Faculty of Information Engineering, Statistics and Informatics

Department of Information Engineering, Electronics and Telecommunications

Doctoral program in
Information and Communications Technologies - Applied Electromagnetics
(Cycle XXIX)

Earthquake damage assessment in urban area from Very High Resolution satellite data

PhD candidate:
Dr. Eng. Roberta Anniballe

Advisor:
Prof. Nazzareno Pierdicca

February 2017

*Al mio papà, alla mia mamma
e a mia sorella Daniela.*

Ad Ernesto.

Contents

List of Figures

List of Tables

Introduction

The use of remote sensing within the domain of natural hazards and disaster management has become increasingly popular, due in part to increased awareness of environmental issues, including climate change, but also to the improvement of geospatial technologies and the ability to provide high quality imagery to the public through the media and internet. As technology is enhanced, demand and expectations increase for near-real-time monitoring and images to be relayed to emergency services in the event of a natural disaster. During a seismic event, in particular, it is fundamental to obtain a fast and reliable map of the damage of urban areas to manage civil protection interventions. Moreover, the identification of the destruction caused by an earthquake provides seismology and earthquake engineers with informative and valuable data, experiences and lessons in the long term. An accurate survey of damage is also important to assess the economic losses, and to manage and share the resources to be allocated during the reconstruction phase.

Satellite remote sensing can provide valuable pieces of information on this regard, thanks to the capability of an instantaneous synoptic view of the scene, especially if the seismic event is located in remote regions, or if the main communication systems are damaged. Many works exist in the literature on this topic, considering both optical data and radar data, which however put in evidence some limitations of the nadir looking view, of the achievable level of details and response time, and the criticality of image radiometric and geometric corrections. The visual interpretation of optical images collected before and after a seismic event is the approach followed in many cases, especially for an operational and rapid release of the damage extension map. Many papers, have evaluated change detection approaches to estimate damage within large areas (e.g., city blocks), trying to quantify not only the extension of the affected area but also the level of damage, for instance correlating the collapse ratio (percentage of collapsed buildings in an area) measured on ground with some change parameters derived from two images, taken before

and after the earthquake. Nowadays, remotely sensed images at Very High Resolution (VHR) may in principle enable production of earthquake damage maps at single-building scale. The complexity of the image forming mechanisms within urban settlements, especially of radar images, makes the interpretation and analysis of VHR images still a challenging task. Discrimination of lower grade of damage is particularly difficult using nadir looking sensors. Automatic algorithms to detect the damage are being developed, although as matter of fact, these works focus very often on specific test cases and sort of canonical situations. In order to make the delivered product suitable for the user community, such for example Civil Protection Departments, it is important to assess its reliability on a large area and in different and challenging situations. Moreover, the assessment shall be directly compared to those data the final user adopts when carrying out its operational tasks. This kind of assessment can be hardly found in the literature, especially when the main focus is on the development of sophisticated and advanced algorithms.

In this work, the feasibility of earthquake damage products at the scale of individual buildings, which relies on a damage scale recognized as a standard, is investigated. To this aim, damage maps derived from VHR satellite images collected by Synthetic Aperture Radar (SAR) and optical sensors, were systematically compared to ground surveys carried out by different teams and with different purposes and protocols. Moreover, the inclusion of a priori information, such as vulnerability models for buildings and soil geophysical properties, to improve the reliability of the resulting damage products, was considered in this study.

The research activity presented in this thesis was carried out in the framework of the APhoRISM (Advanced PRocedures for volcanIc Seismic Monitoring) project, funded by the European Union under the EC-FP7 call. APhoRISM was aimed at demonstrating that an appropriate management and integration of satellite and ground data can provide new improved products useful for seismic and volcanic crisis management.

This thesis is organized in six chapters.

Chapter 1 provides an overview of the role of remote sensing in disaster management and presents the state of the art related to earthquake damage and vulnerability assessment.

Chapter 2 reviews scattering mechanisms determining the SAR response of a building and spectral properties of urban areas in the optical bands. Then, an overview of the

change detection process through Object Based Image Analysis (OBIA) approaches is provided, focusing on specific techniques exploited in this research work.

After an introduction to the A Priori information for Earthquake damage mapping (APE) approach, proposed in APhoRISM to generate maps of damage through the integration of satellite data and a priori information, Chapter 3 provides a summary of the case study investigated in this thesis and describes remotely sensed and reference data exploited during the experiments.

Chapter 4 presents object-oriented change detection methodologies developed to generate earthquake damage maps from VHR satellite images, as well as data fusion approaches implemented for integrating satellite data and a priori information layers. Results are presented and discussed.

In Chapter 5, the Triple Collocation approach, a technique proposed for validating remotely sensed products from three different systems when the source of test data (e.g., ground data) does not represent a reliable reference, is extended to the classification problem. First, the proposed approach is tested on a simulated data set and then applied to the real case of the earthquake damage classification.

Finally, Chapter 6 draws the conclusions of this work and discusses future developments.

1. Earthquake risk and Earth Observation

1.1 Introduction

Natural disasters are becoming more frequent and severe in the modern world and their impact on human lives and economy shows an increasing trend (Figure 1-1). According to the Centre for Research on the Epidemiology of Disasters (CRED), 346 natural disasters were reported worldwide in 2015. They claimed 22773 lives, affected over 98 million others and caused economic damages of US$66.5 billion (CRED, 2016).

The severity of the impact of disasters depends strongly on the level of exposure and vulnerability in the affected area, and evidence indicates that the risk has increased worldwide largely due to the increases in the exposure of people and assets (UNISDR, 2016). Exposure is one of the major drivers of risk, and rapid demographic changes as well as the growing urbanization are key factors, causing more and more people to be exposed to risk (CEOS, 2015). According to the results of 2015 Revision of World Population Prospects released by the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, the world population reached 7.3 billion as of mid- 2015 implying that the world has added approximately one billion people in the span of the last twelve years. Depending on the expected growth rates, world population is estimated to reach 9.7 billion in 2050 and 11.2 billion in 2100 (UN, 2015). Today over half of the world's population (54 per cent.) lives in urban areas. The urbanization, encompassing both the movement of people from rural to urban areas and population growth within cities, results in larger concentrations of exposure, so, when cities are affected by a disaster, losses can be significant. Exposure increases as population grows in hazardous areas, and as improved socioeconomic conditions raise the value of assets (GFDRR, 2016). Cities in the developing world are facing increased

risk of disasters and the potential of economic and human losses from natural hazards is being exacerbated by the rate of unplanned urban expansion and influenced by the quality of urban management.

To effectively reduce the impacts of a disaster a complete strategy for disaster management is required. Earth Observation (EO) data can provide valuable information supporting both research into natural hazards and their causes and operational decision-making related to both planning and response (CEOS, 2012).
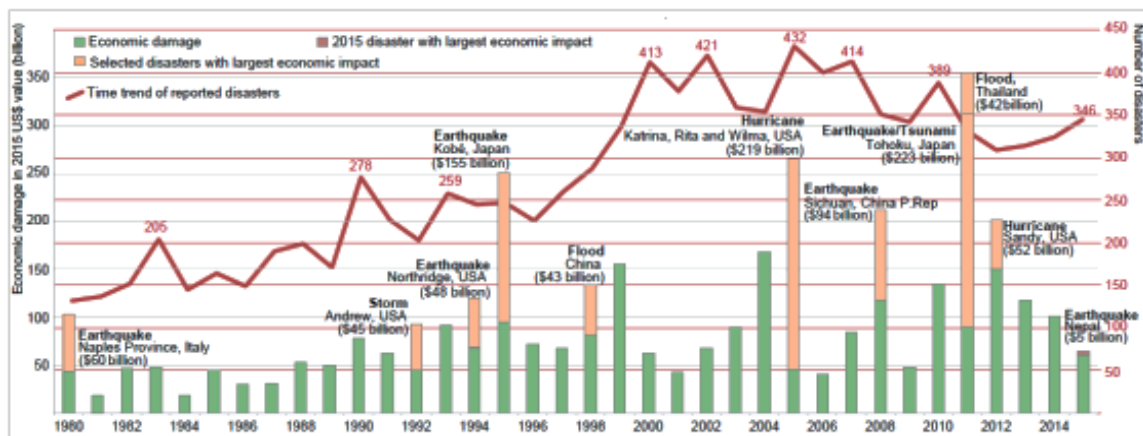
Figure 1-1: Annual reported economic damages and time trend from disasters: 1980-2015 (source: CRED, 2016).

1.2 Defining determinants of Risk: Hazard, Exposure, and Vulnerability

As defined by United Nations- International Strategy for Disaster Risk Reduction a disaster is as a serious disruption of the functioning of a community or a society causing widespread human, material, economic or environmental losses which exceed the ability of the affected community or society to cope using its own resources (UNISDR, 2004-2009).

A disaster is a function of the *risk* process: it results from the combination of exposure to hazards, conditions of vulnerability and insufficient capacity or measures to reduce the potential negative consequences of risk.

