim at creating the CSV files containing statistics of the sample areas. In the "single image" case (Figure 22, Figure 23, Figure 24), statistics are merely the values of the DNs, while in the "multitemporal image" case (Figure 25, Figure 26, Figure 27, Figure 28, Figure 29, Figure 30), DNs of the whole stack are considered and, for each DN, a mean between the values of that DN in all the images is computed. A "mean" image is also performed as an intermediate step.

(a)

B1, B2, B3, B4, B5, target

0.0791598,   0.118760,   0.124161,   0.157488,   0.183836,   0.154629,urban
0.0805990,   0.108997,   0.110691,   0.136829,   0.141449,   0.118370,urban
0.0877954,   0.115506,   0.122665,   0.133385,   0.157596,   0.141270,urban
0.0791598,   0.112251,   0.127155,   0.143715,   0.173744,   0.152720,urban
0.0834776,   0.120387,   0.140625,   0.167818,   0.214113,   0.188979,urban
0.0748420,   0.105743,   0.122665,   0.150601,   0.157596,   0.131728,urban

(b)

@relation TO_2005_spectral_DS_5class

@attribute B1 numeric
@attribute ' B2' numeric
@attribute ' B3' numeric
@attribute ' B4' numeric
@attribute ' B5' numeric
@attribute ' B7' numeric
@attribute ' target' {'urban ','industrial ','water ','vegetation ','bare soil '}

@data

0.07916,0.11876,0.124161,0.157488,0.183836,0.154629,'urban '
0.080599,0.108997,0.110691,0.136829,0.141449,0.11837,'urban '
0.087795,0.115506,0.122665,0.133385,0.157596,0.14127,'urban '
0.07916,0.112251,0.127155,0.143715,0.173744,0.15272,'urban '
0.083478,0.120387,0.140625,0.167818,0.214113,0.188979,'urban '
0.074842,0.105743,0.122665,0.150601,0.157596,0.131728,'urban '

(c)

Figure 19 – Input data in spreadsheet format (a), CSV format (b), ARFF format (c)

Land cover classes are automatically converted in *Urban* (urban and industrial) and *Non Urban* (water, vegetation, bare soil) classes in the 2-class version, or left as they are in the 5-class version. Detailed explanation of the different afforded cases is provided in Figure 22 to Figure 30; Table 14 provides a summary of the procedure used according to the case, while Figure 20 provides the legend used in the above mentioned procedures.

Procedure	Temporal extension of the training set	Attributes	Pre-processing
Figure 22a		Only spectral	Calib, DS
Figure 22b			FLAASH
Figure 23a	Single image	Only indexes	Calib, DS
Figure 23b			FLAASH
Figure 24a		Spectral and indexes	Calib, DS
Figure 24b			FLAASH
Figure 25		Only spectral	Calib, DS
Figure 26			FLAASH
Figure 27	Multitemporal stack	Only indexes	Calib, DS
Figure 28			FLAASH
Figure 29		Spectral and indexes	Calib, DS
Figure 30			FLAASH

Table 14 – Overview of the different applied processing

In general, workflows take into account that:

1. If information is extracted from images (both spectral or index) calibrated or corrected with Dark Subtraction, no intermediate "filtering" process should be performed before extracting the CSV file (Figure 22a, Figure 23a, Figure 25, Figure 27). This is verified also when information is extracted from both spectral and indexes images together (Figure 24a, Figure 29);
2. If information is extracted from spectral images corrected with FLAASH (Figure 22b), DN values must be scaled of a factor 10000 and a preliminary check performed on the DN value: if it is not appropriate (out of the range [0;1]), it must be removed from the output;
3. If information is extracted from the index images corrected with FLAASH (Figure 23b), not appropriate values are already set as null (-999): it's enough to remove them from the output;
4. If information is extracted from both spectral images and index images corrected with FLAASH (Figure 24b), step 1 and 2 are performed;
5. If information is extracted from a multitemporal stack, a "mean" image is produced as intermediate output (Figure 25 to Figure 30); the cell in the "mean" image is set null (-999) if almost one of the DN in the multi-temporal stack is not appropriate (out of the range (0:1) for spectral information with FLAASH (Figure 26), null for indexes with FLAASH (Figure 28), otherwise the mean is computed. When spectral information and indexes are used together, the above mentioned processing are both performed (Figure 24b, Figure 30).

TEXTUAL LEGEND

$X_1...X_n$: input images

S: images containing Spectral Information

I: images containing Indexes

a: calibrated images

b: atmospherically corrected with Darsk Subtraction images

c: atmospherically corrected with FLAASH images

GRAPHIC LEGEND

Image with spectral information calibrated or corrected with Dark subtraction

Image with spectral information corrected with FLAASH

Image with indexes calibrated or corrected with Dark subtraction

Image with indexes corrected with FLAASH

Output in CSV format (training sets)

Implemented procedures

Figure 20 – Textual and graphic legend used in Figure 21 to Figure 29
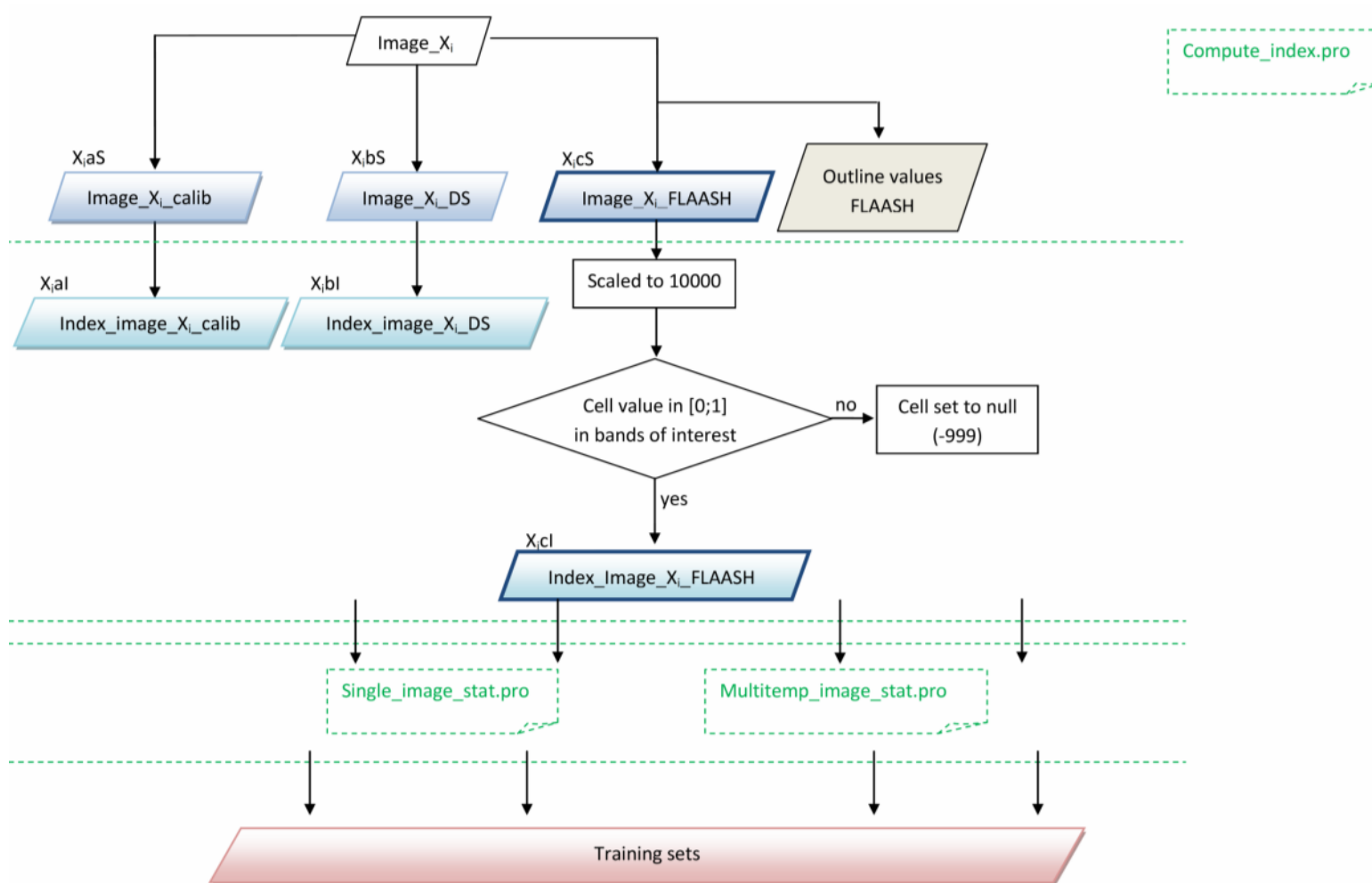
Figure 21 – Overview of the procedures used in order to create training sets to learn the classifier with detail on the methodology used for image pre-processing and indexes computation

Figure 22 – Methodology used from the procedure *Single_image_stat.pro* in order to create training sets with spectral information as attributes; detail on steps used for calibrated or atmospherically corrected with Dark Subtraction images (figure higher left) and for atmospherically corrected with FLAASH images (figure lower right) is provided.
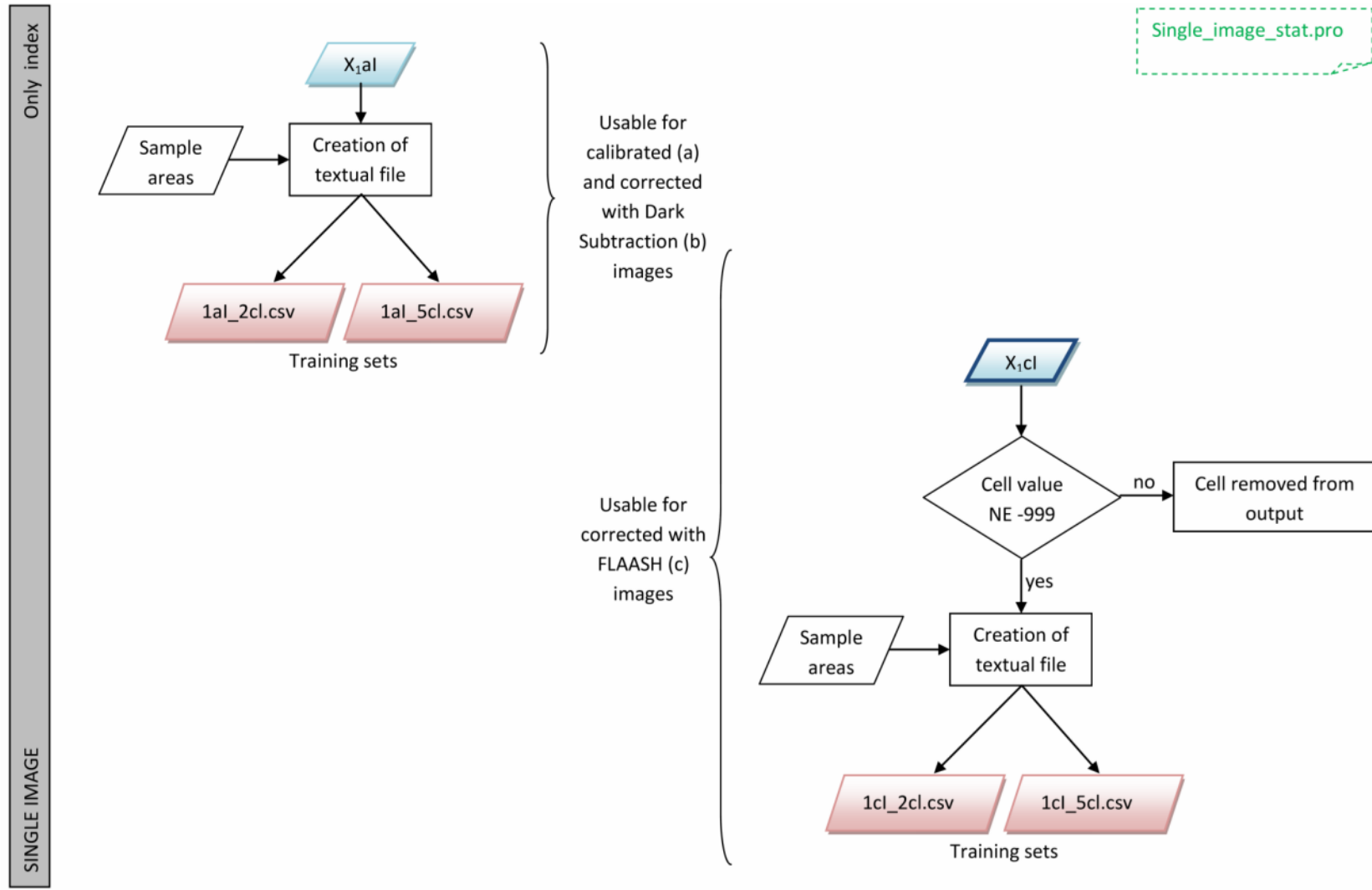
Figure 23 - Methodology used from the procedure *Single_image_stat.pro* in order to create training sets with indexes as attributes; detail on steps used for calibrated or atmospherically corrected with Dark Subtraction images (figure higher left) and for atmospherically corrected with FLAASH images (figure lower right) is provided.

Figure 24 - Methodologies used from the procedure *Single_image_stat.pro* in order to create training sets with spectral information and indexes as attributes; detail on *Single Image* procedure for calibrated or atmospherically corrected with Dark Subtraction images (figure higher left) and for atmospherically corrected with FLAASH images (figure lower right) is provided.

Figure 25 - Methodology used from the procedure *Multitemp_image_stat.pro* in order to create training sets with spectral information as attributes; detail on steps used for calibrated or atmospherically corrected with Dark Subtraction images is provided.

Figure 26 - Methodology used from the procedure *Multitemp_image_stat.pro* in order to create training sets with spectral information as attributes; detail on steps used for atmospherically corrected with FLAASH images is provided.
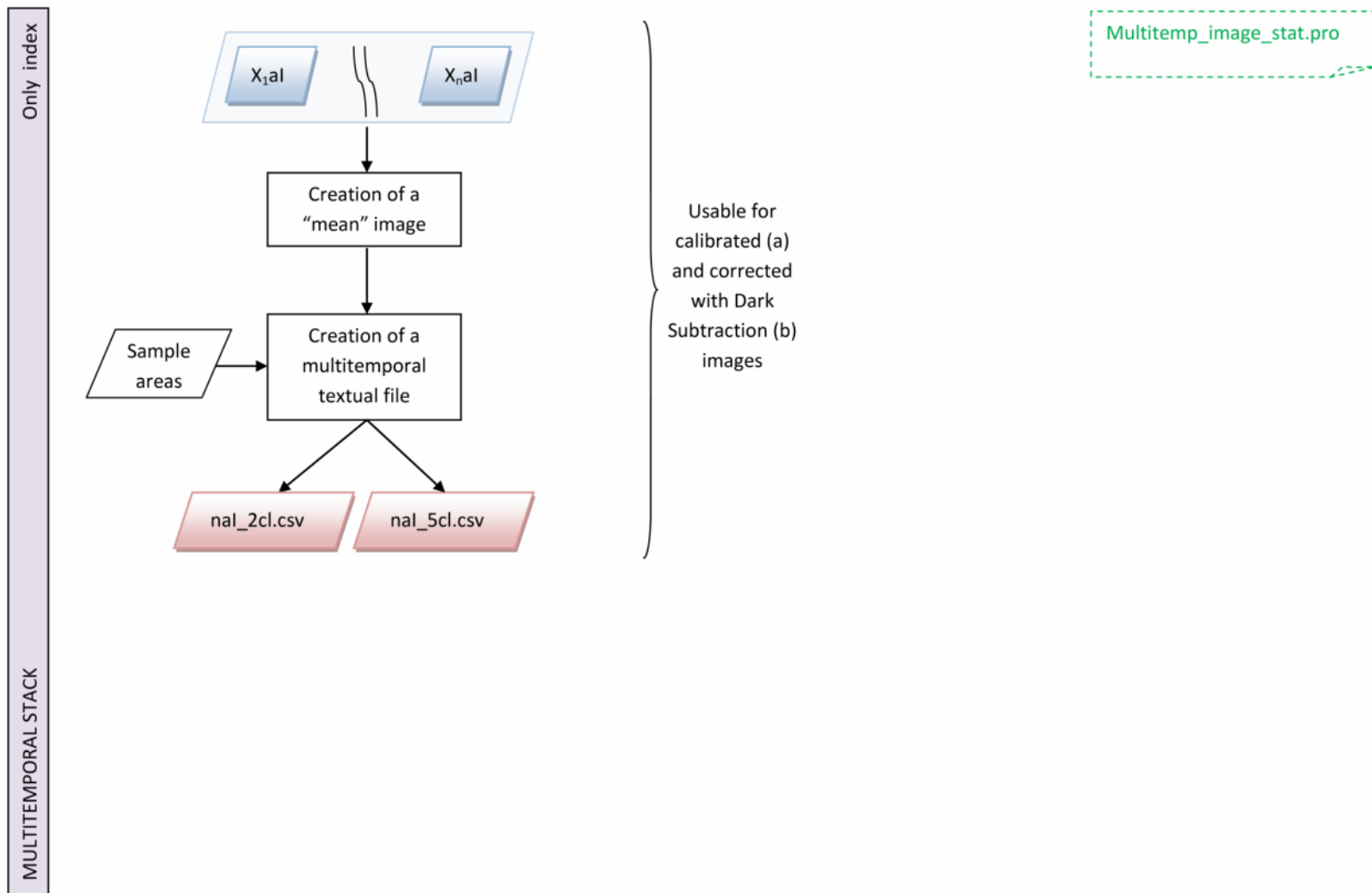
Figure 27 - Methodology used from the procedure *Multitemp_image_stat.pro* in order to create training sets with indexes as attributes; detail on steps used for calibrated or atmospherically corrected with Dark Subtraction images is provided.
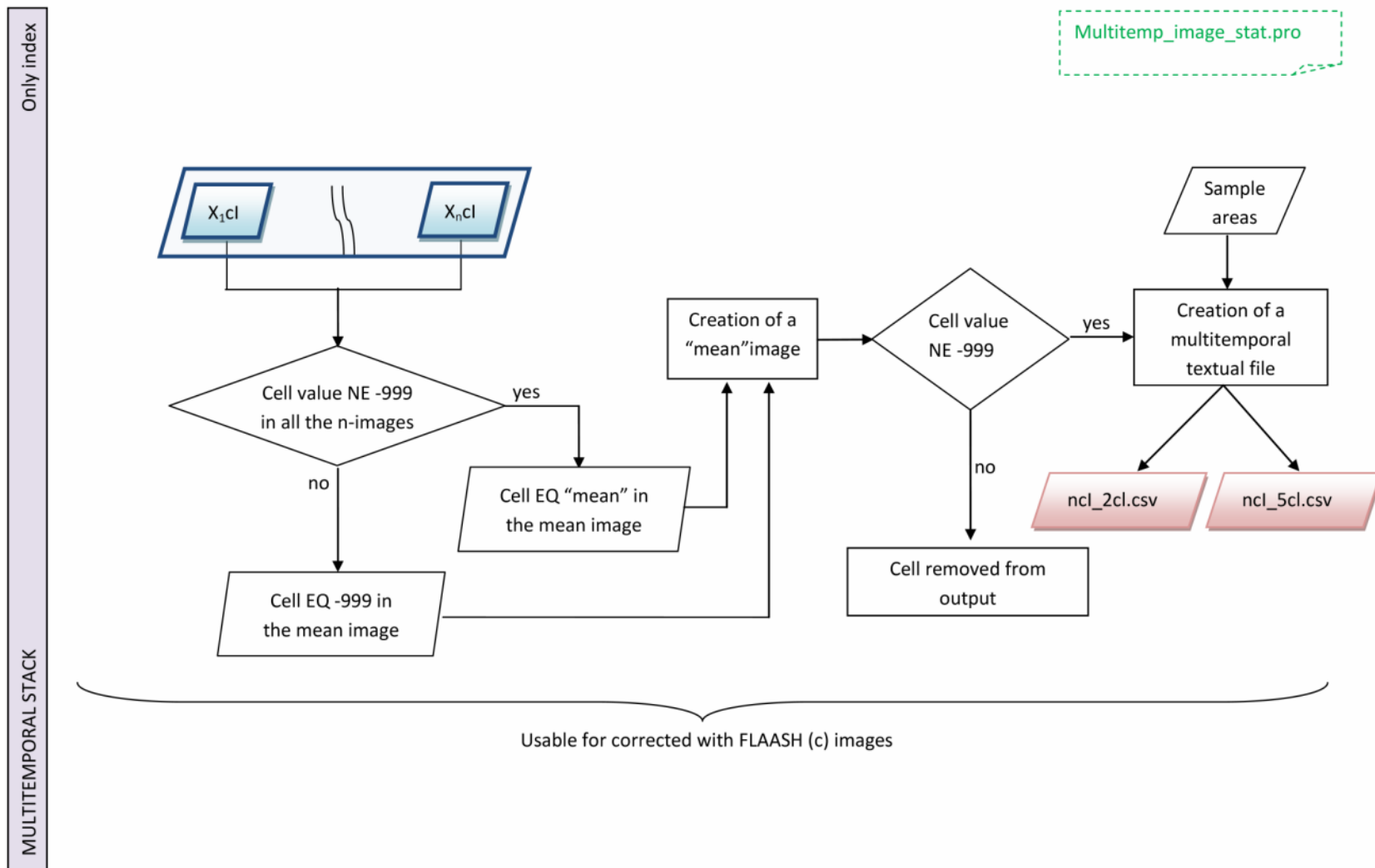
Figure 28 - Methodology used from the procedure *Multitemp_image_stat.pro* in order to create training sets with indexes as attributes; detail on methodologies used for atmospherically corrected with FLAASH images is provided.
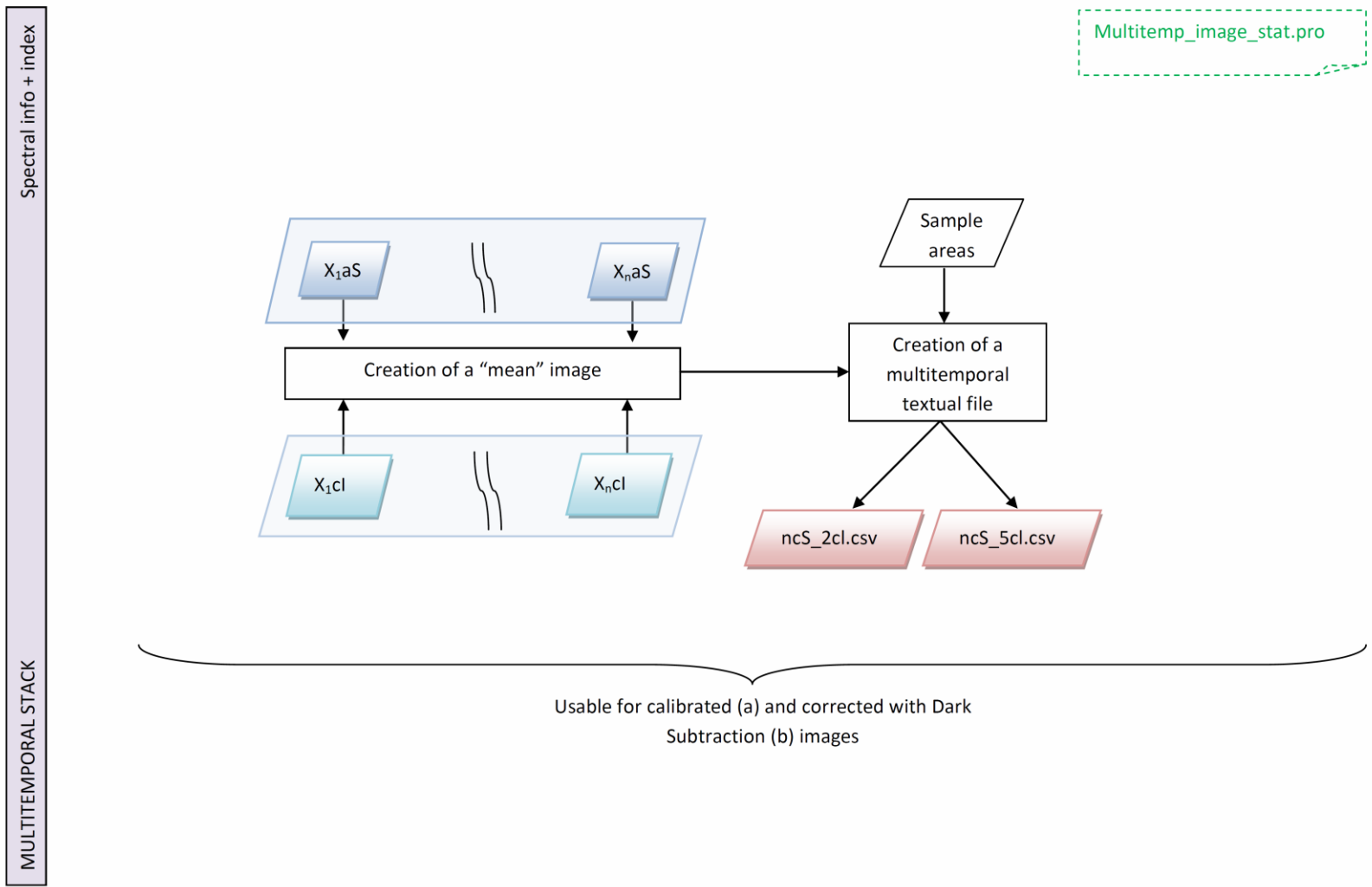
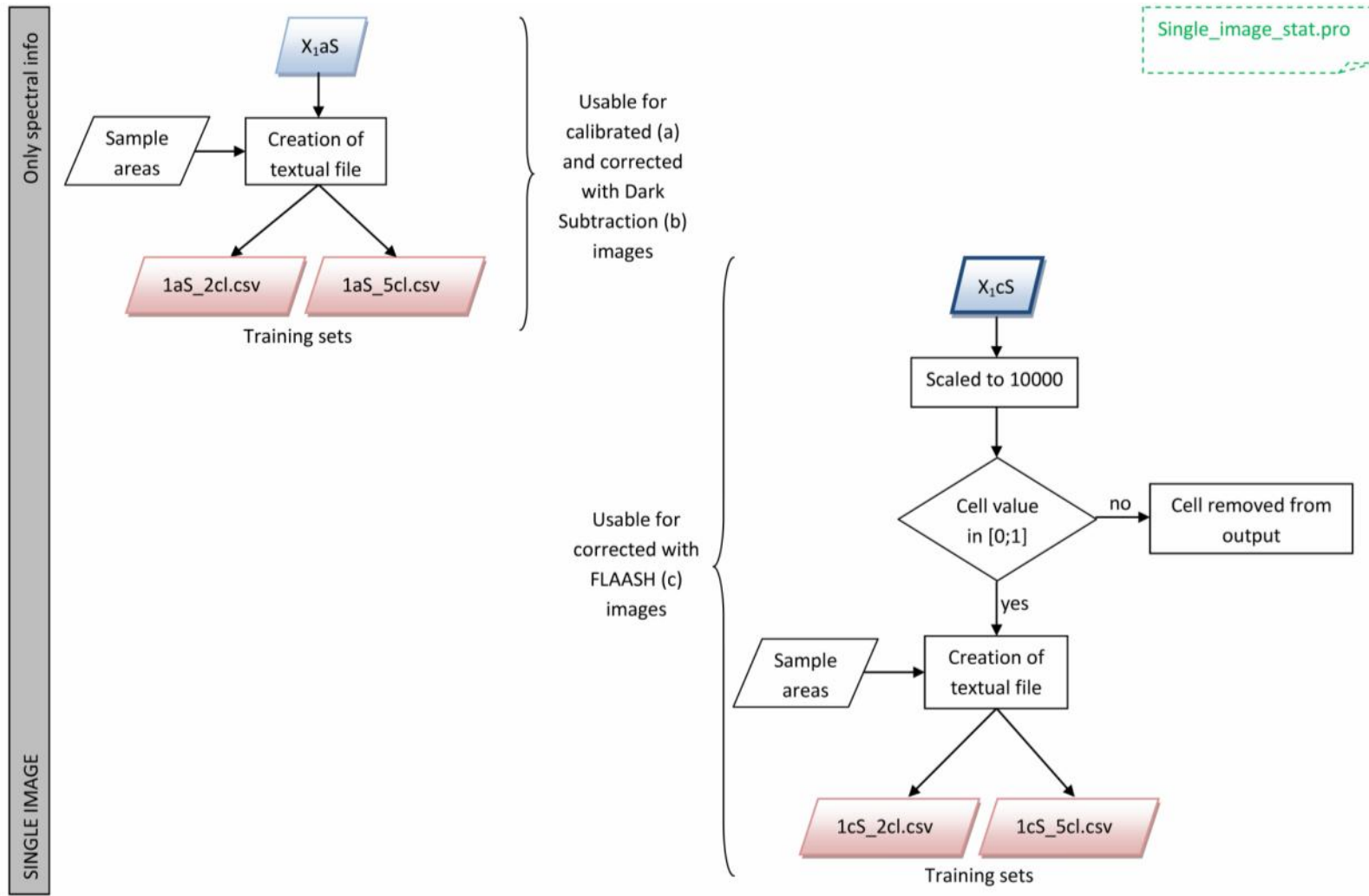**Figure 29 - Methodology used from the procedure *Multitemp_image_stat.pro* in order to create training sets with spectral information and indexes as attributes; detail on steps used for calibrated or atmospherically corrected with Dark Subtraction images is provided.**
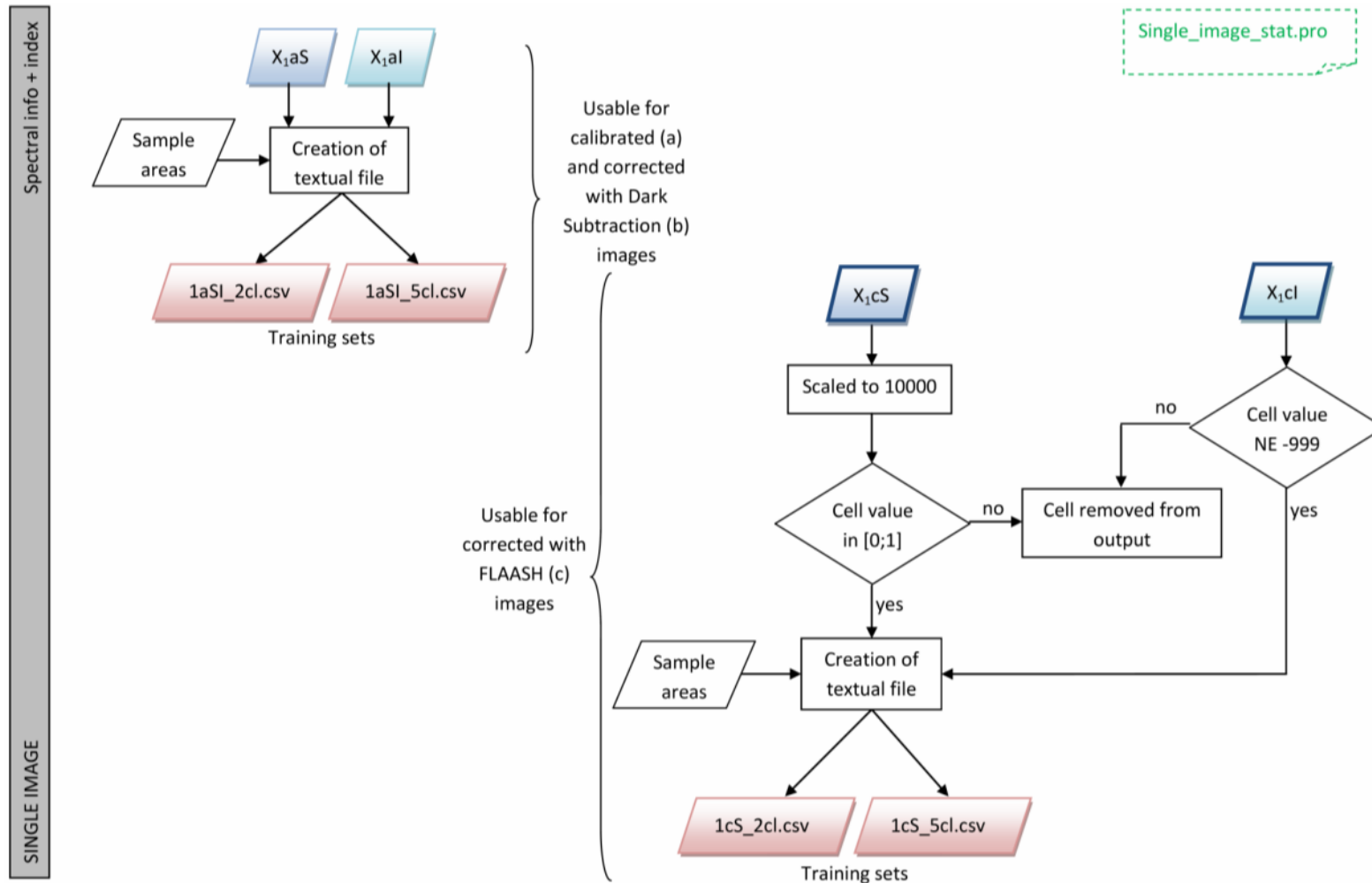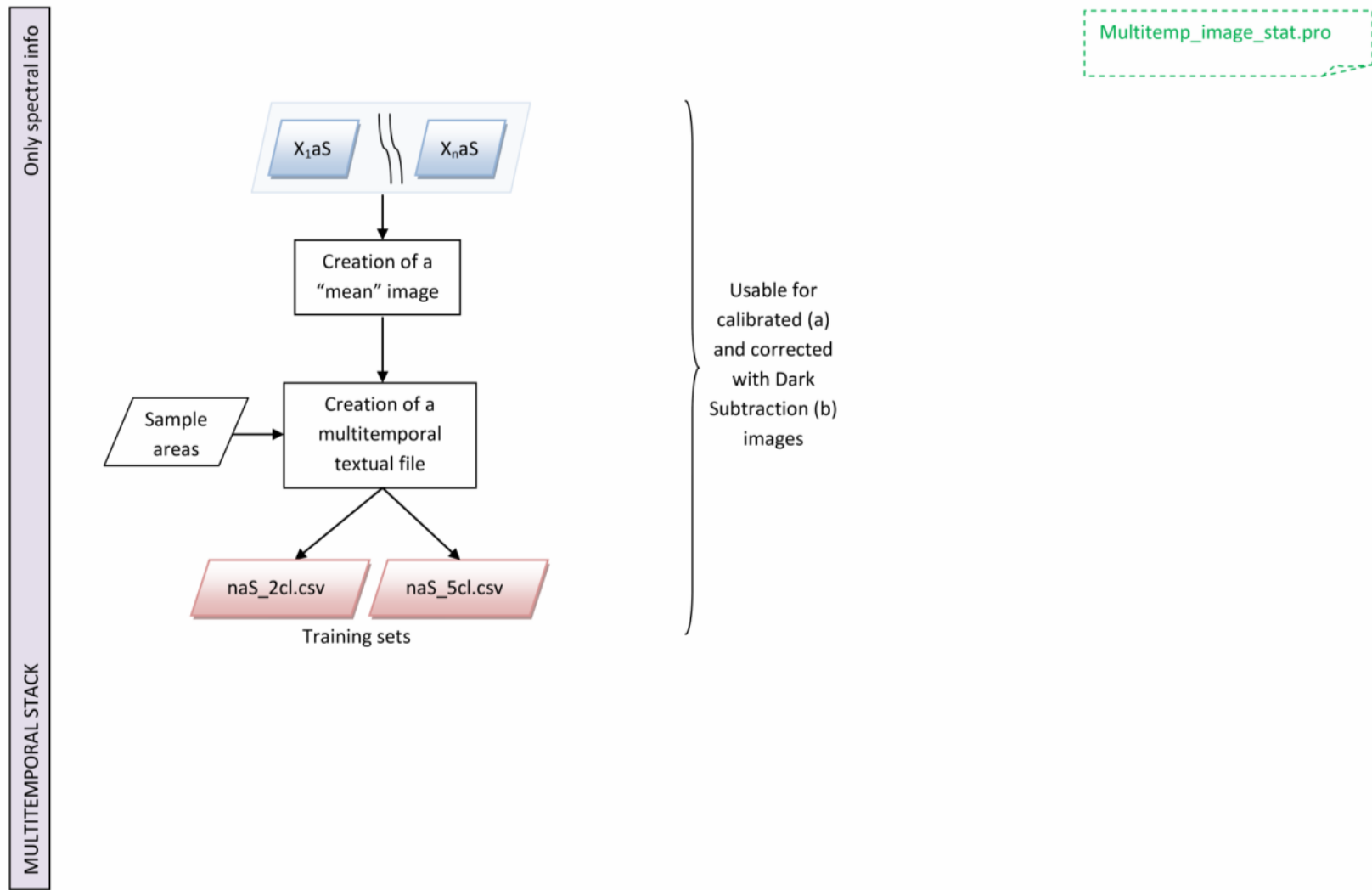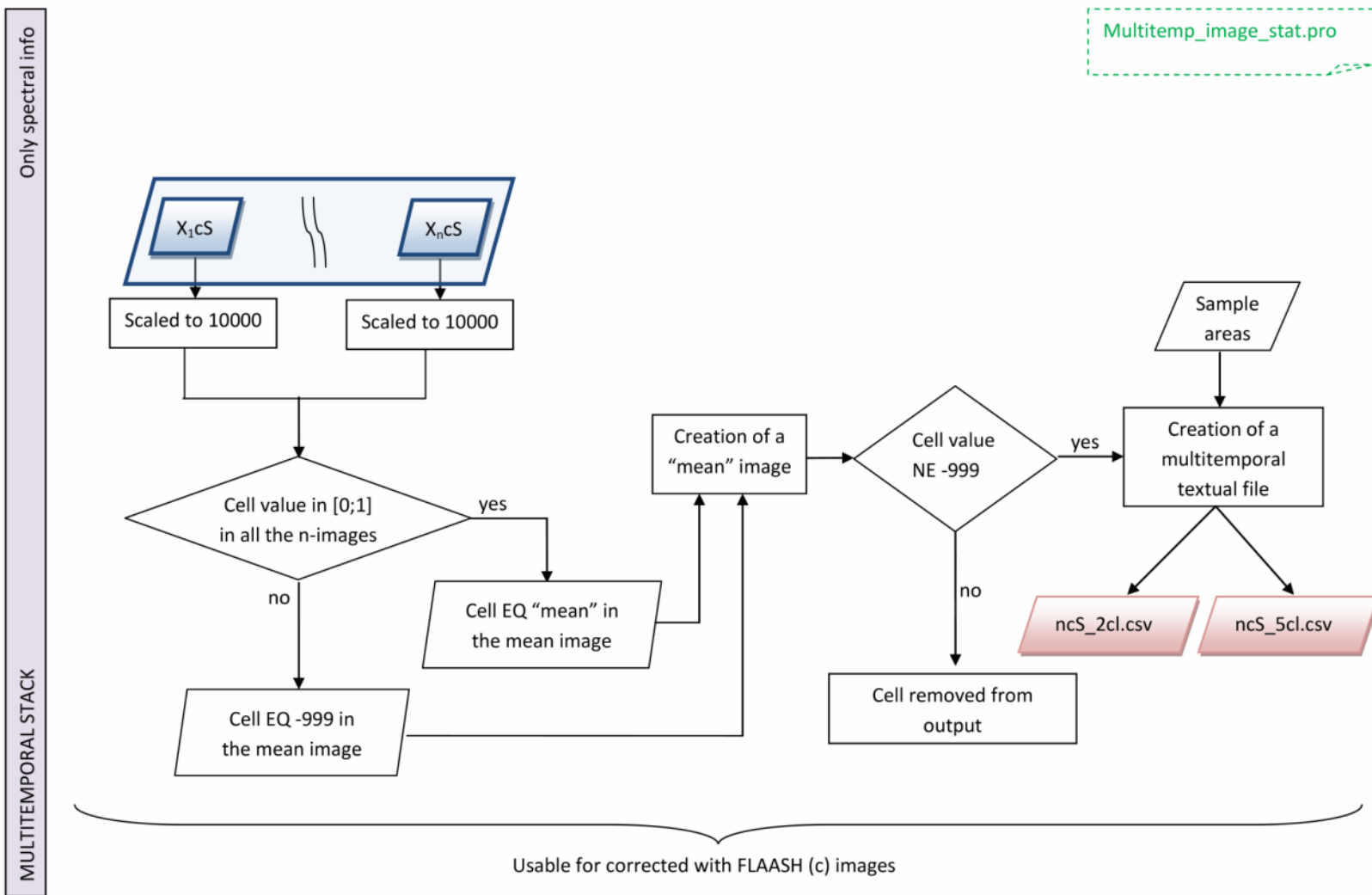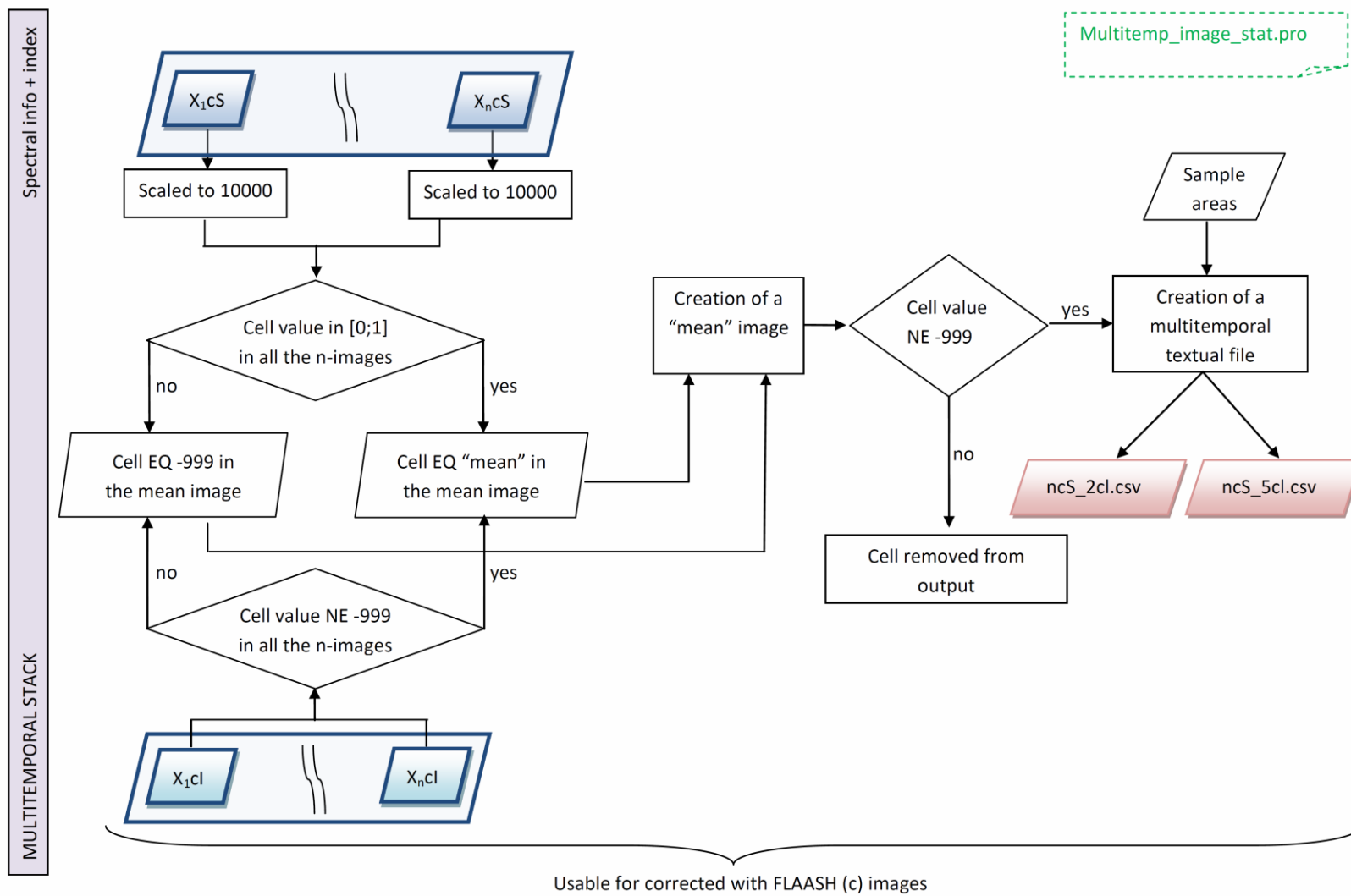
Figure 30 - Methodology used from the procedure *Multitemp_image_stat.pro* in order to create training sets with spectral information and indexes as attributes; detail on steps used for atmospherically corrected with FLAASH images is provided.