The term *hazard* indicates a potential damaging event, phenomenon or human activity that may cause loss of life, injury or other health impacts, property damage, loss of livelihoods and services, social and economic disruption, or environmental degradation (UNISDR, 2004-2009). Hazards can be classified according to their origin. Natural

hazards are natural processes or phenomena in the earth's system that may constitute a damaging event (e.g. earthquakes, volcanic eruptions, hurricanes). Human-induced hazards are those resulting from modifications of natural processes in the earth's system caused by human activities that accelerate/aggravate the damage potential (e.g. land degradation, landslides, forest fires). Human-made hazards originate from technological or industrial accidents, dangerous procedures, infrastructure failures or certain human activities which may cause the loss of life or injury, property damage, social and economic disruption or environmental degradation (e.g. industrial pollution, nuclear activities and radioactivity, toxic wastes, dam failures, transport, industrial or technological accidents such as explosions, fires and oil spills) (Van Westen, 2013). Hazardous events turn into disasters when they affect human or economic systems that cannot withstand their impact (Deichmann et al., 2011). Therefore, disaster risk cannot solely be characterized as a function of a hazard, since elements exposed to a certain hazard can react differently depending on their degree of vulnerability. *Exposure* refers to people, property, systems, or other elements present in hazard zones that are thereby subject to potential losses. The *vulnerability* characterizes the susceptibility of exposed people, assets, and livelihoods to the harmful effects of hazards (UNISDR, 2009). Physical, or structural, vulnerability refers to the damage associated with buildings and infrastructure, which determines asset losses. Social vulnerability refers to people ability to cope with the impacts of asset losses on their livelihoods and security. These impacts, along with losses to public assets, are a focus for governments (GFDRR, 2016). Exposure and vulnerability are two distinct concepts. Exposure is a necessary, but not sufficient, determinant of risk. It is possible to be exposed but not vulnerable (for example by living in a floodplain but having sufficient means to modify building structure and behaviour to mitigate potential loss). However, to be vulnerable to an extreme event, it is necessary to be exposed (IPPC, 2012).

1.3 Disaster Risk Management Cycle and Earth Observation

Disaster risk management (DRM) aims to reduce the loss of life and damage to property from disasters. This involves both disaster risk reduction (DRR), such as pre-event

activities of mitigation and preparedness, and post event phases, as emergency response and recovery (see Figure 1-2) (Cova, 1999; Bitelli and Gusella, 2008; Joyce et al., 2009).

Figure 1-2: Disaster risk management cycle.

Disaster *mitigation* refers to any structural and non-structural measure undertaken to limit the adverse impact of hazards. Mitigation measures include building codes; vulnerability analyses updates; zoning and land use management; building use regulations and safety codes; preventive health care; and public education. *Preparedness* is concerned with actions that are taken in advance of an emergency to develop operational capabilities and facilitate an effective response to an emergency. Preparedness is based on analysis of disaster risk and good linkages with early warning systems and includes activities such as contingency planning, stockpiling of equipment and supplies, development of arrangements for coordination, evacuation and public information, and associated training and field exercises (UNISDR, 2004). The *response* phase involves actions that are taken immediately before, during, or directly after an emergency occurs, to save lives, minimize damage to property and enhance the effectiveness of recovery (Cova, 1999). The main focus of this phase is to carry out efficient search and rescue operations, evacuate potential victims, provide food, water and medical care to those in need. In many cases, hazard monitoring systems ensure authorities are promptly alerted to disaster onset by means of systematic forecasts (e.g., hurricanes). In other cases, such as earthquakes, pre-impact prediction is usually not available, and a prompt assessment of the impact area is essential to direct emergency response resources to the most severely affected areas

(Lindell et al., 2006a). After the catastrophic event, the overall goal of the *recovery* phase is returning the community, business, or other entity to conditions the same or better than existed before the disrupting event. It involves short- (relief and rehabilitation) and long-term activities (reconstruction). In the short term, recovery is an extension of the response phase and it focuses on the immediate tasks of securing the impact area, housing victims, and establishing conditions under which households and businesses can begin the process of recovery (Lindell, 2013).

Remote sensing can provide a valuable source of information throughout the emergency management cycle (Joyce et al., 2009). In the pre-event stages, remote sensing data can be exploited for detecting potentially hazardous areas, such as, for example, areas prone to floods and landslides. Understanding of hazards can support mitigation activities such as land use planning and building codes to reduce disaster risk. Very High Resolution (VHR) optical data can be used for land cover and land use mapping to obtain up-to-date information about buildings, infrastructure and other features that may be threatened by hazard events (Deichmann et al., 2011). This furthermore gives information about human exposure since inhabited structures in hazardous areas imply that people are likely to be exposed to injury or death, for example, when a building collapses after an earthquake. Response activities are primarily focused on protecting life and property during a disaster. Remote sensing can be used here to provide immediate damage assessment and to assist research and rescue activities and evacuation plan. Of course, during the response phase, the temporal relevance of remote sensing data is crucial (Joyce et al., 2009).

The usefulness of EO data for disaster management has been proven by numerous ongoing national and international initiatives (Serpico et al., 2012).

The International Charter on Space and Major Disasters is the main mechanism globally by which countries can access satellite EO data in support of their disaster response activities (CEOS, 2015). The International Charter[1] aims to provide a unified system of space data acquisition and delivery to those affected by natural or man-made disasters worldwide. Starting with only three initial members (European Space Agency (ESA), Centre National d' Études Spatiales (CNES) and Canadian Space Agency (CSA)), the Charter has grown to a group of 16 members. It has been activated over 500 times (1

---

[1] www.disasterscharter.org

August 2016), since it became operational in November 2000, providing meaningful mapping and analysis products to the civil-protection and relief organizations at appropriate scale in time and space.

Similar to the Charter is the Emergency Management Service, of the European Program Copernicus[2], previously known as GMES (Global Monitoring for Environment and Security). Copernicus collects data from EO satellites and in situ sensors such as ground stations, airborne and sea-borne sensors, processes these data and provides users with reliable and up-to-date information through a set of services related to environmental and security issues. The Copernicus program is supported by a family of dedicated satellites, the Sentinels, specifically designed to meet the needs of the Copernicus services and their users. Copernicus offers six different service lines: Emergency Management, Atmosphere Monitoring, Marine Environment Monitoring, Land Monitoring, Climate Change and services for Security applications. The Emergency Management Service has two main components: Early Warning and Mapping. The Early Warning component delivers alerts and risk assessments of floods and forest fires. The Mapping component addresses a wide range of emergency situations resulting from natural or man-made disasters, covering in particular floods, earthquakes, tsunamis, landslides, severe storms, fires, industrial accidents, volcanic eruptions and humanitarian crises. The service is provided in two modules. Rapid Mapping provides high-speed service delivery in the midst of or immediately after catastrophic events or humanitarian crises. Risk & Recovery Mapping is designed for pre- or post-crisis situations in support of recovery, disaster risk reduction, prevention, and preparedness activities.

1.4 Earth Observation for earthquake disaster management

Earthquakes with their unpredictability and devastating nature are a challenge for scientists and organizations around the world. Remote Sensing represents a powerful tool for monitoring deformation of the Earth's crust that is one of the most important parameters for the study of the seismic cycle. It provides important information used to

---

[2] www.copernicus.eu

model the mechanisms of tectonic stress accumulation (in the interseismic phase) and release (in the coseismic and postseismic phases) along fault zones. In each individual seismic cycle, the rapid and large surface displacements that accompany an earthquake are typically preceded by a long period of slow, gradual loading, where strain accumulates on the fault segments (interseismic phase). Geodetic observations of crustal strain during the interseismic phase can make important contributions to the assessment of seismic hazard, as the rate of strain accumulation on a fault may be directly related to the rate of earthquake occurrence. Once sufficient strain has been accumulated such that shear stress on the fault exceeds a fault's frictional strength, an earthquake occurs. This phase of the seismic cycle is referred to as the coseismic phase (Salvi et al., 2012; Tolomei et al., 2015).

Synthetic Aperture Radar (SAR) Interferometric (InSAR) techniques have demonstrated the capacity to obtain information on the dynamics of the deformation occurring during the various phases of the seismic cycle (Tolomei et al., 2015). Large earthquake can cause displacements in the order of decimeters that can be clearly identified in SAR interferogram generated using a pair of images collected before and after the seismic event. Landers earthquake occurred in 1992 was the first seismic event analyzed using the InSAR technique (Massonnet, 1993). Measuring interseismic deformation with InSAR is more challenging because the rates of deformation are small (<3 cm/yr) and the deformation signal can be distributed over tens of kilometers (Salvi et al., 2012).

Focusing on urban areas, in the following sections EO capabilities for earthquake damage and vulnerability assessment are presented.

1.4.1 Vulnerability assessment

Due to the rapid sprawling of built-up areas, unplanned settlements and to the general rapid changes in modern cities and mega-cities, building inventory data and related seismic vulnerability information are often outdated, spatially aggregated and discontinuous and, in many earthquake prone regions of the world, simply inexistent. Conventional approaches to assess the seismic vulnerability of the building inventory involves detailed in situ building-by-building analysis by structural engineers, which could be costly and time consuming (Pittore and Wieland, 2013; Geiß at al., 2016). Remote Sensing technology appears to be a challenging and promising approach to rapid inventory and vulnerability assessment. VHR satellite imagery from optical sensor are

characterized by sub-metric resolution suitable for mapping the urban landscape. They have a great potential to extract different building parameters as vulnerability indicators (Mueller at al. 2006; Deichmann et al., 2011; Harb et al., 2015).

Object based image analysis (OBIA) techniques can be exploited to highlight shape characteristics of the building footprint useful to infer the plan irregularity.

Literature reports several examples that focused on extracting building height as an indicator of both exposure and vulnerability. Mück et al. (2013) derived height information from airborne radar data to calculate the building volume and to identify buildings with more than one storey in order to extract potential shelters for vertical evacuation. In Taubenböck et al. (2008) buildings height was derived from an IKONOS image considering the length of the shadows.

The availability of context information is one of the big advantages of remote sensing data for visual inspection of a given study area. The position of a building in relation to other buildings or objects affects its behaviour in an earthquake event. This parameter can be well observed in VHR satellite images. Once the individual buildings are identified in an image scene, the position and distance to other buildings can be directly measured in the image.

Due to change in building codes along the years, classifying the urban areas or single buildings according to their age is another useful indicator of vulnerability. The age of construction can be estimated from multitemporal data set as, for example, done in Taubenböck et al., (2008) or in Borfecchia et al., (2010). Landsat time series, for example, being available since 1979, can be used to obtain an estimation of building age at block level (Taubenböck, 2012).

1.4.2 Damage assessment

A rapid damage assessment, soon after a seismic event, can address the civil protection interventions towards the most affected areas. Satellite data can be very useful for this purpose, thanks to the wide coverage, the high spatial resolution, and their intrinsic capability to provide a synoptic view over remote regions in the world.

Seismic damage assessment based on EO data can be carried out in two ways. One of them is the mono-temporal approach based on the use of just a satellite image collected after the seismic event. The other one is the multi-temporal approach, where two images, acquired before and after the earthquake respectively, are compared to detect changes

associated with the building damage. The spatial scale of the resulting damage map depends on the spatial resolution of the available EO data. Medium resolution (MR) satellite data allow to detect and quantify damage at city block level. In order to detect damages at single building scale, VHR images are necessary. Both optical and SAR data provide useful information for producing a damage map after an earthquake and several examples can be found in the literature.

SAR sensors represent a valuable source of information for time-critical application, because they can deliver images in almost any weather condition and independently of solar illumination conditions that, instead, affect the observations in the optical spectral range. Serious damages or the buildings collapse after an earthquake modify the observed scenario and its electromagnetic response, thus a damage assessment can be performed comparing pre- and post-disaster SAR images. Multi-temporal SAR images can be used to detect changes in urban areas either looking at the backscattering intensity changes, but also exploiting the information on the phase of the returned signal.

Matsuoka and Yamazaki (1999) investigated the different appearance of undamaged and damaged buildings in SAR imagery. By comparing pre- and post-event ERS and JERS SAR intensity data of the 1995 Hyogoken-Nanbu Earthquake, the authors observed that the SAR backscatter value decrease with increasing damage. Pan et al. (2010) investigated the relationship between building damage level and the differences of the backscattered SAR intensity values in pre- and post-disaster ALOS PALSAR imagery for the 2008 Wenchuan Earthquake. Based on ENVISAT ASAR data, Chini at al. (2013) reported a decrease of the SAR backscattering in damaged urban area due to the 2011 Tohoku (Japan) earthquake and caused by the decrease of the double bounce effect.

An index to estimate the damage level from SAR data combining amplitude changes and the related correlation coefficient has been proposed and applied to different case studies (Aoki et al., 1998; Matsuoka and Yamazaki, 2004, 2002). Hoffman et al. (2004) defined a damage index based on InSAR complex coherence which takes into account the phase of the echoed signal. Yonezawa and Takeuchi (2001) compared changes in the SAR backscattering intensity and phase with the damage observed after the 1995 Kobe (Japan) earthquake. This showed that intensity de-correlation and interferometric coherence reduction behave similarly. Using C-band ERS SAR imagery, Matsuoka and Yamazaki (2000) reported that the coherence is more useful to distinguish slight to moderate damage

levels, while the SAR intensity correlation is more sensitive to large surface changes (e.g. caused by stronger damage).

Stramondo et al. (2006) applied SAR intensity correlation and the coherence, focusing on the 1999 Izmit Earthquake (Turkey) and on the 2003 Bam Earthquake (Iran). It was observed that the intensity correlation and the coherence detect different kinds of changes on the ground. Ito et al. (2000) assessed different SAR change indicators, derived from L-band and C-band sensors, and evaluated the frequency dependent effects of spatial and temporal de-correlations.

Most of the approaches to damage detection from SAR images available in the literature are based on single polarization data. From this point of view an innovative approach is proposed in the work of Li et al. (2012) where, considering the 2010 Yushu (China) earthquake, the potential of polarimetric SAR for extracting collapsed building is investigated. The proposed methodology uses only a post-event acquisition (a RADARSAT-2 image) from which three parameters are generated: the entropy H that indicates the random degree of target scattering, the α parameter, representing the average target scattering mechanism and the circular polarization correlation coefficient, ρ, which is very sensitive to man-made targets. The H and α images are used to remove bare soil, while the circular polarization correlation coefficient is used for discriminating between collapsed and intact buildings. Subsequently, Zhao et al. (2013) improved the H-α-ρ method and replaced the parameter of ρ with the normalized circular-polarization correlation coefficient and, at the same time, the homogeneity texture feature was exploited to solve the problem represented by collapsed buildings and buildings divergent to satellite flight pass that can be being confused.

Nowadays, VHR SAR sensors, in principle, allow producing damage maps at single-building scale. However, the detection of sparse and isolated damages in urban areas by exploiting VHR SAR data is still a challenging task. Despite the capability to observe small details over the acquired scenario, the interpretation and analysis of VHR SAR images in urban environment are not easy tasks. Within urban settlements, the scattering mechanisms are complex and simple change detection analyses or classification procedures developed for MR data can hardly perform. As a matter of fact, new methodologies for damage detection have been proposed. Brett and Guida (2013) developed an approach to earthquake damage mapping based on the detection of changes

in the brightness of double bounce features. First, a ridge-detector is used to extract curvilinear features from the pre-event VHR SAR image. Next, the double-bounce candidates are selected using probability distribution functions derived from a simple geometrical building model. Finally, the amplitude ratio, between the pre- and post- event SAR images, is used to calculate a severity-of-change measure for each candidate object. The fundamental drawback of this approach is that, if no double-reflection line is present in the pre-event image due, for example, to adjacent buildings, even very heavy damages cannot be detected. By taking into account the SAR signature of an isolated building, Marin et al. (2015) developed a method for detecting new and fully destroyed buildings from a pair of VHR SAR images. In the first stage of this methodology, the log - ratio image is used in order to identify areas affected by changes in the backscattering at building scale, and an additional step is performed for detecting areas of change showing a size comparable to buildings. In order to identify new or destroyed buildings, each changed candidate building is analysed by means of four fuzzy rules, aiming at verifying the presence of both regions of increase and decrease in backscattering with specific spatial properties and a specific alignment. The fuzzy logic, as stated by the authors, makes the approach robust with respect to the possible deviations from the ideal model considered for changes. Brunner et al. (2010b) proposed a method to assess the structural status of individual buildings in an urban setting hit by an earthquake based on the use of pre-event VHR optical and post-event detected VHR SAR imagery. The method works under the assumption that the buildings have a rectangular footprint and are isolated. The VHR optical pre-event image is used for extracting buildings parameters such as width, length, height and pitch of the roof. These information are exploited along with the acquisition parameters of the post- event SAR scene to predict the SAR signature of the buildings, assuming them undamaged. This task is performed by means of a ray tracing based SAR simulator. The predicted SAR signature is compared to the actual VHR SAR scene using the Normalized Mutual Information as similarity measurement. Based on the value of this match parameter, the decision if a building belongs to the damaged or undamaged building class is taken. A Bayesian classifier is used for this purpose.

Also optical data can provide valuable information on urban settlement conditions after an earthquake. The spatial resolution of satellite optical sensors has rapidly improved in the last few years, reaching less than 1 m (WorldView, Quickbird, Geo-Eye satellites),

thus becoming a reliable tool for detecting changes of individual buildings. As a matter of the fact, they are currently used in operational context, where the most used approach for damage assessment from space images is still the visual interpretation.

In Stramondo et al. (2006) a good agreement between the collapse ratio (the percentage of collapsed buildings with respect to the total number of buildings within a region) from ground truth data, and the normalized difference between one pre-seismic and one post-seismic panchromatic radiance image, was found. Chini et al. (2009a) developed a method for producing a damage map at single building level from VHR panchromatic data. In this approach the mathematical morphology is used to extract man made structures from the pre-event image. Then damaged buildings are detected by a threshold on the absolute value of normalized difference between pre- and post-seismic images.

Bignami et al. (2011) exploited a pair of Quickbird images collected before and after the 2003 Bam earthquake, to investigate the sensitivity of changes in objects textural features with respect to damage levels. The sensitivity analysis was carried out with respect to the damage level as defined by the European Macroseismic Scale 1998 (EMS-98) (Grünthal, 1998). For this purpose, a total of 367 buildings were selected from a ground based damage map. Instead of the usual textural features extraction approach based on co-occurrence matrix computed within a moving window, in this work, for each building, a co-occurrence matrix was computed, considering only pixels within the footprint of the building itself. For each object, contrast, dissimilarity, entropy and homogeneity were extracted following the formulation given in Haralick (Haralick, 1973). The mean value within damage classes from the ground survey was evaluated and the difference between post-seismic and pre-seismic features was computed as change indicator.

Turker and Sumer (2008) detected buildings damaged after the earthquake occurred in Turkey in 1999, applying a watershed segmentation on a post-event optical image. Exploiting the relationship between the buildings and their shadows they correctly classify as either damaged or undamaged the 80% of the buildings. In Gusella et al. (2005), the number of collapsed buildings of the Bam earthquake in 2003, is quantified starting from the inventory of buildings obtained from a Quickbird images taken before the event. Another work that exploit the OBIA approach is presented in Chesnel et al. (2007). In this work, for each building, the footprint, defined in the reference image by a Geographic Information System (GIS) layer manually drawn, is translated in the crisis

image within a search window. For each offset, the correlation among pixels within the footprint is computed and its maximum value, corresponding to the minimum registration error, is considered as similarity measure to decide whether the building is damaged or not. Rastiveis et al. (2013) proposed a method to automatically estimate the damage rate of a building considering shape and structural properties of the roof in the pre- and post-event image. For each building in the scene, located by means of a vector map, the roof detection in the pre- and post- event image is performed using a segmentation approach based on textural and spectral information. After that, using morphological operators, noisy pixels are removed. Finally, by considering the shape and other structural properties of the intact roof part with its pre-event condition in a fuzzy inference system, the rate of damage for each candidate building is estimated.