3.2.6 Classification

The classification process involved the following steps:

1) Each training set previously created was used as input to train the classifier;
2) For each training set a model was created using the J48 algorithm;
3) The generated model was converted in a procedure that associates the rules to the input image and that can be launched by ENVI software;
4) Procedures were launched and thematic maps created.

Since the first and the second steps had to be performed in the Weka environment, while the fourth had to be performed in ENVI environment, the third step was necessary in order to move from one environment to the other.

Figure 31 shows the methodology and the procedures used for the classification process; training sets (CSV files) generated in the previous steps were used as input files for the classification; classified images were the images considered in the "single image" cases or the reference images (i.e. the same image of the "single image" case) of the multitemporal cases.
 Procedures computed to automate this step are the following:
- Tree extraction from output weka.pro
- IDL codes generation.pro

Tree extraction from output weka

The output generated by j48 algorithm (example in Annex III) is a textual file composed by:

- Run Information; a list of information giving the learning scheme options, relation name, instances, attributes and test mode that were involved in the process
- Classifier model; a textual representation of the classification model that was produced on the full training data.
- Summary; a list of statistics summarizing how accurately the classifier was able to predict the true class of the instances under the chosen test mode.
- Detailed Accuracy By Class; a more detailed per-class break down of the classifier's prediction accuracy.
- Confusion Matrix; Shows how many instances have been assigned to each class. Elements show the number of test examples whose actual class is the row and whose predicted class is the column.

This procedure aimed at extracting only information about the model, that is rules and decision thresholds (i.e. the tree) detected by the algorithm, to be used in the following step.

IDL_codes_generation

The tree extracted from the Weka output was then converted in a IDL procedure. It allowed applying the detected rules to the input image in order to obtain a thematic map.
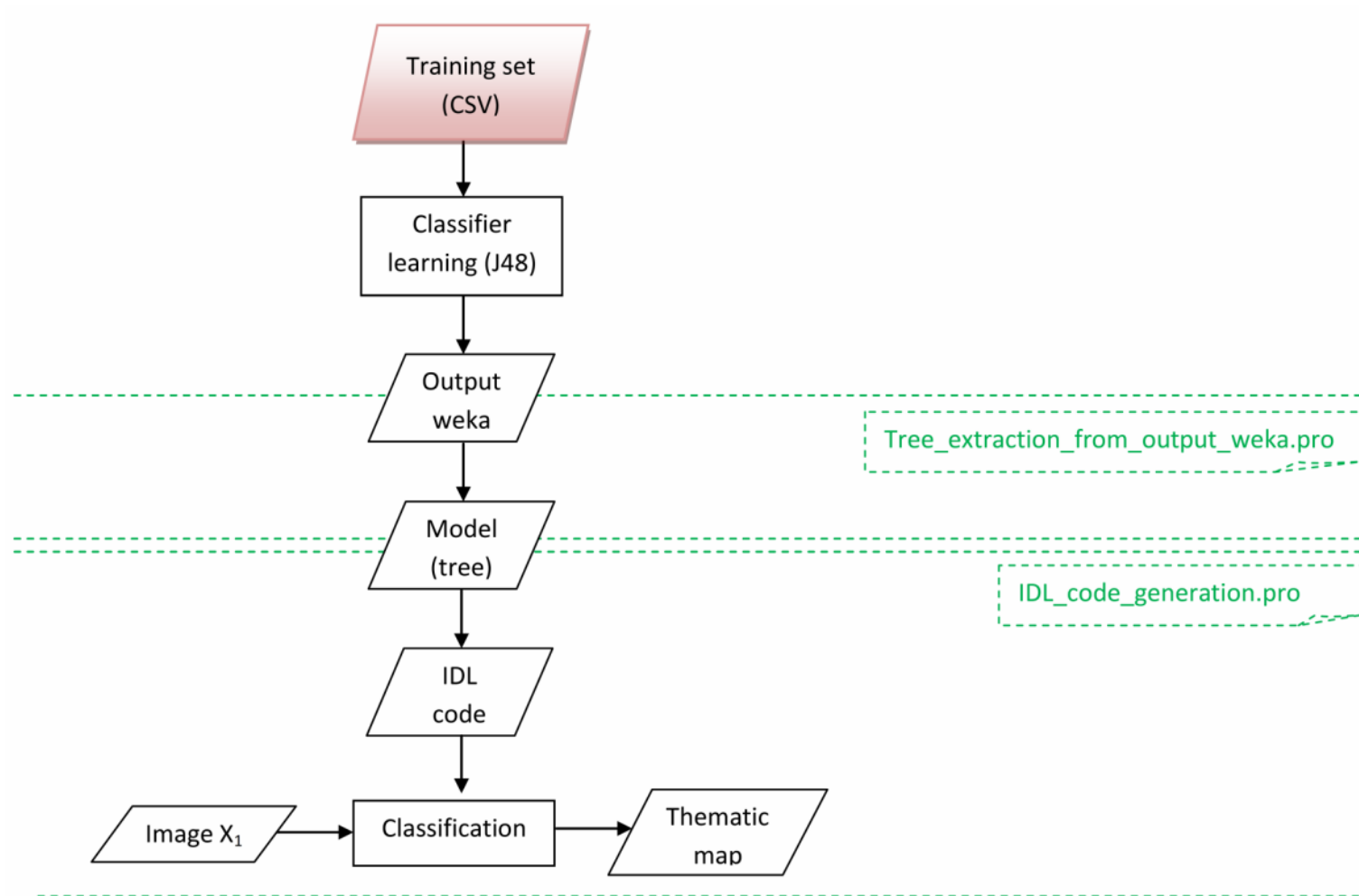
Figure 31 – Methodology and procedures used for the classification process

3.2.7   Post-processing

Post-classification processing is an important step in improving the quality of classifications (Harris and Ventura 1995, Murai and Omatu 1997, Stefanov et al. 2001, Lu and Weng 2004); filters can be applied to reduce the noise that come out from traditional per-pixel classifier ("salt and pepper" effect).

Essentially, two filter have were tested in this work:

- Sieve
- Majority analysis

Sieve

Sieve classes helps in solving the problem of isolated pixels occurring in classification images. It removes isolated classified pixels using blob grouping. The sieve classes method looks at the neighbouring 4 or 8 pixels to determine if a pixel is grouped with pixels of the same class. The four-neighbor region around a pixel consists of the two adjacent horizontal and two adjacent vertical neighbours, while the eight-neighbor region consists of all the immediately adjacent pixels. If the number of pixels in a class that is grouped is less than the value entered, those pixels are removed from the class.

Tests were performed with a group threshold of 4 pixels and 4 or 8 neighbours' pixels. The processing performed with 4 neighbour pixels has a more "filtering effects".

Pixels removed from the class were set to null; a procedure was applied in order to assign the null values to "*Non urban*" class.

Majority analysis

Majority analysis changes spurious pixels within a large class to that class. A kernel size is used in order to replace the center pixel in the kernel with the class value that the majority of the pixels in the kernel has.

The analysis can be applied only to certain classes: the cell value can be changed only if it owns to one of the selected classes, not otherwise. Nevertheless, if the unselected class is the majority class in the kernel, center pixels of selected classes can be changed into an unselected class.

Since the aim was to eliminate areas at low urban density that are difficult to be detected, the class *Urban* was selected; a majority filter with a kernel size 3x3 was applied.

Post-classification processings were applied on both classified images and validation mask; since performed test cases demonstrated a best efficiency of the majority analysis respect on the sieve, a procedure was created in order to post-process all thematic maps with the majority filtering (Figure 32).
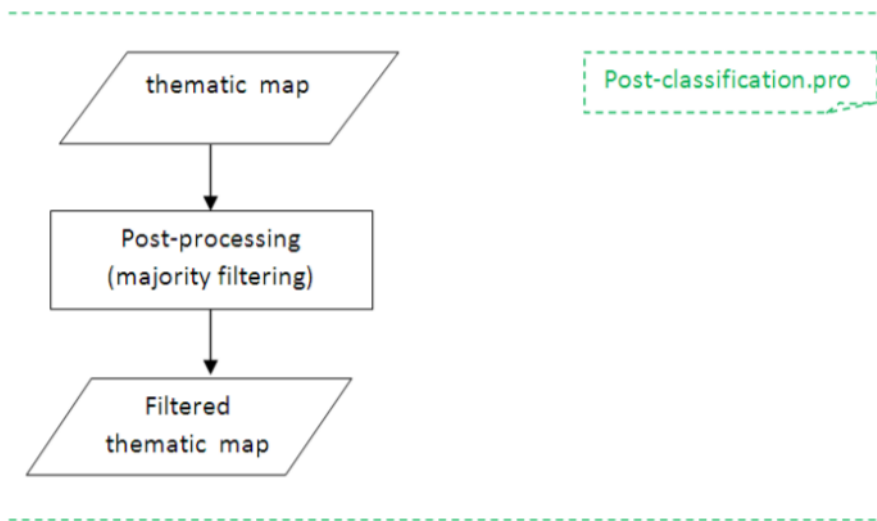
Figure 32 – Post classification (majority filtering) procedure applied

3.2.8   Accuracy assessment

The main objective of accuracy assessment is to derive a quantitative description of the accuracy of the map. This is a nontrivial task, and it is recognized that there is no one universal "best" method of accuracy assessment, but rather a suite of methods of varying value and applicability for any given map and purpose (Strahler et al. 2006). An overall measure of map accuracy (a single statement to provide an index of the general quality of the thematic map) and measures of accuracy on a per-class basis are desired.

The three basic components of an accuracy assessment are:

1) The *sampling Design* used to select the reference sample;

2) The *Response Design* used to obtain the reference land-cover classification for each sampling unit;

3) The *Estimation and Analysis Procedures*.

The *Sampling Design* is the protocol for selecting the locations at which the reference data are obtained. Implementing a probability sampling design contributes to a scientifically defensible accuracy assessment.
The definition of probability sampling focuses on inclusion probabilities, where an inclusion probability is defined as the probability that a particular pixel will be chosen for the sample. Probability sampling requires these inclusion probabilities to be known for all pixels selected in the sample, and nonzero for all pixels in the population (the entire region mapped). Many probability sampling designs have been developed; simple random, stratified random cluster and systematic sampling are all probability sampling design where inclusion probabilities do not have

to be computed explicitly because they are already taken into account in the standard estimation formulas.

The reference or "true" classification must be obtained for each sampling unit interpreting higher resolution data, using ground visits or other available ancillary data. The method used to determine this reference classification is called *Response Design*. Response design can be subdivided into two components: the *Evaluation Protocol*, which consists of procedures used to collect information contributing to the reference classification determination, and the *Labelling Protocol*, which assigns a land cover classification to the sampling unit based on the information obtained from the evaluation protocol.

The land cover classifications from the map are compared to the reference classifications, and the extent to which these two classifications agree is defined as map accuracy (Stehman and Czaplewski 1998).

The *Analysis and Estimation Protocol* applied to the reference sample data constitutes the third main component of an accuracy assessment. An error matrix effectively summarizes the key information obtained from the sampling and response designs providing a great wealth of information on a classification. It may be used to provide overall and per class summary metrics of land cover classification accuracy; moreover it is relatively easy to interpret and is familiar to both the map user and producer communities.

In this research work, no Sampling Design was performed due to the availability of ancillary data on the whole area of interest.

Concerning Response Design, since information to determine the reference classification were already available, no Evaluation Protocol was needed; conversely, the Labelling protocol was created in order to convert CTRN in a reference data suitable for validation. It's described in paragraph 3.2.8.1.

The confusion matrix and the derived measures of accuracy, thus resulting in the Analysis and Estimation Protocol, are instead depicted in paragraph 3.2.8.2.

3.2.8.1   Creation of the validation mask (Response Design)

The performed Response Design mainly focused on the definition of the Labelling Protocol, that is necessary in order to determine rules for assigning one or more reference classifications to each sampling unit. Reference data used for validation was Piedmont Region base cartography at scale 1:10.000 in vector format, described in section 2.2.3.

Some processing were needed in order to convert the cartography, in vector format and subdivided in a huge amount of different categories, into a raster with only elements of Urban/Non Urban. The following steps were considered:

1) Labelling protocol: attribution of CTRN categories to the correct class (Urban/Non Urban). This step was performed automatically, when possible, and manually, when it was not possible to attribute the whole category to a final class;

2) Conversion into raster format and re-sampling (Rasterization and resampling).

Labelling protocol

All areal features in group 1 were assigned to Urban class, apart from some exceptions (Codes: 2.01.25, 2.01.32, 2.01.35) where some processing where needed in order to correctly allocate features in the right class.

In particular (Table 15):

- Areas related to railways (code: 2.01.25, examples in Figure 33): features included in this sub-group were, in general, areas under the jurisdiction of railways authority (infrastructures, railways, storages); generally they were paved areas, but they could also be completely vegetated areas. Since only few features were coded in this group, they were manually checked and assigned to the right class. Only the feature represented in Figure 33a was assigned to Non Urban class since it represented a vegetated area;

- Sports facilities (code: 2.01.35): features included in this sub-groups could be soccer fields, tennis courts, stadium or indoor sports arenas; all features were manually checked and assigned to Non Urban class apart from indoor sports areas or stadium coverage;

- Not paved areas (code: 2.01.32, example in Figure 34): they represented one of the class with higher number of features; all areas not officially attributed to paved areas became part of this class. They could be both internal courtyards, both areas near to houses or vegetated areas (Figure 34). In order to correctly assign these features, a buffer of 10 meters was applied to buildings (houses, industries, churches, castles, etc); all not paved areas that fall in this buffer were considered as Urban (example in Figure 35).

GROUP 1: Streets and railways		
Assignment	**Name**	**Code (2.01....)**
Elements assigned to Urban class without any modification	Streets (paved and unpaved, tracks, under construction, intersections), traffic islands, manoeuvre areas	01, 02, 03, 04, 041, 05, 06, 07, 08, 42, 33
	Squares and parking	30
	Gas stations	41
	Paved areas	31
Elements assigned to Urban class after some processing	Areas related to railways*	25
	Not paved areas**	32
	Sport facilities*	35

Table 15 – Assignment of CTRN group 1 features to Urban and Non urban class

* Manually assigned at Urban or Not Urban class through visual interpretation

** Assigned to Urban class within a buffer of 10 meters from buildings
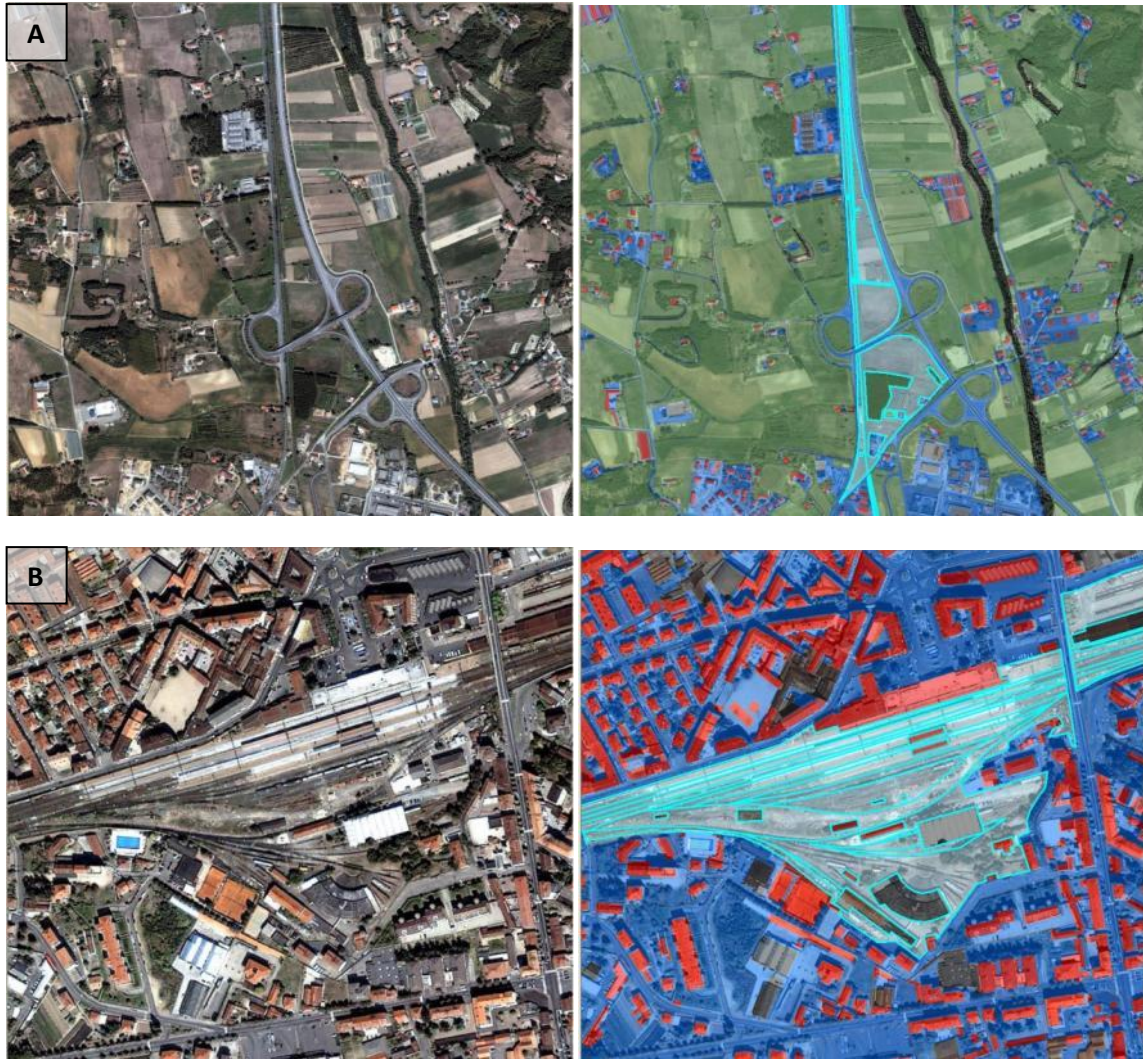
Figure 33 – Example of features coded 2.01.25 (areas related to railways): in one case (A) it's a vegetated area, in another (B) it contains buildings, railways, paved and unpaved areas.
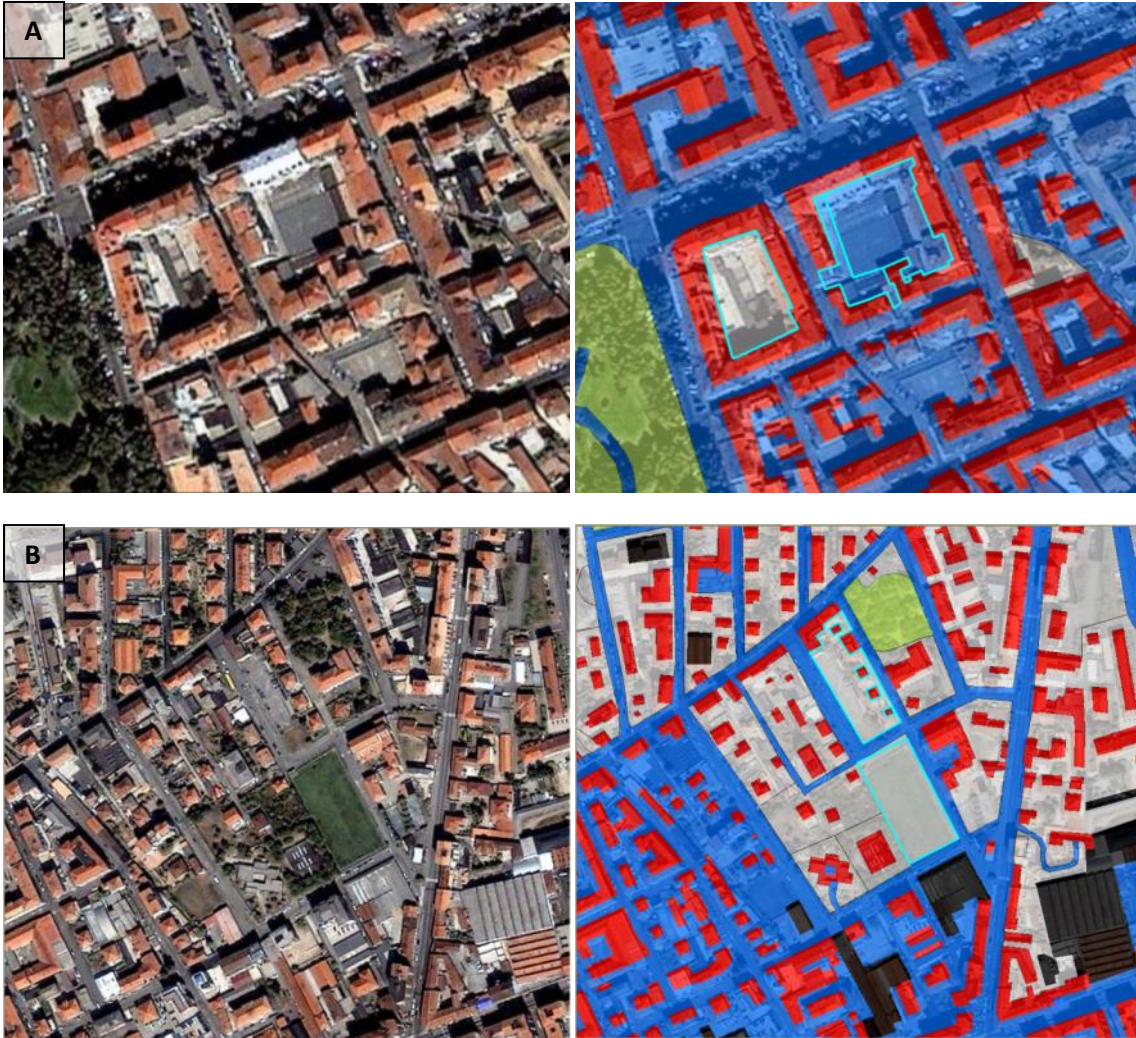
Figure 34 – Examples of a not paved areas (code: 2.01.32), in grey; in the image at the top (A) the feature is a courtyard, and it's not vegetated; in the lower one (B) one feature is a green area while the other is a non-vegetated area in front of buildings.

Figure 35 – Example of the buffer applied to buildings in order to assign Not Paved areas to urban class when close to buildings

Areal feature in group 2 were assigned to Urban class, apart from one exceptions: flights (codes: 2.02.22). They could either be paved or vegetated fights so they were manually checked in order to assign each feature to the right class.

GROUP 2: Buildings		
Assignment	**Name**	**Code (2.02....)**
Elements assigned to Urban class without any modification	Houses, industries, railway stations, religious buildings, historic buildings	01, 011, 012, 013, 014
	Distribution substation, transformer room, silos and tanks	015, 26, 027
	Sheds, greenhouses	07, 071
	Other	04, 06, 09
Elements assigned to Urban class after some processing	Flights*	21

Table 16 - Assignment of CTRN group 2 features to Urban and Non urban class

* Manually assigned at Urban or Not Urban class through visual interpretation

The sole areal elements in Group 5 (Table 17) were cemeteries; they were assigned to Urban class apart from some more vegetated ones (Sassi, Monumentale, Parco, all in Turin).

GROUP 5: Ground Elements		
Assignment	**Name**	**Code (2.05....)**
Elements assigned to Urban class after some processing	Cemetery*	06

Table 17 - Assignment of CTRN group 5 features to Urban and Non urban class

* Manually assigned at Urban or Not Urban class through visual interpretation

All not mentioned above CTRN features were included in Non Urban class; in general, Non Urban areas were:

- All features in group 3 (Water and handiworks related to water);
- All features in group 6 (Ground Elements);
- All features in group 7 (Vegetation);
- The sole areal feature in group 9 (not defined areas, code: 2.09.80);
- All features manually checked (as mentioned above);

A complete list of entities making part of CTRN, together with their codes, can be consulted in Annex II (CTRN Piemonte codes list).

Rasterization and resampling

The conversion of the new vector dataset into raster format and sub-sequent re-sampling was performed in one step, using the Conversion tool provided by ArcGis[©]; the conversion was based on the following parameters:

- The field used in order to assign values to the output raster; in this case the just created Urban or Non Urban value;
- The method to determine how the value has to be assigned to a cell when more than one features falls within a cell; a maximum combined area was chosen, in order to combine features that falls in a cell and to assign a value to the cell considering the combined feature with the largest area;
- The alignment of the resulting raster; in order to perfectly align resulting classification and validation mask, a snap was performed between the two datasets during the execution;
- Cell size of raster dataset; a 30 m spatial resolution was considered.

3.2.8.2 Measures of accuracy (Analysis and Estimation procedures)

A traditional procedure of verification of the classification accuracy is obtained from the confusion matrix analysis. The confusion matrix (or error matrix) represents a contingency table in which the diagonal entries represent correct classifications, or agreement between the map and reference data, and the off-diagonal entries represent misclassifications, or lack of agreement between the map and reference data. One benefit of a confusion matrix is that it is easy to see if the system is confusing two classes, i.e. commonly mislabelling one as another.

The use of the confusion matrix in accuracy assessment applications is based on a number of important assumptions. In particular, it is assumed that each pixel can be allocated to a single class in both the ground truth and the thematic map and that these two data sets have the same spatial resolution and are perfectly registered.

An example of confusion matrix for a two-class classifier is provided in Figure 23; it reports on the lines classes of the produced thematic map, and on the columns the ground truth classes. The diagonal elements in this matrix indicate numbers of sample for which the classification results agree with the reference data. Off diagonal elements in each row present the numbers of sample that were misclassified by the classifier,while the off-diagonal elements in each column are those samples being omitted by the classifier.

Ground truth		
	Class 1 (Urbano)	Class 2 (Non urbano)
Classification results — Class 1 (Urbano)	a_{11}	a_{12}
Class 2 (Non urbano)	a_{21}	a_{22}

Table 18 – Example of confusion matrix to assess the classification accuracy

Various summary measures can be derived from the error matrix to describe accuracy; population parameters which represent well-defined probabilities of either correct classifications or various misclassifications are here described.

- *Overall accuracy (OA)*: it is the overall proportion of correctly classified area and it's given by the sum of elements in the diagonal divided by the sum of all the elements. It represents the probability that a randomly selected point is classified correctly on the map. In a two-class classification, it can be expressed as:

$$OA = \frac{a_{11} + a_{22}}{a_{11} + a_{12} + a_{21} + a_{22}}$$

Overall Accuracy

Measures can also be derived for each class:

- User's accuracy (UA) for class *i*: it is defined by the ratio between pixel correctly classified as *i* and the total number of pixels that have been assigned to that class by the classification. It represents the probability that a randomly selected point classified as category *i* by the map, belongs to class i in the reference data. With reference to Table 18, it can be expressed as:

$$UA(class_1) = \frac{a_{11}}{a_{11} + a_{12}}$$

User's Accuracy

$$UA(class_2) = \frac{a_{22}}{a_{22} + a_{21}}$$

- Producer's Accuracy (PA) for class i: it is defined by the ratio between pixels correctly classified as i and the total number of pixels that belongs to class i in the reference. It represents the probability that a randomly selected point classified as category i by the reference data is classified as category i by the map. With reference to Table 18, it can be expressed as:

$$PA(class_1) \frac{a_{11}}{a_{11} + a_{21}}$$

Producer's Accuracy

$$PA(class_2) \frac{a_{22}}{a_{22} + a_{12}}$$

- *Commission Error (CE)* for class i: it's defined by the ratio between pixels that have been wrongly classified as i by the classifier and the total number of pixels that have been assigned to that class.

It represents the probability that a randomly selected pixel classified as i, belongs to another class in the ground truth. With reference to Table 18, it can be expressed as:

$$CE(class_1) = 1 - UA = \frac{a_{12}}{a_{11} + a_{12}}$$

Commission Error

$$CE(class_2) = 1 - UA = \frac{a_{21}}{a_{22} + a_{21}}$$

- *Omission Error (OE)* for class i, it's defined by the ratio between pixels that belongs to class i in the reference but have been wrongly classified in the map. It represents the probability that a randomly selected pixel in class i in the reference, has been assigned to another class in the map. With reference to Table 18, it can be expressed as:

$$OE(class_1) = 1 - PA = \frac{a_{21}}{a_{11} + a_{21}}$$

Omission Error

$$CE(class_2) = 1 - PA = \frac{a_{12}}{a_{22} + a_{12}}$$

Often used in the field of Information Retrieval and Machine Learning, rarely in Remote Sensing, is the *f-measure*, defined as a harmonic mean of Precision and Recall; Precision (P) and Recall (or True Positive Rate, TP) measures are only related to the considered "positive" class, while other measures are related to the "negative" class (false positive rate, true negative rate, false negative rate). Since we were interested in classifying one specific land type (urban), the urban class was considered the "positive" class, and the concept of f-measure was thus adapted to our analysis in this way:

$$f - measure = \frac{2 * P * TP}{P + TP} = \frac{2 * PA(cl_1) * UA(cl_1)}{PA(cl_1) + UA(cl_1)}$$

This provides a combined measure of Producer's and User's Accuracy.

The procedure *confusion_matrix.pro* (Figure 36) was created in order to carry out all the steps of the Accuracy Assessment; the procedure combine classes if the input is a 5-class classification (class are assigned to Urban or Non Urban class) and extracts measures of accuracy comparing the classification and the validation mask. The procedure was applied to both post-processed and not post-processed images: in the case where post-processed images where used, also the validation mask was post-processed in the same way (Figure 37).
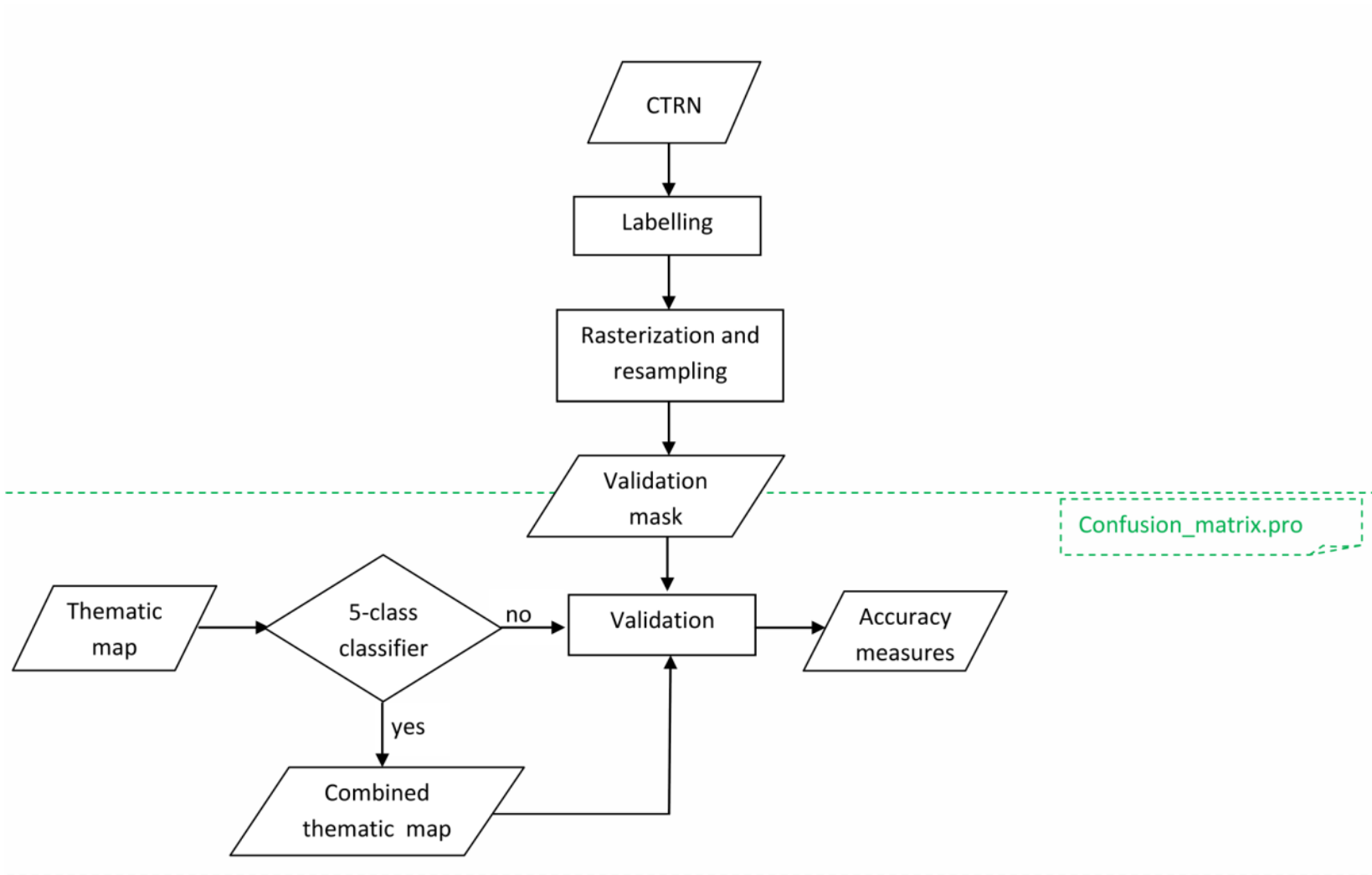
Figure 36- Methodology and procedure used to perform the Accuracy Assessment on not post-processed images
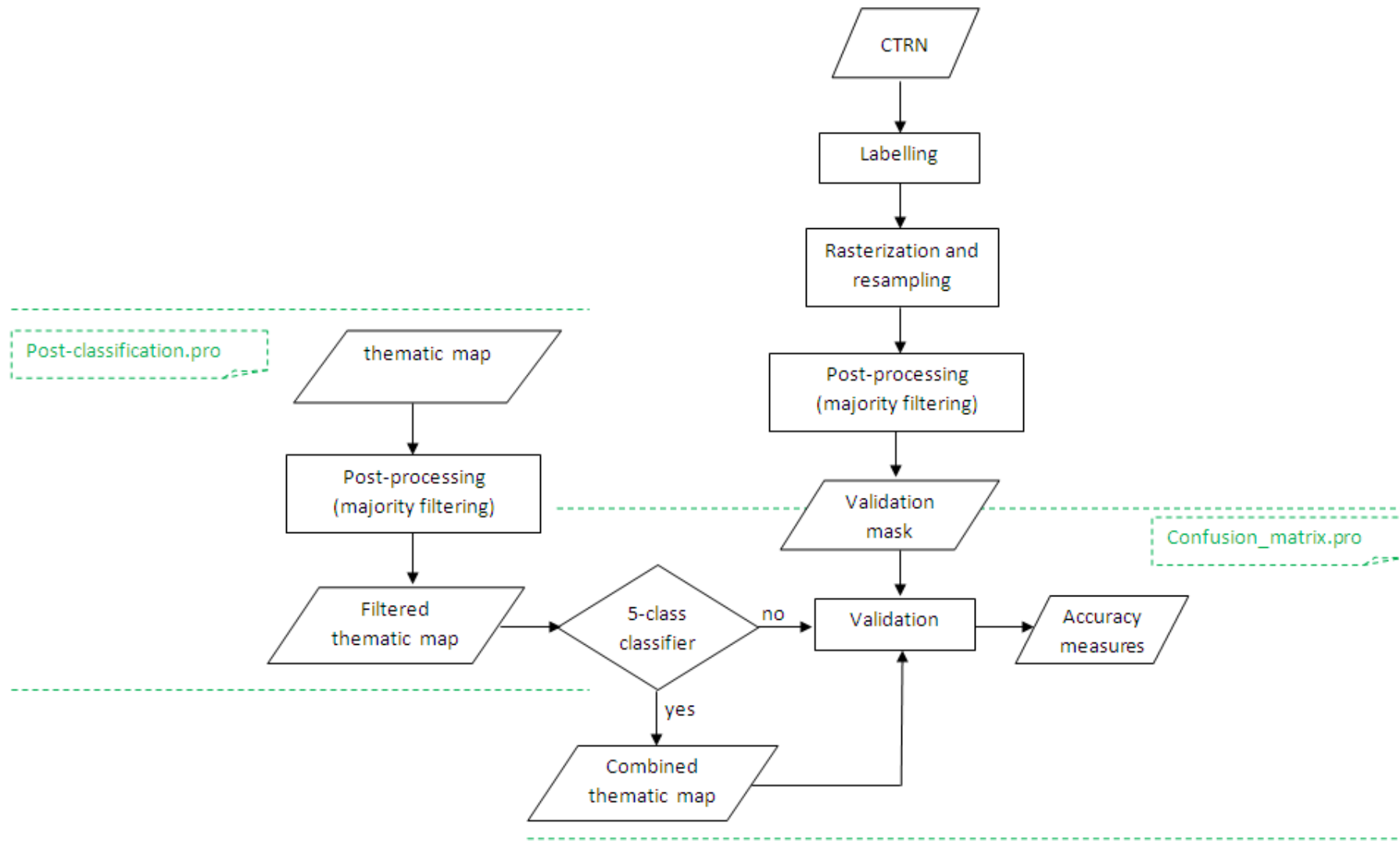
Figure 37 - Methodology and procedures used to perform the Accuracy Assessment on post-processed images

3.3 Data collection and analysis

3.3.1 Study area

The Piedmont region, located in the north-west part of Italy, was chosen as study area.

The proposed methodology, described in section 3.2, was applied to two different geographic areas, one including Torino administrative center, the other including Asti, one of the Region district. We thus refer to:

- Torino case study
- Asti case study

Data used for the training phase are all part of Landsat ETM+ slc-off collection, described in section 2.2.1.2; data used for the classification process are instead part of GLS 2005 collection, described in section 2.2.2; data for validation is the CTRN, described in section 2.2.3, with different update according to the analyzed area.

All data used in the case study are shown in Table 19.

Case study	Data for training set (date and path/row)	Reference data	Data for validation
Torino	06.10.1999 (194/29)	GLS 2005 date: 02.07.2005 path/row: 194/29	CTRN 2005
	24.08.2001 (194/29)		
	28.09.2002 (194/29)		
	30.11.1999 (195/29)		
	05.06.2000 (195/29)		
	30.07.2001 (195/29)		
	02.07.2005 (194/29)		
Asti	06.10.1999 (194/29)	GLS 2005 date: 02.07.2005 path/row: 194/29	CTRN 2004
	24.08.2001 (194/29)		
	28.09.2002 (194/29)		
	23.11.1999 (194/29)		
	01.05.2000 (194/29)		
	21.06.2001 (194/29)		
	02.07.2005 (194/29)		

Table 19 – Data used for the analysis

The dimension of the investigated area was influenced by the availability of the CTRN; 5 section of the CTRN were investigated in Torino case study, 61 in Asti case study, for a total of about 270 km^2 in Torino case study and 2.250 km^2 in Asti case study.

A representation of the investigated area is provided in Figure 38.
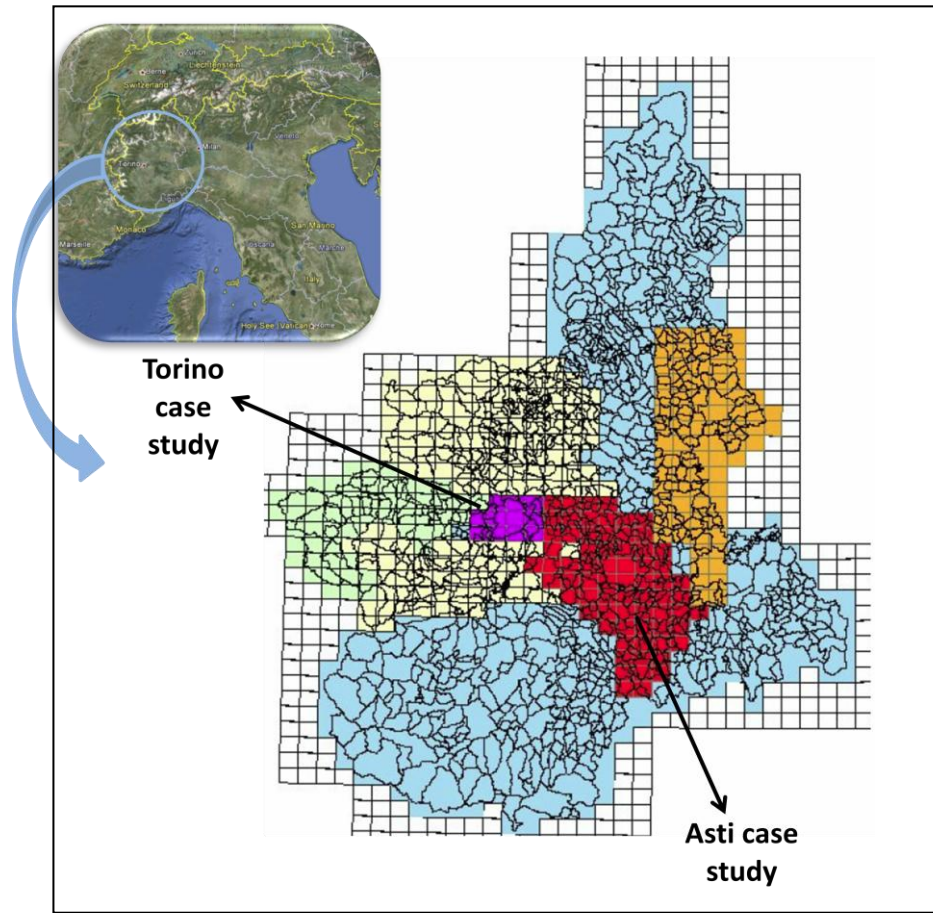
Figure 38 – Investigated area

4   Results

4.1   Introduction

Results are expressed in terms of accuracies reachable by each classifier; accuracies, obtained by comparing the resulting classifications with the validation mask, are expressed in terms of Overall Accuracy (OA), Producer's and User's Accuracy (PA and UA) and f-measure (reference to the Accuracy Assessment described in section 3.2.8).

In this chapter, a general overview about classifiers performances is initially provided (section 4.2); a summary of best classifiers and rules that constitute them is herein included. A qualitative analysis is then given in order to evaluate the spatial distribution of errors (section 4.3). The different testes variables are then evaluated (section 4.4): pre and post processing applied to images, as well other features, that are attributes, training of the classifier on multiple land cover types, use of a multitemporal stack.

4.2   General overview

The methodology was tested on two different cases, as described in section 3.3.1:

- Torino case study;
- Asti case study.

The validation performed on Torino case study made use of 5 CTRN sections, and was centered on the area of Torino city and suburbs. The choice of the area to validate was uniquely due to the update of CTRN: these 5 sections were the only ones with an update suitable for our aims. Since Asti region was entirely update in 2004, all available CTRN sections were used (61); it resulted a validated area with a medium large urban area (Asti town, about ten times smaller than Torino in terms of inhabitants), many small towns, and endless sparse houses. Since in this case the urban density is very different from the other case, another subset was considered for the validation, in order to recreate a context more similar to the previous one; although the validation on the subset made use of 4 CTRN sections, that was more or less the same used for Torino, it must be taken into consideration that, again, the situation is very different: in the first case the analysis was centered in the metropolitan area of Torino and near surrounding, in the second case the analysis was centered on the town, the surrounding and, again, a huge part of countryside, with a different level of urban density.

Results (Table 20) show that best classifiers in Torino allowed reaching Overall Accuracy round 81%, while in Asti case study Overall Accuracy round 96% - 97% were reached.

Overall Accuracy provides information about the whole classification, and it is representative of how all classes were classified; nevertheless, it happens that, if a class is badly classified but just few elements own to that class, this doesn't stand out. If information about a single class are needed, it must refer to class accuracy measures, such as Producer's and User's Accuracy. Since

these two measures can have different trends, f-measure is a good way in order to have univocal information about reached accuracies. In Torino case study f-measure of 77% were reached, against just 48% in Asti case study. With the decreasing of the dimension of the validation subset, f-measure increased till to 66%; this is representative of the problem in detecting sparse urban. Results shown in Table 20 are related to post-processed classifications and are representative of best reached accuracies.

	TORINO	ASTI	ASTI subset
N° validated pixel (km^2)	300.000 (270 km^2)	2.500.000 (2.250 km^2)	200.000 (180 km^2)
N° used CTRN sections	5	61	4
Best *OA*	81%	97%	96%
Best *f-measure*	77 %	48%	66%

Table 20 – Results: general overview

Table from 23 to 28 shows results obtained from each classifier, for classification with and without post processing. Table 23 and Table 24 are related to Torino case study, while Table 25, Table 26,

Table 27 and Table 28 are related to Asti case study, with or without post-processing, with the validation performed on the whole area or only on a subset. In the tables grey cells represent the three best f-measures and the three best OA; the green square underlines the best classifiers that are the ones that provided both better f-measure and OA. Results in bold types are instead related to procedure described in section 4.3; the orange squares underlines classifiers that provided good results as described in section 4.3.

A summary of best classifiers, subdivided according to the case study, is provided in Table 21 (pp stands for post-processed classifications):

Case study / Classifier	TO	TO – pp	AT	AT - pp	AT subset	AT subset - pp
1_index_DS_2cl	X					
1_index_DS_5cl	X	X				
1_spectral_index_FLAASH_5cl			x		x	
1_ index_FLAASH_2cl				x		
1_ spectral_index_FLAASH_2cl				x		
1_ spectral_index_DS_2cl					x	x

Table 21 – Summary of best classifiers, shared according to the case study

These are classifiers that provided, at the same time, both best f-measure and OAs.

It' noticeable how all best classifiers make use of indexes, both alone or together with spectral information; moreover, all considered classifiers exploit information from just one image and are obtained from an atmospherically corrected image.

The whole trees are shown in Figure 39 to Figure 44; rules for the detection of built-up areas extracted from the following trees are reported in Table 22.
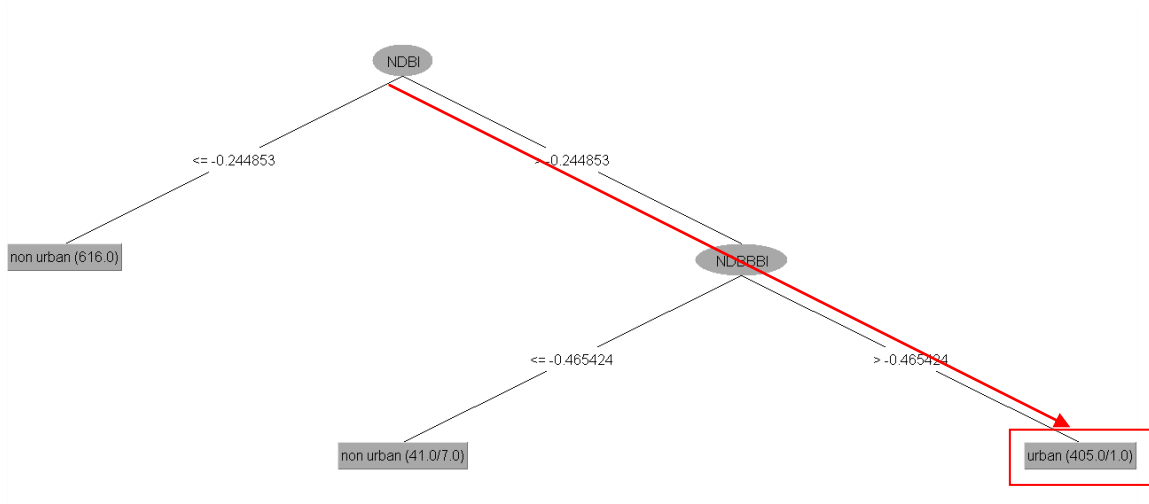
Figure 39 – Classifier "1 index DS 2class"
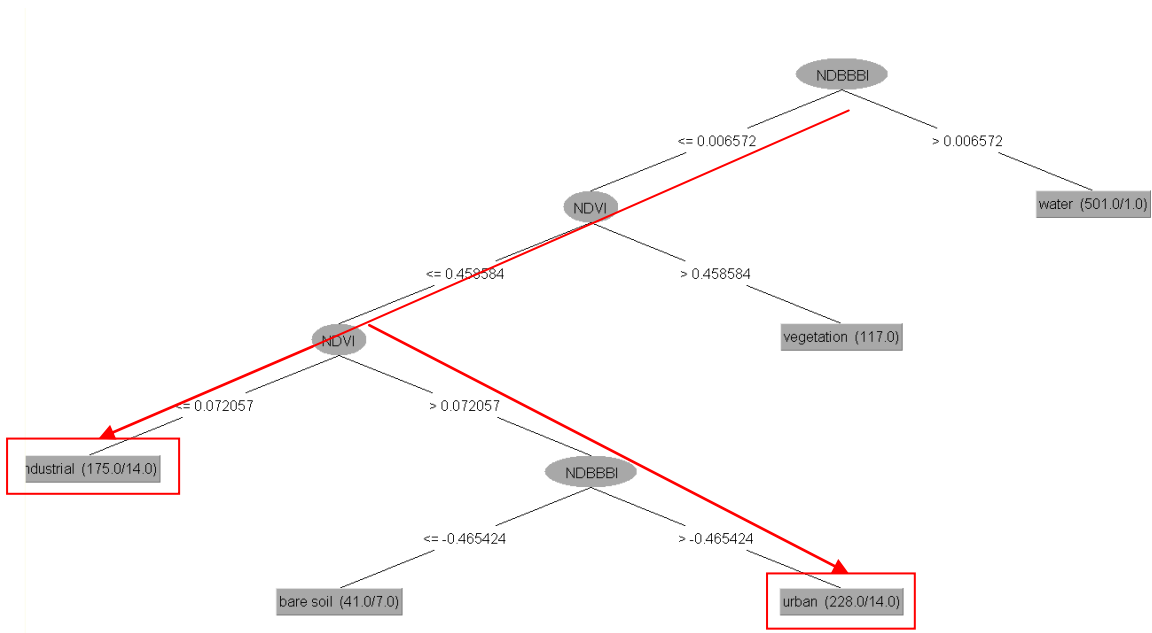
Figure 40 – Classifier "1 index DS 5class"
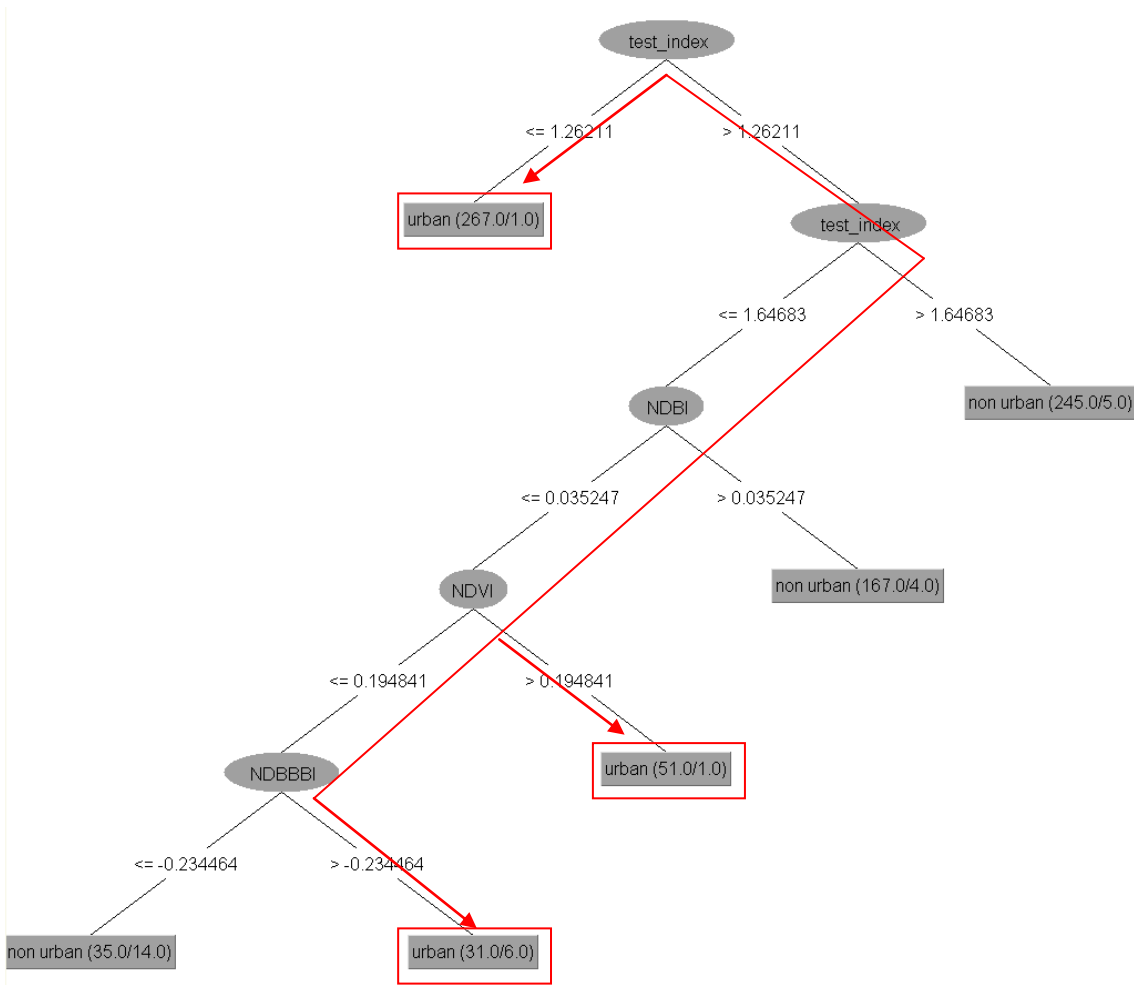
Figure 41 – Classifier "1 index flash 2class"
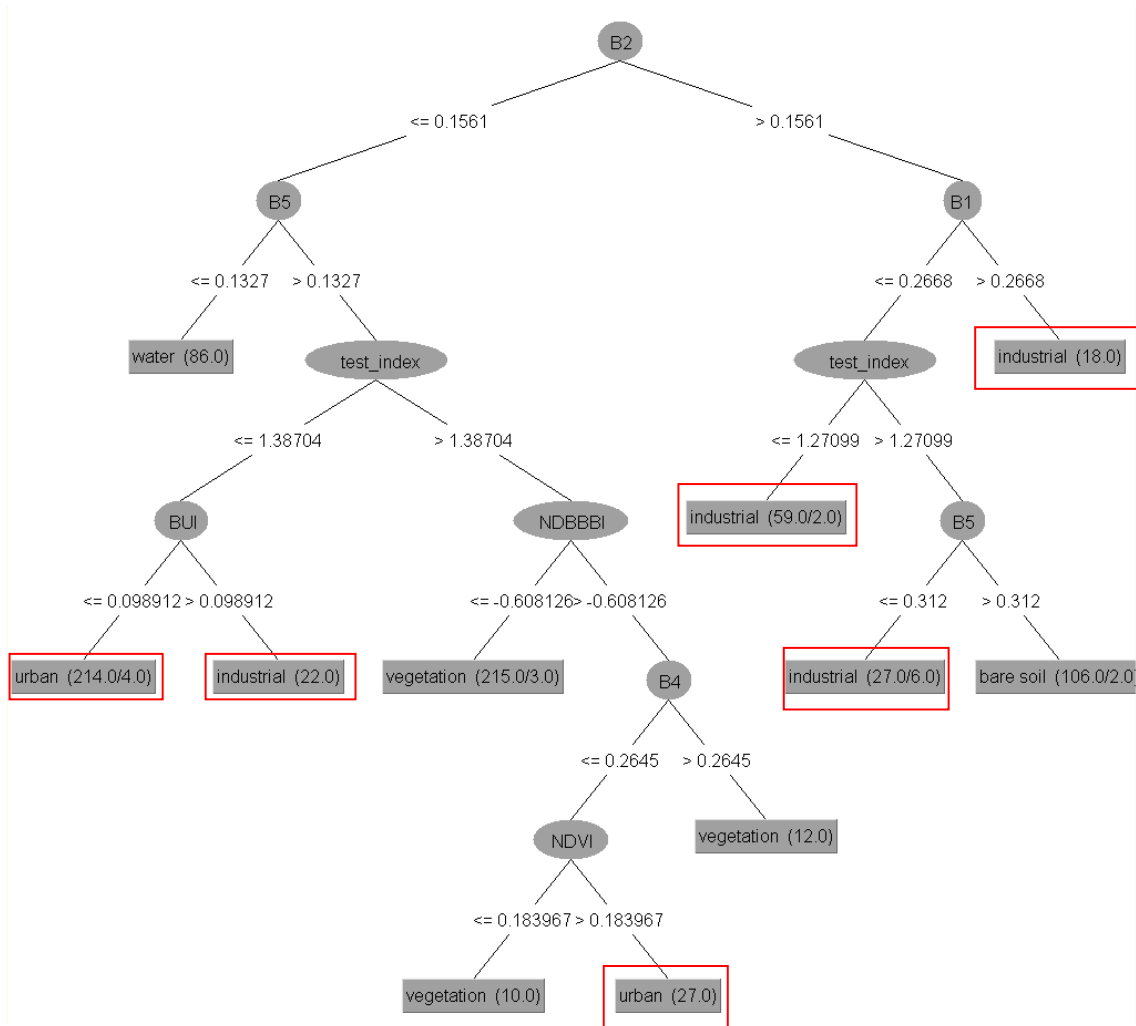
Figure 42 – Classifier "1 spectral index flash 5 class"
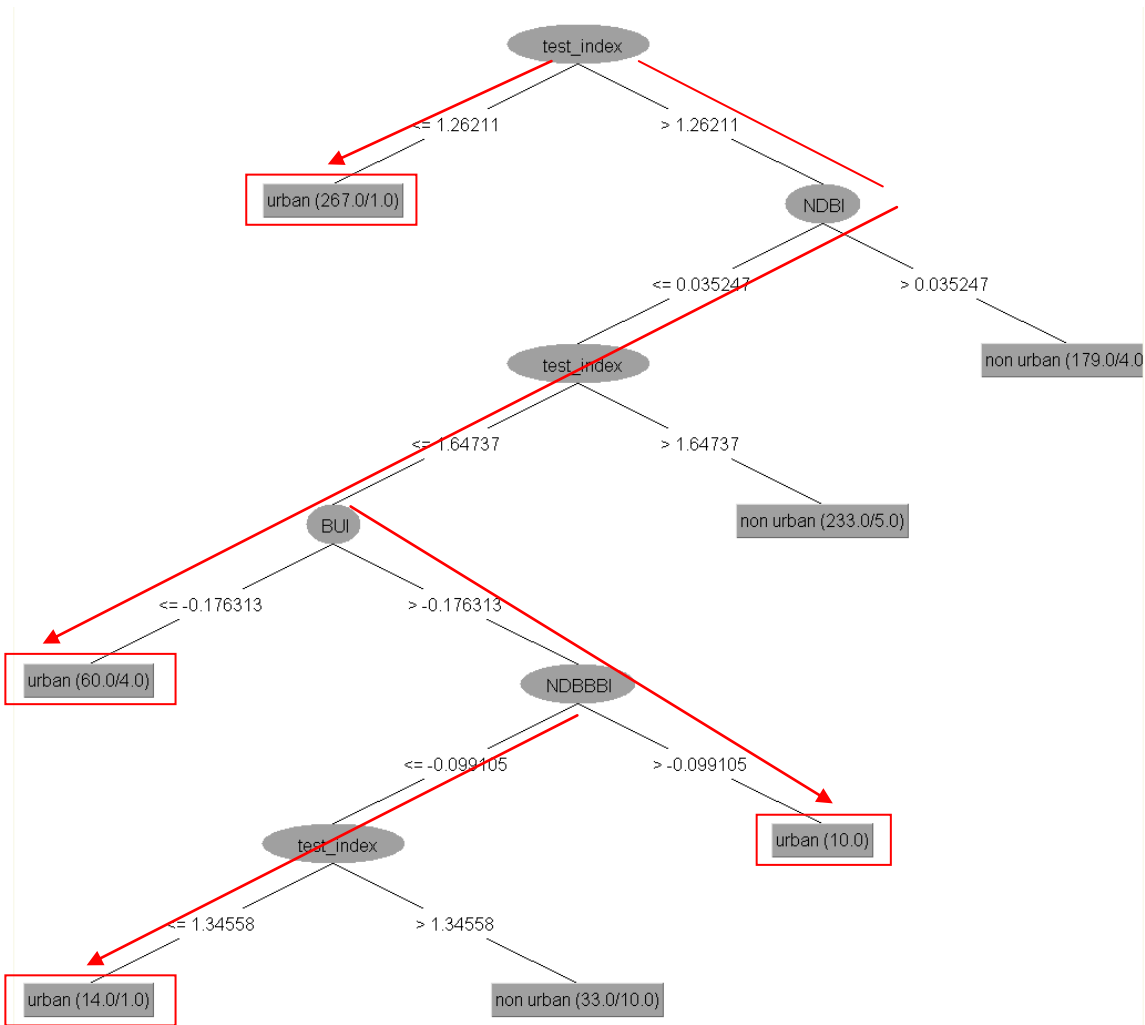
Figure 43 – Classifier: "1 spectral index FLAASH 2 class"
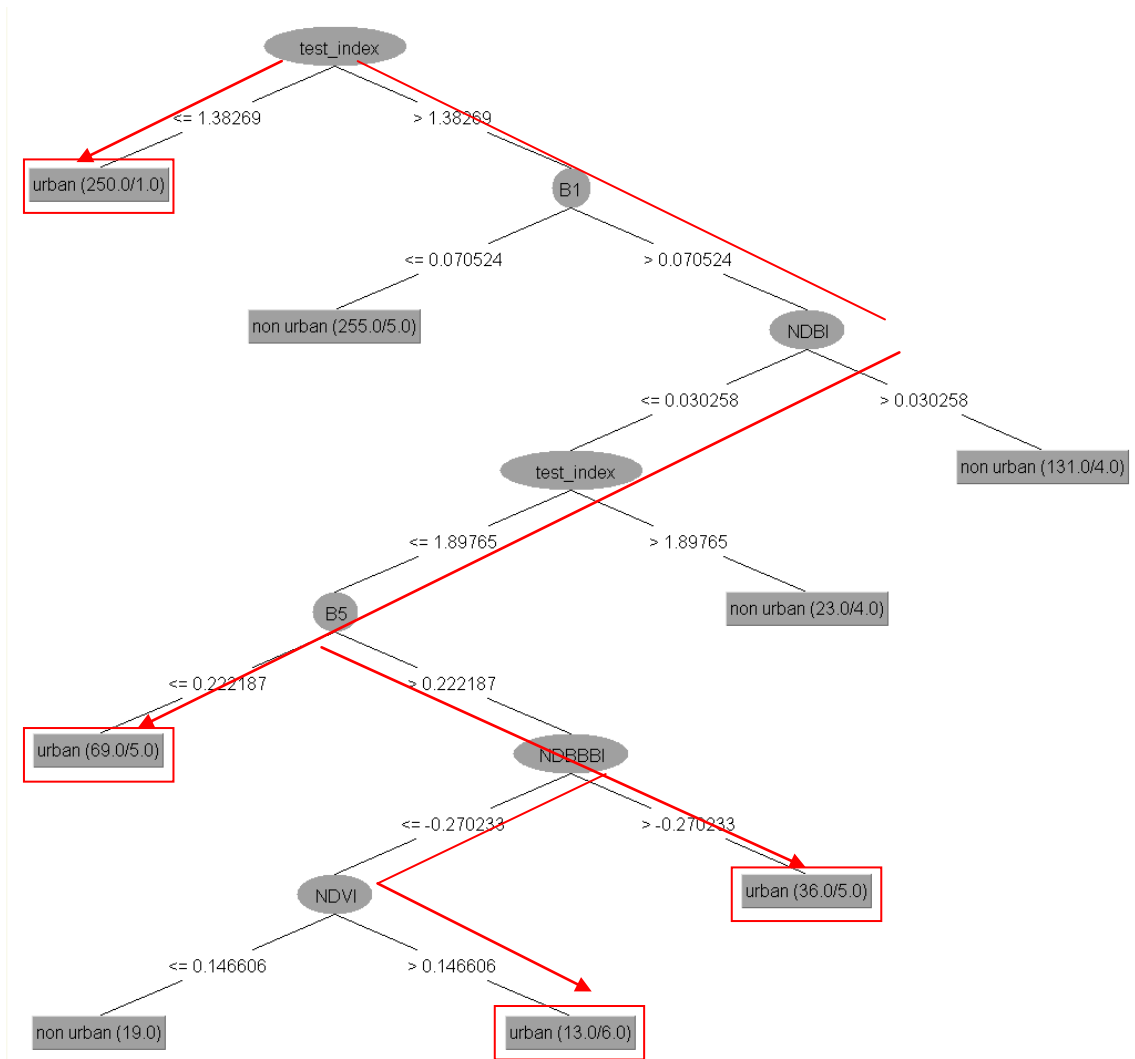
Figure 44 – Classifier "spectral index DS 2 class"

Classifier	Rules for the extraction of urban thematism
1 index DS 2class	**NDBI** ≥ -0.244853 and **NDBBBI** > -0.465424
1 index DS 5class	-0.465424 ≤ **NDBBBI** ≤ 0.006572 and 0.072057 ≤ **NDVI** ≤ 0.458584 or **NDBBBI** ≤ 0.006572 and **NDVI** ≤ 0.072057
1 index FLAASH 2class	**Test_index** ≤ 1.26211 or 1.26211 ≤ **Test_index** ≤ 1.64683 and **NDBI** ≤ 0.035247 and **NDVI** ≤ 0.194841 and **NDBBBI** ≥ -0.234464 or 1.26211 ≤ **Test_index** ≤ 1.64683 and **NDBI** ≤ 0.035247 and **NDVI** > 0.194841
1 spectral index FLAASH 5 class	**B2** > 0.1561 and **B1** > 0.2668 or **B2** > 0.1561 and **B1** ≤ 0.2668 and **test-index** ≤ 1.27099 or **B2** > 0.1561 and **B1** ≤ 0.2668 and **test-index** > 1.27099 and **B5** ≤ 0.312 or **B2** ≤ 0.1561 and **B5** > 0.1327 and **test-index** ≤ 1.38704 or **B2** ≤ 0.1561 and **B5** > 0.1327 and **test-index** > 1.38704 and **NDBBBI** > -0.608126 and **B4** ≤ 0.2645 and **NDVI** > 0.183967
1 spectral index FLAASH 2 class	**Test_index** ≤ 1.26211 or **Test_index** > 1.26211 and **NDBI** ≤ 0.035247 and **test_index** ≤ 1.64737 and **BUI** ≤ -0.176313 or **Test_index** > 1.26211 and **NDBI** ≤ 0.035247 and **test_index** ≤ 1.64737 and **BUI** > -0.176313 and **NDBBBI** > -0.099105 or **Test_index** > 1.26211 and **NDBI** ≤ 0.035247 and **test_index** ≤ 1.64737 and **BUI** > -0.176313 and **NDBBBI** ≤ -0.099105 and **test_index** ≤ 1.34558
1 spectral index DS 2 class	**Test_index** ≤ 1.38269 or **Test_index** > 1.38269 and **B1** > 0.070524 and **NDBI** ≤ 0.030258 and **Test_index** ≤ 1.89765 and **B5** ≤ 0.222187 or **Test_index** > 1.38269 and **B1** > 0.070524 and **NDBI** ≤ 0.030258 and **Test_index** ≤ 1.89765 and **B5** > 0.222187 and **NDBBBI** > -0.270233 or **Test_index** > 1.38269 and **B1** > 0.070524 and **NDBI** ≤ 0.030258 and **Test_index** ≤ 1.89765 and **B5** > 0.222187 and **NDBBBI** ≤ -0.270233 and NDVI > 0.146606

Table 22 – Rules for the extraction of urban thematism in the best classifiers

			OA	URBAN			NON URBAN	
				PA	UA	f-measure	PA	UA
2 CLASS	**TO single image**	index_calib	69.85%	91.35%	59.03%	71.71%	54.40%	89.73%
		index_DS	79.86%	86.90%	71.26%	78.30%	74.79%	88.81%
		index_flaash	72.72%	91.54%	61.73%	73.73%	59.18%	90.68%
		spectral_calib	59.58%	97.30%	50.88%	66.82%	32.45%	94.36%
		spectral_DS	59.58%	97.30%	50.88%	66.82%	32.45%	94.36%
		spectral_flaash	59.46%	97.34%	50.80%	66.76%	32.21%	94.40%
		spectral_index_calib	59.58%	97.30%	50.88%	66.82%	32.45%	94.36%
		spectral_index_DS	58.28%	97.63%	50.07%	66.19%	29.98%	94.61%
		spectral_index_flaash	59.46%	97.34%	50.80%	66.76%	32.21%	94.40%
	TO multitemporal	index_calib	77.02%	87.37%	67.38%	76.08%	69.58%	88.45%
		index_DS	78.55%	72.09%	75.52%	73.77%	83.20%	80.56%
		index_flaash	77.37%	69.94%	74.42%	72.11%	82.71%	79.28%
		spectral_calib	75.54%	68.69%	71.67%	70.15%	80.47%	78.14%
		spectral_DS	63.79%	98.00%	53.68%	69.37%	39.19%	96.46%
		spectral_flaash	77.45%	81.33%	69.77%	75.11%	74.65%	84.76%
		spectral_index_calib	79.20%	65.30%	81.29%	72.43%	89.19%	78.14%
		spectral_index_DS	78.55%	72.09%	75.52%	73.77%	83.20%	80.56%
		spectral_index_flaash	77.37%	69.94%	74.42%	72.11%	82.71%	79.28%
5 CLASS	**TO single image**	index_calib	71.14%	85.93%	61.01%	71.36%	60.51%	85.67%
		index_DS	80.49%	85.62%	72.64%	78.60%	76.81%	88.13%
		index_flaash	72.61%	90.11%	61.85%	73.35%	60.03%	89.41%
		spectral_calib	67.41%	96.62%	56.46%	71.27%	46.41%	95.02%
		spectral_DS	67.41%	96.62%	56.46%	71.27%	46.41%	95.02%
		spectral_flaash	58.74%	97.55%	50.36%	66.42%	30.84%	94.60%
		spectral_index_calib	68.23%	90.50%	57.66%	70.44%	52.21%	88.43%
		spectral_index_DS	67.93%	89.21%	57.52%	69.95%	52.63%	87.15%
		spectral_index_flaash	59.28%	90.49%	50.75%	65.03%	36.84%	84.34%
	TO multitemporal	index_calib	77.72%	66.78%	76.91%	71.49%	85.58%	78.18%
		index_DS	78.36%	74.98%	73.74%	74.36%	80.79%	81.79%
		index_flaash	79.76%	70.95%	78.59%	74.58%	86.10%	80.47%
		spectral_calib	78.59%	69.12%	77.29%	72.98%	85.39%	79.36%
		spectral_DS	72.63%	92.61%	61.48%	73.90%	58.27%	91.64%
		spectral_flaash	77.50%	79.73%	70.41%	74.78%	75.90%	83.89%
		spectral_index_calib	78.59%	69.12%	77.29%	72.98%	85.39%	79.36%
		spectral_index_DS	73.21%	92.48%	62.56%	74.64%	58.88%	91.33%
		spectral_index_flaash	77.32%	79.79%	70.12%	74.64%	75.55%	83.87%

Case study: Torino   **Post-processing: NO**

Best classifiers

Table 23 –Classification without post-processing for Torino case study: validation results

			OA	URBAN			NON URBAN	
				PA	UA	f-measure	PA	UA
2 CLASS	**TO single image**	index_calib	70.10%	92.54%	56.71%	70.32%	56.18%	92.39%
		index_DS	**80.42%**	87.42%	69.38%	**77.36%**	76.07%	90.70%
		index_flaash	74.03%	92.72%	60.49%	**73.21%**	62.43%	93.25%
		spectral_calib	57.90%	97.64%	47.57%	63.97%	33.25%	95.78%
		spectral_DS	57.90%	97.64%	47.57%	63.97%	33.25%	95.78%
		spectral_flaash	57.73%	97.67%	47.47%	63.89%	32.96%	95.80%
		spectral_index_calib	57.90%	97.64%	47.57%	63.97%	33.25%	95.78%
		spectral_index_DS	56.33%	97.93%	46.65%	63.20%	30.53%	95.97%
		spectral_index_flaash	57.73%	97.67%	47.47%	63.89%	32.96%	95.80%
	TO multitemporal	index_calib	78.20%	88.79%	66.00%	**75.72%**	71.63%	91.15%
		index_DS	**80.15%**	72.52%	74.85%	**73.66%**	84.88%	83.28%
		index_flaash	78.91%	69.88%	73.67%	71.73%	84.51%	81.89%
		spectral_calib	76.33%	66.69%	70.05%	68.33%	82.31%	79.94%
		spectral_DS	62.42%	98.58%	50.47%	66.76%	39.99%	97.84%
		spectral_flaash	78.91%	81.55%	69.00%	**74.75%**	77.28%	87.10%
		spectral_index_calib	**80.94%**	65.23%	81.28%	72.38%	90.68%	80.79%
		spectral_index_DS	**80.15%**	72.52%	74.85%	**73.66%**	84.88%	83.28%
		spectral_index_flaash	78.91%	69.88%	73.67%	71.73%	84.51%	81.89%
5 CLASS	**TO single image**	index_calib	**71.98%**	86.68%	59.15%	**70.32%**	62.86%	88.39%
		index_DS	**80.67%**	86.78%	69.95%	**77.46%**	76.87%	90.36%
		index_flaash	73.23%	91.60%	59.83%	72.38%	61.84%	92.23%
		spectral_calib	66.60%	97.26%	53.51%	69.04%	47.59%	96.55%
		spectral_DS	66.60%	97.26%	53.51%	69.04%	47.59%	96.55%
		spectral_flaash	56.76%	97.86%	46.90%	63.41%	31.27%	95.93%
		spectral_index_calib	68.26%	91.17%	55.17%	68.74%	54.05%	90.80%
		spectral_index_DS	68.03%	89.78%	55.06%	68.26%	54.55%	89.58%
		spectral_index_flaash	58.15%	90.75%	47.56%	62.41%	37.92%	86.86%
	TO multitemporal	index_calib	**79.14%**	68.81%	74.70%	71.63%	85.54%	81.55%
		index_DS	**79.58%**	76.31%	72.01%	**74.10%**	81.60%	84.74%
		index_flaash	**81.51%**	71.61%	78.25%	**74.78%**	87.65%	83.27%
		spectral_calib	**79.62%**	70.73%	74.69%	**72.66%**	85.14%	82.42%
		spectral_DS	71.00%	94.22%	57.39%	71.33%	56.60%	94.05%
		spectral_flaash	77.51%	81.81%	66.86%	**73.59%**	74.85%	86.90%
		spectral_index_calib	**79.62%**	70.73%	74.69%	**72.66%**	85.14%	82.42%
		spectral_index_DS	71.00%	94.22%	57.39%	71.33%	56.60%	94.05%
		spectral_index_flaash	77.34%	81.87%	66.60%	**73.45%**	74.53%	86.89%