2. Background and Methodologies

2.1   Electromagnetic response of urban areas

The characteristics of urban areas images acquired by SAR and optical sensors greatly differ. The SAR system is an active sensor that measures the backscatter of a transmitted signal, typically in a narrow microwave frequency band, providing information on both amplitude and phase of the returned signal. Backscattering is primarily determined by geometrical and dielectric properties of the target and the transmit/receive configuration of the SAR sensor (Brunner, 2010b). Optical sensors measure the radiometric properties of reflected sunlight in spectrally distinct regions of the visual and near-infrared spectrum or integrated in a single panchromatic band. The material properties of objects, the illumination conditions of the scene and the sensor perspective determine the radiometric and geometric appearance of distinct targets in optical imagery.

In the development of automatic image analysis and change-detection techniques, sensor properties such as, geometrical, spectral and radiometric resolutions have to be taken into account to properly extract information from the considered data. The purpose of the following sections is to introduce relevant properties of urban areas mapped by SAR and optical sensors.

2.1.1 Radar response of buildings

In this section, the main scattering mechanisms determining the SAR response of a typical urban structure (Franceschetti et al., 2003; Brunner et al., 2010a, 2010b; Guida et al., 2010; Marin et al., 2015) are analysed. For this analysis, we refer to a scene composed by an isolated flat-roof building modelled as a rectangular structure with uniform surfaces and flat surrounding, as the one shown in

Figure 2-1.

By proceeding on the SAR image at constant azimuth from near to far range, we first

identify a bright stripe corresponding to the superimposition of backscattering from ground, building wall and roof that are sensed at the same time being generated from scatterers equidistant from the SAR. This is a typical example of the so-called layover effect that arises whenever the slope of the observed surface is greater than the SAR incidence angle θ. In the ground range plane the length *l* of the area affected by layover is related to the incidence angle θ and the building height *h* through the following equation:

$$l = h \, \cot(\theta) \qquad\qquad (2.1)$$
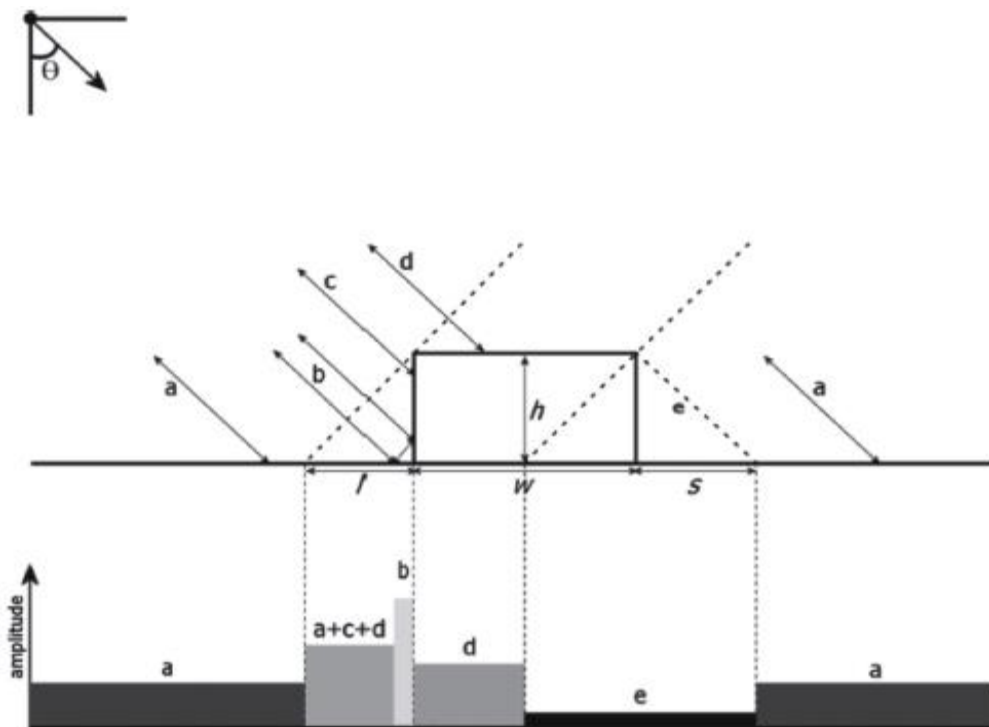
Figure 2-1: Scattering mechanisms from a flat-roof building with width w at height h: a) backscattering from ground; b) double bounce; c) backscattering from vertical wall; d) backscattering from roof; e) shadow. Symbols l and s denote the length of the layover and shadow areas in ground projected image space. Θ is the SAR incidence angle. Different gray levels in the backscattering profile represent different amplitudes. (Brunner et al., 2010b).

At the dihedral reflector composed by building wall and ground several double reflections occur. The superimposition of such contributions due to the fact that double-bounce ray paths all have the same length, equal to the distance between SAR sensor and the dihedral intersection, results in a bright line located in correspondence of the base of the building wall. The strength of the double bounce line depends on the building height. The higher

is the building, the stronger is the double bounce contribution (Brunner et al., 2010a). It also depends on the aspect angle, i.e. the angle φ that defines the orientation of the building with respect to the azimuth direction. When the building wall facing the radar is almost parallel to the azimuth direction (φ=0°), the double bounce contribution has a strong signature. An empirical study on the relationship between double bounce and the orientation of building (Ferro et al., 2011) pointed out that the strength of the double bounce drops off significantly if the aspect angle increases from φ = 0° up to 10° , whereas it decays moderately for higher angles.

The bright stripe due to the double bounce effect is followed by a grey area corresponding to the return from the building roof. Finally, a dark area can be observed due to the shadowing effect that occurs because the region of ground immediately behind the building is prevented from being illuminated. Shadow effect is more pronounced for large SAR looking angle as can be inferred from the equation:

$$l = h \ \tan(\theta) \qquad (2.2)$$

which provide the length *s* of the shadow area as a function of building height and incidence angle.

It is worth to note that, on the base of the relationship between SAR incidence angle and geometric parameter of the building, namely height and width along the range direction, the presented scenario, may change. In order to observe the backscattering profile described so far, the building width *w* along the range direction has to be greater than the limit value defined as:

$$w_b = \frac{h}{\tan(\theta)} \qquad (2.3)$$

where *h* is the building height and *θ* is the incidence angle.

If $w \leq w_b$, the whole roof contribution is sensed in the layover area before the double bounce with the difference that in the case $w < w_b$, the layover area is splitted in two regions, the one corresponding to backscattering contributions from ground, building wall and roof, the other corresponding to contributions from only building wall and ground.

The scattering effects of a gable roof building are different from what is observed for a flat roof building. The main difference is the presence of a second bright feature due to the direct backscattering from the part of the roof that is oriented towards the sensor. It appears closer to the sensor than to the double-bounce contribution (Brunner et al., 2010).

It is worth to note that double bounce and layover areas, as well as dark regions due to shadow effect, can be identified separately if the SAR image resolution is higher than the dimension of the building itself, while this distinction is not possible if the resolution is lower. As a matter of fact, an urban area imaged by a SAR sensor with a resolution in the order of tenths of meters, appears uniformly brighter than other land cover classes since the double-bounce effects are the dominant scattering component in the SAR resolution cell (i.e., the image pixel).

2.1.2 Spectral signature of urban areas in the optical bands

Urban areas are characterized by a complex spatial arrangement of spectrally heterogeneous land cover types which, typically, include man-made structures (buildings, transportation net and parking), green spaces (parks, gardens), bare soil zones and water bodies.

Figure 2-2 shows reflectance curves for some land cover types found in urban areas derived from the Santa Barbara urban spectral library (Herold et al., 2003)
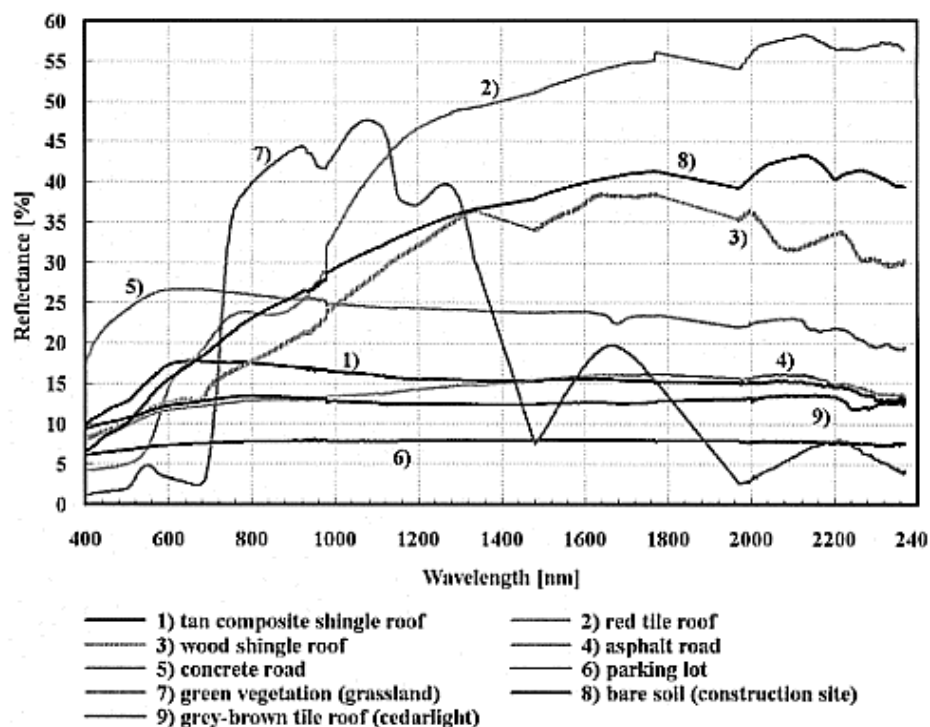
Figure 2-2: Example spectra of typical urban land cover from the Santa Barbara spectral library (Herold et al., 2003).