Case study: Torino   **Post-processing: YES**

Best classifiers	Good classifiers

Table 24 - Classification with post-processing for Torino case study: validation results

			OA	URBAN			NON URBAN	
				PA	UA	f-measure	PA	UA
2 CLASS	AT single image	index_calib	94.18%	41.41%	38.86%	40.10%	96.78%	97.10%
		index_DS	95.70%	27.46%	59.19%	37.52%	99.06%	96.51%
		index_flaash	95.01%	38.01%	46.39%	41.78%	97.83%	96.97%
		spectral_calib	95.14%	26.97%	47.16%	34.32%	98.51%	96.47%
		spectral_DS	95.14%	26.94%	47.15%	34.29%	98.51%	96.47%
		spectral_flaash	95.26%	26.19%	49.26%	34.20%	98.67%	96.44%
		spectral_index_calib	93.93%	43.00%	37.40%	40.01%	96.45%	97.16%
		spectral_index_DS	95.29%	30.24%	49.96%	37.68%	98.50%	96.62%
		spectral_index_flaash	95.01%	38.95%	46.46%	42.38%	97.78%	97.01%
	AT multitemporal	index_calib	90.56%	52.42%	25.51%	34.32%	92.44%	97.52%
		index_DS	80.65%	68.85%	15.34%	25.09%	81.23%	98.14%
		index_flaash	72.72%	69.60%	11.25%	19.36%	72.87%	97.98%
		spectral_calib	95.39%	6.99%	58.34%	12.48%	99.75%	95.60%
		spectral_DS	95.39%	8.78%	56.56%	15.20%	99.67%	95.68%
		spectral_flaash	94.80%	16.65%	38.09%	23.17%	98.66%	95.99%
		spectral_index_calib	95.66%	16.38%	65.86%	26.23%	99.58%	96.02%
		spectral_index_DS	95.31%	11.13%	50.82%	18.26%	99.47%	95.77%
		spectral_index_flaash	95.20%	13.19%	46.61%	20.56%	99.25%	95.86%
5 CLASS	AT single image	index_calib	95.40%	31.19%	51.87%	38.95%	98.57%	96.67%
		index_DS	92.69%	43.42%	30.55%	35.87%	95.13%	97.15%
		index_flaash	91.86%	49.96%	28.90%	36.62%	93.93%	97.44%
		spectral_calib	94.74%	29.37%	41.62%	34.44%	97.97%	96.56%
		spectral_DS	94.74%	29.37%	41.62%	34.44%	97.97%	96.56%
		spectral_flaash	94.68%	30.15%	41.08%	34.78%	97.86%	96.60%
		spectral_index_calib	94.14%	33.12%	36.51%	34.73%	97.15%	96.71%
		spectral_index_DS	88.48%	49.02%	20.18%	28.59%	90.42%	97.29%
		spectral_index_flaash	95.40%	34.19%	51.74%	41.17%	98.43%	96.80%
	AT multitemporal	index_calib	83.89%	54.41%	15.49%	24.12%	85.35%	97.43%
		index_DS	92.59%	40.64%	29.29%	34.04%	95.15%	97.01%
		index_flaash	90.88%	36.26%	21.82%	27.24%	93.58%	96.75%
		spectral_calib	74.78%	26.27%	5.38%	8.93%	77.17%	95.49%
		spectral_DS	94.77%	12.85%	34.86%	18.78%	98.81%	95.83%
		spectral_flaash	94.35%	16.85%	31.28%	21.90%	98.17%	95.98%
		spectral_index_calib	95.30%	5.84%	50.99%	10.48%	99.72%	95.54%
		spectral_index_DS	92.05%	19.45%	18.05%	18.72%	95.64%	96.01%
		spectral_index_flaash	91.52%	21.14%	17.27%	19.01%	95.00%	96.06%

Case study: Asti   **Post-processing: NO**

Best classifiers

Table 25 - Classification without post-processing for Asti case study: validation results

			OA	URBAN			NON URBAN	
				PA	UA	f-measure	PA	UA
2 CLASSI	AT single image	*index_calib*	**97.42%**	38.60%	56.49%	**45.87%**	99.13%	98.23%
		index_DS	**97.71%**	23.83%	83.09%	37.03%	99.86%	97.83%
		index_flaash	**97.71%**	35.18%	68.39%	**46.46%**	99.53%	98.14%
		spectral_calib	**97.52%**	31.64%	62.23%	**41.95%**	99.44%	98.04%
		spectral_DS	**97.52%**	31.60%	62.23%	**41.91%**	99.44%	98.04%
		spectral_flaash	**97.55%**	30.38%	64.11%	**41.22%**	99.50%	98.00%
		spectral_index_calib	**97.30%**	42.28%	52.75%	**46.94%**	98.90%	98.33%
		spectral_index_DS	**97.59%**	35.31%	63.25%	**45.32%**	99.40%	98.14%
		spectral_index_flaash	**97.73%**	37.14%	68.24%	**48.10%**	99.50%	98.19%
	AT multitemporal	index_calib	95.89%	50.62%	34.57%	**41.09%**	97.21%	98.54%
		index_DS	85.07%	74.39%	12.91%	22.00%	85.38%	99.13%
		index_flaash	76.51%	73.35%	8.37%	15.02%	76.61%	99.00%
		spectral_calib	**97.27%**	6.85%	68.21%	12.44%	99.91%	97.36%
		spectral_DS	**97.31%**	8.94%	69.13%	15.83%	99.88%	97.41%
		spectral_flaash	**97.06%**	18.97%	45.18%	26.72%	99.33%	97.68%
		spectral_index_calib	**97.51%**	16.40%	78.30%	27.11%	99.87%	97.62%
		spectral_index_DS	**97.30%**	12.46%	61.16%	20.70%	99.77%	97.51%
		spectral_index_flaash	**97.25%**	14.94%	55.29%	23.53%	99.65%	97.57%
5 CLASSI	AT single image	*index_calib*	**97.41%**	35.44%	56.92%	**43.69%**	99.22%	98.14%
		index_DS	96.49%	47.24%	39.85%	**43.23%**	97.92%	98.46%
		index_flaash	95.25%	53.59%	30.64%	38.99%	96.47%	98.62%
		spectral_calib	**97.16%**	35.56%	49.85%	**41.51%**	98.96%	98.14%
		spectral_DS	**97.16%**	35.56%	49.85%	**41.51%**	98.96%	98.14%
		spectral_flaash	**97.05%**	36.34%	47.28%	**41.10%**	98.82%	98.16%
		spectral_index_calib	96.93%	39.04%	45.05%	**41.83%**	98.61%	98.23%
		spectral_index_DS	91.59%	58.50%	18.62%	28.25%	92.55%	98.71%
		spectral_index_flaash	**97.34%**	40.55%	53.93%	**46.29%**	98.99%	98.28%
	AT multitemporal	index_calib	87.26%	58.00%	12.44%	20.49%	88.11%	98.63%
		index_DS	95.19%	48.55%	29.08%	36.37%	96.55%	98.47%
		index_flaash	93.81%	40.63%	20.32%	27.09%	95.36%	98.22%
		spectral_calib	75.79%	27.82%	3.43%	6.11%	77.18%	97.35%
		spectral_DS	96.76%	16.96%	35.11%	22.87%	99.09%	97.62%
		spectral_flaash	96.32%	21.99%	29.76%	25.30%	98.49%	97.75%
		spectral_index_calib	**97.18%**	7.70%	51.71%	13.40%	99.79%	97.38%
		spectral_index_DS	95.07%	21.46%	18.33%	19.77%	97.22%	97.70%
		spectral_index_flaash	94.70%	22.45%	17.00%	19.35%	96.81%	97.72%

Case study: Asti   **Post-processing: YES**

Best classifiers	Good classifiers

Table 26 - Classification with post-processing for Asti case study: validation results

			OA	URBAN			NON URBAN	
				PA	UA	f-measure	PA	UA
2 CLASS	**AT single image**	index_calib	91.77%	48.69%	58.99%	53.34%	96.38%	94.61%
		index_DS	92.89%	39.49%	75.20%	51.79%	98.61%	93.84%
		index_flaash	92.36%	45.03%	65.16%	53.25%	97.42%	94.31%
		spectral_calib	92.82%	43.65%	70.84%	54.02%	98.08%	94.21%
		spectral_DS	92.81%	43.62%	70.84%	54.00%	98.08%	94.20%
		spectral_flaash	93.15%	42.62%	76.01%	54.61%	98.56%	94.13%
		spectral_index_calib	91.57%	50.24%	57.32%	53.55%	96.00%	94.74%
		spectral_index_DS	93.18%	48.59%	71.73%	57.94%	97.95%	94.68%
		spectral_index_flaash	92.50%	47.28%	65.54%	54.93%	97.34%	94.52%
	AT multitemporal	index_calib	89.26%	59.91%	45.76%	51.89%	92.40%	95.56%
		index_DS	80.72%	70.84%	29.38%	41.54%	81.78%	96.32%
		index_flaash	74.38%	67.79%	22.56%	33.85%	75.09%	95.61%
		spectral_calib	91.46%	17.25%	75.57%	28.09%	99.40%	91.82%
		spectral_DS	91.82%	21.56%	77.78%	33.76%	99.34%	92.21%
		spectral_flaash	92.12%	34.27%	68.53%	45.70%	98.32%	93.32%
		spectral_index_calib	92.73%	31.43%	82.56%	45.53%	99.29%	93.12%
		spectral_index_DS	92.27%	26.24%	80.79%	39.61%	99.33%	92.64%
		spectral_index_flaash	92.28%	28.17%	77.83%	41.36%	99.14%	92.80%
5 CLASS	**AT single image**	index_calib	92.84%	45.12%	70.20%	54.93%	97.95%	94.34%
		index_DS	89.49%	53.86%	46.25%	49.77%	93.30%	94.97%
		index_flaash	90.80%	61.72%	52.02%	56.46%	93.91%	95.82%
		spectral_calib	93.08%	48.44%	70.72%	57.50%	97.85%	94.66%
		spectral_DS	93.08%	48.44%	70.72%	57.50%	97.85%	94.66%
		spectral_flaash	92.99%	48.42%	69.84%	57.19%	97.76%	94.66%
		spectral_index_calib	92.65%	50.78%	65.43%	57.18%	97.13%	94.86%
		spectral_index_DS	88.50%	63.53%	43.51%	51.65%	91.17%	95.90%
		spectral_index_flaash	93.39%	50.13%	73.04%	59.45%	98.02%	94.84%
	AT multitemporal	index_calib	85.27%	50.94%	33.03%	40.07%	88.95%	94.43%
		index_DS	90.11%	47.86%	48.82%	48.34%	94.63%	94.43%
		index_flaash	88.79%	45.12%	42.51%	43.77%	93.47%	94.09%
		spectral_calib	72.31%	44.17%	16.07%	23.57%	75.32%	92.65%
		spectral_DS	91.84%	29.83%	67.72%	41.41%	98.48%	92.91%
		spectral_flaash	91.78%	34.82%	63.64%	45.01%	97.87%	93.35%
		spectral_index_calib	91.17%	14.13%	72.10%	23.63%	99.41%	91.54%
		spectral_index_DS	89.77%	34.01%	46.08%	39.14%	95.74%	93.13%
		spectral_index_flaash	89.35%	34.98%	43.64%	38.83%	95.16%	93.19%

Case study: Asti subset   **Post-processing: NO**

Best classifiers

Table 27 - Classification without post-processing for Asti case study: validation results performed on the subset

			OA	URBAN			NON URBAN	
				PA	UA	f-measure	PA	UA
2 CLASS	**AT single image**	index_calib	**95.64%**	51.83%	75.62%	61.50%	98.80%	96.61%
		index_DS	**95.66%**	40.85%	88.16%	55.83%	99.60%	95.90%
		index_flaash	95.74%	47.13%	81.71%	59.78%	99.24%	96.30%
		spectral_calib	**95.97%**	52.56%	80.68%	**63.66%**	99.09%	96.67%
		spectral_DS	**95.96%**	52.52%	80.67%	**63.62%**	99.09%	96.66%
		spectral_flaash	**96.11%**	50.96%	85.24%	**63.79%**	99.36%	96.57%
		spectral_index_calib	95.57%	54.68%	72.58%	62.37%	98.51%	96.79%
		spectral_index_DS	**96.12%**	57.62%	78.90%	**66.60%**	98.89%	97.01%
		spectral_index_flaash	95.95%	51.47%	81.51%	**63.09%**	99.16%	96.59%
	AT multitemporal	index_calib	94.92%	63.78%	61.84%	62.79%	97.16%	97.39%
		index_DS	85.20%	76.50%	28.00%	40.99%	85.83%	98.07%
		index_flaash	78.45%	70.62%	19.51%	30.57%	79.01%	97.39%
		spectral_calib	94.29%	17.88%	86.20%	29.61%	99.79%	94.40%
		spectral_DS	94.67%	23.67%	88.91%	37.39%	99.79%	94.78%
		spectral_flaash	**95.24%**	40.27%	78.29%	53.19%	99.20%	95.84%
		spectral_index_calib	**95.35%**	34.71%	89.92%	50.09%	99.72%	95.50%
		spectral_index_DS	**95.09%**	30.87%	88.74%	45.80%	99.72%	95.24%
		spectral_index_flaash	**95.13%**	33.05%	85.58%	47.69%	99.60%	95.38%
5 CLASS	**AT single image**	index_calib	**95.60%**	53.54%	73.84%	62.08%	98.63%	96.72%
		index_DS	94.56%	60.19%	59.35%	59.77%	97.03%	97.13%
		index_flaash	94.20%	68.80%	55.52%	61.45%	96.03%	97.71%
		spectral_calib	**96.01%**	58.26%	76.73%	**66.23%**	98.73%	97.05%
		spectral_DS	**96.01%**	58.26%	76.73%	**66.23%**	98.73%	97.05%
		spectral_flaash	**95.89%**	58.12%	75.11%	**65.53%**	98.61%	97.03%
		spectral_index_calib	95.86%	60.65%	73.19%	66.33%	98.40%	97.20%
		spectral_index_DS	91.98%	74.94%	44.30%	55.68%	93.21%	98.10%
		spectral_index_flaash	**96.01%**	59.96%	75.57%	66.86%	98.60%	97.16%
	AT multitemporal	index_calib	89.01%	52.71%	31.20%	39.20%	91.63%	96.42%
		index_DS	93.54%	57.17%	51.77%	54.33%	96.16%	96.89%
		index_flaash	92.12%	50.88%	42.76%	46.47%	95.09%	96.41%
		spectral_calib	73.37%	48.97%	12.42%	19.81%	75.12%	95.34%
		spectral_DS	94.66%	37.28%	69.00%	48.41%	98.79%	95.63%
		spectral_flaash	94.53%	43.51%	63.55%	51.66%	98.20%	96.02%
		spectral_index_calib	94.01%	16.45%	74.82%	26.97%	99.60%	94.30%
		spectral_index_DS	93.51%	41.12%	52.15%	45.98%	97.28%	95.82%
		spectral_index_flaash	93.17%	41.39%	49.05%	44.90%	96.90%	95.83%

Case study: Asti subset	Post-processing: YES

Best classifiers	Good classifiers

Table 28 - Classification with post-processing for Asti case study: validation results performed on the subset

4.3 Qualitative analysis of classification error (Spatial distribution of error/ quality control)

The erroneous allocations made by a classification are typically not randomly distributed over the region (Congalton 1988; Steele et al. 1996). Often there is a distinct of thematic errors arising from the sensor's properties (Foody 1988) and/or the ground conditions, with, for example, errors spatially correlated at the boundaries of classes (Congalton 1988; Edwards and Lowell 1996; Steele et al. 1996; Vieira and Mather 2000). Much of the error occurring at boundaries is associated with misregistration of the data sets and mixed pixels.

Accuracy measures derived from the error matrix provide information on the quality of the map as a whole but cannot be used to characterize distinct areas of the map (Mayaux et al. 2006); as a consequence, quality control can be used in order to describe the spatial distribution of errors.

The performed quality control was aimed at detecting more recursive errors in the resulting "good" classifiers. The questions were:

- Where were my errors located?
- Were there some land covers usually misclassified each-other?
- Is it possible to detect at least the presence of a small-medium built-up area, in spite of classification errors?

The analysis was performed on the resulting "good-classifiers", taking into consideration only post-processed classifications. Good classifiers were pinpointed, for each case study, in the following way: first, nine best OA and nine best f-measures were selected; second, the accuracy percentage (both OA and f-measure) of the ninth classifier was considered and round down of, at maximum, 1%: this value represented the threshold detected not to exclude any classifier with accuracy measure very similar to the first nine classifiers; finally, among selected classifiers, those that were selected both for best OA and f-measures, were considered the "good classifiers". The number of selected good-classifiers together with the selected thresholds is shown Table 29.

Table 24, Table 26 and Table 28 show, in orange, selected good classifiers and, in bold type, best f-measures and OA detected with the above mentioned criterion.

Case study	Number of selected "good classifier"	OA threshold	f-measure threshold
Torino – post processed	6	79%	73%
Asti – post processed	13	97%	41%
Asti subset – post processed	10	95%	63%

Table 29 – Number of "good" classifiers selected for the qualitative analysis and threshold fixed to determine "good classifiers"

The procedure created in order to perform the quality control, used, in input, both the "good classifiers" and the validation mask and gave, in output, an error mask: the information associated to each pixel was the membership to the Urban or Non Urban class, together with the number of times in which that pixel was correctly classified by the "good classifiers". The concept was represented with two color ramps, one from white to dark grey for the Non Urban class, the other from red to yellow for the Urban class. Non Urban pixels correctly classified from all classifiers appear white, and the colour became more and more grey according to the percentage of classifiers that mis-classify that pixel: dark grey pixels were misclassified by all the classifiers (that is, classified as Urban). In the same way, Urban pixels correctly classified from all classifiers appear red, and the colour became more and more yellow according to the percentage of classifiers that mis-classify that pixel: yellow pixels were misclassified by all the classifiers (that is, classified as Non Urban). An example of the error mask legend is provided in Figure 45.
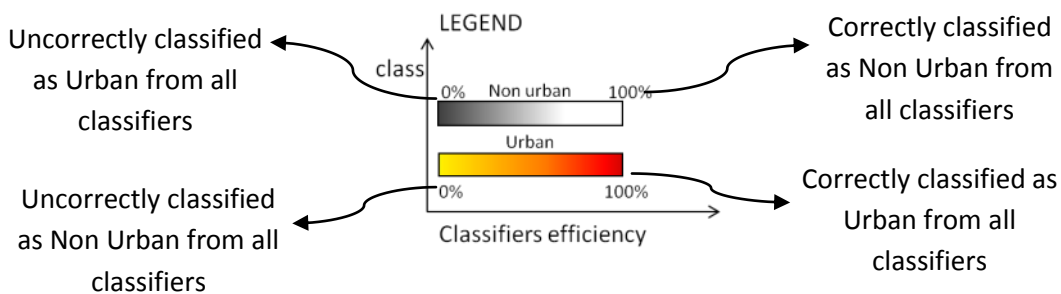

Figure 45 – Legend to interpret the error mask

The aim was to evaluate where most recurring errors were located. The analysis, performed both for Torino and Asti case study, is described in sections 4.3.1 and 4.3.2.

4.3.1   Torino case study

The error mask built for Torino case study provides results shown in

Figure 46; the blue line represents the district boundaries, while the validated area is the non-black one. The analysis shows that dense built-up area is correctly classified by almost all classifiers; most recursive classification errors are hooped in green and a zoom is provided in Figure 47 to Figure 51. A background imagery was used in the zoom (Source: Ikonos, Date: 23.02.2001) in order to provide a quick understanding of the land cover, but it is not the one used for the mask construction and is 4 years prior the date of the classified image.
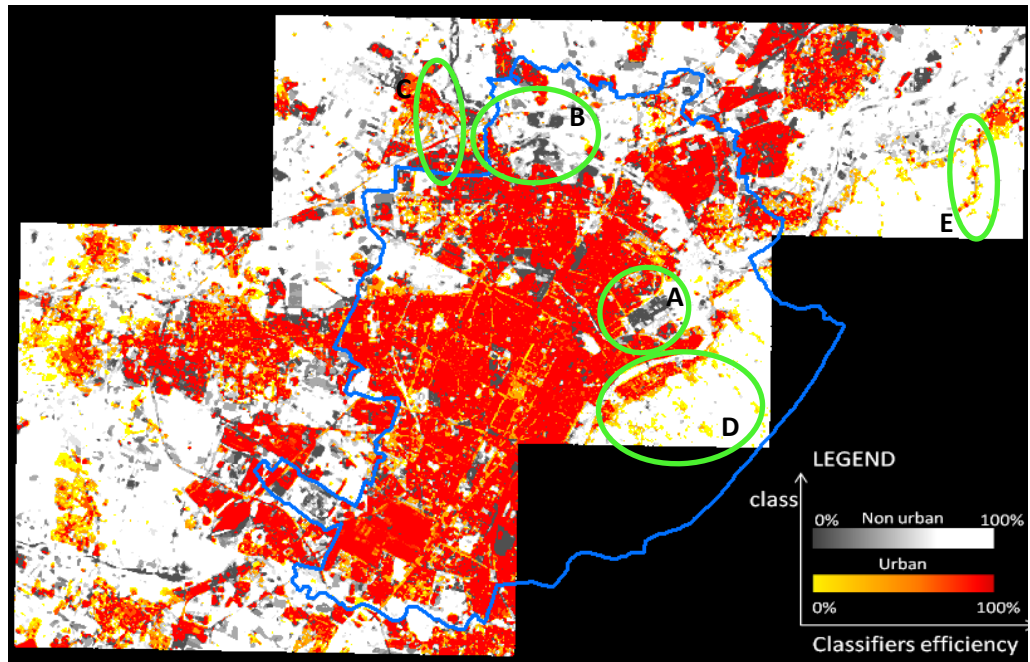
Figure 46 – Quality control performed on Torino case study: number of times in which each pixel is correctly classified by "good" classifiers.

Figure 47 and Figure 48 provide example of Non Urban areas that were often incorrectly classified as Urban.

Zoom A (Figure 47) represents a cemetery, while Zoom B (Figure 48) shows areas where excavation were performed or quarries are present (lower green circles) or, even, cultivated fields that were misclassified (higher green circles).

Zoom A
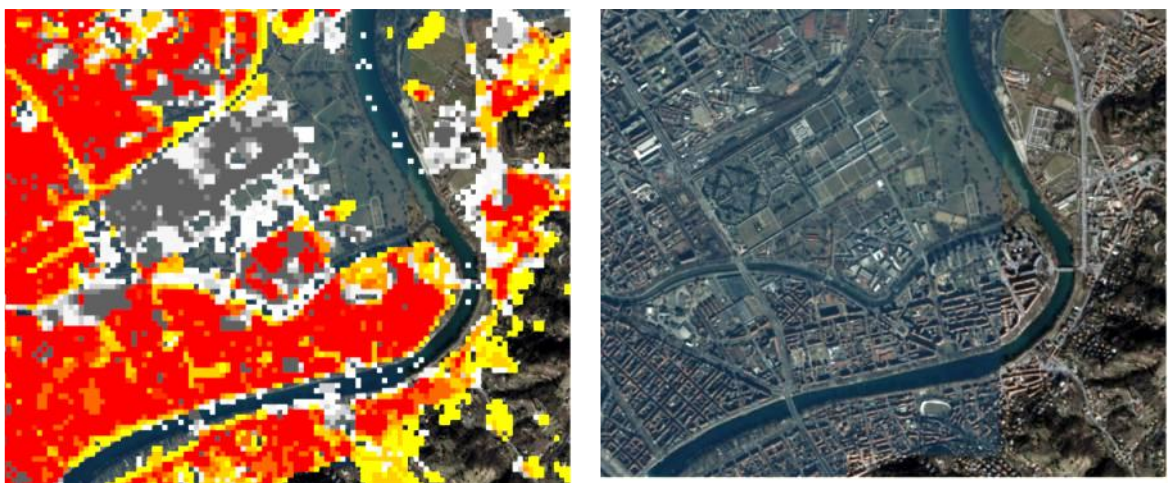
Figure 47 – Area often incorrectly classified as urban: cemetery

Zoom B

Figure 48 – Areas often uncorrectly classified as urban: quarries or excavation site (lower circles) and cultivated fields (higher circles)

Zoom C (Figure 49) provides example of bare soil areas that were classified as urban: the confusion between bare soil and urban is a typical and frequent problem in urban classification.

Zoom C
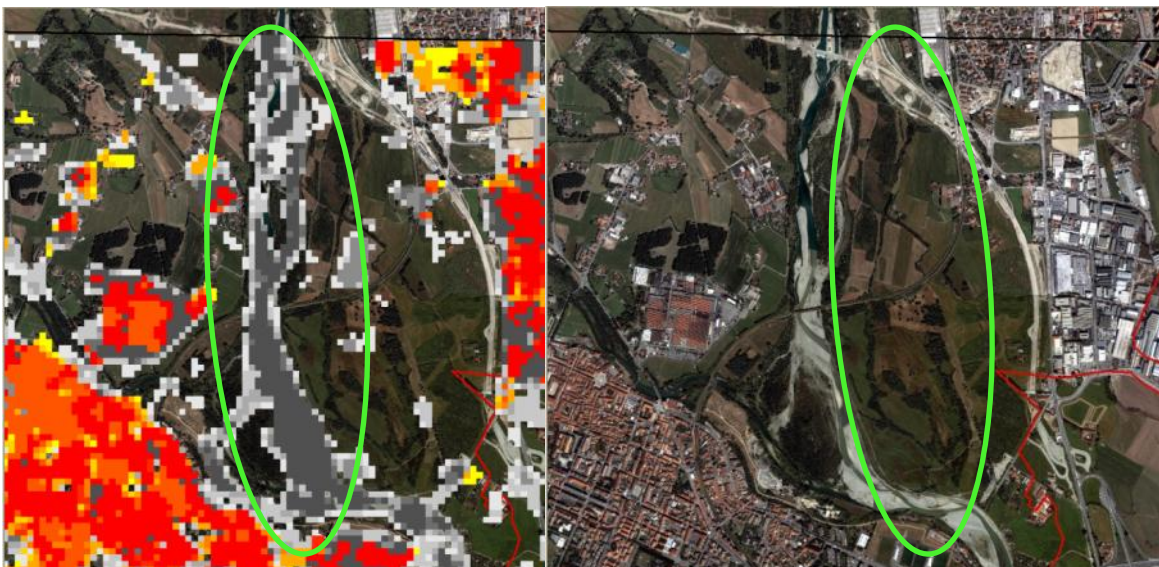
Figure 49 – Areas often uncorrectly classified as urban: bare soil

Zoom D and E (Figure 50 and Figure 51) shows examples where Urban areas were not detected; this is the case of sparse urban.

Zoom D
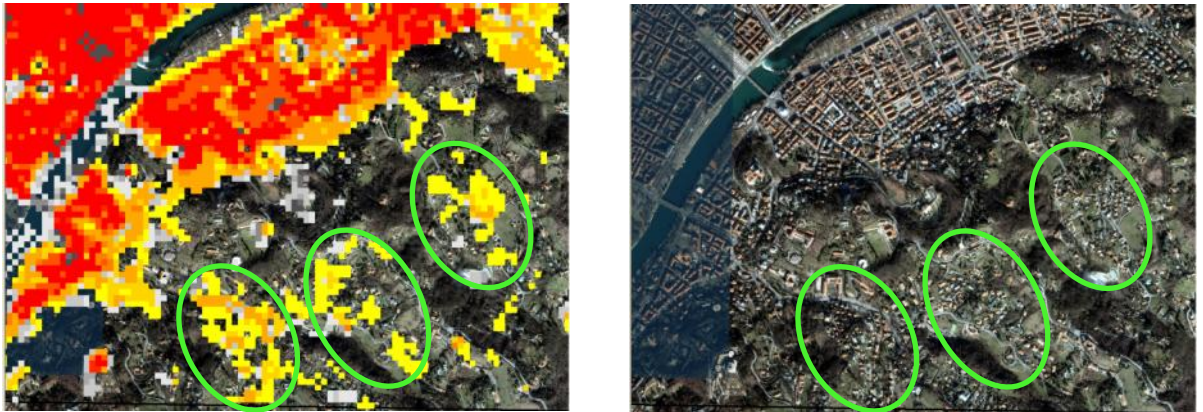
Figure 50 – Urban areas often not detected: sparse houses.

Figure 51 shows an example of a small built up area; it states that at least a small core part was always detected, while mis-classified pixels were located at the boundaries.

Zoom E
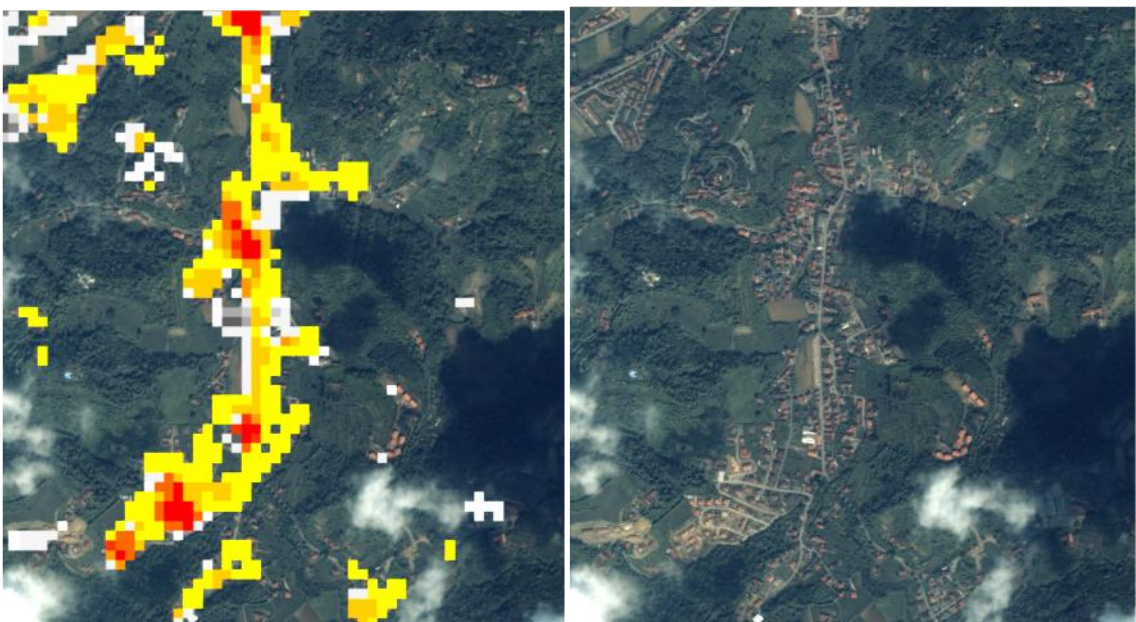
Figure 51 – Urban areas often not detected: low density built up area

4.3.2 Asti case study

In Asti case study, the investigated area was very different from the previous one: the extent was larger and arranged by many small-medium built-up areas, unlike Torino case study where the analysis was centred on the city and near surroundings.

The aforementioned quality analysis was here extended thanks to the availability of a municipality boundaries vector layer; the "error mask" was here superimposed to the vector layer in order to extract errors statistics for each municipality.

The adopted Piedmont region municipality boundaries vector layer provides, among other things, information about the dimension of each municipality.

759 districts own to Piedmont region; among them, 676 are very small (less than 250.000 m^2, that means, more or less, a dimension of 500 m x 500 m), 77 are small (between 250.000 and 1.200.000 m^2, that is, more or less, a dimension of 1 km x 1 km), 3 are medium-small (between 1.200.000 and 3.200.000 m^2, that is, more or less, a dimension of 1.7 km x 1.7 km) and 1 is medium (Asti district).

In the performed analysis, only small and medium small built-up areas were considered. The analysis was performed on the "good-classifiers", as discussed in section 4.3. The number of "good-classifiers" on which the analysis was performed is reported in Table 29. For each district it was calculated:

- The number of pixel detected as Urban in the validation mask;
- The number of pixel detected as Urban in more then 80% of cases;
- The number of pixel detected as Urban in more then 60% of cases

Results are summarized in Annex IV; they are ordered according to the number of pixel in the validation mask.

The outcome allowed identifying places where classifiers had good performances, and places (town) where this didn't happened. Once places where bad performances occourred were identified, a visual analysis was performed in order to verify causes of bad classification.

Figure 52 and Figure 54 shows example of good performances; districts of about 500.000 m^2 and 250.000 m^2 had 42.5% and 49.5% of pixels correctly classified in 80% of cases.

On the contrary, Figure 53 and Figure 55 are example of bad performances; districts of about 540.000 m^2 and 400.000 m^2 had just 19.7% and 1% of pixels correctly classified in 80% of cases.

The analysis allowed to verify how the core part of a small town can be identified if the built-up area density is quite high; on the contrary, where the built up area is sparser, the classifier has difficulty in detecting urban pixels. The application of post-processing techniques can help in recreating a more homogeneus texture, making use of clump or majority filtering procedures.
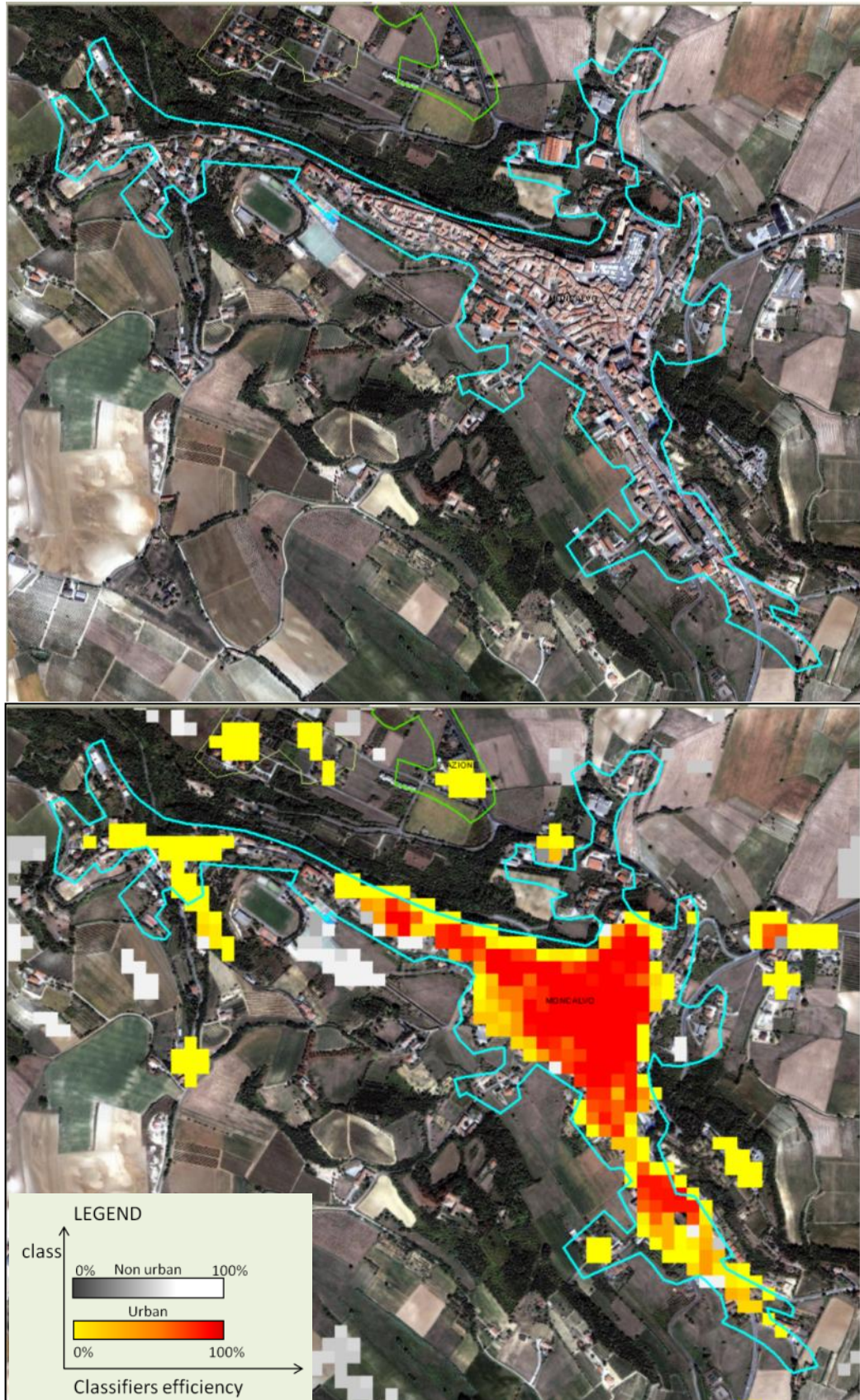
Figure 52 – Moncalvo, 528.400 m^2. 42.5% of pixels were correctly classified as Urban by the majority of "good classifiers". The dense built-up area was correctly classified in almost all cases.
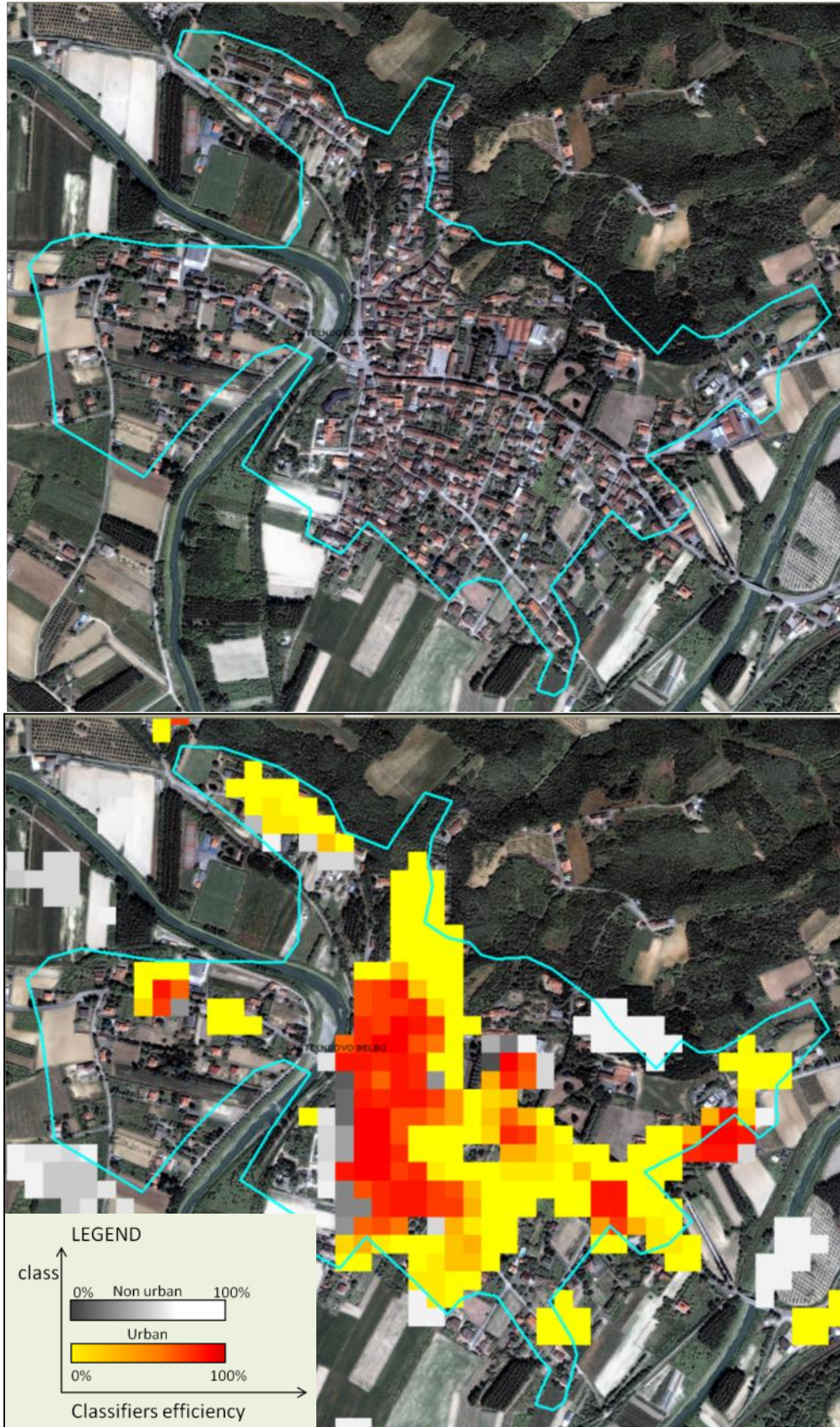
Figure 53 – Castelnuovo Belbo, 541.600 m^2. Only 19.7% of pixels were correctly classified in 80% of cases. This can be due to the high presence of vegetation in the whole area (lower built-up density)

Figure 54 – Rocchetta Tanaro, 263.257 m². 49.5% of pixels were correctly classified as Urban by the majority of "good classifiers". Errors were located where the presence of vegetation was higher.

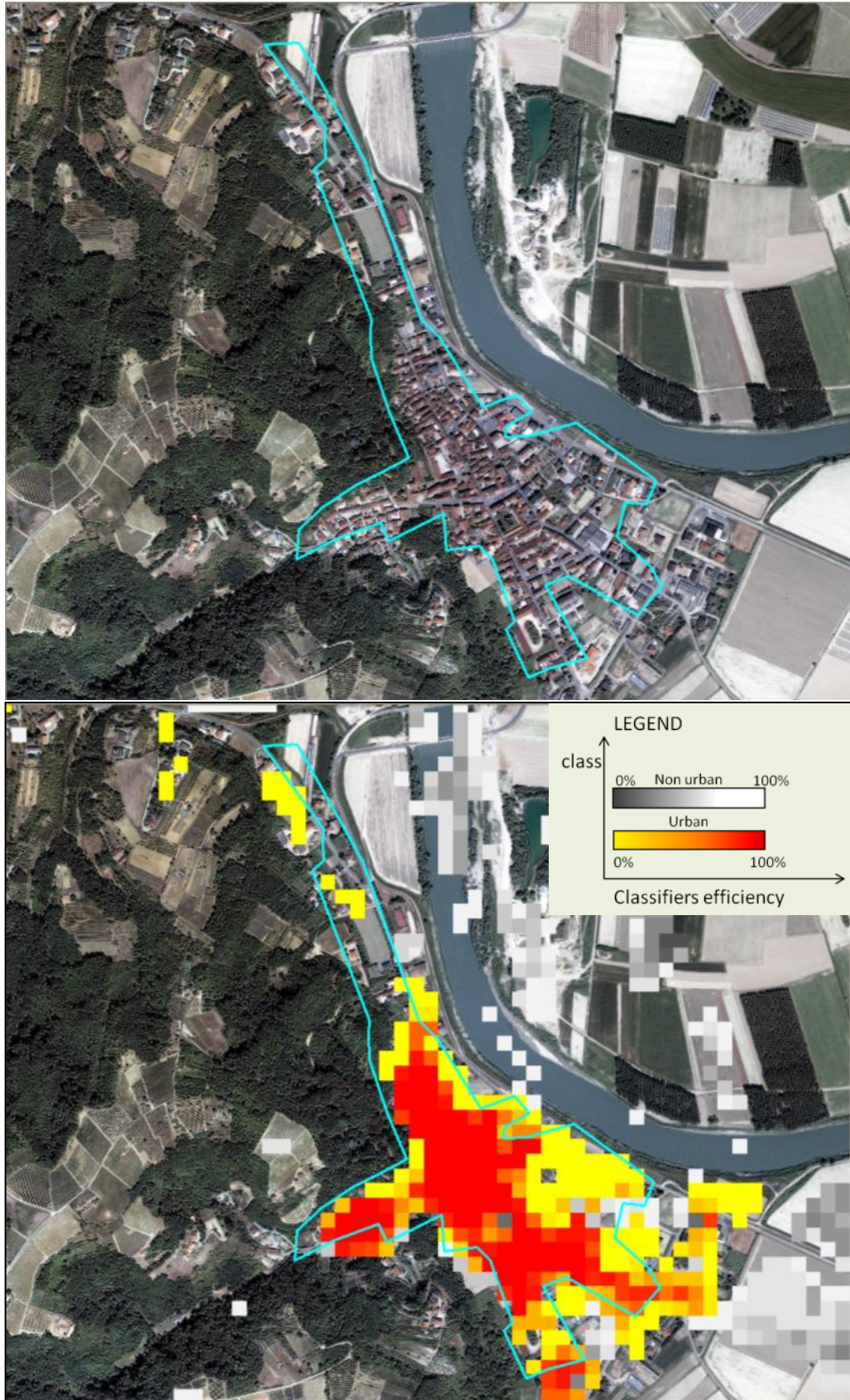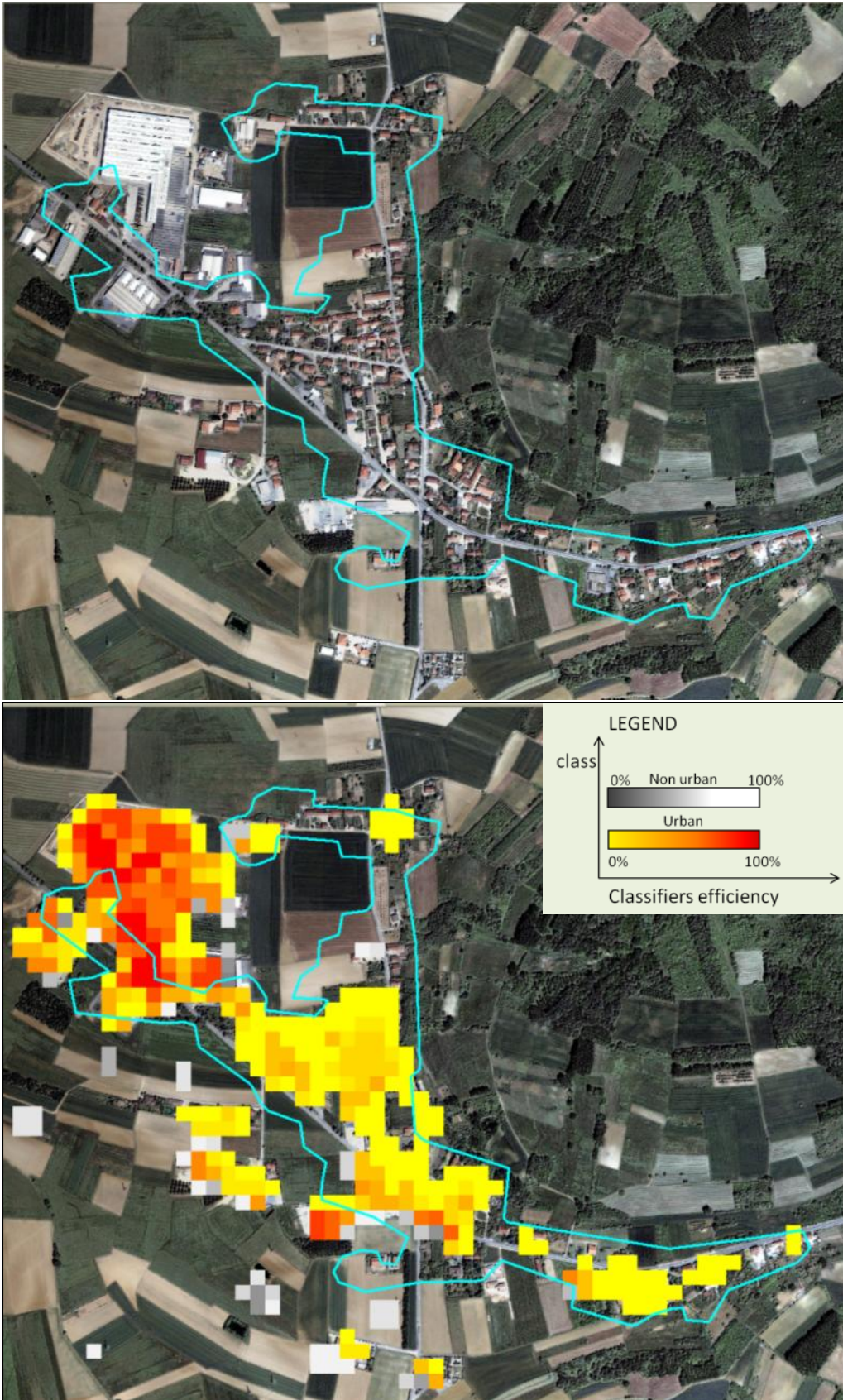**Figure 55 - Dusino, 391.200 m^2. Only 1% of pixels were correctly classified by almost all "good classifiers". This is a clear example of where urban procedure classification fails.**

4.4 Analysis of tested variables

In order to evaluate the best variables to use for the classification process, the different obtained classifiers were compared by analyzing one by one each variable. For example, classifiers obtained:

- with the same input features, from images with the same radiometric pre-processing but different post processing, or
- from images with the same radiometric pre-processing, the training set learned on the same number of classes, post processed in the same way but with the use, or not, of the multitemporal stack, or
- from images with the same radiometric pre-processing, post processed in the same way, the training set learned with the same temporal extension but on different number of classes, and so on.

Results, available in Table 23 to Table 28, are combined in the following graphs according to examined variable.

4.4.1 Post-processing

Results analysis starts from the comparison between accuracy reached by the thematic maps as obtained from the classification process, and the ones that were post-processed with a majority filter with kernel size 3 x 3. Despite post-processing is only the final step of the classification process, results about its employment are placed before. The choice is aimed at simplifying the interpretation of the variables examined in the following sections.

In order to evaluate improvement due to the application (or not) of the post-processing, the analyzed parameters in Asti case study state what follow:

- Overall accuracy: post-processed classifications provided better results in terms of Overall Accuracy (Figure 56). Taking into consideration all classifications, Overall Accuracy in post-processed thematic maps was, on average, 2.64% higher than other, up to a maximum of 5.34%. This trend occurred in 100% of cases (Table 30).
- f-measure: post-processed classifications provided better results in terms of f-measure (Figure 57) in 78% of cases. The mean improvement obtained in these cases was 4.70%, up to a maximum of 7.64% (Table 30). In remaining 22% of cases (8 occurrences), where classifications without post-processing provided better results, f-measure was very scarce: in 2 cases lower than 15%, in 5 cases between 15% and 30%, in one case between 30% and 40% (better f-measure were round 48%).
- Producer's Accuracy: post-processed classifications provided better results in terms of PA (Figure 58) in 81% of cases. The mean improvement obtained in these cases was 4.05%, up to a maximum of 9.48% (Table 30). In remaining 19% of cases (7 occurrences), where classifications without post-processing provided better results, f-measure was very low: in

2 cases lower than 30%, in 2 cases between 30% and 40%, in one case between 40% and 55% (better PA are round 75%).

- User's Accuracy: post-processed classifications provided better results in terms of UA (Figure 59) in 75% of cases. The mean improvement provided in these cases was 10.36%, up to a maximum of 23.9% (Table 30). In remaining 25% (9 occurrences), classifications without post-processing provided better results, but with very low UA: in 8 cases lower than 30%, in 1 case between 30% and 40% (better UA were more than 80%).

Post-processed thematic maps were particularly better than others in 2-class Single Image Classifiers.

From the analysis resulted that the application of a majority filtering on the performed classifications generally improved all accuracy measures. The major improvement was in User's Accuracy: the wrong allocation of pixels in the urban class was rather reduced with the employment of post-processing.

The behaviour observed on the subset was approximately the same observed on the whole validation; a summary of the results is provided in Table 30.

Figure 56 - Overall Accuracy comparison in the thematic maps obtained with or without post-processing. On the x-axis, classifier names are shortened as follows: 1: single image classifier; n: multitemporal classifier; I: only indexes; S: only spectral information; IS: both indexes and spectral information; 2cl: 2 classes; 5cl: 5 classes.
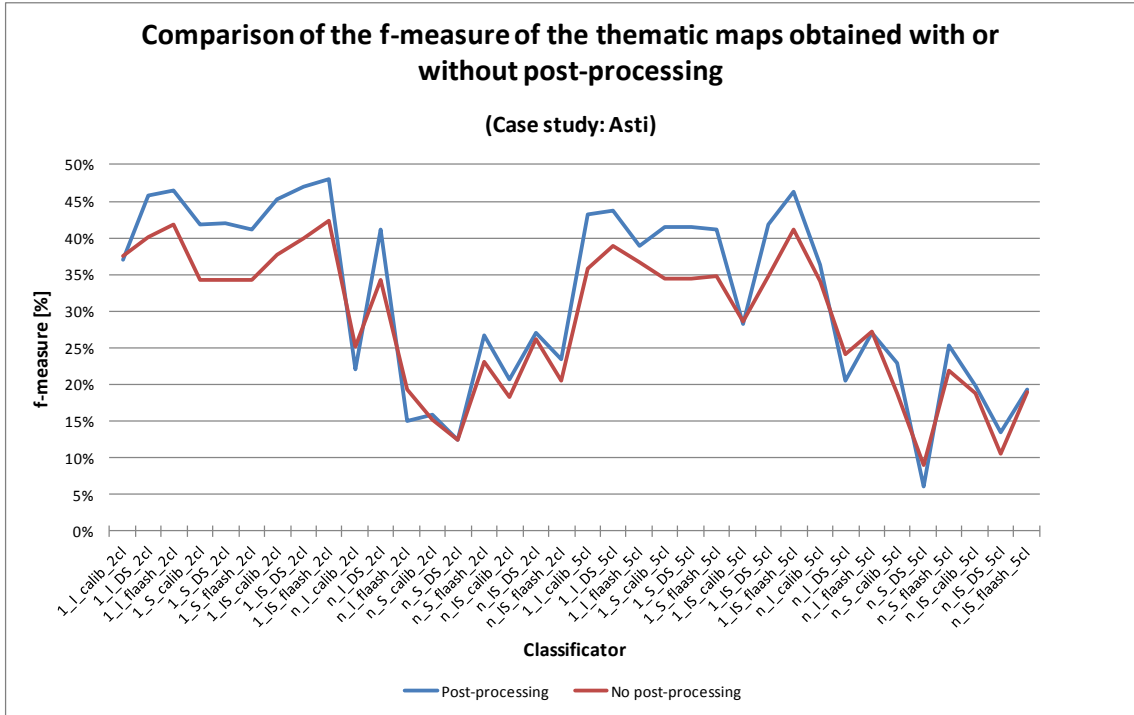
Figure 57 - f-measure comparison in the thematic maps obtained with or without post-processing. On the x-axis, classifier names are shortened as follows: 1: single image classifier; n: multitemporal classifier; I: only indexes; S: only spectral information; IS: both indexes and spectral information; 2cl: 2 classes; 5cl: 5 classes.
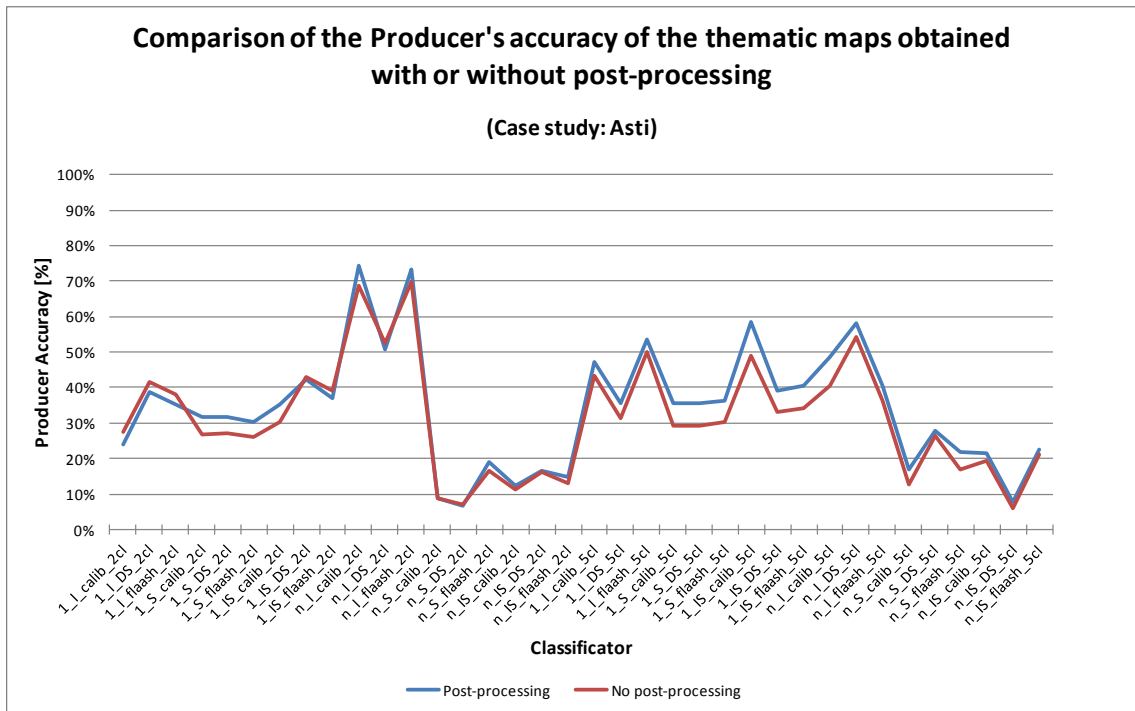
Figure 58 - PA comparison in the thematic maps obtained with or without post-processing. On the x-axis, classifier names are shortened as follows: 1: single image classifier; n: multitemporal classifier; I: only indexes; S: only spectral information; IS: both indexes and spectral information; 2cl: 2 classes; 5cl: 5 classes
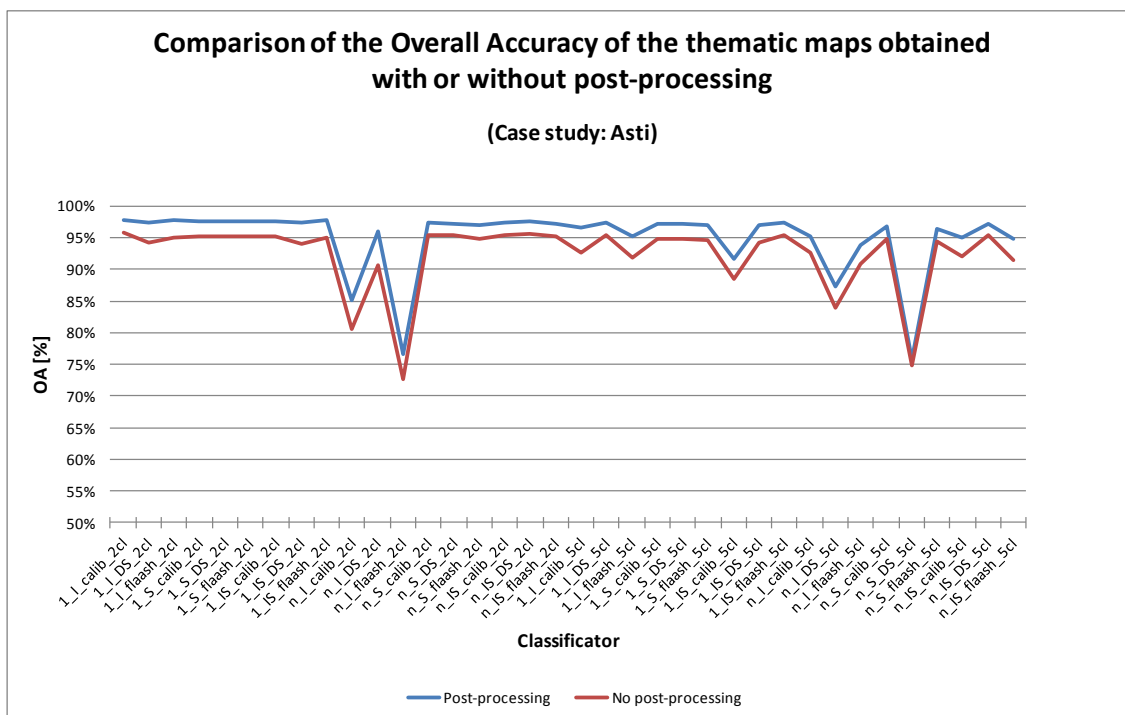
125
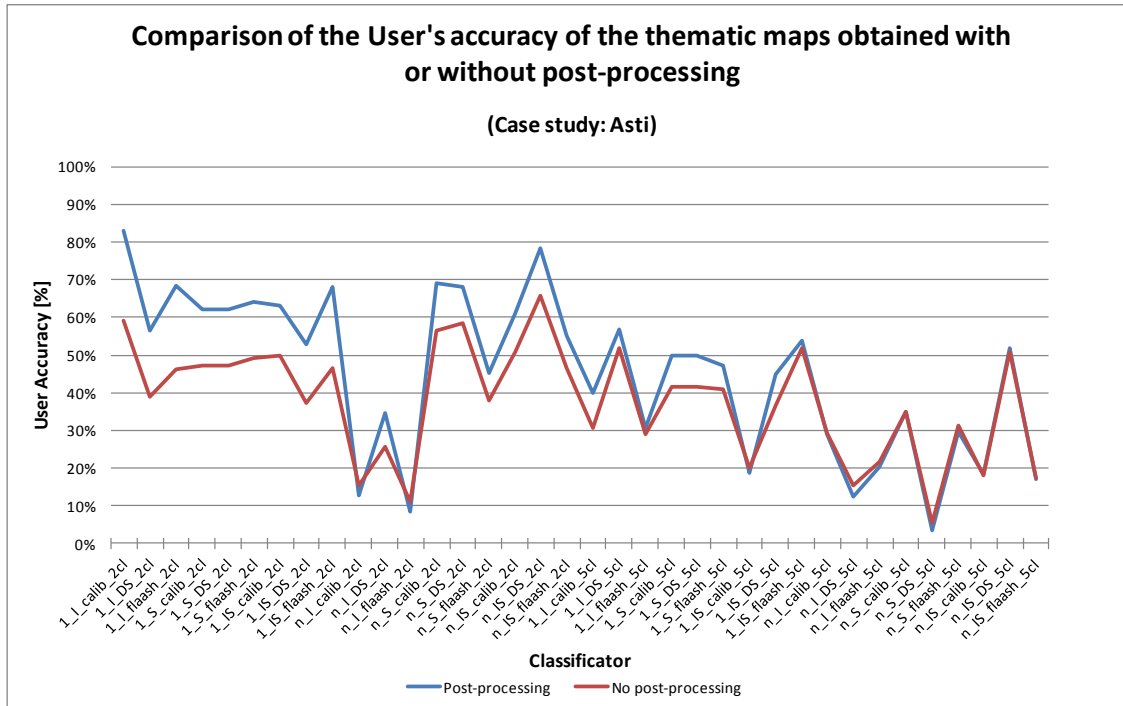
Figure 59 - UA comparison in the thematic maps obtained with or without post-processing. On the x-axis, classier names are shortened as follows: 1: single image classier; n: multitemporal classifier; I: only indexes; S: only spectral information; IS: both indexes and spectral information; 2cl: 2 classes; 5cl: 5 classes

The above mentioned analysis was performed for Torino case study too.

The analyzed parameters stated what follows:

- Overall accuracy: post-processed classifications provided better results in terms of Overall Accuracy in 64% of cases (Figure 60). When post-processed classifications   provided better result, Overall Accuracy was, on average, 0.95% better; on the other hand, when classifications without post-processing provided better result, Overall Accuracy was, on average, 1.57% better (Table 30).

- f-measure: classifications without post-processing provided better results in terms of f-measure in 94% of cases (Figure 61). When post-processed classifications provided better result, f-measure was, on average, 0.18% better; on the other hand, when classifications without post-processing provided better result, f-measure was, on average, 1.53% better (Table 30). In both cases, improvements provided by the performed choice are very low.

- Producer's Accuracy: post-processed classifications provided better results in terms of PA (Figure 62) in 89% of cases. When post-processed classifications provided better results, PA was, on average, 0.91% better; on the other hand, when classifications without post-processing provided better result, Overall Accuracy was, on average, 0.55% better (Table 30). In both cases, improvements provided by the performed choice are very low.

- User's Accuracy: classifications without post-processing always provided better results in terms of UA (Figure 63). Taking into consideration all classification, UA in thematic maps without post-processing was, on average, 2.36% higher than other (Table 30).

From the analysis it resulted that the application of a majority filtering on the resulting classifications caused an improvement in Producer's Accuracy and a decrease of User's Accuracy. The increase or decrease of the different measures of accuracy was, anyway, quite low for the considered case study.
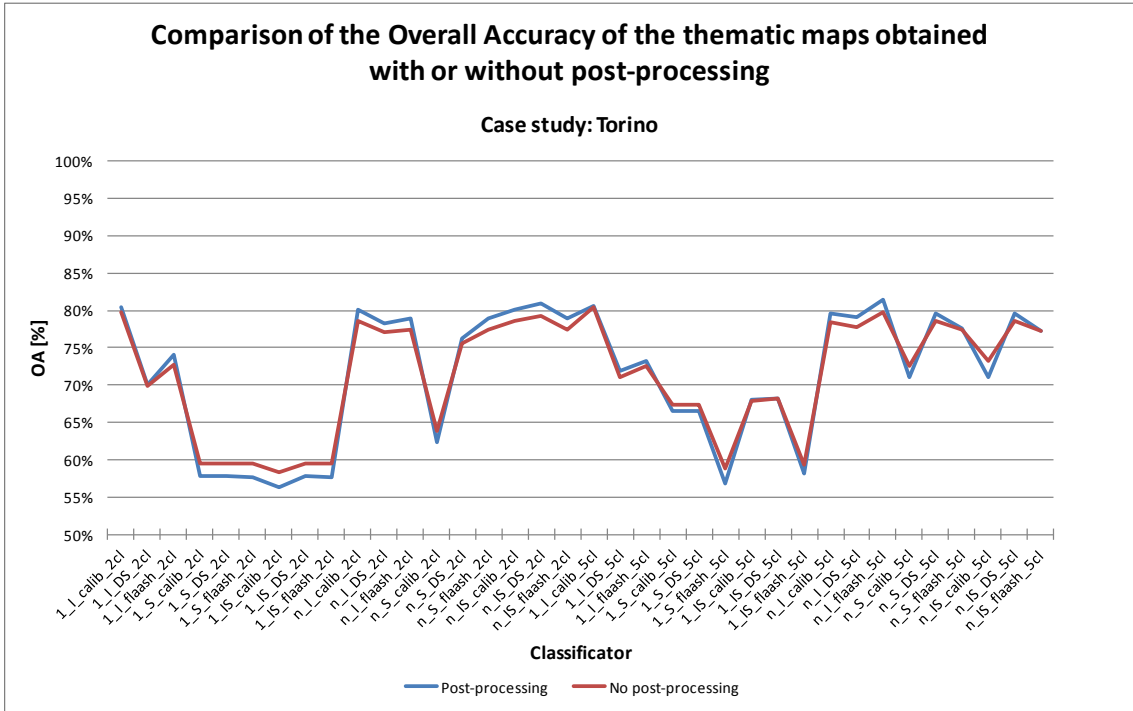
Figure 60 - Comparison of the different thematic maps with or without post-processing on the basis of Overall Accuracy
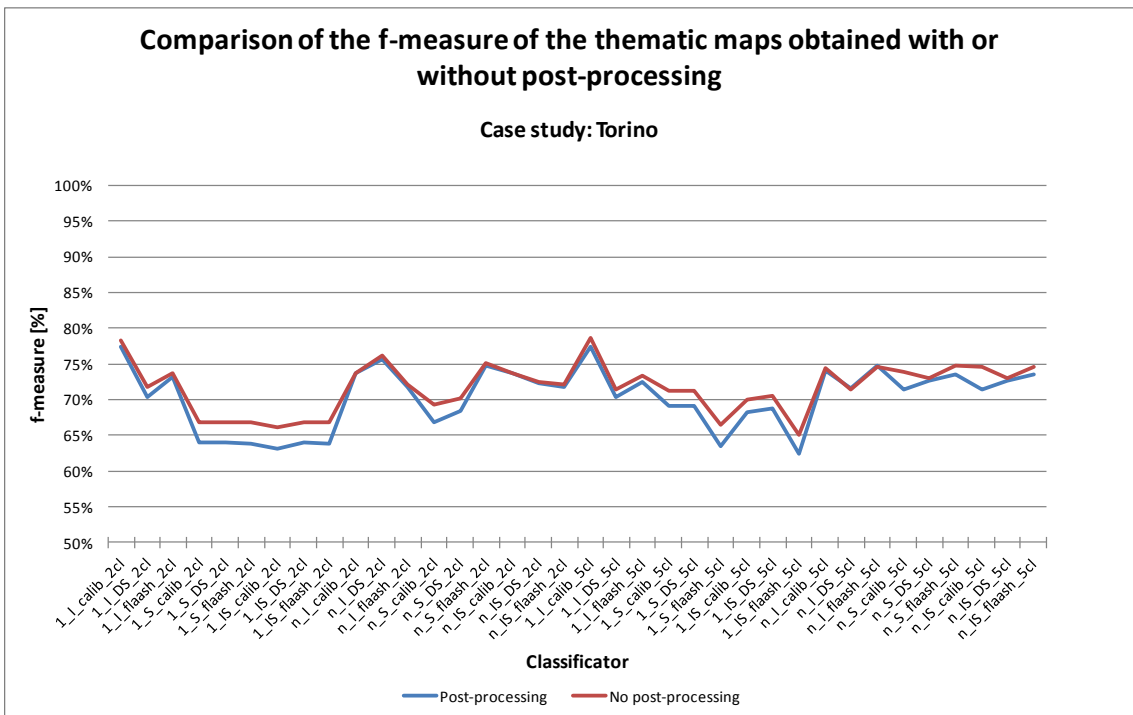
Figure 61 - Comparison of the different thematic maps with or without post-processing on the basis of the f-measure
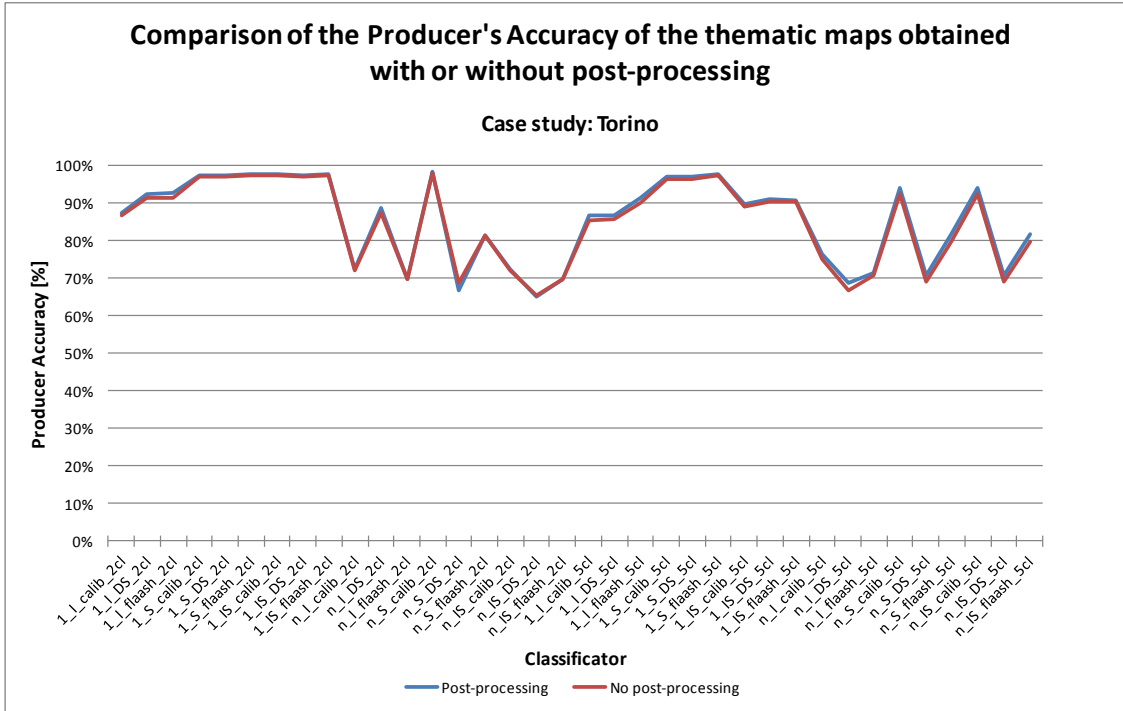
Figure 62 - Comparison of the different thematic maps with or without post-processing on the basis of the Producer's Accuracy
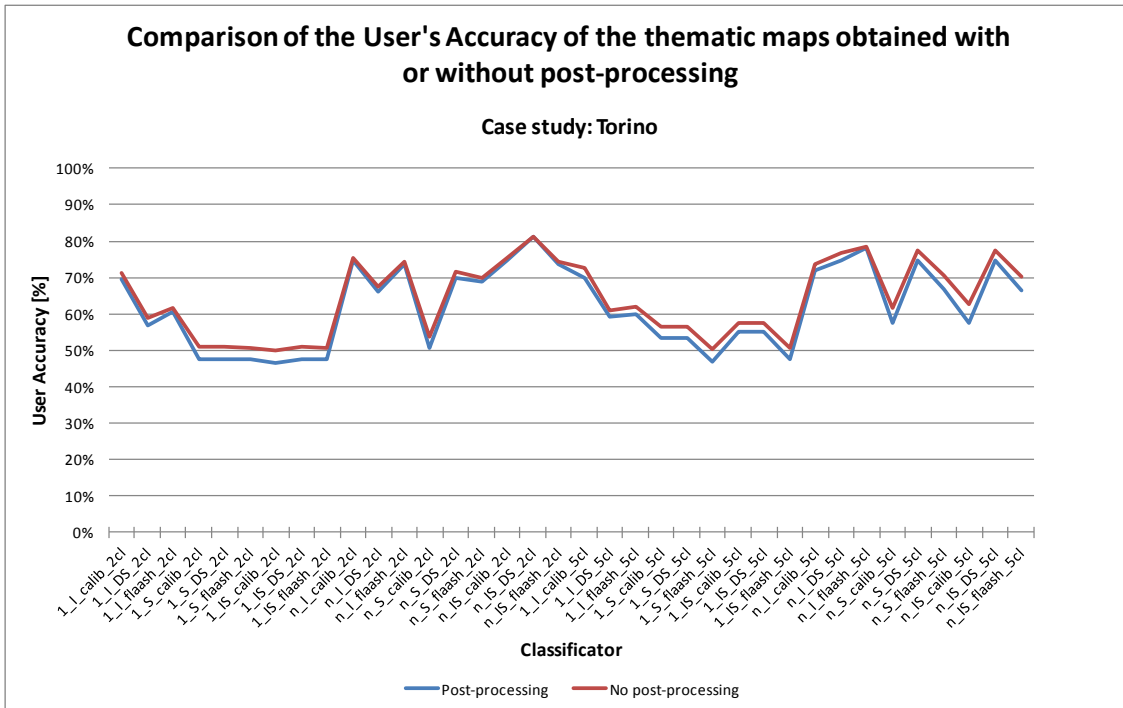
Figure 63 - Comparison of the different thematic maps with or without post-processing on the basis of the User's Accuracy

The application of post-processing caused, in general, two connected effects: the increase of the Minimum Mapping Unit both in the validation mask and in the classification itself.

In Asti case study, the improvement in the use of post-processing was evident: where medium density urban land cover and very sparse urban land cover is present, it's necessary to increase the Minimum Mapping Unit, because single urban pixels are difficult to detect and their identification can also be unuseful.

Where urban density is very high, the application of a post-processing can, instead, be avoided.

Table 30 provides a summary of the above mentioned results. In particular, it's aimed at enhancing:

- *Better results with post-processing*: number of times (expressed as % on the whole) in which classifications with post-processing provided better results than without post-processing;
- *Better results without post-processing*: number of times (expressed as % on the whole) in which classifications without post-processing provided better results than without post-processing;

For both of these groups:

- *Mean improvement*: Mean improvement of considered measure of accuracy calculated only on classification that provided better results with post-classification (or without). Minus sign is insered when results without post-processing were better than others;

In general:

- Mean: mean improvement calculated on all classifications;
- Max: maximum improvement (in cases in which post-processed classifications are better than other);
- Min: minimum improvement (in cases in which not post-processed classifications are better than other).

	OA			f-measure			PA			UA		
	TO	AT	AT subset	TO	AT	AT subset	TO	AT	AT subset	TO	AT	AT subset
Better results with post-processing	64%	100%	100%	6%	78%	89%	89%	81%	100%	0%	75%	86%
mean improvement	0.95%	2.64%	3.27%	0.18%	4.70%	6.89%	0.91%	4.05%	6.12%	-	10.36%	8.11%
Better results without post-processing	36%	0%	0%	94%	22%	11%	11%	19%	0%	100%	25%	14%
mean improvement	-1.57%	-	-	-1.53%	-1.86%	-2.11%	-0.55%	-1.96%	-	-2.36%	-1.71%	-2.00%
General improvement												
mean	0.04%	2.64%	3.27%	-1.44%	3.24%	5.89%	0.75%	2.88%	6.12%	-2.36%	7.34%	6.70%
max	1.75%	5.34%	5.66%	0.21%	7.64%	10.91%	2.09%	9.48%	11.41%	-0.01%	23.90%	16.64%
min	-2.20%	1.01%	1.06%	-3.31%	-4.34%	-3.76%	-2%	-3.64%	0.63%	-5.18%	-3.05%	-3.66%

Table 30 – Information extracted from the comparison between the same classifiers, with or without post-processing. The comparison was performed taking into consideration 4 measures of accuracy: Overall Accuracy, f-measure, Producer's Accuracy, User's Accuracy

4.4.2 Radiometric pre-processing

In order to detect which kind of radiometric pre-processing allowed to obtain better accuracy in the classification phase, if any, classifiers learned with the same number of classes, the same extension of the training set, the same attributes, but different kind of radiometric pre-processing were compared. As mentioned in section 3.2.3, images were calibrated into reflectance values, or depured from atmospheric effects with Dark Subtraction or with FLAASH.