Focusing on the roof spectra, a great variability among signatures associated with elements belonging to the same semantic class, can be observed. Compared with other

roofing materials, red tile roofs and wood shingle roofs show distinct spectral signatures with a significant reflectance increase in the NIR (Near Infrared) and SWIR (Shortwave Infrared) wavelength regions. Due to the iron-oxide content, red tile roofs signature exhibits absorption features in the visible and NIR, while, the spectrum of roofs with wood shingles shows lignocellulose absorption features in the SWIR, typical of non-photosynthetic vegetation. Conversely, other roof types have near constant, low reflectance, with no significate absorption features and they share similar spectral characteristics with other impervious surfaces such as asphalt roads and parking lots. Similarity can also be observed between wood shingle roofs and bare soil spectra (Roberts and Herold, 2004; Herold et al. 2003, 2004). Some urban land cover types are thus difficult to separate especially at coarse spectral resolution.

Besides the spectral resolution, the discrimination among land covers depends on the sensor spatial resolution. In remotely sensed images, the measured spectral reflectance of a pixel is the integration of contributions from all the objects within the Ground Instantaneous Field Of View (GIFOV) of an optical sensor. Consequently, spectral heterogeneity at scales comparable to the GIFOV results in a preponderance of spectrally mixed pixels. This is the case of an urban area imaged by medium resolution sensors, such, for example, the Landsat TM/ETM+ and the Spot HRV sensors, characterized by spatial resolutions of 30 and 20 meters, respectively (Small, 2003).

2.2   Image change detection approaches

In the context of remote sensing, change detection is the process that leads to the identification of changes occurred on the Earth surface by jointly processing two or more images acquired on the same geographical area at different times (Bruzzone and Bovolo, 2013).

The latest generation of space-borne SAR and optical sensor provides images with metric or submetric resolution allowing the identification of individual urban structures and thus their changes. In metric resolution images, a single pixel represents a small area of the acquired scene, usually smaller than the size of the objects of interest. Hence, the information of a single pixel is strongly linked to the information carried by its neighbours which are part of the same object, pointing out the necessity to perform an object-oriented analysis, although the measureable statistics typically exploited for classification

purposes are necessarily pixel-based (Blaschke, 2010; Chini, 2014a). This aspect is very interesting, because it makes it possible to focus on a single object, for example a building or also other man-made structures, having the advantage of a valuable amount of pixels to characterise the object itself.

As opposed to traditional pixel-based change detection methods which implicitly assume that neighbouring pixels are relatively independent to each other, OBIA techniques recognize that important semantic information is not always represented in single pixels but in meaningful image objects and in their contextual relations. Usually, object-based classification techniques comprises two steps: the image segmentation and the object classification. Image segmentation sub-divides the image into groups of contiguous pixels called objects or segments that correspond to meaningful features or targets in the field (Blaschke, 2010). Segmentation is based on the spectral information and local patterns or textural information that are inherent to groups of neighbouring pixels. As a result of the above step, the object-based classifications may consider a wide range of variables such as mean reflectance, texture, shape, size of objects, and can potentially produce more accurate and detailed maps than conventional classification strategies (Mathieu and Aryal, 2007; Blaschke, 2010; Chini et al., 2014b).

A complete change detection algorithm, as shown in Figure 2-3, is composed of several steps in cascade that will be detailed in the following with particular attention to specific techniques exploited for this research work.
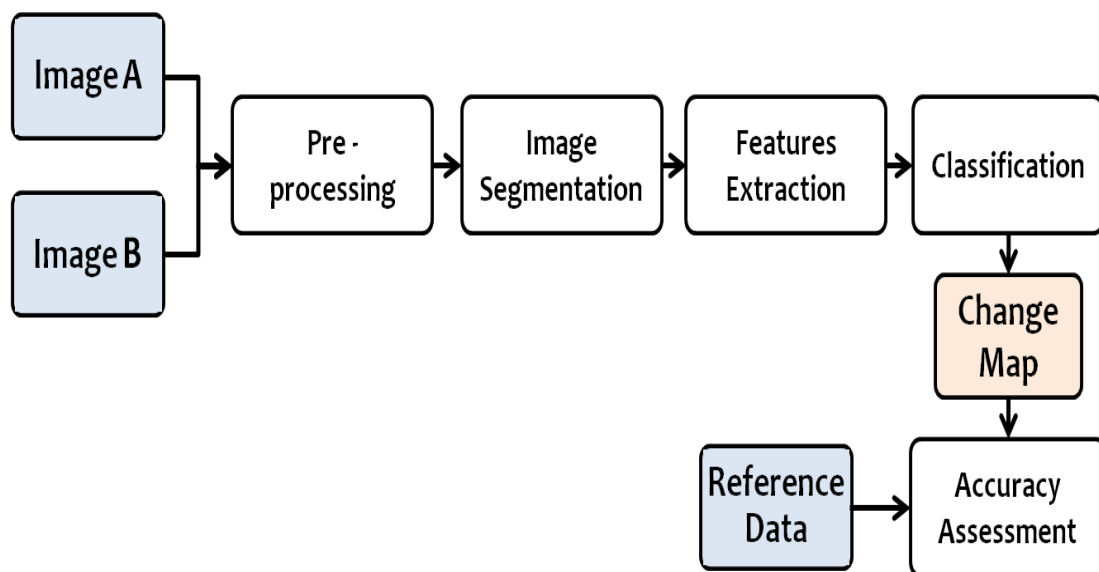
Figure 2-3: Block diagram of a standard object based change detection approach.

2.2.1 Image pre-processing

The image pre-processing for change detection aims at generating two images that are as much similar as possible unless in changed areas (Chen, 2007).

An accurate co-registration of the images to be compared is the first indispensable processing step to properly perform change detection as it allows the derivation of a pair of images where corresponding pixels are associated to the same position on the ground. The second critical requirement for successful change detection is that images collected at different times are comparable in terms of radiometric characteristics (Coppin et al. 2004).

This usually does not happen for optical images for several reason, such as, for example, variations in solar illumination conditions, atmospheric scattering and absorption, changes in the sensor-target-illumination geometry. Therefore, in order to use multi-date images for quantitative analysis based on radiometric information, as in the case of change detection, radiometric divergences have to be corrected (Mas, 1999). Two approaches to radiometric correction are possible: absolute and relative. The absolute approach converts image Digital Numbers (DNs) in ground reflectance values and requires the use of ground measurements at the time of data acquisition for atmospheric correction and sensor calibration. The relative approach consider one image as reference and adjusts the DN values of the other image to match the reference. Relative Radiometric Normalization methods may be further subdivided into three groups: statistical adjustments, histogram matching, and linear regression normalization. The statistical adjustments approach includes methods based on the linear adjustment of two images to resemble each other in terms of their dynamic range (minimum and maximum DN values), statistical mean and standard deviation, or other possible statistical variables. Histogram matching is a radiometric enhancement technique that transform an image so that its histogram correspond as closely as possible to the one of a reference image. It involves determining a lookup table that convert the histogram of one image to resemble that of another. It is, therefore, a useful technique for matching image data of the same scene acquired at different dates with slightly different sun angles or atmospheric effects. Linear regression normalization includes different methods such for example the radiometric normalization based on image regression. This approach involves relating

each pixel of the subject image with that in the reference image, band by band, to produce a linear equation through a least-squares regression (Yang and Lo, 2000).

SAR has the advantage of being less affected by atmospheric condition. Nevertheless, SAR images should be calibrated and normalized if a proper change detection analysis is to be conducted. Radiometric calibration involves corrections to compensate for near range/far range differences caused by the antenna pattern, and normalization of radar cross section to the actual illuminated area within the resolution cell, which differs because of local incidence angle modulation due to topography. Radiometric normalization equalizes the backscattered energy in the range direction by a modified cosine correction, common to all pixels, in order to partially account for the change of backscattering coefficient with incidence angle (which in fact should be adapted to the target category) (Pierdicca et al. 2014).

2.2.2 Image segmentation

The segmentation technique appears as one of the main approaches to overcome the paradigm to consider the pixel as a reference unit for the classification analysis, transforming an image by merging pixels that belong to homogeneous areas. In a segmented image, the geometric structures of objects are depicted, making it possible to base subsequent statistical analysis on objects composed of several pixels rather than on single pixels.

To single out spatial homogeneity and identify objects with different spatial extension, the mathematical morphology can be exploited (Benediktsson et al., 2005, 2003; Chini et al., 2009a; Pesaresi and Benediktsson, 2001; Pulvirenti et al., 2011).

Erosion and dilation are the basic operators of the mathematical morphology (Soille, 2003). They are applied to images by a structuring element (SE), which is a kind of filtering window of fixed shape, corresponding to specific geometrical figures, in order to evaluate how structures in the image match that specific element. The two morphological operators derived from erosion and dilation are the opening and closing ones. The function of opening is to dilate an eroded image in order to recover the original image as much as possible. On the contrary, the function of closing is to erode a dilated image in order to recover the initial shape of image structures that are dilated. In particular, they are used to isolate bright (opening) and dark (closing) structures in the image. For bright and dark, we mean brighter and darker with respect to other

neighbouring structures. The main characteristic of the filtering process performed by the opening and closing operators is that not all the structures within the original image are recovered when these operators are applied sequentially. The size of the SE with respect to the size of the actual structures of the scene influences the output of the filtering operation, so that it may be necessary to use SEs with different sizes and shapes for the characterisation of the various objects. Mathematical morphology can be exploited in a multi-scale approach based on a range of different SE sizes to investigate the various spatial domains in a scene (Pesaresi and Benediktsson, 2001).

2.2.3 Change detection features extraction

Once objects are extracted by image segmentation, they can be described and compared according to various features that include colour, texture, shape, and context properties in several forms. The simplest and more common methods for detecting changes from multi-temporal SAR and optical imagery are Image Differencing and Image Rationing. Due to the multiplicative nature of the speckle, the ratio operator is often preferred when a pair of SAR images is considered. Both image difference and rationing assume that a change in the scene appear as a modification of the mean value of the image evaluated locally, i.e. considering only pixels pertained to a particular image object.