Post-processed classifications were compared; OA and f-measure were used for the comparison.

Results state that:

- In Asti case study, calibrated images provided better results then other in term of Overall Accuracy in 4 cases, corrected with Dark Subtraction images provided better results then other in 4 cases, while corrected with FLAASH images were better in 3 cases (Figure 64). For what f-measure is concerned, calibrated images provided better results in 2 cases, corrected with Dark Subtraction images provided better results then other in 4 cases, while corrected with FLAASH images were better in 5 cases (Figure 65).
- In Torino case study, calibrated images provided better results then other in term of Overall Accuracy in 5 cases, corrected with Dark Subtraction images provided better results then other in 3 cases, while corrected with FLAASH images were better in 2 cases (Figure 66). For what f-measure is concerned, calibrated images provided better results in 3 cases, corrected with Dark Subtraction images provided better results then other in 3 cases, while corrected with FLAASH images were better in 4 cases (Figure 67).

From the analysis it results that the distribution of best accuracies is independent from the type of applied radiometric pre-elaboration.

Cases study without post-processing and "Asti subset" case study were examined but not reported since the trend was very similar to the one described above.
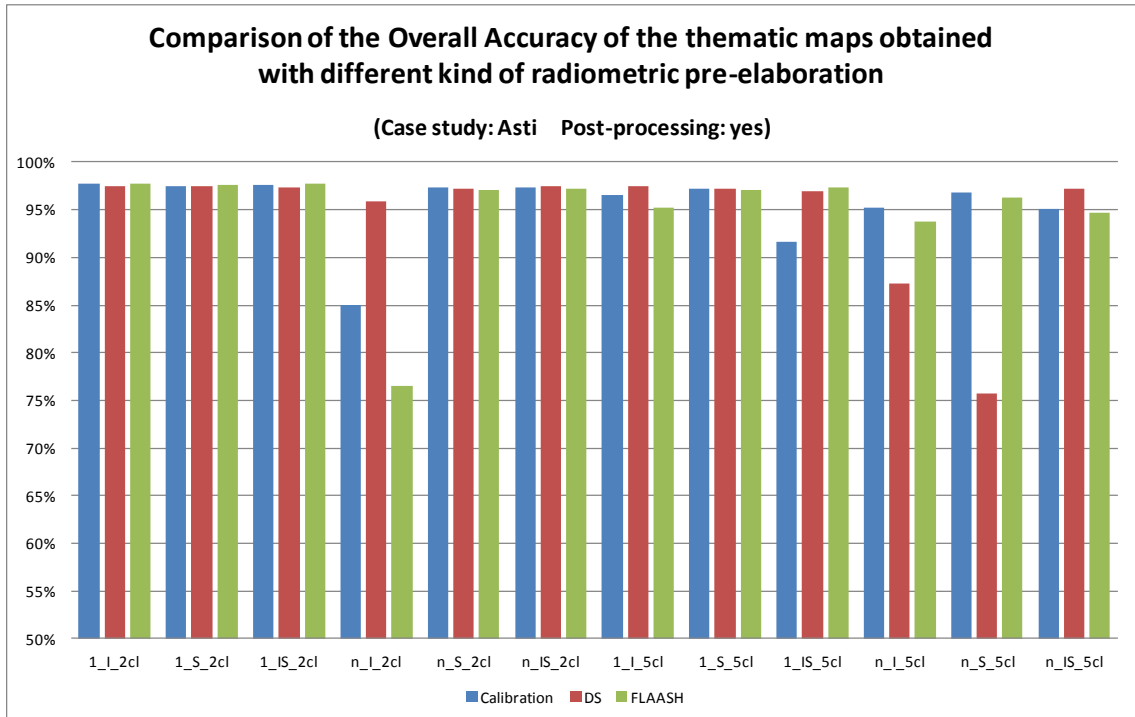
Figure 64 – Overall Accuracy comparison in thematic maps obtained with different kind of radiometric pre-processing
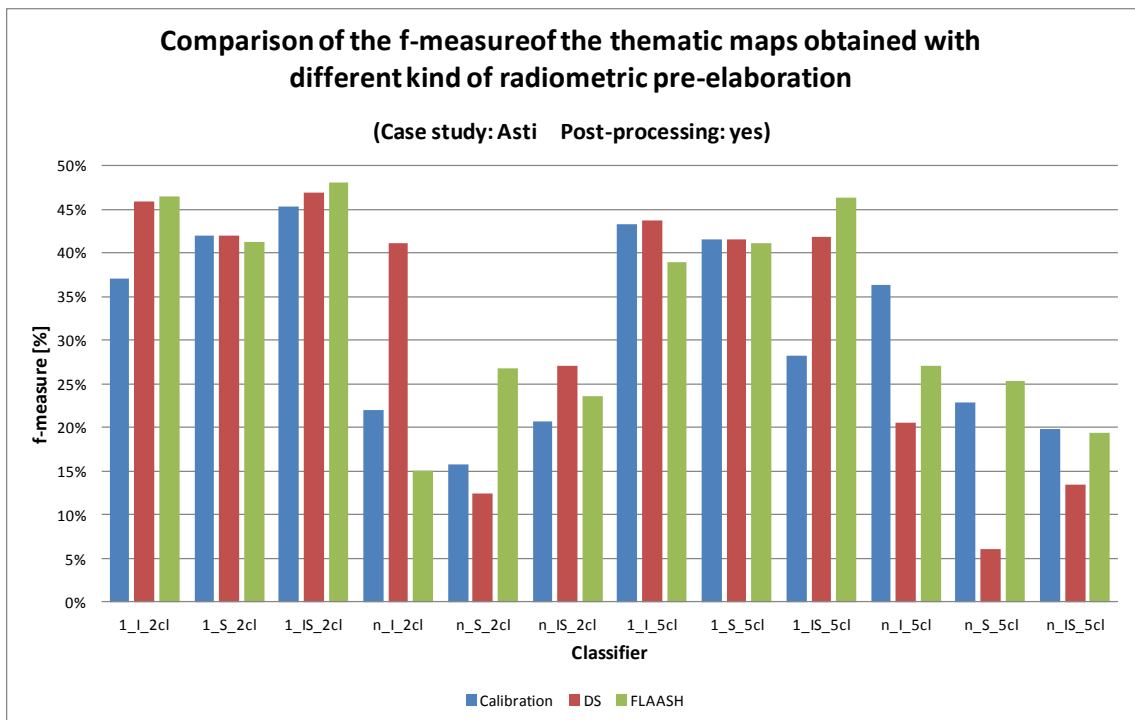
Figure 65 – f-measure comparison in thematic maps obtained with different kind of radiometric pre-processing
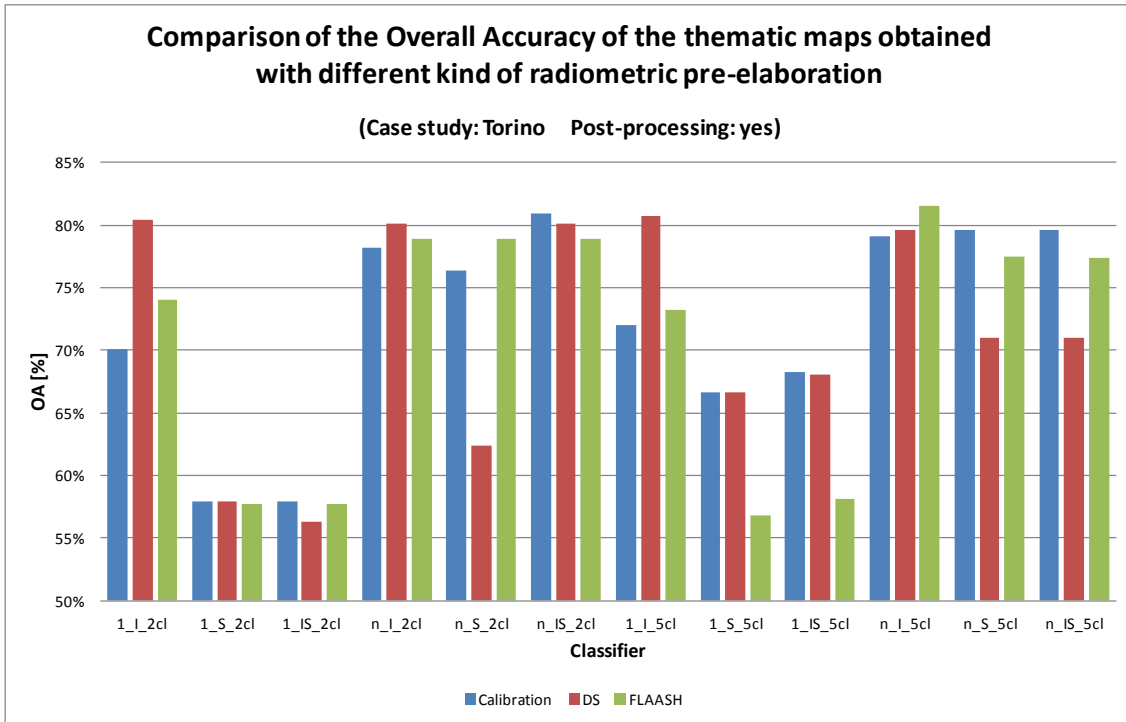
Figure 66 - Overall Accuracy comparison in thematic maps obtained with different kind of radiometric pre-processing
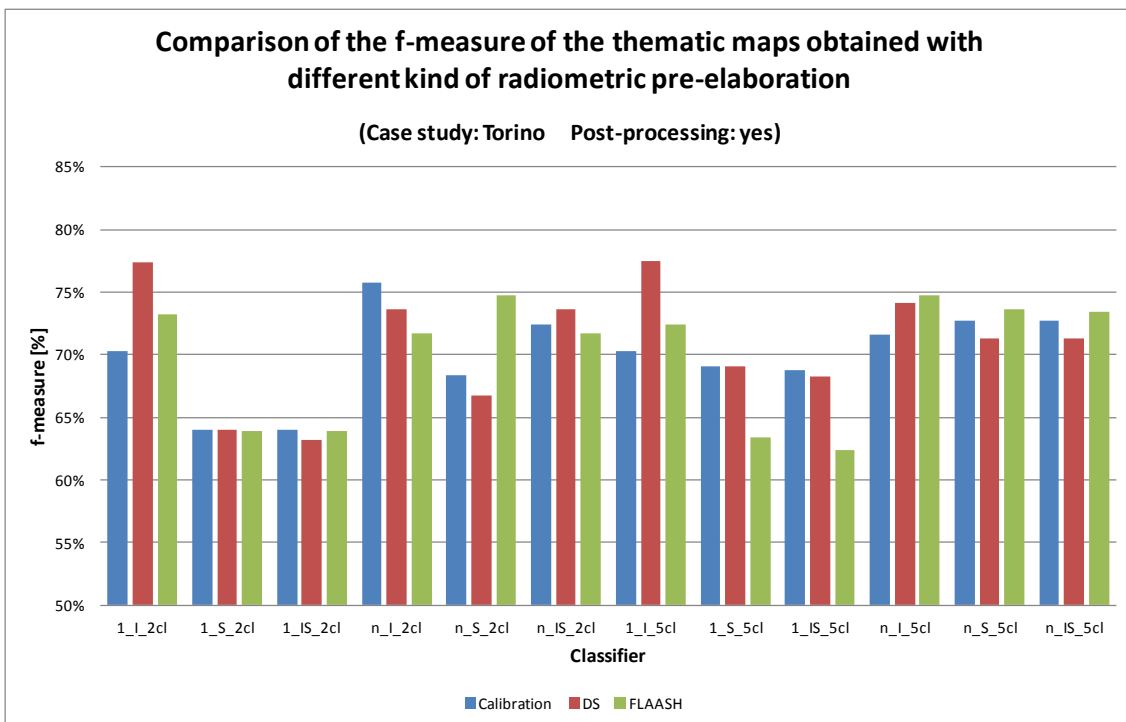
Figure 67 - f-measure comparison in thematic maps obtained with different kind of radiometric pre-processing

4.4.3 Number of classes

In order to understand if it's better to learn the classifier on two or more classes, accuracies reached from the classification obtained with the two kinds of learned-classifiers were compared. Overall Accuracy and f-measure were used for the comparison. The influence of the chosen number of classes was evaluated on both post-processed and not-post processed classifications and, in Asti case study, on both the whole area and the subset.

Results (Table 31) state that:

- In Asti case study, OA of "2-class" classifiers was better in 89% of cases (Figure 68), while f-measure was better in 78% of cases (Figure 69) in post-processed classifications.
  Considering only a subset, OA of "2-class" classifiers was better in 66% of cases (Figure 68), while f-measure was better in 33% of cases (Figure 69).
  When no post-processing was applied, the trend was similar to the above mentioned (graphs were not provided but results are summarized in Table 31), with, in general, a reduced number of occurrences: OA of "2-class" classifiers was better in 72% of cases, while f-measure was better in 61% of cases when the validation was performed on the whole area; with the validation performed on a subset, OA of "2-class" classifiers was better in 55% of cases, while f-measure was better in 44% of cases.

- In Torino case study, OA of "5-class" classifiers was better in 61% of cases (Figure 70), while f-measure was better in 61% of cases (Figure 71).
  When no post-processing was applied, OA of "5-class" classifiers was better in 61% of cases, while f-measure was better in 67% of cases (results are summarized in Table 31).
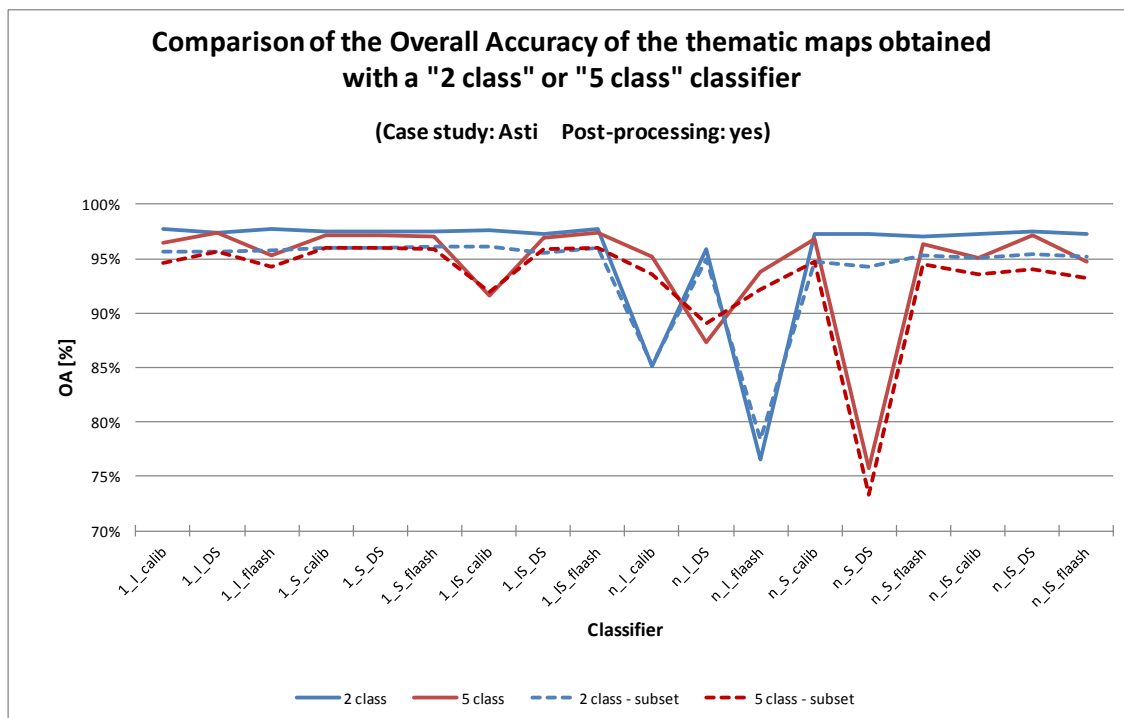
Figure 68 - Overall Accuracy comparison in post-processed thematic maps obtained with a "2-class" or "5-class" classifiers
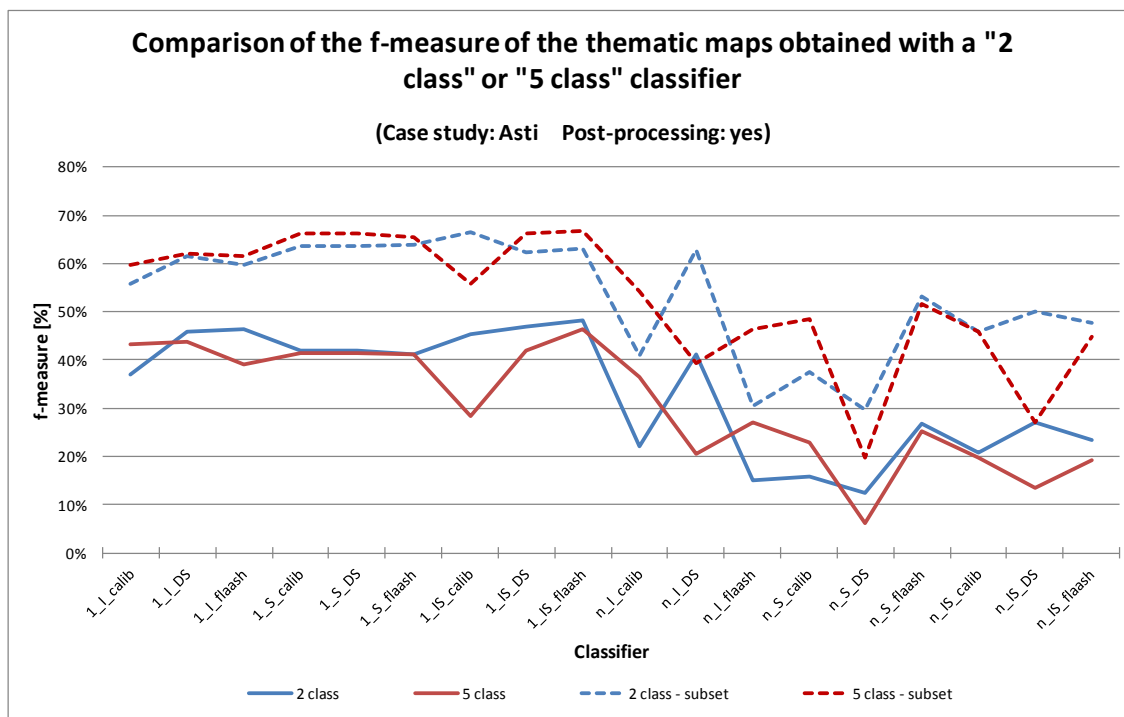
Figure 69 – f-measure comparison in post-processed thematic maps obtained with a "2-class" or "5-class" classifiers
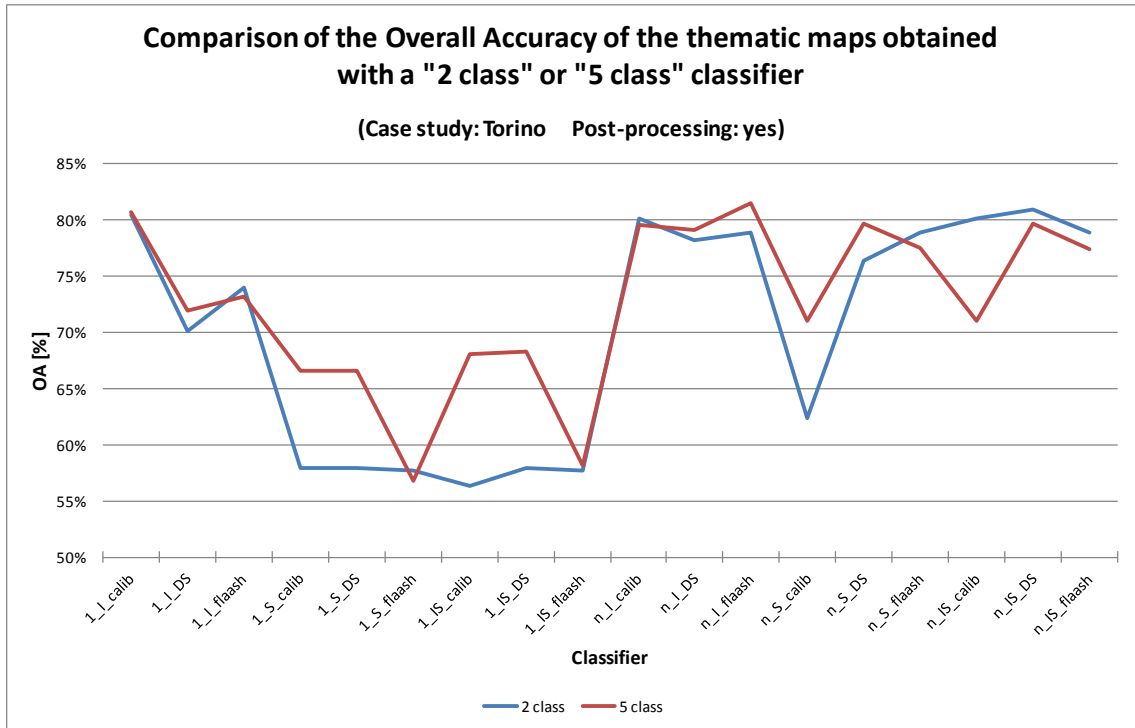
Figure 70 - Overall Accuracy comparison in post-processed thematic maps obtained with a "2-class" or "5-class" classifiers
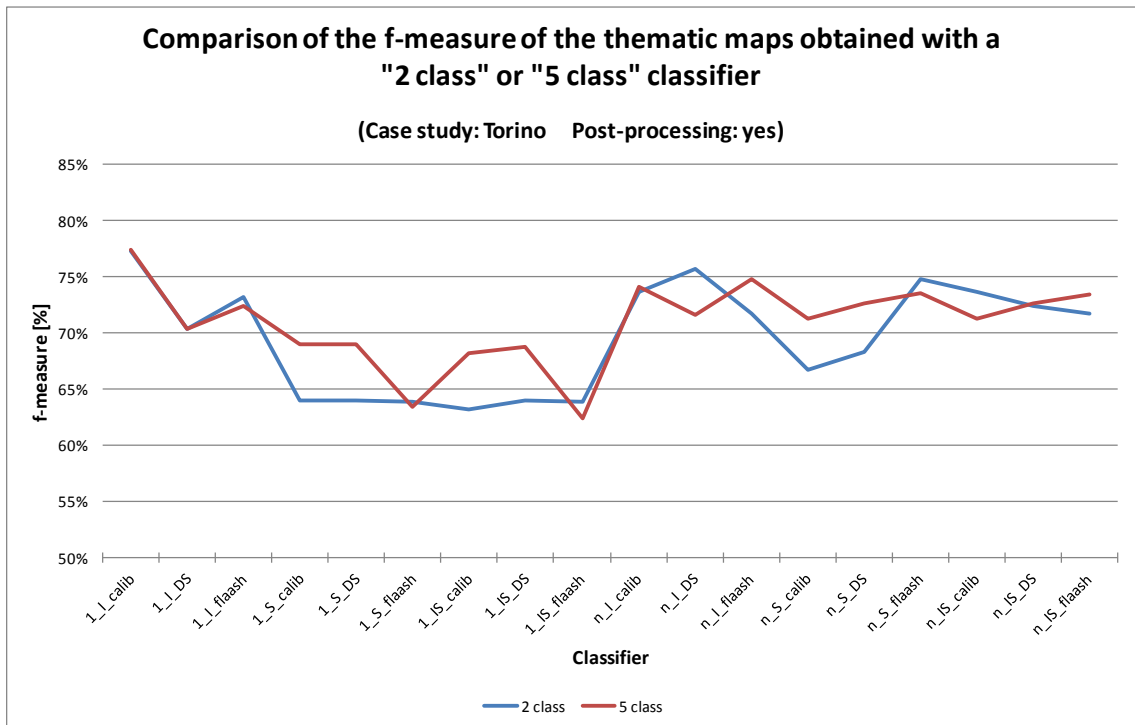
Figure 71 - f-measure comparison in post-processed thematic maps obtained with a "2-class" or "5-class" classifiers

Table 31 provides a summary of the above mentioned behaviour, with information about the mean improvement provided in the different cases. In particular, it's enhanced:

- *Better results with 2 class*: number of times (expressed as % on the whole) in which classifiers learned with 2-class provided better results than when learned with 5-class;
- *Better results with 5 class*: number of times (expressed as % on the whole) in which classifiers learned with 5-class provided better results than when learned with 2-class;

For both these groups:

- *Mean improvement*: Mean improvement of considered measure of accuracy calculated only on classifications that provide better results with 2 class (or with 5);

From a more detailed analysis it can be gathered that:

- 2-class classifiers provided better results both in terms of Overall Accuracy and f-measure in Asti case study; the improvement obtained with the 2-class classifiers was particularly evident in post-processed classifications (better results with post-processing in 89% (OA) and 78% (f-measure) of cases). In classifications without post-processing, the statement was true in a reduced number of times (72% (OA) and 61% (f-measure) of cases).
  The mean improvement obtained when 5-class classifiers resulted best, is quite high (OA in post-processed case: -13.71%): this is due to two high gaps in two cases ("TO_multitemp_index_calib" and "TO_multitemp_index_flaash"). Otherwise, all other differences are round equal to zero.
  When a subset was considered for validation, the difference in using a 5-class classifier or a 2-class classifier was reduced; 2-class classifiers were better then other in 66% (taking into consideration OA) and 34% (taking into consideration f-measure) of cases in post-processed classifications, and in 55% (OA) and 44% (f-measure) of cases in classifications without post-processing
- Classifiers learned with a 5-class training set provided better results both in terms of Overall Accuracy and f-measure in Torino case study; 5-class classifiers were better than other in 61% of cases in post-processed classifications (for both OA and f-measure), and in 61% (OA) and 67% (f-measure) of cases when no post-processing was applied.

In general, the use of a 2-class classifier was better in Asti case study, particularly in classifications without post-processing, while the use of a 5-class classifier resulted better in Torino case study.

	Post processing: YES						Post processing: NO					
	OA			f-measure			OA			f-measure		
	TO	AT	AT subset	TO	AT	AT subset	TO	AT	AT subset	TO	AT	AT subset
Better results with 2-class	39%	89%	66%	39%	78%	34%	39%	72%	55%	33%	61%	44%
mean improvement	2.25%	3.01%	3.29%	1.48%	5.84%	11.96%	1.03%	3.85%	4.03%	1.29%	5.08%	6.28%
Better results with 5-class	61%	11%	34%	61%	22%	66%	61%	28%	45%	67%	39%	56%
mean improvement	5.22%	13.71%	3.74%	3.13%	9.92%	5.11%	4.63%	6.39%	3.42%	2.58%	3.10%	4.69%

Table 31 - Information extracted from the comparison between classifiers learned with same temporal extension of the training set, same performed radiometric pre-elaboration and same features but with different number of classes. The comparison was performed taking into consideration Overall Accuracy and f-measure, both for post-processed and not post-processed classifications.

4.4.4 Use of a multitemporal stack

In this section, classifiers learned with the same number of classes, the same kind of features and the same radiometric pre-elaborations but different extension of the training set were compared, both for post-processed and not post-processed classifications.

Results (Table 32) state that:

- In Asti case study, OA of "single image" classifiers was better in 83% of cases (Figure 72), with a mean improvement of 5.01%, while f-measure was better in 100% of cases (Figure 73) in post-processed classifications.

  Considering only a subset, OA of "single image" classifiers was better in 94% of cases (Figure 72), with a mean improvement of 4.36%, while f-measure was better in 94% of cases (Figure 73).

  When no post-processing was applied, the trend was similar to the above mentioned (graphs were not provided but results are summarized in Table 32), with, in general, a reduced number of occurrences: OA of "single image" classifiers was better in 55% of cases, with a mean improvement of 7.82%, while f-measure was better in 100% of cases when the validation was performed on the whole area; with the validation performed on a subset, OA of "single image" classifiers was better in 83% of cases, with a mean improvement of 5.03%, while f-measure was better in 83% of cases.

- In Torino case study, OA of "multitemporal" classifiers was better in 89% of cases (Figure 74), while f-measure was better in 83% of cases (Figure 75).

  When no post-processing was applied, OA of "multitemporal" classifiers was better in 89% of cases, while f-measure was better in 83% of cases (results are summarized in Table 32).

Graphs for the above mentioned case study without post processing were not been provided since they were very similar to the post processed ones in terms of differences between accuracies reachable with a single image or a multitemporal stack.
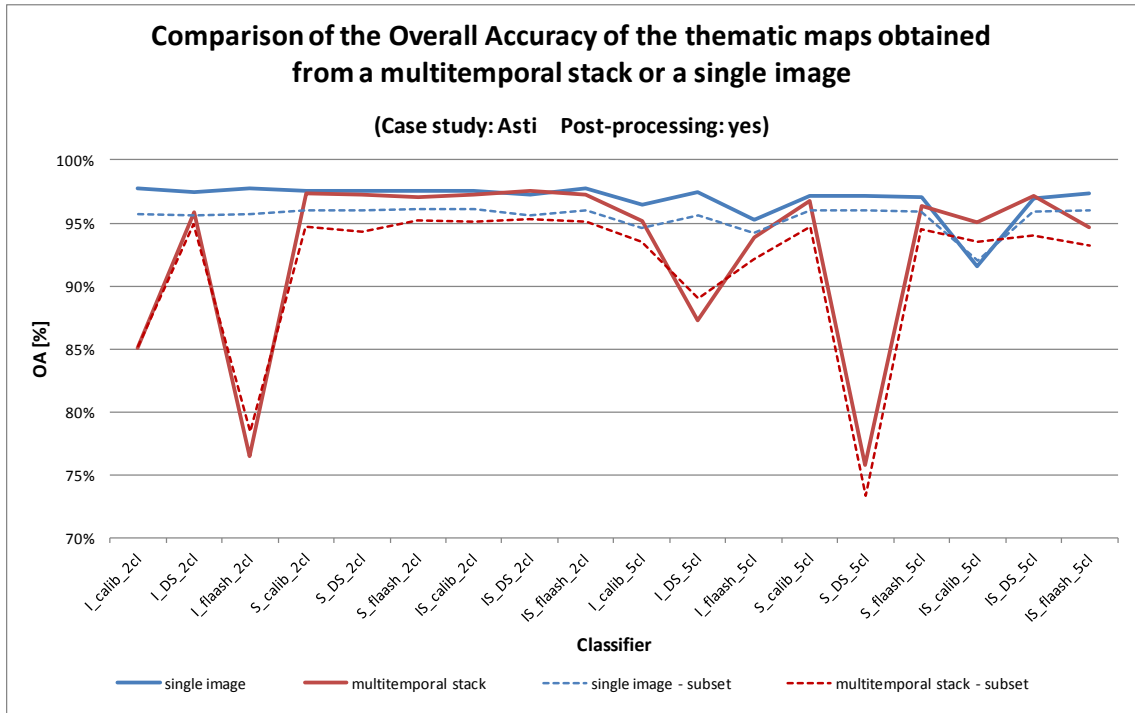
Figure 72 - Overall Accuracy comparison in post-processed thematic maps obtained with a "single image" or a "multitemporal" classifier
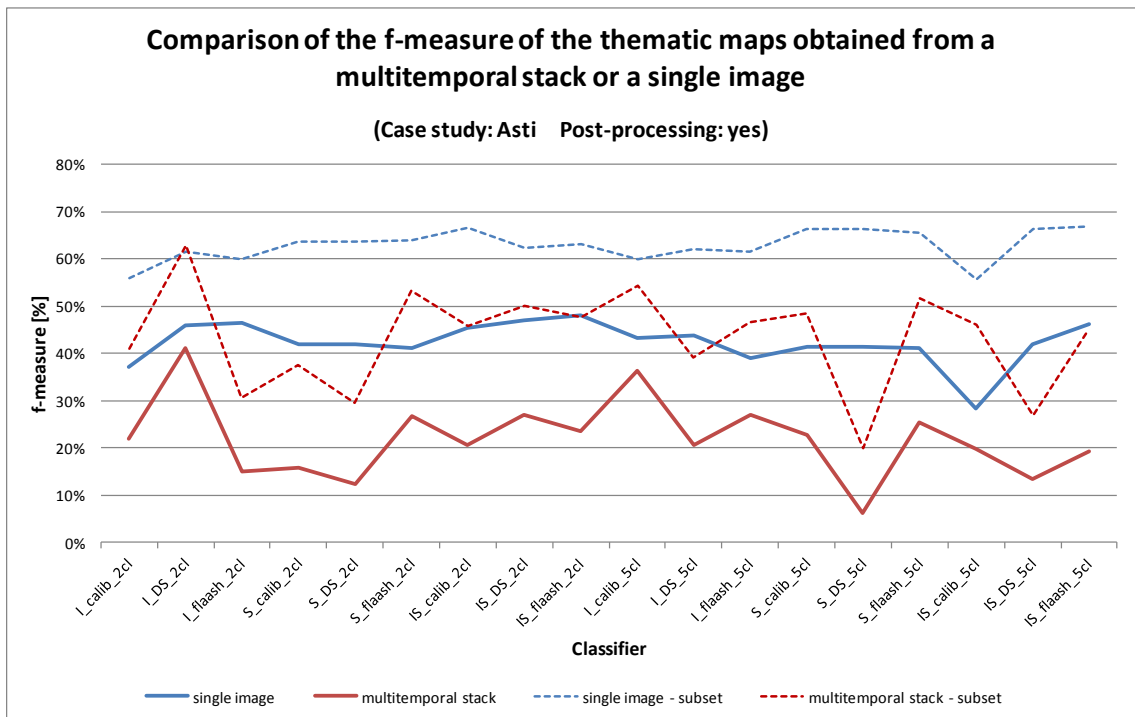
Figure 73 – f-measure comparison in post-processed thematic maps obtained with a "single image" or a "multitemporal" classifier
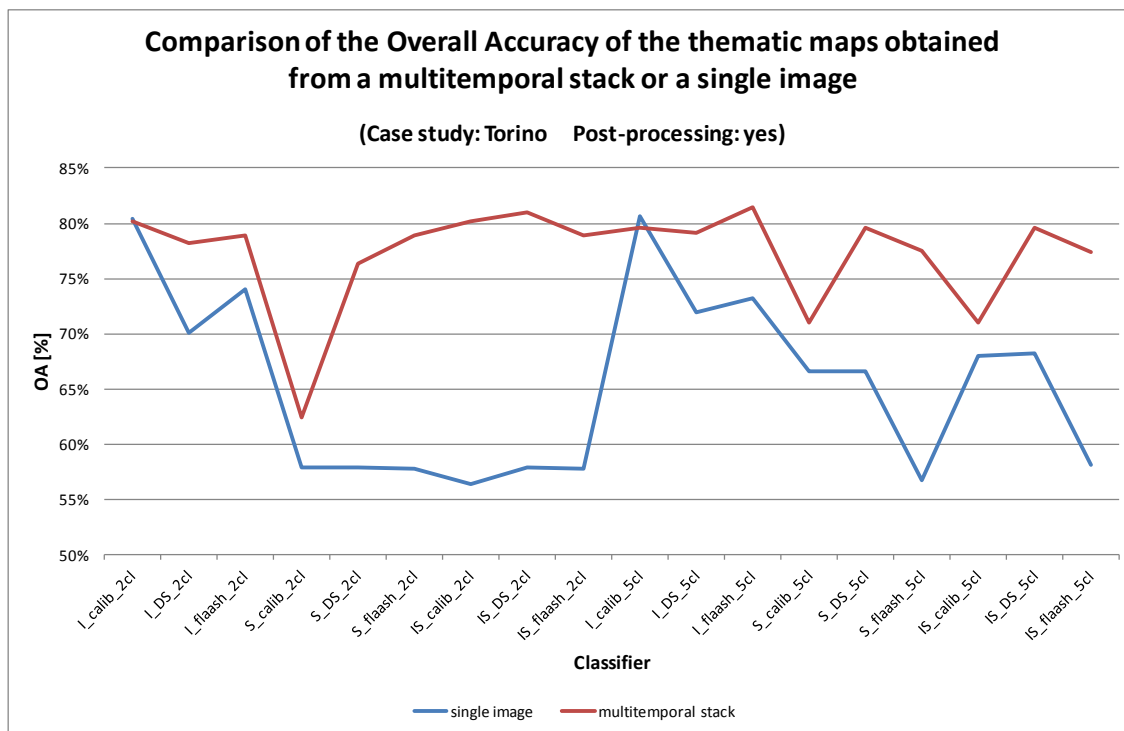
Figure 74 - Overall Accuracy comparison in post-processed thematic maps obtained with a "single image" or a "multitemporal" classifier
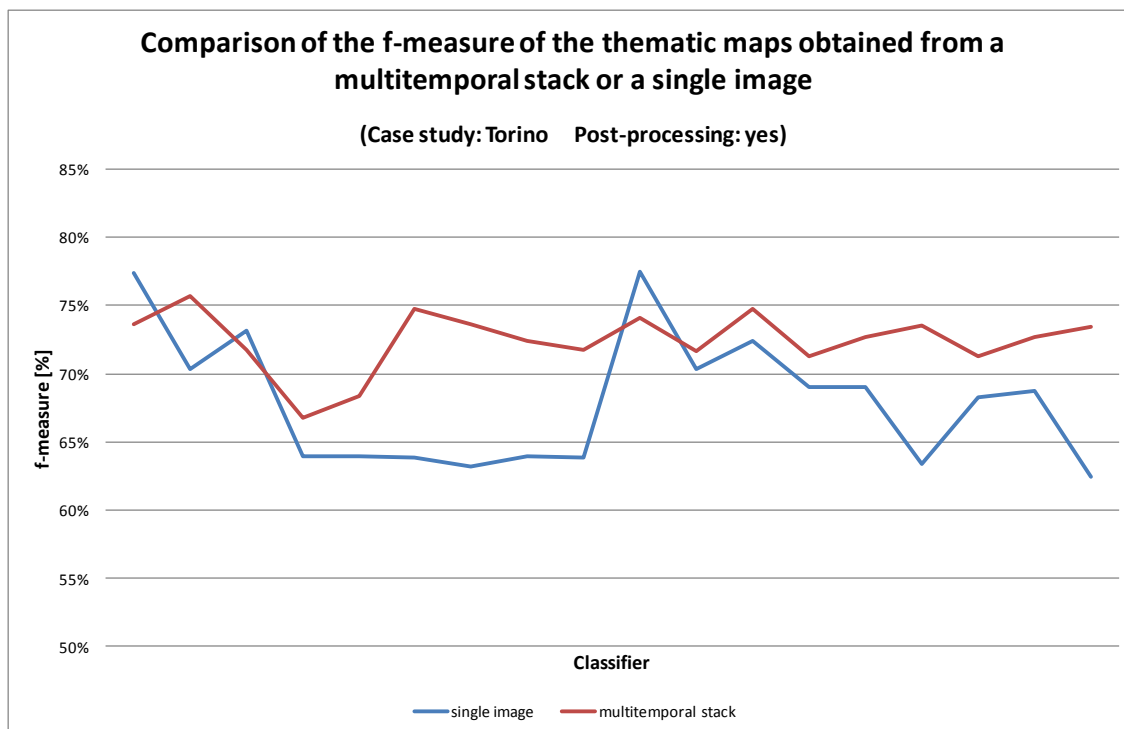
Figure 75 – f-measure comparison in post-processed thematic maps obtained with a "single image" or a "multitemporal" classifier

The use of a single image to learn the classifier allowed to obtain better performance in Asti case study, both in term of OA and f-measure; the improvement reached with the use of a single image

140

was about 5.01% (OA), till to a maximum of 21.38%. The behaviour was the same when the validation was performed only on the subset. Generally, results provided from the "single image" classifiers were more stable, while with the use of a multitemporal stack some peaks of very low OA occurred ("AT_singleimage_index_flaash_2class" and "AT_singleimage_spectral_DS_5class"). The trend was confirmed also when the f-measure was considered: in 100% of cases "single image" classifier provided better f-measure with a mean increase of about 20.33%! %. The behaviour was the same when the validation was performed only on the subset.