2.2.3.1 Change detection metrics based on textural characteristics

Texture, a representation of the spatial relationship of grey levels in an image, is an important characteristic for the automated or semi-automated interpretation of digital imagery.

Texture analysis provide supplementary information about image properties useful for discriminating objects with similar radiometric properties. In the remote sensing literature, many examples of the use of textural parameters have been proposed for the extraction of quantitative information of building density (Karathanassi et al., 2000) or for the recognition of different urban patterns (Zhang et al., 2003) from optical imagery. When changes occur in the land cover, textures also change. Therefore, texture can be exploited to improve change detection reliability.

Many approaches to texture analysis exist, usually categorized into geometrical, statistical, model-based and signal processing methods (Tuceryan and Jain, 1993).

Statistical methods (Srinivasan and Shobha, 2008) analyse the spatial distribution of grey levels by computing local features at each point in the image, and deriving a set of

statistics from the distributions of the local features. Depending on the number of pixels defining the local feature, statistics are classified into first-order (one pixel), second-order (two pixels) and higher-order (three or more pixels) statistics. First-order statistics describe texture properties based on individual pixel values, without considering the spatial relationship between image pixels. They can be computed from the local histogram of pixel intensities in the image. Variance and mean value of grey levels evaluated locally in the image are examples of the first-order statistics. Second- and higher-order statistics estimate properties of two or more pixel values occurring at specific locations relative to each other. The Gray Level Co-occurrence Matrix (GLCM) method is a way of extracting second order statistical texture features.

The GLCM described in the original Haralick (Haralick et al., 1973) paper is a symmetrical matrix whose element *G(i,j)* of the GLCM counts how many times two pixels with grey levels *i* and *j* occur in the image separated by a given distance *d* along a given direction θ.

Various texture measures can be derived from the GLCM. Among the 14 parameters originally proposed in Haralick et al., six parameters, namely Energy, Contrast, Variance, Correlation, Entropy and Inverse Difference Moment (also known as Homogeneity) are considered to be the most relevant, some of which are strongly correlated with each other (Cossu, 1988).

Let g(*i,j*) be the element (*i,j*) of the normalized GLCM (i.e., the one having sum of its elements is equal to 1), the aforementioned features are defined as follows:

$$Contrast = \sum_{i,j=1}^{N_g} |i - j|^2 g(i,j) \qquad (2.4)$$

$$Correlation = \frac{\sum_{i,j=1}^{N_g} (ij)g(i,j) - \mu_X \mu_Y}{\sigma_X \sigma_Y} \qquad (2.5)$$

$$Energy\ or\ Angular\ Second\ Moment = \sum_{i,j=1}^{N_g} g(i,j)^2 \qquad (2.6)$$

$$Homogeneity = \sum_{i,j=1}^{N_g} \frac{1}{1 + (i-j)^2}\, g(i,j) \qquad (2.7)$$

$$Entropy = -\sum_{i,j=1}^{N_g} g(i,j) \log g(i,j) \qquad\qquad (2.8)$$

$$Variance = \sum_{i,j=1}^{N_g} (i-\mu)^2 g(i,j) \qquad\qquad (2.9)$$

where N_g is the number of distinct grey levels in the image, while μ_x, μ_y and σ_x, σ_y, are mean values and standard deviations of the GLCM row and column sums.

Energy, also known as *Angular Second Moment*, measures textural uniformity, i.e., pixel pairs repetitions. High energy values occur when the image segment under consideration is homogeneous (only similar grey levels are present) or when it is texturally uniform (the grey level distribution has a periodic form so that the shift vector always falls on the same (*i,j*) grey-level pair). In these cases, few elements of GLCM will be greater than 0 and close to 1, while many elements will be close to 0, leading to energy measurement close to its maximum, equal to 1. *Variance* is a measure of the heterogeneity and it is strongly correlated to first order statistics such as standard deviation. It quantifies the dispersion around the mean. *Contrast* take into account the difference between grey levels of neighbour pixels. A low contrast image segment is not necessarily characterized by a low variance value, but the low contrast corresponds to low spatial frequencies. *Homogeneity* assumes higher values for smaller DN differences in pair elements. Therefore, this parameter is sensitive to the presence of near diagonal elements in the GLCM. *Entropy* measures the disorder in an image. When the image is not uniform, many GLCM elements have very small values, which implies that Entropy is very large. If we consider a window with completely random grey tones, the histogram for such a window is a constant function, i.e., all *g(i,j)* are the same, and Entropy reaches its maximum.

Correlation is expressed by the correlation coefficient between two random variables i and j, and it is a measure of the linear-dependencies between values within the image. High Correlation values imply a linear relationship between the grey levels of pixel pairs. Thus, Correlation is uncorrelated with Energy and Entropy, i.e. to pixel pair repetitions (Baraldi and Parmiggiani, 1995; Pacifici et al., 2009).

2.2.3.2 Statistical similarity measures from Information Theory

Recently in the literature, statistical similarity measures such as Kullback-Leibler (KL) Divergence and Mutual Information (MI) have been successfully applied to highlight

change information in multitemporal images. The MI is a commutative measure of the difference between the joint probability distribution $f_{X,Y}(\mathbf{x},\mathbf{y})$ and the marginal probability distribution $f_X(\mathbf{x})$ and $f_Y(\mathbf{x})$ of the random variables X and Y, respectively. It measures how much knowing one of the two variables reduces the uncertainty about the other and it is defined as (Erten et al., 2012):

$$MI(\boldsymbol{X},\boldsymbol{Y}) = \int\limits_{Y} \int\limits_{X} \log\left(\frac{f_{X,Y}(\mathbf{x},\mathbf{y})}{f_X(\mathbf{x})f_Y(\mathbf{x})}\right) f_{X,Y}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} \qquad (2.10)$$

The KL Divergence (Kullback and Laibler, 1951), also called Relative Entropy, is a measure of the difference between two probability density functions (pdfs) and then it can be used for assessing the similarity of pixel intensity distributions of two objects. Given two n-dimensional random vectors X and Y with pdfs $f_X(\mathbf{x})$ and $f_Y(\mathbf{x})$, respectively, the KL Divergence between the two distributions or, equivalently, the relative entropy of X with respect to Y, is given by:

$$KL(\boldsymbol{X},\boldsymbol{Y}) = \int f_X(\mathbf{x}) \log\left(\frac{f_X(\mathbf{x})}{f_Y(\mathbf{x})}\right) d\mathbf{x} \qquad (2.11)$$

The KL Divergence is always non-negative, and equals zero only if $f_X(\mathbf{x}) = f_Y(\mathbf{x})$, but it is not a true metric because it is non symmetric. A symmetric version, called KL Distance (KLD), can be defined as:

$$KLD(\boldsymbol{X},\boldsymbol{Y}) = KL(\boldsymbol{X},\boldsymbol{Y}) + KL(\boldsymbol{Y},\boldsymbol{X}) \qquad (2.12)$$

In order to estimate the KLD or the MI, the pdfs of the two variables to be compared have to be known. For two multivariate normal distributions, with mean vectors $\boldsymbol{\mu}_X$, $\boldsymbol{\mu}_Y \in \mathbb{R}^n$, covariance matrices $\mathbf{C}_X$, $\mathbf{C}_Y \in \mathbb{R}^{n \times n}$, and cross-covariance matrix $\mathbf{C}$, it is possible to derive a closed-form expression for both MI and KLD:

$$\begin{aligned} KLD(\boldsymbol{X},\boldsymbol{Y}) = \frac{1}{2}[tr(\boldsymbol{C}_X^{-1}\boldsymbol{C}_Y) + tr(\boldsymbol{C}_Y^{-1}\boldsymbol{C}_X)] \\ + \frac{1}{2}[(\boldsymbol{\mu}_X - \boldsymbol{\mu}_Y)^T\boldsymbol{C}_X^{-1}(\boldsymbol{\mu}_X - \boldsymbol{\mu}_Y) \\ + (\boldsymbol{\mu}_X - \boldsymbol{\mu}_Y)^T\boldsymbol{C}_Y^{-1}(\boldsymbol{\mu}_X - \boldsymbol{\mu}_Y)] - n \end{aligned} \qquad (2.13)$$

$$MI(\boldsymbol{X},\boldsymbol{Y}) = -\frac{1}{2}\log\left(\frac{det(\boldsymbol{C})}{det(\boldsymbol{C}_X)det(\boldsymbol{C}_Y)}\right) \qquad (2.14)$$

In the case of two univariate Gaussian distributions, with means μ_x, μ_Y and variances σ_X^2, σ_Y^2, and covariance σ_{XY}^2, equations (2.13) and (2.14) simplify to the following

expressions:

$$KLD(X,Y) = \frac{1}{2}\left(\frac{\sigma_X^2}{\sigma_Y^2} + \frac{\sigma_Y^2}{\sigma_X^2}\right) + \frac{1}{2}\left[\frac{(\mu_X - \mu_Y)^2}{\sigma_X^2} - \frac{(\mu_X - \mu_Y)^2}{\sigma_Y^2}\right] - 1 \qquad (\,2.15\,)$$

$$MI(X,Y) = -\frac{1}{2}\log\left(\frac{\sigma_X^2\sigma_Y^2 - \sigma_{XY}^2}{\sigma_X^2\sigma_Y^2}\right) = -\frac{1}{2}\log(1 - \rho_{XY}^2) \qquad (\,2.16\,)$$

where ρ_{XY} is the correlation coefficient. It can be noted that, in the univariate case (2.16), the MI reduces to a function of the correlation coefficient.

Inglada and Mercier (2007) derived an expression for the KLD that does not require a parametric assumption about the two pdfs to be compared. Assuming that the density to be approximated is not too far from a Gaussian pdf, they demonstrated that it is possible to model the shape of a statistical distribution using the infinite Edgeworth series expansion of cumulants truncated at a given order. In this work, the expansion of the two pdfs truncated to order of four has been introduced into the KLD expression, leading to the so-called cumulant-based approximation of the KLD (CKLD). The use of the Edgeworth series expansion allows involving in the comparison step statistics of higher order, such as skewness and kurtosis that provide information about symmetry and flatness of a pdf. The CKLD is useful for detecting change in SAR images that typically do not assume a Gaussian distribution.

2.2.3.3 Correlation based change metrics

The use of correlation coefficient to detect changes stems from the consideration that grey levels from the same object observed at two different times tend to be highly correlated if the objects does not change, and less correlated when changes have occurred in the surface target.

Given two objects extracted from SAR data, the correlation can be estimated considering backscattering intensity values as well as complex values describing both amplitude and phase of the radar returns (Chini et al., 2009a; Stramondo et al., 2006). The complex correlation coefficient ρ_C, also known as interferometric coherence, is mathematically obtained through the following equation:

$$\rho_C = \frac{E(s_1 s_2^*)}{\sqrt{E(s_1 s_1^*)E(s_2 s_2^*)}} \qquad (\,2.17\,)$$

where s_1 and s_2 are the corresponding complex pixel values of the two considered images,

the symbol * denotes the complex conjugate and finally $E(\cdot)$ represents the expected value.

The intensity correlation ρ_I is defined as:

$$\rho_I = \frac{\left|E\left[\left(I_1 - E(I_1)\right)\left(I_2 - E(I_2)\right)\right]\right|}{\sqrt{E\left(I_1 - E(I_1)\right)^2 E((I_2 - E(I_2)))^2}} \qquad (\,2.18\,)$$

where e I_1 and I_2 are the corresponding values of the pixel intensity I = |s|.

Although these two parameters can measure how much the object has changed between the two SAR acquisitions, they bring slightly different information about the changes in the scene. The intensity correlation is more related to the change in the statistics of magnitude of the radar return, while the complex coherence is mostly influenced by the phase difference between radar returns, a distinctive parameter measured by a coherent sensor such as a SAR, and it is particularly related to the spatial arrangement of the scatterers within the pixel and thus to their possible random displacements. The interferometric coherence is an important change indicator especially in urban areas, where man-made structures show themselves coherent in time unless sensible changes have occurred (Chini et al., 2015a). However it is worth to recall that the increase of spatial baseline, in particular its perpendicular component, implies a decrease of the coherence, thus reducing the capability of this feature to discriminate between changed and unchanged targets, especially in urban areas where the geometrical complexity of structures accentuate this effect (Bignami et al., 2004).

2.2.4 Classification approaches

Change detection can be carried out according unsupervised or supervised classification techniques. Unsupervised approaches identify regions of change without any additional information besides the raw images considered. Conversely, the supervised approach to change detection requires the availability of a suitable training set for classifier learning. The description provided in the following focuses on supervised classification techniques used in this research, including the Bayesian classification approach and Support Vector Machines.

2.2.4.1 Bayesian Classifiers

The basic problem in classification is to formulate a decision rule, $d(\mathbf{x})$: $\mathbb{R}^d \rightarrow \{\omega 1, \omega_2, \ldots, \omega_C\}$, for classifying a d-dimensional observation $\mathbf{x}$ into one of C competing classes.

In the Bayesian framework, a classification problem is solved predicting the most likely class for each observed features vector. Formally (Flusser et al., 2016), given a set of C possible classes $\Omega = \{\omega_1, \omega_2, \ldots, \omega_C\}$ and an input sample described by a d-dimensional features vector $\mathbf{x} = \{x_1, x_2, \ldots, x_d\}$ representing the observed values of d attributes, the Bayesian classifier assign the sample to the class $\omega^* \in \Omega$ such that:

$$\omega^* = argmax_{\omega_k \in \Omega}\big(P(\omega_k \mid \mathbf{x})\big) \qquad\qquad (2.19)$$

i.e to the class with the Maximum A Posteriori (MAP) probability $P(\omega_k \mid \mathbf{x})$.

According to the Bayes rule, the posterior probability for each class can be expressed as:

$$P(\omega_k \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid \omega_k)\, P(\omega_k)}{p(\mathbf{x})} = \frac{p(\mathbf{x} \mid \omega_k)\, P(\omega_k)}{\sum_{c=1}^{C} p(\mathbf{x} \mid \omega_c)\, P(\omega_c)} \qquad (2.20)$$

being $P(\omega_k)$ the prior probability of the class ω_k, $p(\mathbf{x} \mid \omega_k)$ the class-conditional pdf also known as likelihood function, and $p(\mathbf{x})$ the evidence. The latter acts as normalization factor ensuring that the sum of posterior probabilities, $P(\omega_k \mid \mathbf{x})$ over all values of ω_k equals one.

If the naive Bayes approximation is made, i.e., if it is assumed that the features are conditionally independent given the class, the likelihood function $p(\mathbf{x} \mid \omega_k)$ can be factorized as follows:

$$p(\mathbf{x} \mid \omega_k) = \prod_{i=1}^{d} p(x_i \mid \omega_k), \qquad k = 1,2,\ldots C \qquad (2.21)$$

In practice, only the numerator of equation (2.20) needs to be evaluated, since the evidence is the same for all the classes, and therefore it does not influence the decision process.

Prior probabilities can be empirically estimated from the relative proportion of each class in the training set. Such an approach, however, implies the assumption that the frequency distribution within the training set is representative of the entire domain from which new observations are made. If this assumption does not hold all classes are typically assumed to be equally likely (Yager, 2006). In this case, the classification task reduce to the problem of determining the class that maximizes the likelihood and the resulting classifier is known as Maximum Likelihood Classifier.

Likelihood functions can be estimated from training data using either parametric or non-parametric approaches. The parametric approach to density estimation assumes the shape of the pdf to be known. Such assumptions significantly simplify the problem, since only

the parameters of the chosen family of functions need to be determined. For example, in the case of a Gaussian distribution, the density estimation process reduces to determining the mean vector $\boldsymbol{\mu}$ and the covariance matix $\boldsymbol{\Sigma}$ from the available data set.

If a specific functional form does not provide a good representation of the true density, non-parametric approaches, can be used. The simplest methodology for non-parametric density estimation is represented by the histogram. Data range is divided into non-overlapping regular bins, and the estimated density within each bin is assigned a uniform value, proportional to the number of observations that fall within (Tabak and Turner, 2013). The shape of the resulting distribution is not smooth and dependent on the number of bins as well as the origin of the bins. Kernel Density Estimation (KDE) of a pdf, also known as Parzen window estimation, can be viewed as a generalization of the histogram that does not require a partition of the data range arbitrarily chosen. The KDE of a pdf is obtained as the sum of d-dimensional kernel functions, $K(\cdot)$, (where d is the dimension of the features space) placed on each data point as shown in the following equation (Parzen, 1962; Silverman 1986):

$$f(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad \mathbf{x} \in \mathbb{R}^d \qquad (\ 2.22\ )$$

where n is the number of samples used to estimate $f(\mathbf{x})$ and h is the so called bandwidth parameter.

2.2.4.2 **Support Vector Machine**

A Support Vector Machine (SVM) is a supervised machine learning algorithm for solving classification, as well as, regression problems.

Conceptually, given a binary classification problem, a SVM maps input patterns into a feature space of higher dimension where an optimal hyperplane is constructed in order to separate patterns belonging to different classes (Cortes and Vapnik, 1995; Vapnik, 1999), as sketched in Figure 2-4.

A hyperplane in the feature space is described by the following equation:

$$\mathbf{w} \cdot \boldsymbol{\Phi}(\mathbf{x}) + \mathbf{b} = \mathbf{0} \qquad (\ 2.23\ )$$

where $\Phi: \mathbb{R}^d \rightarrow \mathbb{R}^H$ is a vector function which maps the d-dimensional input vector $\mathbf{x}$ into a H-dimensional (with H>d) feature space, $\mathbf{w} \in \mathbb{R}^H$ is a vector perpendicular to hyperplane and b is a scalar term.
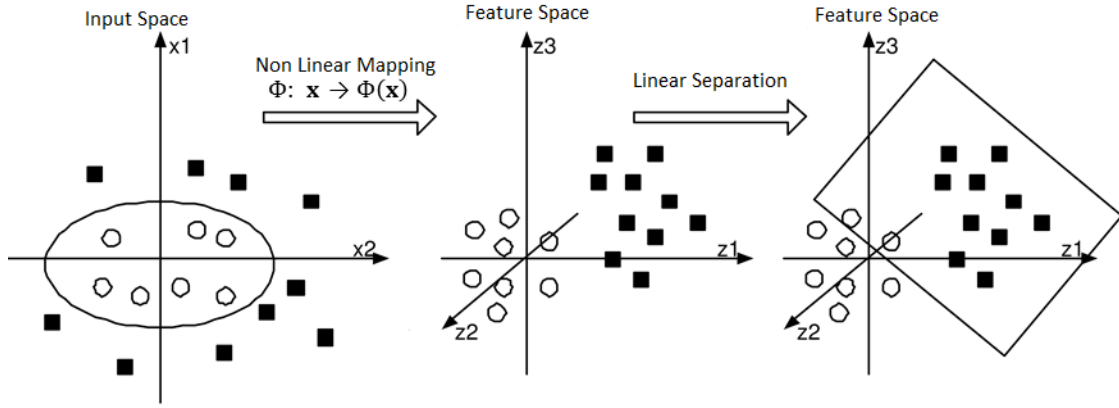
Figure 2-4: The idea of the SVM: mapping of the input patterns into a higher dimension feature space where an optimal hyperplane is constructed in order to separate patterns belonging to different classes.