In Torino case study, the behaviour was exactly opposite: "multitemporal" classifiers provided better accuracies both in term of Overall Accuracy and f-measure. In 89% of cases, "multitemporal" classifiers provided better OA of, on average, 13.27%, till to a maximum of 23.82%! The difference was particularly evident in 2-class classifiers that used only spectral information or both indexes and spectral information. The f-measure was affected in the same way for what the number of cases is concerned, but with lower differences: about 5.86%, with a peak of 11.04%.

The trend was confirmed also in classifications without post-processing: Overall accuracy of single image classifiers was better only in 55% of cases, only because in 5 cases multitemporal classifiers were better than other of less then 0.2%.

	Post processing: YES						Post processing: NO					
	OA			f-measure			OA			f-measure		
	TO	AT	AT subset	TO	AT	AT subset	TO	AT	AT subset	TO	AT	AT subset
Better results with single image	11%	83%	94%	17%	100%	94%	11%	55%	83%	17%	100%	100%
mean improvement	0.68%	5.01%	4.36%	2.85%	20.33%	20.93%	1.72%	7.82%	5.03%	3.47%	15.78%	15.78%
Better results with multitemporal stack	89%	17%	6%	83%	-	6%	89%	45%	17%	83%	-	-
mean improvement	13.27%	1.32%	1.52%	5.86%	-	1.29%	11.90%	0.90%	1.02%	4.53%	-	-

Table 32 - Information extracted from the comparison between classifiers learned with same number of classes, same performed radiometric pre-elaboration and same features but with different temporal extension of the training set are provided. The comparison has been performed taking into consideration Overall Accuracy and f-measure, both for post-processed and not post-processed classifications.

4.4.5 Attributes

The last analysis was performed on the attributes in order to understand:

- The correlation, if any, between considered classifier (index, spectral, spectral and index) and reached accuracy;
- Which attributes better allows separating classes (and are chosen by the algorithm).

The first question was addressed by comparing accuracies reached by classifiers learned with the same number of classes, the same extension of the training set, the same radiometric pre-processing and post-processing, but different attributes. Classifiers learned with only spectral information, only indexes, or both were compared in Figure 76 to Figure 79.

Results state that:

- In Asti case study, "Spectral" classifiers and "Index and spectral" classifiers provided very similar results in terms of Overall Accuracy, better than the ones provided by "Index" classifiers (Figure 76).
  By the comparison of f-measure, it resulted that none of the classifiers was clearly better than others (Figure 77). In general, it can be stated that in classifiers where higher accuracy were reached (that is single image classifier), the behaviour of "spectral" and "index and spectral" classifier was quite similar.
- In Torino case study, "index classifiers" provided better results in most of the cases, both in terms of Overall Accuracy (Figure 78) and f-measure (Figure 79).

From the above-mentioned results, it's difficult to detect, in unambiguous way, which classifiers were better than other; the second part of the analysis was therefore focused on the single attributes. Which attributes were chosen by the J48 algorithm, on the basis of the training set, to be part of the classifier? Attributes and thresholds were chosen in order to maximize the GainRATIO, the splitting criterion used by the algorithm, as described in Section 2.1.2.

Table 35 and Table 36 show, in detail, all attributes used by each classifier; Table 33 and Table 34, instead, provide a summary of each attribute occurrence according to the information used by the classifier (indexes, spectral information or both).

Results in Table 33 and Table 34 state that:

- Between considered indexes, BUI was used a very limited number of times, both in Asti and Torino case study (respectively 4 and 3 times);
- NDVI and NDBBBI were widely used in both case study (respectively 15 and 17 in Asti case study, 14 and 10 in Torino case study);
- Test-index was considered quite often, more in Asti than in Torino case study (respectively 16 and 13 times);
- NDBI was considered quite often but only in Asti case study (14 times versus 5);
- Between considered spectral information, band 1 was widely used in both cases (respectively, 21 and 19 times);

- Band 4, 5 and 7 were widely used in both cases (21, 15, 15 times and 10, 12, 12 times), while band 2 and 3 were less used than other (12 and 5, 3 and 11).

Type of attribute used by the classifier	Spectral Information						Indexes				
	B1	B2	B3	B4	B5	B7	NDBI	NDVI	BUI	NDBBBI	test_index
Indexes							10	10	2	10	10
Spectral information	12	6	1	12	9	12					
Indexes and spectral information	9	6	4	9	6	3	4	5	2	7	6
TOT	21	12	5	21	15	15	14	15	4	17	16
							Case study: Asti				

Table 33 – Summary of attributes occurrence according to the used classifier (Asti case study)

Type of attribute used by the classifier	Spectral Information						Indexes				
	B1	B2	B3	B4	B5	B7	NDBI	NDVI	BUI	NDBBBI	test_index
Indexes							5	8	3	9	10
Spectral information	12	2	7	6	7	6					
Indexes and spectral information	7	1	4	4	5	6	0	6	0	1	3
TOT	19	3	11	10	12	12	5	14	3	10	13
							Case study: Torino				

Table 34 - Summary of attributes occurrence according to the used classifier (Torino case study)

	Classifier	Spectral Information						Indexes				
		B1	B2	B3	B4	B5	B7	NDBI	NDVI	BUI	NDBBBI	test_index
2 CLASS — AT single image	index_calib							X			X	X
	index_DS							X	X			X
	index_flaash							X	X		X	X
	spectral_calib	X			X	X	X					
	spectral_DS	X			X	X	X					
	spectral_flaash	X		X	X	X	X					
	spectral_index_calib	X		X	X						X	
	spectral_index_DS	X			X			X	X		X	X
	spectral_index_flaash							X		X	X	X
2 CLASS — AT multitemporal	index_calib								X		X	X
	index_DS							X				X
	index_flaash							X	X		X	
	spectral_calib	X			X		X					
	spectral_DS	X			X		X					
	spectral_flaash	X			X		X					
	spectral_index_calib	X			X						X	X
	spectral_index_DS	X				X						X
	spectral_index_flaash				X			X			X	
5 CLASS — AT single image	index_calib							X	X	X	X	X
	index_DS								X		X	X
	index_flaash							X	X	X	X	X
	spectral_calib	X	X		X	X	X					
	spectral_DS	X	X		X	X	X					
	spectral_flaash	X	X		X	X	X					
	spectral_index_calib	X	X		X	X	X					
	spectral_index_DS	X	X		X	X		X				X
	spectral_index_flaash	X	X		X	X			X	X	X	X
5 CLASS — AT multitemporal	index_calib							X	X		X	X
	index_DS							X	X		X	X
	index_flaash							X	X		X	
	spectral_calib	X	X		X	X	X					
	spectral_DS	X	X		X	X	X					
	spectral_flaash	X	X		X	X	X					
	spectral_index_calib		X	X	X		X		X		X	
	spectral_index_DS	X	X	X	X		X		X			
	spectral_index_flaash	X	X	X	X				X			
Case study: Asti												

Table 35 – Detail on attributes used by each classifier (Asti case study)

144

	Classifier	Spectral Information						Indexes				
		B1	B2	B3	B4	B5	B7	NDBI	NDVI	BUI	NDBBBI	test_index
2 CLASS — TO single image	index_calib							X			X	
	index_DS							X			X	X
	index_flaash							X			X	X
	spectral_calib	X				X	X					
	spectral_DS	X				X	X					
	spectral_flaash	X				X	X					
	spectral_index_calib		X				X				X	
	spectral_index_DS	X				X	X					
	spectral_index_flaash	X				X	X					
TO multitemporal	index_calib								X			X
	index_DS											X
	index_flaash								X			X
	spectral_calib	X	X	X	X							
	spectral_DS	X	X	X	X							
	spectral_flaash	X		X	X	X						
	spectral_index_calib								X			X
	spectral_index_DS			X					X			X
	spectral_index_flaash								X			X
5 CLASS — TO single image	index_calib								X	X		
	index_DS							X	X		X	X
	index_flaash							X	X		X	X
	spectral_calib	X		X		X						
	spectral_DS	X		X		X						
	spectral_flaash	X		X		X	X					
	spectral_index_calib	X		X		X			X			
	spectral_index_DS	X		X		X			X			
	spectral_index_flaash			X		X	X		X			
TO multitemporal	index_calib								X	X	X	X
	index_DS								X	X	X	X
	index_flaash								X	X	X	X
	spectral_calib	X			X		X					
	spectral_DS	X			X		X					
	spectral_flaash	X		X	X							
	spectral_index_calib	X			X		X					
	spectral_index_DS	X			X		X					
	spectral_index_flaash	X		X	X							

Case study: Torino

Table 36 - Detail on attributes used by each classifier (Torino case study)

145

Figure 76 – Overall Accuracy comparison in post-processed thematic maps obtained with "index", "spectral" or "spectral and index" classifiers

Figure 77 – f-measure comparison in post-processed thematic maps obtained with "index", "spectral" or "spectral and index" classifiers
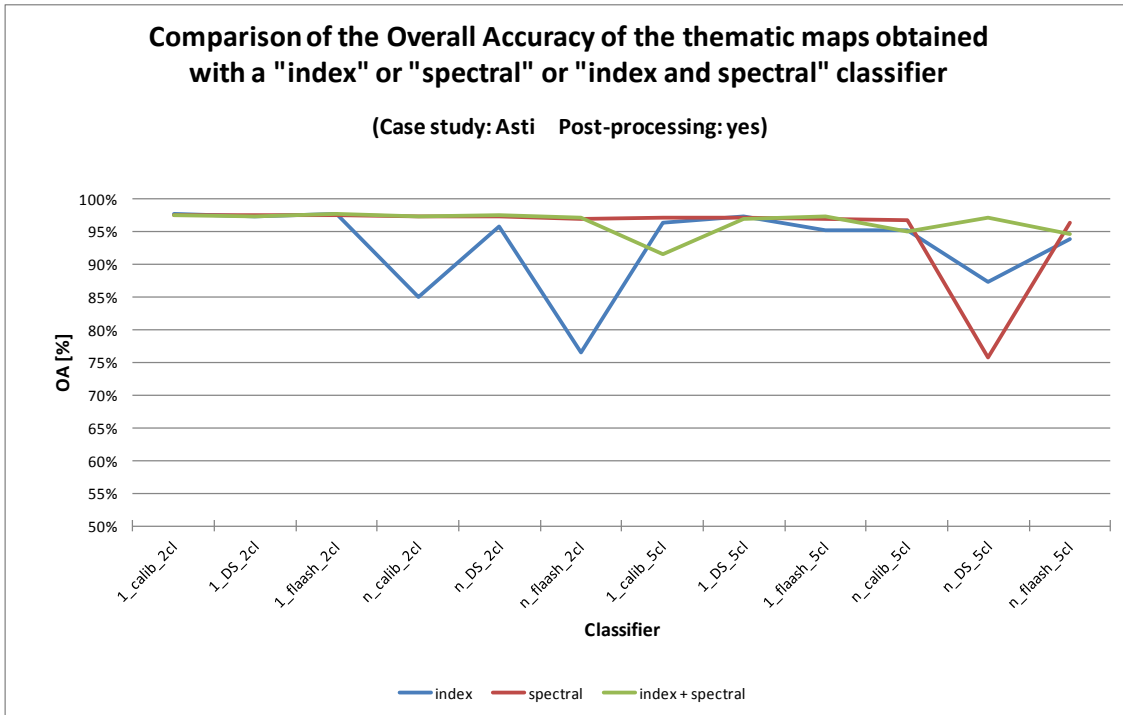
Figure 78 - Overall Accuracy comparison in post-processed thematic maps obtained with "index", "spectral" or "spectral and index" classifiers
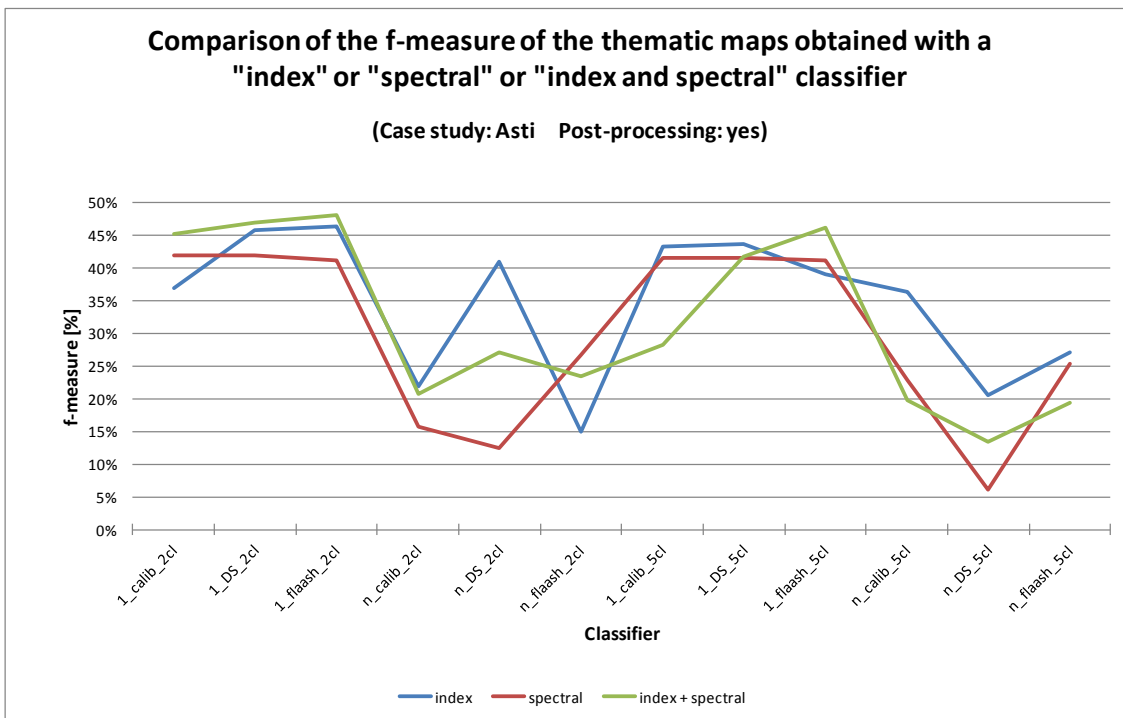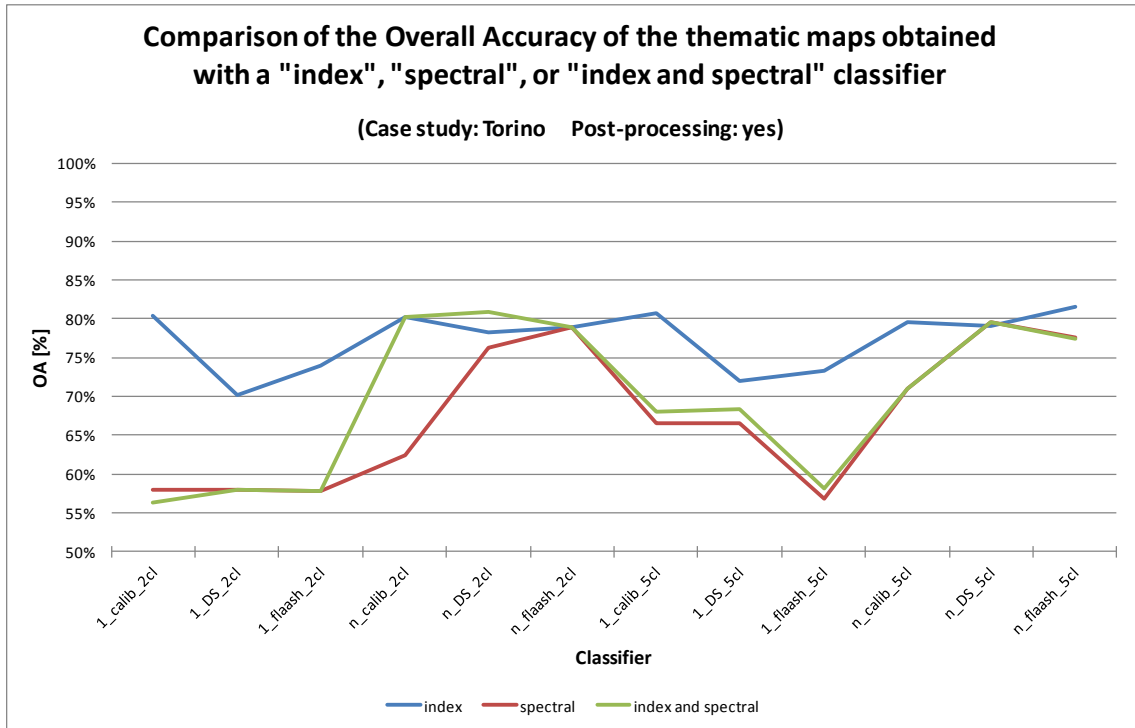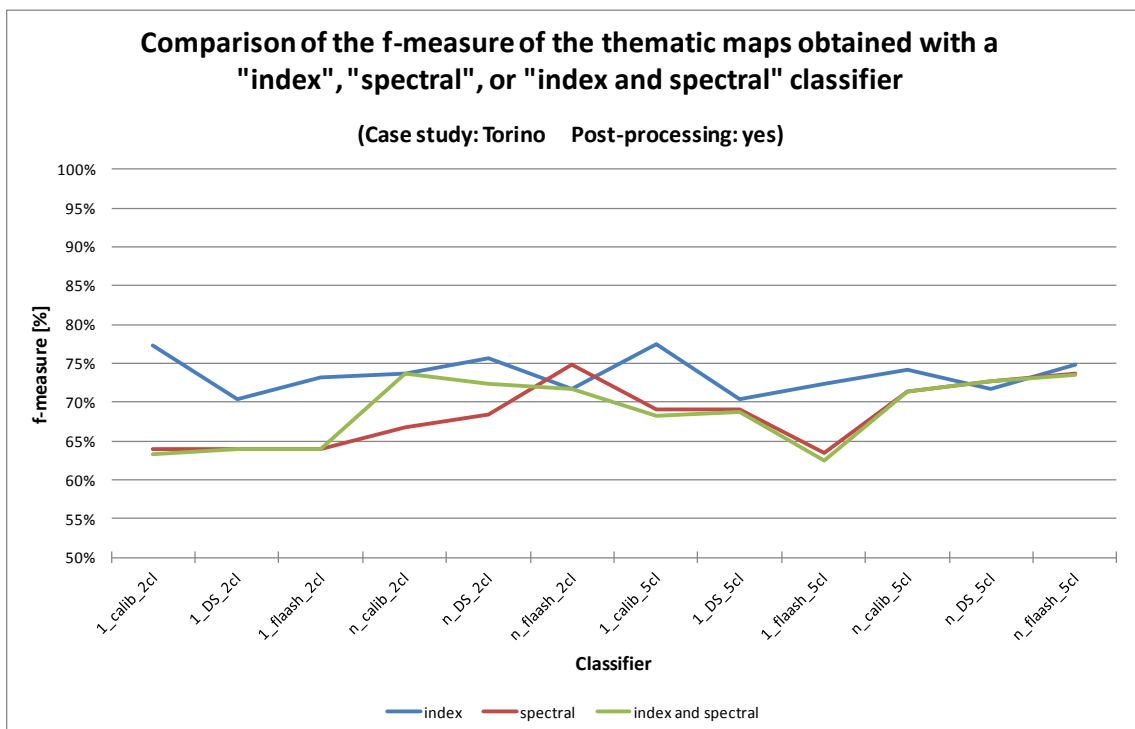

Figure 79 - f-measure comparison in post-processed thematic maps obtained with "index", "spectral" or "spectral and index" classifiers

4.5 Discussion

The aim of this work was the evaluation of the usability of Decision Tree classifiers in order to automate the process of urban extraction from Landsat data.

In particular, the focus was on variables to use for the classification process: different kinds of radiometric pre-processing were tested, as well as different post-processing. The training set to be used for the learning phase of the classifier was deeply analyzed: is it better to learn the classifier on the two class of interest (Urban and Non Urban), or it's better to learn the classifier or more classes, and then to extract only the rules of interest? Which solution enables a better division between classes? Then, what if statistics from a multitemporal stack are extracted, instead than from the single image? is it possible to detect more "general" rules? Finally, which attributes can be used? Is it better to use only spectral information, only attributes (that can be derived from the image spectral information and that increase the sensitivity toward desired information), or best performances can be reached only when they are used together? Furthermore, which attribute or band is more frequently selected from the algorithm to be part of the model? This work was aimed at providing an answer to these questions. The classifiers were generated for two different cases study, named "Torino case study" and "Asti case study", from the name of the greater town depicted in the scene.

From a general point of view, best Overall Accuracies were high in both cases: round 81% in Torino case study, round 96%-97% in Asti case study. By the analysis of the f-measure, that is representative of how the Urban class was detected, it can be deduced that the Urban class was detected better in Torino case study (best f-measure: 77%) than in Asti case study (best f-measure: 48%). The different behaviours can be attributed to the different urban land cover consistency: very dense in the first case, with a density from medium to low in the second case. This observation was confirmed by the following analysis: when the dimension of the validated area was reduced, and mainly focused on the greater town depicted in the scene (Asti), f-measure increased from 48% to 66%. The qualitative analysis later performed was aimed at detecting more recursive errors in the detection of built-up areas: cemeteries, excavation sites and bare soil were areas more frequently confused with the urban class. The analysis performed in Asti case study with the integration of the municipality boundaries vector layer, allowed to extend the investigation to each single municipality: results stated that major errors were located at the built-up area boundaries, and that the same built-up area detection is related to the built-up area density: the more dense it is, the more it's easy to detect.

In the second part of the analysis, the different tested variables were compared.

Three radiometric pre-processing were compared: calibration into reflectance value, atmospheric correction with Dark Subtraction and atmospheric correction with FLAASH. Obtained accuracies showed that there was no trend in any of the tested pre-processing: no-one of the applied corrections could be considered better than others.

Different post-processing were tested on some selected classifiers; after that, the majority filtering, that obtained better results in the performed tests, was applied to the whole set of data. The performances improvement is evident in Asti case study, where the application of a majority

filtering on the performed classifications generally improved all accuracy measures, particularly for what User's Accuracy is concerned, less in Torino case study, where the improvement given by the post-processing application happened in a reduced number of times and, in general, with a lower efficiency: the post processing is particularly needed where medium density urban land cover and very sparse urban land cover are present, otherwise it can be avoided.

The classifiers were then learned using 2 classes or 5 classes, and using a single image or a multitemporal stack. Obtained results diverged a lot: 2-class classifiers were better than others in Asti case study, particularly in post-processed classifications, less when the validation was performed only on a subset; on the contrary, in Torino case study accuracies were higher when 5-class classifiers were considered in about 61% of cases. For what the use of a multitemporal stack was concerned, the situation was quite similar: in Asti case study single image classifiers provided better results than others in most of the cases, sometimes in a consistent way (the f-measure of post-processed images was, on average, 20% better in single image classifiers with respect to 5-class classifiers), while in Torino case study multitemporal classifiers were better than other in most of the cases and with Overall Accuracy better, on average, of about 13%. Other test are necessary in order to decide which is the best approach to use.

Finally, for what the use of different attributes was concerned, again, results differ in the two considered cases. "Index classifiers" were better in Torino case study, while "spectral" and "spectral and index" classifiers provided higher accuracy in Asti case study and, in most cases, very similar each other. Consideration could be performed, instead, on the attributes that were more often chosen by the algorithm to be part of the classifier: some indexes were considered a very few times (BUI), others were widely used in both cases (NDVI and NDBBBI). Among considered bands, band 1 was the most considered one, against band 2 and 3 that were less considered. This can be due to the high correlation between the three bands: one of the three can be considered representative for the whole set. Band 4, 5 and 7 were considered quite often in both cases.

Classifiers that provided better results both in terms of OA and f-measure were: "1 index DS 2class" and "1 index DS 5class" in Torino case study, "1 index FLAASH 2class", "1 spectral index FLAASH 5class", "1 spectral index FLAASH 2class" and "1 spectral index DS 2class" in Asti case study. They are all derived from a single image, and makes use of indexes; most of all, they all have as source radiometric pre-elaborated images.
If, from one side, it was not possible to detect a clear trend for what the use of a multitemporal stack and the use of different pre-elaboration was concerned, this results evaluation give some more information about efficient input variables.

In order to better evaluate the results and to understand the potentiality of the method, a comparison was performed with the Corine Land Cover 2006.

Corine Land Cover is a land cover data available for the whole Europe at scale 1:100.000; in its last version, it was derived from IRS p6 satellite data at 23 meters resolution and SPOT data at 20 meters (multi-spectral) and 10 meters (panchromatic) spatial resolution (see section 1.2.2). Beyond the output scale and the input images spatial resolution, also the Minimum Mapping

Units (MMA) has to be considered. In the last version, the minimum mapping unit is set to 5 ha, that is about 55 pixel at 30 meters resolution (Landsat ETM+ spatial resolution).

The analysis was performed in the following way:

1) Two of the classifications that provided the best performances in Asti case study ("1 index FLAASH 2 class" and "1 spectral index FLAASH 5 class") were post-processed in order to obtain a MMA comparable with the CLC2006's one. A Clump procedure with an Operator Size 3x3 was initially applied to Urban class, then a Sieve with 50 as Group Minimum Threshold and 8 neighbour was used to complete the post-processing.

2) The validation mask derived from the CTRN (as described in 3.2.8.1) was post-processed following the same steps as described in 1).

3) Artificial Surfaces were extracted from CLC 2006, and then the vector layer was rasterized at 30 meters resolution in order to match with the classified data.

4) The two post-processed classifications were compared with the new validation mask as well as the refined CLC2006 compared with the new validation mask.

The following accuracies were reached (Table 37):

Validated layer	OA	PA	UA	f-measure
CLC2006	97.5%	45.57%	74.89%	56.66%
1 index FLAASH 2 class	97.5%	50.51%	72.16%	59.42%
1 spectral index FLAASH 5 class	97.35%	42.75%	72.10%	53.67%

Table 37 – Comparison between accuracy reached by the resulting classifiers and CLC2006 (pre-processed as above-described)

Results are very encouraging to the obtained classifications. They state that the provided method is able to produce accuracies comparable with the ones provided by CLC2006 that is completely obtained from photo-interpretation, for what the interpretative part is concerned.

Conclusion and future development

This thesis was inserted in the framework of algorithms development for the classification of urban areas from medium spatial resolution data.

Decision Tree classifier was investigated as classification techniques, thus it allows to extract rules that can be later applied to different scenes. In particular, the aim was to evaluate which steps to perform in order to obtain a good classification procedure, mainly focusing on processing that can be applied to images and on training set features.

The training set was evaluated on the basis of the number of classes to use for its creation, together with the temporal extension of the training set and input attributes, while images were submitted to different kind of radiometric pre and post-processing. The aim was the evaluation of the best variables to set for the creation of the training set, to be used for the classifier generation. Different procedures were created in order to perform all the above mentioned analyses; they mainly concern the automatic creation of training sets with all the possible combinations, the transformation of the rules generated by the algorithm into classification procedures that allow to generate the thematic maps, the automatic application of post processing and the automatic generation of confusion matrix in order to provide accuracy measures.

The analysis of training set input features proved that:

-for what the use of two or more classes for the training set generation and the use of a multitemporal stack is concerned, trends diverged a lot in the two applied cases study; other tests are therefore necessary in order to give an answer to this issue;

- for what the use of input attributes is concerned, indexes revealed to be very important for the land cover extraction; particularly, indexes such as NDVI and NDBBBI demonstrated their ability in separating classes;

The analysis of different radiometric pre-elaboration and post processing, stated that:

- for what the use of different radiometric pre-elaboration is concerned, it's not so evident the better efficiency of one respect to the others;

- for what the use of different post-processing is concerned, the application of a majority filtering with a small window (3x3) allowed to increase classification accuracies thus creating a more homogeneous texture in a land cover otherwise very fragmented.

Best reached Overall Accuracies were round 81% in Torino case study, round 96-97% in Asti case study. By the analysis of f-measure, instead, that is representative of how the urban class was classified, it resulted 77% accuracy in Torino case study and 48% in Asti case study: this value increased from 48% to 66% when the validation was performed only on the subset where the major town was depicted, thus demonstrating that major problems occurred in areas with low built-up density.

In order to analyze the method potentiality, a comparison was performed with Corine Land Cover, a land cover dataset available in Europe at scale 1:100.000. In order to perform the comparison, the generated validation mask and the classifications were adapted in order to be compliant with the Minimum Mapping Units used to generate CLC2006. Results obtained by the comparison of the adapted validation mask with the CLC2006 and by the comparison of the new adapted validation mask and the adapted classification, produced encouraging results since very similar results were obtained in the two cases. This is very positive, taking into account that CLC2006 was derived from photo-interpretation while our datasets were produced in a more automatic way.

The performed qualitative analysis demonstrated that the most recursive errors are located in area with lower built-up density and that cemeteries, excavation sites and bare soil are the more frequently confused areas.

Finally, for what classifiers that provided higher accuracies is concerned, it results that they are all derived from a single image, makes use of indexes and have, as source, radiometrically corrected images: if, from on side, it was difficult to detect a clear trend when the analysis was performed on all classifiers, on the other side it's quite clear which variables were used by the best classifiers.

Since this work constitute just a first phase for the creation of an algorithm for urban land cover extraction, many steps are still necessary. Future developments follow these directions:
1. application of the generated methodology to other cases study in order to evaluate better trends for what the use of different number of classes and temporal extension of the training set is concerned;
2. application of the generated rules to other scenes in order to evaluate reachable accuracies; in a first phase, this can be performed in areas where a cartography with features similar to the ones provided by the Digital Regional Cartography Map is available, in order to exploit the validation procedure already generated;
3. consider, in the algorithm creation, other features, such as texture measures;
4. consider, in the methodology creation, other pre-processing such as segmentation techniques that could allow to obtain more homogeneous areas.

The automation of feature extraction procedures is always a challenge. The challenge become greater if the feature to be extracted is built-up areas: the presence of mixed pixels, typical or urban environment, makes more difficult to identify common features, particularly if only spectral information is considered.

The possibility of having a procedure for the extraction of this thematism, also if not with very high accuracies, is anyway a good starting point specially in application such as emergency mapping, where information need to be extracted quickly and the elimination, also if partially, of manual digitalization procedures, can be considered a great breakthrough.

Annex

Annex I – Techniques used for the classification of each GLC2000 Regional Map

Regional Windows	Methods of mapping
South America	Classification based on multi-resolution satellite data where each source of data used contributes to mapping a specific ecosystem or land cover type.
Africa	Multi-sensor approach (SPOT VGT, radar and DMSP data). Synthetic images and NDVI profiles classified using unsupervised clustering.
Northern Eurasia	Clustering of seasonal mosaics
Asia	Supervised classification using monthly NDVI composites as an input
South Asia	ISODATA clustering carried out on the maximum NDVI composite layer of the above nine-month data to obtain the land use / land cover map. Clusters assigned to various classes on the basis of ground truth and reference information available for the region. Stable night-time lights (derived from DMSP data) used to discriminate urban areas
South East Asia	Unsupervised digital clustering method. Mapping and labelling of the classes supported by Landsat TM image interpretation and field records.
North East Europe	Unsupervised classification; 40 classes created, by majority clustering, and then labelled according to relatively accurate reference data
Europe	Unsupervised classification, based on the simultaneous use of spectral and temporal patterns, applied to the composites that where were created using the mean compositing method and then corrected to remove coastal values and to discriminate permanent snow from temporary snow cover.

Regional Windows	Methods of mapping
North West Europe	Classification was based on thresholding with CART (Classification And Regression Tree). Thresholds defined for each class and for each sub region. For water, urban areas and wetlands a mask was derived from the PELCOM land cover database (Mücher et al. 2001)
Souther Europe	Two-month mean composites created for the red, NIR and SWIR bands. Composites used as input data for an unsupervised classification with, initially, 110 clusters. Clusters labelled according to CORINE data as main reference data set. PELCOM and Landsat TM Quicklooks used as complementary ancillary data.
China	Classification undertaken by using an unsupervised clustering method; labelling of the classes supported by a 1:1000000 land use map in China along with expert opinion
North America	The mapping procedure for transforming satellite observations acquired by the SPOT4/VEGETATION sensor into land cover information includes: 1) conversion of daily data into ten-day composites; 2) post- seasonal correction and refinement of apparent surface reflectance in ten day composite images; 3) extraction of land cover information from the composite images.
Australia	3-month average synthesis of the Red, NIR and SWIR channels clustered in 100 statistical classes. Labelling supported by the Vegetation Map of Australia. Composition of each cluster computed in a GIS and cluster assigned to the main land-cover class within the cluster.
New Zeland	Unsupervised digital clustering performed. Spectral clusters labeled and regrouped to eleven land cover classes, using the Waikato Region Landcover Database (LCDB) from the year 1997 as a reference.
Greenland and Iceland	The mapping method includes detection of pixels contaminated by clouds/snow and defective sensor detectors; a synthesis of spectral channels' mosaics; and hybrid supervised and unsupervised land cover classification with use of these mosaics

Regional Windows	Methods of mapping
Solomon Island	Unsupervised ISODATA classification used to create 25 clusters per image. Clusters assigned to classes by referring to freely available Landsat TM imagery, and to ancillary information describing the vegetation of the islands
New Caledonia and Vanuatu	ISODATA classification separating 20 Clusters. Clusters labeled by using ancillary data.
Fijian Islands	Unsupervised ISODATA classification performed, creating 25 clusters per image. Clusters assigned to classes using Landsat TM Quicklooks, and ancillary landcover data from the Management Services Devision of the Fiji Forestry Department.
Hawaiian Islands	Classification divided into 3 separate parts. Each part classified separately, using an unsupervised ISODATA classification, creating 25 clusters which were then assigned to classes using Landsat TM imagery, and ground truth data obtained from the Hawaiian land cover analysis project
Northern Africa and South-Western Asia	Data classification entirely based on the analysis of transformed NDVI temporal signal

Annex II – CTRN Piemonte code list

ENTITA'

CODICE	NOME	TOPO-LOGIA	SIMB	COPERT.	FILE

Gruppo 01   STRADE E FERROVIE

CODICE	NOME	TOPO-LOGIA	SIMB	COPERT.	FILE
2.01.01	Tronco di strada asfaltata	AREA	0	Sxxxxxx	PAT
2.01.02	Tronco di carreggiata autostradale	AREA	0	Sxxxxxx	PAT
2.01.03	Tronco di strada non asfaltata	AREA	0	Sxxxxxx	PAT
2.01.04	Tronco di strada campestre, mulattiera, percorso pedonale o ciclabile isolato, sentiero	AREA	0	Sxxxxxx	PAT
2.01.041	Tratturo	AREA	0	Sxxxxxx	PAT
2.01.042	Tronco di strada non rappresentabile a misura	LINEA	-	Sxxxxxx	AAT [*]
2.01.05	Tronco di strada in costruzione	AREA	0	Sxxxxxx	PAT
2.01.06	Tronco di strada in galleria	AREA	0	Sxxxxxx	PAT
2.01.07	Zona di intersezione fra strade sovrapposte	AREA	0	Sxxxxxx	PAT
2.01.08	Zona di intersezione fra strade e corsi d'acqua sovrapposti	AREA	0	Sxxxxxx	PAT
2.01.09	Passo, colle, valico	LINEA	-	SxxxxxxL	AAT [*]
2.01.20	Tronco di ferrovia a scartamento ordinario, elettrificato	LINEA	-	Sxxxxxx	AAT [*]
2.01.21	Tronco di ferrovia a scartamento ordinario, non elettrificato	LINEA	-	Sxxxxxx	AAT [*]
2.01.22	Tronco di ferrovia a scartamento ridotto, elettrificato, tramvia	LINEA	-	Sxxxxxx	AAT [*]
2.01.23	Tronco di ferrovia a scartamento ridotto, non elettrificato	LINEA	-	Sxxxxxx	AAT [*]
2.01.24	Tronco di ferrovia in galleria	LINEA	-	Sxxxxxx	AAT [*]
2.01.25	Aree di pertinenza delle ferrovie	AREA	0	Sxxxxxx	PAT
2.01.30	Piazze e parcheggi	AREA	0	Sxxxxxx	PAT
2.01.31	Aree urbane pavimentate	AREA	0	Sxxxxxx	PAT
2.01.32	Aree urbane non pavimentate	AREA	0	Sxxxxxx	PAT
2.01.33	Aree di manovra per la circolazione	AREA	0	Sxxxxxx	PAT
2.01.35	Impianti sportivi	AREA	0	Sxxxxxx	PAT
2.01.41	Benzinai, stazioni di servizio	AREA	0	Sxxxxxx	PAT
2.01.42	Spartitraffico autostradale	AREA	0	Sxxxxxx	PAT

Gruppo 02   FABBRICATI, ARREDO URBANO E MANUFATTI TECNICI

CODICE	NOME	TOPO-LOGIA	SIMB	COPERT.	FILE
2.02.01	Unità volumetrica di edificio residenziale	AREA	71, 0	Sxxxxxx	PAT
2.02.011	Unità volumetrica di edificio industriale	AREA	716, 0	Sxxxxxx	PAT
2.02.012	Unità volumetrica di stazione ferroviaria per passeggeri e merci	AREA	712, 0	Sxxxxxx	PAT
2.02.013	Unità volumetrica di chiesa o edificio religioso	AREA	711, 0	Sxxxxxx	PAT

2.02.014	Unità volumetrica di castello o edificio di valore storico-artistico	AREA	714, 0	Sxxxxxx	PAT
2.02.015	Unità volumetrica di stazione o sottostazione di trasformazione	AREA	715, 0	Sxxxxxx	PAT
2.02.04	Unità volumetrica di edificio sovrastante strada	AREA	71, 0	Sxxxxxx	PAT
2.02.05	Unità volumetrica di edificio sovrastante corsi o specchi d'acqua	AREA	71, 0	Sxxxxxx	PAT
2.02.06	Tettoia o pensilina	AREA	72	Sxxxxxx	PAT (**)
2.02.07	Unità volumetrica di baracca	AREA	73	Sxxxxxx	PAT (**)
2.02.071	Unità volumetrica di serra	AREA	731	Sxxxxxx	PAT (**)
2.02.08	Unità volumetrica di edificio a sbalzo	AREA	71, 0	Sxxxxxx	PAT
2.02.09	Unità volumetrica di acquedotto pensile	AREA	0	Sxxxxxx	PAT
2.02.10	Limite di edificio in costruzione	LINEA	-	SxxxxxxL	AAT (*)
2.02.21	Scale e gradonate	AREA	0	Sxxxxxx	PAT
2.02.22	Rampe	AREA	0	Sxxxxxx	PAT
2.02.23	Spallette di ponte di qualsiasi natura	LINEA	-	SxxxxxxL	AAT (*)
2.02.24	Imbocco di galleria stradale o ferroviaria	LINEA	-	SxxxxxxL	AAT (*)
2.02.25	Manufatto tecnico complesso	LINEA	-	SxxxxxxL	AAT (*)
2.02.26	Unità volumetrica di cabina elettrica	AREA	0	Sxxxxxx	PAT
2.02.27	Unità volumetrica di silos, contenitore, cisterna	AREA	0	Sxxxxxx	PAT
2.02.28	Tralicci e supporti delle linee aree elettriche e telefoniche	LINEA	-	SxxxxxxL	AAT (*)

Gruppo 03   ACQUE E MANUFATTI INERENTI LE ACQUE

2.03.01	Tronco di corso d'acqua rappresentabile come superficie	AREA	0	Sxxxxxx	PAT
2.03.02	Specchio d'acqua	AREA	0	Sxxxxxx	PAT
2.03.021	Risaia	AREA	80	Sxxxxxx	PAT
2.03.022	Palude	AREA	81	Sxxxxxx	PAT
2.03.03	Tronco di corso d'acqua sovrapposto ad altro corso o specchio d'acqua	AREA	0	Sxxxxxx	PAT
2.03.04	Tronco di corso d'acqua non rappresentabile a misura	LINEA	-	Sxxxxxx	AAT (*)
2.03.05	Salto d'acqua di qualsiasi natura	LINEA	-	Sxxxxxx	AAT (*)
2.03.06	Canale di scolo, canaletta irrigua	LINEA	-	Sxxxxxx	AAT (*)
2.03.10	Manufatti vari inerenti le acque	LINEA	-	SxxxxxxL	AAT (*)
2.03.11	Imbocco di corso d'acqua sotterraneo	LINEA	-	SxxxxxxL	AAT (*)
2.03.12	Vasca o piscina	AREA	0	Sxxxxxx	PAT
2.03.13	Pozzo	PUNTO	40	SxxxxxxP	PAT
2.03.14	Sorgente	PUNTO	41	SxxxxxxP	PAT
2.03.15	Isola	AREA	77	Sxxxxxx	PAT

Gruppo 04   LINEE E CONDOTTE PER TRASPORTO DI ENERGIA, MATERIALI FLUIDI E SOLIDI, PERSONE

2.04.01	Tratta di elettrodotto aereo	LINEA	-	SxxxxxxL	AAT (*)

2.04.021	Tratta di condotta per materiali liquidi (oleodotti sopraelevati e interrati)	LINEA	-	SxxxxxxL	AAT [*]
2.04.022	Tratta di condotta per materiali gassosi (metanodotti sopraelevati e interrati)	LINEA	-	SxxxxxxL	AAT [*]
2.04.03	Tratta di linea di trasporto a cavo	LINEA	-	SxxxxxxL	AAT [*]
2.04.04	Tratta di teleferica per materiali	LINEA	-	SxxxxxxL	AAT [*]
2.04.05	Tratta di acquedotto	LINEA	-	SxxxxxxL	AAT [*]
2.04.06	Tratta di condotta forzata	LINEA	-	Sxxxxxx	AAT [*]

Gruppo 05   DIVISIONI DEL TERRENO

2.05.01	Muro isolato, divisorio o di sostegno	LINEA	-	Sxxxxxx	AAT [*]
2.05.02	Muro a secco, gabbionata di sostegno	LINEA	-	Sxxxxxx	AAT [*]
2.05.03	Recinzione (non in muratura)	LINEA	-	Sxxxxxx	AAT [*]
2.05.04	Divisione permanente del terreno non altrimenti classificata	LINEA	-	Sxxxxxx	AAT [*]
2.05.05	Divisione non materializzata del terreno, per discontinuità di pavimentazione/superficie, strada in disuso	LINEA	-	Sxxxxxx	AAT [*]
2.05.06	Cimitero	AREA	83	Sxxxxxx	PAT

Gruppo 06   FORME TERRESTRI

2.06.01	Scarpata con rivestimento o terrapieno	AREA	0	Sxxxxxx	PAT
2.06.02	Scarpata senza rivestimento	AREA	0	Sxxxxxx	PAT
2.06.021	Trincea	AREA	0	Sxxxxxx	PAT
2.06.03	Calanco o frana	LINEA	-	SxxxxxxL	AAT [*]
2.06.04	Scarpata generica non rappresentabile a misura	LINEA	-	Sxxxxxx	AAT [*]
2.06.10	Zona rocciosa	AREA	76	Sxxxxxx	PAT
2.06.11	Zona sabbiosa o pietrosa	AREA	77	Sxxxxxx	PAT
2.06.12	Ghiacciaio	AREA	82	Sxxxxxx	PAT
2.06.13	Morena mediana e morena marginale	AREA	76	Sxxxxxx	PAT
2.06.14	Morene deposte laterali, morene deposte frontali	AREA	76	Sxxxxxx	PAT
2.06.20	Affioramento roccioso	LINEA	-	Sxxxxxx	AAT [*]
2.06.21	Spuntone roccioso	PUNTO	34	SxxxxxxP	PAT

Gruppo 07   VEGETAZIONE

2.07.01	Albero isolato	PUNTO	42	SxxxxxxP	PAT
2.07.02	Filare di alberi	LINEA	-	Sxxxxxx	AAT [*]
2.07.03	Zona di coltivazione arborea	AREA	78	Sxxxxxx	PAT
2.07.031	Zona di coltivazione erbacea	AREA	0	Sxxxxxx	PAT
2.07.032	Prati stabili	AREA	0	Sxxxxxx	PAT
2.07.033	Pascoli	AREA	0	Sxxxxxx	PAT
2.07.034	Zona con vegetazione rada o assente	AREA	0	Sxxxxxx	PAT
2.07.04	Zona boscata	AREA	79	Sxxxxxx	PAT
2.07.05	Siepe	LINEA	-	Sxxxxxx	AAT [*]

2.07.06	Giardini e orti urbani	AREA	0	Sxxxxxx	PAT
2.07.065	Parchi urbani	AREA	0	Sxxxxxx	PAT
2.07.071	Vigneto	AREA	84	Sxxxxxx	PAT
2.07.072	Frutteto	AREA	85	Sxxxxxx	PAT

Gruppo 08 ALTIMETRIA

2.08.01	Curva di livello direttrice	LINEA	-	SxxxxxxL	AAT (*)
2.08.02	Curva di livello ordinaria	LINEA	-	SxxxxxxL	AAT (*)
2.08.03	Curva di livello ausiliaria	LINEA	-	SxxxxxxL	AAT (*)
2.08.04	Curva di livello di determinazione incerta	LINEA	-	SxxxxxxL	AAT (*)
2.08.05	Linea di discontinuità (*break line*)	LINEA	-	SxxxxxxL	AAT (*)
2.08.06	Punto quotato al suolo	PUNTO	43	SxxxxxxP	PAT
2.08.07	Punto quotato isolato su copertura di edifici	PUNTO	0	SxxxxxxP	PAT

Gruppo 09 ELEMENTI GENERALI DI RAPPRESENTAZIONE

2.09.01	Simbolo di direzione (freccia)	PUNTO	44	SxxxxxxP	PAT
2.09.02	Simbolo di direzione (spina di pesce)	PUNTO	45	SxxxxxxP	PAT
2.09.03	Simbolo di elettrificazione ("saetta")	PUNTO	46	SxxxxxxP	PAT
2.09.04	Simbolo di acque	PUNTO	47	SxxxxxxP	PAT
2.09.05	Trattino di riferimento quote su edifici	PUNTO	48	SxxxxxxP	PAT
2.09.06	Simbolo di passo, colle, valico	PUNTO	481	SxxxxxxP	PAT
2.09.07	Simbolo di edificio religioso (croce)	PUNTO	147	SxxxxxxP	PAT
2.09.08	Simbolo di cappella	PUNTO	148	SxxxxxxP	PAT
2.09.09	Pilone, tabernacolo, croce isolata,	PUNTO	149	SxxxxxxP	PAT
2.09.10	Monumento	PUNTO	150	SxxxxxxP	PAT
2.09.11	Aeroporto	PUNTO	151	SxxxxxxP	PAT
2.09.12	Idroscalo	PUNTO	152	SxxxxxxP	PAT
2.09.13	Torre, ciminiera, campanile	PUNTO	153	SxxxxxxP	PAT
2.09.14	Miniera	PUNTO	154	SxxxxxxP	PAT
2.09.15	Pozzo di petrolio o metano	PUNTO	155	SxxxxxxP	PAT
2.09.16	Discariche e rottamai	PUNTO	156	SxxxxxxP	PAT
2.09.17	Cabina di trasformazione	PUNTO	157	SxxxxxxP	PAT
2.09.18	Antenna per telecomunicazioni	PUNTO	169	SxxxxxxP	PAT
2.09.19	Centrale idroelettrica, termoelettrica, nucleare, con edificio esterno o sotterraneo.	PUNTO	158	SxxxxxxP	PAT
2.09.20	Campeggio	PUNTO	159	SxxxxxxP	PAT
2.09.21	Tendone pressurizzato	PUNTO	160	SxxxxxxP	PAT
2.09.22	Mulino	PUNTO	161	SxxxxxxP	PAT
2.09.23	Termine di seggiovia/funivia	PUNTO	170	SxxxxxxP	PAT
2.09.40	Distributore di carburante	PUNTO	163	SxxxxxxP	PAT
2.09.51	Crocicchi	LINEA	-	SxxxxxxL	AAT (*)
2.09.52	Indicazioni di cornice	LINEA	-	SxxxxxxL	AAT (*)
2.09.60	Mezzeria dei tronchi stradali	LINEA	-	SxxxxxxL	AAT (*)

2.09.70	Mezzeria dei tronchi di corso d'acqua rappresentabili come superficie	LINEA	-	SxxxxxxL	AAT [*]
2.09.80	Area non definita	AREA	0	Sxxxxxx	PAT

Gruppo 10   LIMITI AMMINISTRATIVI

2.10.01	Limite di stato	LINEA	-	SxxxxxxL	AAT [*]
2.10.02	Limite di regione	LINEA	-	SxxxxxxL	AAT [*]
2.10.03	Limite di provincia	LINEA	-	SxxxxxxL	AAT [*]
2.10.04	Limite di comune	LINEA	-	SxxxxxxL	AAT [*]
2.10.05	Area soggetta a censura militare	LINEA	-	SxxxxxxL	AAT [*]

Gruppo 11   RETE DI INQUADRAMENTO

2.11.01	Vertice della rete geodetica IGM - I, II, III e IV ordine	PUNTO	50	SxxxxxxP	PAT
2.11.02	Vertice catastale di rete o sottorete	PUNTO	51	SxxxxxxP	PAT
2.11.03	Vertice di raffittimento planimetrico	PUNTO	511	SxxxxxxP	PAT
2.11.04	Punto di triangolazione aerea	PUNTO	52	SxxxxxxP	PAT
2.11.05	Punto fiduciale catastale verificato con triangolazione aerea	PUNTO	53	SxxxxxxP	PAT
2.11.06	Caposaldo di livellazione geometrica IGM - I, II, III e IV categoria	PUNTO	54	SxxxxxxP	PAT
2.11.07	Caposaldo di raffittimento	PUNTO	54	SxxxxxxP	PAT

[*]   Il nome dell'Entità, di tipo lineare, viene introdotta nel campo ENT_LIN della copertura corrispondente

[**]   Le linee di vestizione delle Entità 2.02.06 (Tettoia o pensilina), 2.02.07 (Unità volumetrica di baracca), 2.02.071 (Unità volumetrica di serra), 2.06.01 (Scarpata con rivestimento o terrapieno) e 2.06.02 (Scarpata senza rivestimento) vengono introdotte nella AAT della copertura SxxxxxxL

Annex III – Example of algorithm output generated from Weka

=== Run information ===

      Scheme: weka.classifiers.trees.J48 -C 0.25 -M 20

      Relation: TO_2005_index_calib_5class

      Instances: 1014

      Attributes: 6

           NDVI

            NDBBBI

            NDBI

            BUI

            Sara

            target

      Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

      J48 pruned tree

      NDBBBI <= 0.08903

      | NDBI <= -0.162264

      | | NDBBBI <= -0.510709: vegetation (117.0)

      | | NDBBBI > -0.510709: water (21.0)

      | NDBI > -0.162264

      | | NDVI <= 0.096346: industrial (184.0/9.0)

      | | NDVI > 0.096346

      | | | Sara <= 1.27775: urban (205.0/2.0)

      | | | Sara > 1.27775

      | | | | NDBI <= 0.0304: urban (31.0/5.0)

      | | | | NDBI > 0.0304: bare soil (28.0)

      NDBBBI > 0.08903: water (428.0)

      Number of Leaves :   7

      Size of the tree :   13

      Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances   981   96.7456 %

Incorrectly Classified Instances   33   3.2544 %

Kappa statistic   0.9537

Mean absolute error   0.0196

Root mean squared error   0.1107

Relative absolute error   6.9735 %

Root relative squared error   29.5088 %

Coverage of cases (0.95 level)   97.929 %

Mean rel. region size (0.95 level)   22.5641 %

Total Number of Instances   1014

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.945	0.015	0.949	0.945	0.947	0.986	urban
0.966	0.011	0.95	0.966	0.958	0.985	industrial
0.987	0.012	0.985	0.987	0.986	0.993	water
0.991	0.002	0.983	0.991	0.987	0.999	vegetation
0.794	0.003	0.9	0.794	0.844	0.944	bare soil
Weighted Avg. 0.967	0.011	0.967	0.967	0.967	0.989	

=== Confusion Matrix ===

```
  a   b   c   d   e   <-- classified as
222   6   4   0   3 |   a = urban
  4 170   2   0   0 |   b = industrial
  1   3 446   2   0 |   c = water
  0   0   1 116   0 |   d = vegetation
  7   0   0   0  27 |   e = bare soil
```

Annex IV – Results of qualitative analysis performed in Asti case study for each town

Town	Area (m²)	Urban pixel (in the validation mask)	Percentage of pixels correctly classified in more than 80% of cases	Percentage of pixels correctly classified in more than 60% of cases
NIZZA MONFERRATO	2534052	1589	44.2%	53.9%
SAN DAMIANO	1208636	716	40.6%	49.3%
VILLANOVA D'ASTI	1592700	704	29.4%	38.8%
CASTAGNOLE DELLE LANZE	1154600	599	28.9%	35.6%
VILLAFRANCA	913300	502	20.9%	30.7%
STAZIONE	1121562	450	29.8%	40.9%
VALFENERA	654373	394	20.6%	28.2%
MONTEGROSSO	876446	381	24.9%	30.4%
PIANO - MOLINI D'ISOLA	979500	344	21.5%	25.9%
COSTIGLIOLE	773910	336	20.5%	28.0%
BOGLIETTO	579800	319	36.4%	53.3%
CALLIANO	521000	298	19.5%	27.9%
MONTECHIARO	473000	292	20.2%	24.0%
MONCALVO	528400	268	42.5%	51.5%
MOMBERCELLI	805200	251	37.5%	48.6%
MONTEMAGNO	414600	250	30.0%	41.2%
COCCONATO	444500	249	8.4%	10.0%
CASTELLO D'ANNONE	425840	241	49.8%	58.9%
CASTELNUOVO BELBO	541600	238	19.7%	31.9%
GRANA	401300	232	21.6%	25.4%
SAN PAOLO SOLBRITO	394000	229	11.4%	17.9%
REFRANCORE	494500	228	19.3%	27.2%
BALDICHIERI	308927	222	23.4%	28.8%
GHIARE - MADONNA	496356	221	16.7%	29.9%
VIARIGI	397600	216	26.9%	40.3%
TONCO	450000	214	22.4%	25.7%
CASTAGNOLE MONFERRATO	475374	206	27.2%	34.0%
ROCCHETTA TANARO	263257	206	49.5%	58.7%
DUSINO	391200	196	1.0%	5.6%
AGLIANO	443981	187	12.8%	18.7%
BRUNO	333200	186	6.5%	22.0%

Town	Area (m²)	Urban pixel (in the validation mask)	Percentage of pixels correctly classified in more than 80% of cases	Percentage of pixels correctly classified in more than 60% of cases
QUARTO	704968	182	19.2%	24.7%
SAN MICHELE	392240	175	14.3%	20.0%
CISTERNA D'ASTI	436100	170	17.1%	20.6%
MOMBARUZZO	337400	151	18.5%	31.1%
GRAZZANO BADOGLIO	367200	150	16.0%	24.0%
SAVI	339700	148	0.0%	0.0%
MOTTA	430700	142	23.2%	30.3%
VILLA SAN SECONDO	322100	141	7.1%	12.8%
CUNICO	289800	132	3.0%	8.3%
FERRERE	325500	130	18.5%	36.2%
CALAMANDRA	275500	129	13.2%	21.7%
SAN MARZANOTTO	382000	125	6.4%	8.0%
SAN MATTEO	629100	124	4.0%	10.5%
MONTIGLIO	370100	124	11.3%	13.7%
SCURZOLENGO	296100	122	15.6%	18.9%
CALLIANETTO	378600	117	0.9%	5.1%
CRIVELLE	355900	106	0.0%	0.0%
VILLA	293606	106	1.9%	1.9%
MONGARDINO	267400	103	13.6%	18.4%
SAN PIETRO	826631	95	0.0%	0.0%
VALENZANI	330800	94	6.4%	16.0%
VAGLIERANO	362000	93	23.7%	36.6%
CELLARENGO	407500	89	7.9%	41.6%
MONTEMARZO	599400	87	40.2%	52.9%
STAZIONE DI PORTACOMARO	270500	87	21.8%	32.2%
ANNUNZIATA	296877	81	0.0%	0.0%
SAN MARTINO ALFIERI	275800	78	3.8%	7.7%
GORZANO	376700	76	0.0%	0.0%
REPERGO	369700	70	1.4%	5.7%
REVIGLIASCO	344200	70	21.4%	22.9%
SERRAVALLE	371300	67	7.5%	7.5%
TORRAZZO	571600	63	9.5%	11.1%
SAN GRATO	482966	62	0.0%	0.0%
CAMERANO	342600	57	0.0%	1.8%
SANT ANNA	250442	56	0.0%	3.6%
SAN MARZANOTTO	278900	53	17.0%	20.8%
VINCHIO	309700	50	10.0%	18.0%
SAN CARLO	373285	38	0.0%	0.0%

Town	Area (m^2)	Urban pixel (in the validation mask)	Percentage of pixels correctly classified in more than 80% of cases	Percentage of pixels correctly classified in more than 60% of cases
CASORZO	607500	30	0.0%	0.0%
MADONNA DI VIATOSTO	384460	22	0.0%	0.0%
SERRA	322758	21	0.0%	0.0%
MOMBARONE	281800	21	0.0%	0.0%
VASCAGLIANA	277000	20	0.0%	0.0%
BRICCOLINO	336800	19	0.0%	0.0%
CORTE	278859	19	0.0%	0.0%
CASTIGLIONE	397200	15	0.0%	0.0%
PALUCCO	285000	9	22.2%	44.4%
BUTTIGLIERA	452380	0	0.0%	0.0%
CASTELNUOVO DON BOSCO	450200	0	0.0%	0.0%

Bibliography

Adams, J. B., D. E. Sabol, V. Kapos, R. A. Filho, D. A. Roberts, M. O. Smith, and A. . GILLESPIE. 1995. Classification of multispectral images based on fractions of endmembers: application to land cover change in the Brazilian Amazon. Remote Sensing of Environment 52:137–154.

Ajmar, A., P. Boccardo, F. Giulio Tonolo, and C. Veloso. 2010a. Earthquake damage assessment using remote sensing imagery. The Haiti case study. Geoinformation for Disaster and Risk Management Examples and Best Practices 32–37.

Ajmar, A., F. Giulio Tonolo, and F. Perez. 2010b. Flood Mapping in Support of Humanitarian Organizations. In O. Altan, R. Backhaus, P. Boccardo, & S. Zlatanova, eds., Geoinformation for Disaster and Risk Management Examples and Best Practices. JB GIS, Copenhagen, Denmark.

Angel, S., S. C. Sheppard, D. L. Civco, R. Buckley, A. Chabaeva, L. Gitlin, A. Kraley, et al. 2005. *The Dynamics of Global Urban Expansion*. Washington D.C.

Arino, O., D. Gross, F. Ranera, M. Leroy, P. Bicheron, C. Brockman, P. Defourny, et al. 2007. GlobCover: ESA service for global land cover from MERIS. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium p. 2412-2415 (pp. 2412–2415). IEEE-Inst Electrical Electronics Engineers Inc.

Arino, O., J. R. Perez, V. Kalogirou, P. Defourny, and F. Achard. 2009. Globcover 2009 4–6.

ASPRS ACCURACY STANDARDS FOR LARGE-SCALE MAPS. 1990.Photogrammetric Engineering & Remote Sensing 56:1068 – 1070.

Bailey, K. 1994. Typologies and taxonomies: an introduction to Classification Techniques. Sage University paper series on Quantitative Applications in the Social Sciences, series no. 07-102.

Baraldi, a., V. Puzzolo, P. Blonda, L. Bruzzone, and C. Tarantino. 2006. Automatic Spectral Rule-Based Preliminary Mapping of Calibrated Landsat TM and ETM+ Images. IEEE Transactions on Geoscience and Remote Sensing 44:2563–2586.

Bartholomé, E., and A. Belward. 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing 26:1959–1977.

Binaghi, E., P. A. Brivio, P. Ghezzi, and A. Rampini. 1999. A fuzzy set accuracy assessment of soft classification. Pattern Recognition Letters 20:935–948.

Binaghi, E., P. Madella, M. G. Montesano, and A. Rampini. 1997. Fuzzy contextual classification of multisource remote sensing images. IEEE Transactions on Geoscience and Remote Sensing 35:326 – 339.

Blaschke, T., S. Lang, E. Lorup, J. Strobl, and P. Zeil. 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. Environmental information for planning, politics and the public 2:555–570.

Boccardo, P., and F. Giulio Tonolo. 2012. Remote-sensing techniques for natural disaster impact assessment. In X. Yang & J. Li, eds., Advances in mapping from remote sensor imagery (pp. 387–414). Taylor & Francis Group, Boca Raton, FL.

Boccardo, P., F. Perez, F. Giulio Tonolo, P. Pasquali, S. Steffenino, A. Facello, R. Vigna, et al. 2012. *Call for tenders No ENTR/2009/27 Lot 2 - Implementation of an initial GMES service for geospatial reference data access covering areas outside Europe. D10.1 – D10.2 – D10.3 - Task 10 Analysis of non-European reference data availability (M30)* (p. 281). Torino.

Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson. 2002. Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. Global Biogeochemical Cycles 16:1021.

Bontemps, S., P. Defourny, E. Van Bogaert, O. Arino, V. Kalogirou, and J. R. Perez. 2011. *GLOBCOVER 2009 - Products Description and Validation Report*.

Brivio, P. A., G. Lechi, and E. Zilioli. 2006. Principi e metodi di telerilevamento. (CittàStudi, ed.).

Chen, D., and D. . Stow. 2002. The effect of training strategies on supervised classification at different spatial resolution. Photogrammetric Engineering and Remote Sensing 68:1155–1162.

Congalton, R. G. 1988. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric Engineering & Remote Sensing 54:587–592.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37:35–46.

Cortijo, F. J., and N. . De La Blanca. 1998. Improving classical contextual classification. International Journal of Remote Sensing 19:1591–1613.

Danko, D. 1992. The digital chart of the world project. Photogrammetric engineering and remote sensing 58:1125–1128.

Defourny, P., S. Bontemps, V. Bogaert, and J. Weber. 2010. *Globcorine 2009*.

Defries, R., M. C. Hansen, and J. Townshend. 1995. Global Discrimination of Land Cover Types from Metrics Derived from AVHRR Pathfinder Data. Remote Sensing of Environment 54:209–222.

Defries, R. S., and J. C. Chan. 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sensing of Environment 74:503–515.

Defries, R. S., M. Hansen, J. R. G. Townshend, and R. Sohlberg. 1998. Global land cover classification at 8km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. International Journal of Remote Sensing 19:3141–3168.

DeFries, R., and J. R. G. Townshend. 1994. NDVI-derived land cover classifications at a global scale. International Journal of Remote Sensing 15:3567–3586.

Di Gregorio, A., and L. J. M. Jansen. 2005. Land cover classification system (LCCS): classification concepts and user manual. For software version 2.0. Translator (Vol. 8, pp. 1–5). Food & Agriculture Organization of the United.

Dymond, J. R., and J. D. Shepherd. 1999. Correction of the topographic effect in remote sensing. IEEE Transactions on Geoscience and Remote Sensing 37:2618–2620.

Edwards, G., and K. Lowell. 1996. Modeling uncertainty in photointerpreted boundaries. Photogrammetric Engineering and Remote Sensing 62:377–391.

EEA. 2007. *CLC2006 technical guidelines*. Copenhagen, Denmark.

Epstein, J., K. Payne, and E. Kramer. 2002. Techniques for mapping suburban sprawl. Photogrammetric Engineering and Remote Sensing 68:913–918.

Fayyad, U., and K. Irani. 1992. On the handling of continuous-valued attributes in decision tree generation. Machine Learning 8:87–102.

Felbier, A., T. Esch, A. Roth, W. Heldens, H. Taubenböck, and M. Schwinger. 2012. The Urban Footprint Processor – Concept and Implementation of a Processing Chain within the TanDEM-X Mission. EUSAR (pp. 15–18).

Flygare, A. 1997. A comparison of contextual classification methods using Landsat TM. International Journal of Remote Sensing 18:3835–3842.

Foody, G. M. 1988. The effects of viewing geometry on image classification. International Journal of Remote Sensing 9:1909–191.

Foody, G. M. 1996. Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. International Journal of Remote Sensing 17:1317–1340.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment 80:185–201.

Foody, G. M., and D. . Cox. 1994. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. International Journal of Remote Sensing 15:619–631.

Friedl, M., and C. Brodley. 1997a. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61:399–40.

Friedl, M., and C. Brodley. 1997b. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61:399–409.

Friedl, M., C. E. Brodley, and A. H. Strahler. 1999. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. IEEE Transactions on Geoscience and Remote Sensing 37:969–977.

Friedl, M., D. . McIver, J. C. . Hodges, X. . Zhang, D. Muchoney, A. . Strahler, C. . Woodcock, et al. 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83:287–302.

Friedl, M., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang. 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment 114:168–182.

Fritz, S., E. Bartholomé, A. Belward, A. Hartley, H. Eva, P. Mayaux, S. Bartalev, et al. 2003. *Harmonisation , mosaicing and production of the Global Land Cover 2000 database (Beta Version)* (p. 41).

Gomarasca, M. 2009. Basics of Geomatics. (Springer, ed.) (p. 656).

Gu, D., and A. Gillespie. 1998. Topographic normalization of Landsat TM images of forest based on subpixel sun-canopy-sensor geometry. Remote Sensing of Environment 64:166–175.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. 2009. The Weka Data Mining Software: An Update. SIGKDD Exmplorations 11.

Hansen, M. C., R. DeFries, J. Townshend, and R. Sohlberg. 2000. Global land cover classification at 1 km spatial resolution using a classi cation tree approach. International Journal of Remote Sensing 21:1331–1364.

Hansen, M., R. Dubayah, and R. Defries. 1996. Classification trees: an alternative to traditional land cover classifiers. International Journal of Remote Sensing 17:1075–1081.

Harris, P. M., and S. J. Ventura. 1995. The integration of geographic data with remotely sensed imagery to improve classification in an urban area. Photogrammetric Engineering and Remote Sensing 61:993–998.

He, C., and D. Xie. 2007. IMPROVING THE NORMALIZED DIFFERENCE BUILD-UP INDEX TO MAP URBAN BUILD-UP AREAS BY USING A SEMIAUTOMATIC SEGEMENTATION APPROACH 69:2007.

Hubert-Moy, L., A. Cotonnec, L. Le Du, A. Chardin, and P. Perez. 2001. A comparison of parametric classification procedures of remotely sensed data applied on different landscape units. Remote Sensing of Environment 75:174–187.

Huete, A., K. Didan, T. Miura, R. E, X. Gao, and L. G. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83:195–213.

Hughes, G. F. 1968. On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory 14:55–63.

Kartikeyan, B., A. Sarkar, and K. . Majumder. 1998. A segmentation approach to classification of remote sensing imagery. International Journal of Remote Sensing 19:1695–1709.

Keuchel, J., S. Naumann, M. Heiler, and A. Siegmund. 2003. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. Remote Sensing of Environment 86:530–541.

Kohavi, R., and J. R. Quinlan. 1999. Decision Tree Discovery. Discovery 3.

Landgrebe, D. . 2003. Signal Theory Methods in Multispectral Remote Sensing. John Wiley and Sons, Hoboken, NJ.

Lawrence, R., A. Bunn, S. Powell, and M. Zmabon. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sensing of Environment 90:331–336.

Lee, J. I. N. A., S. S. Lee, and K. H. Chi. n.d. Development of an urban classification method using a built-up index. Selected topics in power systems and remote sensing 39–43.

Lisini, G., F. D. Acqua, G. Trianni, P. Gamba, D. Elettronica, V. Ferrata, and I.- Pavia. 2005. Comparison and combination of multiband classifiers for Landsat urban land cover mapping. October 2823–2826.

Lotsch, A., Y. Tian, M. Friedl, and R. B. Myneni. 2003. Land cover mapping in support of LAI and FPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors. International Journal of Remote Sensing 24:1997–2016.

Loveland, T. R., and A. S. Belward. 1997. The IGBP-DIS global 1km land cover data set, DISCover: first results. International Journal of Remote Sensing 18:3289–3295.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing 21:1303–1330.

LU, D., E. MORAN, and M. BATISTELLA. 2003. Linear mixture model applied to Amazonian vegetation classification. Remote Sensing of Environment 87:456–469.

Lu, D., and Q. Weng. 2004. Spectral Mixture Analysis of the Urban Landscape in Indianapolis with Landsat ETM + Imagery. Photogrammetric Engineering & Remote Sensing 70:1053–1062.

Lu, D., and Q. Weng. 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28:823–870.

Magnussen, S., P. Boudewyn, and M. Wulder. 2004. Contextual classification of Landsat TM images to forest inventory cover types. International Journal of Remote Sensing 25:2421 – 2440.

Mannan, B., J. Roy, and A. . Ray. 1998. Fuzzy ARTMAP supervised classification of multi-spectral remotely-sensed images. International Journal of Remote Sensing 19:767–774.

Masek, J. G. 2005. White Paper on Use of Gap-Filled Products for the Mid-Decadal Global Land Survey (MDGLS).

Maselli, F., A. Rodolfi, and C. Conese. 1996. Fuzzy classification of spatially degraded Thematic Mapper data for the estimation of sub-pixel components. International Journal of Remote Sensing 17:537–551.

Mather, P. . 2004. Computer Processing of Remotely-Sensed Images: An introduction (3rd edn.). John Wiley & Sons, Chichester.

Matthews, E. 1983. Global vegetation and land-use - new high resolution databases for climate studies. Journal of climate and applied meteorology 22:474–487.

Maucha, G., G. Büttner, and B. Kosztra. 2010. *European validation of GMES FTS Soil Sealing Enhancement data*.

Mayaux, H. Eva, J. Gallego, A. H. Strahler, M. Herold, S. Agrawal, S. Naumov, et al. 2006. Validation of the global land cover 2000 map. IEEE Transactions on Geoscience and Remote Sensing 44:1728–1739.

Murai, H., and S. Omatu. 1997. Remote sensing image analysis using a neural network and knowledge-based processing. International Journal of Remote Sensing 18:811–828.

Myneni, R. B., S. Hoffman, Y. Knyazikhin, J. L. Privette, J. Glassy, Y. Tian, Y. Wang, et al. 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sensing of Environment 83:214–231.

Oetter, D. R., W. B. Cohen, M. Berterretche, T. K. Maiersperger, and R. E. Kennedy. 2000. Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data. Remote Sensing of Environment 76:139–155.

Otukei, J. R., and T. Blaschke. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. International Journal of Applied Earth Observation and Geoinformation 12:S27–S31.

Pal, M., and P. M. Mather. 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment 86:554–565.

Pesaresi, M., and M. Halkia. 2012. *Global Human Settlement Layer / Urban Atlas Integration : Feasibility Report*. Ispra, Italy.

Pohl, C., and J. L. Van Genderen. 1998. Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. International Journal of Remote Sensing 19:823–854.

Price, K. P., X. Guo, and J. M. Stiles. 2002. Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. International Journal of Remote Sensing 23:5031–5042.

Quinlan, J. R. 1993. C4.5: Programs for Machine Learning. (M. Kaufmann, ed.)Morgan Kaufmann San Mateo California (Vol. 1, p. 302). Morgan Kaufmann.

Quinlan, R., and R. Rivest. 1989. Inferring decision trees using the minim um description length principle. Information and Computation 80:227–248.

R., R. 1997. Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. International Journal of Remote Sensing 18:1099.

Rashed, T., J. R. Weeks, M. S. Gadalla, and A. . Hill. 2001. Revealing the anatomy of cities through spectral mixture analysis of multispectral satellite imagery: a case study of the Greater Cairo region, Egypt. Geocarto International 16:5–15.

Rastogi, R., and K. Shim. 2000. PUBLIC : A Decision Tree Classifier that Integrates Building and Pruning. Data Mining and Knowledge Discovery 4:315–344.

Ravan, S. 2010. Spatial data to complement the use of space-based information for disaster management. In O. Altan, R. Backhaus, P. Boccardo, & S. Zlatanova, eds., Geoinformation for Disaster and Risk Management Examples and Best Practices. JB GIS, Copenhagen, Denmark.

Ricotta, C., and G. C. Avena. 1999. The influence of fuzzy set theory on the areal extent of thematic map classes. International Journal of Remote Sensing 20:201–205.

Roberts, D. A., M. Gardner, R. Church, S. Ustin, G. Scheer, and R. . Green. 1998. Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models. Remote Sensing of Environment 65:267–279.

Running, S. W., T. R. Loveland, L. L. Pierce, R. Nemani, and E. R. Hunt. 1995. A remote sensing based vegetation classification logic for global land cover analysis. Remote Sensing of Environment 51:39–48.

S., E. 1996. Landsat TM-based forest damage assessment: correction for topographic effects. Photogrammetric Engineering and Remote Sensing 62:151.

Schaaf, C. B., F. Gao, A. H. Strahler, W. Lucht, X. Li, T. Tsang, N. C. Strugnell, et al. 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sensing of Environment 83:135–148.

Schneider, A., M. Friedl, and D. Potere. 2009. A new map of global urban extent from MODIS satellite data. Environmental Research Letters 4:044003.

Schneider, A., M. Friedl, and D. Potere. 2010. Mapping global urban areas using MODIS 500-m data: New methods and datasets based on "urban ecoregions". Remote Sensing of Environment 114:1733–1746.

Schneider, A., D. K. McIver, M. Friedl, and C. E. Woodcock Mapping urban areas using coarse resolution remotely sensed data. , 4 IEEEISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas Cat No01EX482 2623–2625 (2003). Ieee.

Shalan, M. A., M. K. Arora, and S. K. Ghosh. 2003. An evaluation of fuzzy classifications from IRS 1C LISS III imagery: a case study. International Journal of Remote Sensing 24:3179–3186.

Shen, W., G. Wu, Z. Sun, W. Xiong, Z. Fu, and R. Xiao. 2011. Study on classification methods of remote sensing image based on decision tree technology. 2011 International Conference on Computer Science and Service System (CSSS) 4058–4061.

Steele, B. M., J. C. Winne, and R. L. Redmond. 1996. Estimation and Mapping of Misclassification Probabilities for Thematic Land Cover Maps. Remote Sensing of Environment 66:192–202.

Stefanov, W. L., M. S. Ramsey, and P. R. Christensen. 2001. Monitoring urban land cover change : An expert system approach to land cover classification of semiarid to arid urban centers. Remote Sensing of Environment 77:173–185.

Stehman, S. V., and R. L. Czaplewski. 1998. Design and Analysis for Thematic Map Accuracy Assessment : Fundamental Principles. Science 344:331–344.

Stone, T. A., P. Schlesinger, R. A. Houghton, and G. M. WOODWELL. 1994. A map of the vegetation of South America based on satellite imagery. Photogrammetric engineering and remote sensing 60:541–551.

Strahler, A. H., L. Boschetti, G. M. Foody, M. Friedl, M. C. Hansen, M. Herold, P. Mayaux, et al. 2006. *Global Land Cover Validation : Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps*. GOFC-GOLD.

Stuckens, J., P. R. Coppin, and M. E. Bauer. 2000. Integrating Contextual Information with per-Pixel Classification for Improved Land Cover Classification. Remote Sensing of Environment 71:282–296.

Tateishi, R., M. A. Bayaer, H. Ghar, H. Al-Bilbisi, J. Tsendayush, S. A, A. Kasimua, et al. 2003. A new global land cover map, GLCNMO 2000:1369–1372.

Teillet, P. M., B. Guindon, and D. . Goodenough. 1982. On the slope-aspect correction of multispectral scanner data. Canadian Journal of Remote Sensing 8:84–106.

Thenkabail, P. S., E. A. Enclona, M. S. Ashton, C. Legg, and M. J. De Dieu. 2004. Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. Remote Sensing of Environment 90:23–43.

Tokola, T., J. Sarkeala, and M. Van Der Linden. 2001. Use of topographic correction in Landsat TM-based forest interpretation in Nepal. International Journal of Remote Sensing 22:551–563.

Tuceryan, M., and A. K. Jain. 1993. Texture analysis. In C. Chen, L. Pau, & P. Wang, eds., Handbook of pattern recognition and computer vision Singapore:World Scientific.

V., M. 1998. The use of census data in urban image classification. Photogrammetric Engineering and Remote Sensing 64:431.

Verstraete, M. M., and B. Pinty. 1996. Designing Optimal Spectral Indexes for Remote Sensing applications. IEEE Transactions on Geoscience and Remote Sensing 34:1254–1265.

Vieira, C. A. O., and P. M. Mather. 2000. Visulisation of measures of clas- sifier reliability and error in remote sensing. In G. B. M. Heuvelink & L. M. J. P. M, eds., 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences (pp. 701–708).

Villa, P. 2012. A hybrid multi-step approach for urban area mapping in the Province of Milan, Italy. European Journal of Remote Sensing 333–347.

Wan, Z., Y. Zhang, Q. Zhang, and Z. Li. 2002. Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. Remote Sensing of Environment 83:163–180.

Wilson M., and Henderson-Sellers A. 1984. A global archive of land cover and soils data for use in general circulation climate models. Journal of climatology 5:119–143.

Witten, I., and E. F. Frank. 2005. Data Mining: Pratical Machine Learning Tools and Techiniques - 2nd edition. Machine Learning (p. 525). Morgan Kaufmann Publishers, San Francisco.

Woodcock, C. ., and S. Gopal. 2000. Fuzzy set theory and thematic maps: accuracy assessment and area estimation. International Journal of Geographic Information Science 14:153–172.

Wu, D., and J. Linders. 2000. Comparison of three different methods to select features for discriminating forest cover types using SAR imagery. International Journal of Remote Sensing 21:2089–2099.

Zha, Y., J. Gao, and S. Ni. 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing 24:583–594.

Zhang, J., and R. . Kirby. 1999. Alternative criteria for defining fuzzy boundaries based on fuzzy classification of aerial photographs and satellite images. Photogrammetric Engineering and Remote Sensing 65:1379–1387.

Zhang, Q., J. Wang, X. Peng, P. Gong, and P. Shi. 2002. Urban built-up land change detection with road density and spectral information from multitemporal Landsat TM data. International Journal of Remote Sensing 23:3057–3078.