Given a set of N training samples (x_i, y_1), ..., (x_N, y_N), where $\mathbf{x_i} \in \mathbb{R}^d$ is the d-dimensional feature vector representing the ith training sample, and $y_i \in \{-1;1\}$ is the class label associated with $\mathbf{x_i}$, a SVM constructs the best separating hyperplane solving the following constrained optimization problem:

$$\min_{w,b,\xi} \frac{1}{2}\mathbf{w} \cdot \mathbf{w} + C \sum_{i=1}^{N} \xi_i \qquad (2.24)$$

$$subject\ to: \begin{cases} y_i(\mathbf{w} \cdot \Phi(\mathbf{x_i}) + b) \geq 1 - \xi_i, \ \forall\ i = 1,..N \quad (a) \\ \xi_i \geq 0, \ \forall\ i = 1,..N \qquad\qquad\qquad\qquad (b) \end{cases} \qquad (2.25)$$

where the slack variables ξ_i and the regularization parameter C are introduced to deal with non linearly separable training data.

When the training samples are linearly separable in the feature space, $\xi_i = 0\ \forall\ i = 1,..,N$. In this case the SVM algorithm constructs the so called maximum margin hyperplane which is the hyperplane that separates the training data without errors (i.e. the one satisfying the constraints $y_i(\mathbf{w} \cdot \Phi(\mathbf{x_i}) + b) \geq 1$, $\forall\ i = 1,..,N$) and maximizes the distance between itself and the closest training vectors of each class. Such a distance, called margin, is equals to $2/(\mathbf{w} \cdot \mathbf{w})$ and it is maximized by minimizing $(\mathbf{w} \cdot \mathbf{w})/2$. Greater is the margin, greater is the SVM generalization ability.

If training samples cannot be separated without error, the parameters $(\mathbf{w}, b)$ that minimize the functional in the (2.24) under the constraints defined by (2.25), determine the hyperplane that minimize the training error, measured through the sum of the slack variables ξ, and separates the rest of the elements with the maximum margin. For each

33

training sample lying on the wrong side of the decision boundary (i.e those with associated a slack variable $\xi_i > 1$) but also for the training samples falling within the margin on the correct side of the decision boundary (i.e the samples for which $0 < \xi_i \leq 1$) a penalty is introduced, increasing the objective function by $C\xi_i$. A too small C value, determines many misclassifications. Conversely, a too large C value may lead to overfitting problems. The user parameter C controls the tradeoff between the two goals of the optimization problem.

The constrained minimization problem defined by the equations (2.24) and (2.25) is usually translated into the following dual problem:

$$\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i\, \alpha_j\, y_i y_j\, \Phi(\mathbf{x_i})\cdot\Phi(\mathbf{x_j}) \qquad (\,2.26\,)$$

$$subject\ to: \begin{cases} 0 \leq \alpha_i \leq C, & \forall\, i = 1,..N \quad (a) \\ \sum_{i=1}^{N} \alpha_i\, y_i = 0, & \forall\, i = 1,..N \quad (b) \end{cases} \qquad (\,2.27\,)$$

where the variables α_i are Lagrange multipliers corresponding to the constraints defined by (2.24a).The solution $\boldsymbol{\alpha}^*$ of such dual problem is related to the parameter $\mathbf{w*}$ defining the optimal hyperplane trough the following relation:

$$\mathbf{w}^* = \sum_{i=1}^{N_s} \alpha_i^*\, y_i \Phi(\mathbf{x_i}) \qquad (\,2.28\,)$$

where the summation is limited to the training vectors for which $\alpha_i \neq 0$, called support vectors. They are the only ones that determine the final decision function, i.e. removing a non support vector from the training set, does not change the solution found by the SVM algorithm.

As for the scalar term b* of the optimal hyperplane, it is determined from the KKT (Karush Kuhn Tucker) complementarity conditions (Burges, 1998).

The decision function corresponding to the optimal hyperplane can be written as:

$$f(\mathbf{x}) = sign(\mathbf{w}^* \cdot \Phi(\mathbf{x}) + b^*) = sign\left(\sum_{i=1}^{N_s} \alpha_i^*\, y_i \Phi(\mathbf{x_i}) \cdot \Phi(\mathbf{x}) + b^* \right) \qquad (\,2.29\,)$$

It can be note that both the dual optimization problem (2.26-2.27) and the decision function (2.29) depend on the images of the samples in the feature space only through

their dot product. By replacing all dot products $\Phi(\mathbf{x_i}) \cdot \Phi(\mathbf{x_j})$ with a kernel function K: $\mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, such that:

$$K(\mathbf{x_i}, \mathbf{x_j}) = \Phi(\mathbf{x_i}) \cdot \Phi(\mathbf{x_j}) \qquad (2.30)$$

the mapping of the training data $\mathbf{x}$ in the feature space by means of the function Φ does not need to be explicitly performed.

When the dot products are replaced with the kernel function, the decision function (2.29) became:

$$f(\mathbf{x}) = sign(\mathbf{w}^* \cdot \Phi(\mathbf{x}) + b^*) = sign\left(\sum_{i=1}^{N_s} \alpha_i^* \, y_i K(\mathbf{x_i}, \mathbf{x}) + b^*\right) \qquad (2.31)$$

where the parameters $\boldsymbol{\alpha}^*$ are retrieved maximizing the functional:

$$W(\boldsymbol{\alpha}) = \max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \, \alpha_j \, y_i y_j K(\mathbf{x_i}, \mathbf{x_j}) \qquad (2.32)$$

subject to the constraints defined by (2.27)

Valid kernel functions are those satisfying the Mercer's theorem (Mercer, 1909). The most commonly used are the following (Hsu et al., 2010):

1. linear: $K(\mathbf{x_i}, \mathbf{x_j}) = \mathbf{x_i} \cdot \mathbf{x_j}$

2. polynomial: : $K(\mathbf{x_i}, \mathbf{x_j}) = (\gamma \mathbf{x_i} \cdot \mathbf{x_j} + r)^d, \gamma > 0$

3. radial basis function (RBF): $K(\mathbf{x_i}, \mathbf{x_j}) = exp\left(-\gamma \|\mathbf{x_i} \cdot \mathbf{x_j}\|^2\right), \gamma > 0$

4. sigmoid: $K(\mathbf{x_i}, \mathbf{x_j}) = tanh(\gamma \mathbf{x_i} \cdot \mathbf{x_j} + r), \gamma > 0$

Here $\gamma, r, d,$ are the kernel parameters, and they have to be tuned, together with the regularization parameter C, during the model selection phase.

A cross-validation (CV) approach is typically used for this purpose. The training set is split into k disjoint subsets of approximately equal size. Each subset is used for validating the SVM trained on remaining k-1 subsets. At the end of the kth iteration, the classification performances are evaluated. Various parameters values are tried and the ones giving the best classification performance are chosen. Usually the percentage of data correctly classified during the CV procedure is considered as objective function. A grid search in the parameter space can be performed, which is an exhaustive method but, on the other hand, it can be time consuming. An alternative approach for carrying out the

parameters optimization is the use of a gradient descendent algorithm in order to find the parameters giving the minimum CV error (Chapelle et al., 2002).

In the standard SVM algorithm introduced so far, the regularization parameter C equally penalizes misclassification of positive and negative samples. When faced with an unbalanced dataset where the number of negative instances is significantly higher than that of positive ones, SVMs that follow the formulation of equations (2.24-2.25) tend to produce a decision boundary severely skewed towards the minority class (Akbani et al., 2004; Batuwita, and Palade, 2013).

To cope with this issue, Veropoulos et al. (1999) proposed an approach that uses different misclassification costs, C^+ and C^-, for positive and negative classes.

In this method, the SVM optimization problem defined by equation (2.24) and (2.25) is modified by replacing the penalty term, $C \sum_{i=1}^{N} \xi_i$, with the sum of two terms, one for each class:

$$C \sum_{i=1}^{N} \xi_i \; \rightarrow \; C^+ \sum_{i\,|\,y_i=1} \xi_i + C^- \sum_{i\,|\,y_i=-1} \xi_i$$

By assigning a higher cost C^+ to the misclassification of samples from the minority class, here assumed to be the positive one, the effect of class imbalance could be reduced. That is, the modified SVM algorithm would not tend to skew the separating hyperplane towards the minority class samples to reduce the total misclassifications as they are now assigned with a higher misclassification cost.

As stated at the beginning of this section, SVMs are designed to solve binary classification problems. In order to address multiclass problem two methods are typically adopted: the One-Against-One (OAO) strategy and the One-Against-All (OAA) strategy (Melgani and Bruzzone, 2004; Pierdicca et al., 2014). The OAA approach constructs one SVM for each class. The mth SVM discriminates between samples belongings to m^{th} class, considered as positive class, from the samples in all remaining classes, considered as negative class. The class assigned to a test sample $\mathbf{x}_i$ is the one corresponding to the SVM with the highest value of the discriminant function $f_m(\mathbf{x}_i) = \mathbf{w}_m \cdot \Phi(\mathbf{x}_i) + b_m$. The OAO strategy, instead, uses one SVM for each pair of classes. Given a test sample, for each classification, one vote is given to the winning class and the sample is assigned to the class having most votes.

2.2.5 Accuracy assessment

An effective and widely used approach for assessing performances of a change detection algorithm with respect to reference data consists in the analysis of the classification Confusion Matrix (CM) (Congalton and Green, 1999; Foody, 2002; Foody, 2010; Powers, 2011). A CM is a square k×k array, where k is the number of classes, with rows and columns that usually represent predicted and true class labels, respectively. Each element n_{ij} of the CM denotes the number of objects from the actual class j that were assigned to the class i. Thus, the ideal CM is a k×k identity matrix.

Table 2-1 shows the CM for a binary classification problem where an object have to be assigned to one out of the two classes "change" and "no-change", which we refer to as positive and negative. The main diagonal entries indicate the number of True Positive and True Negative, i.e. the number of positive and negative cases for which classification and reference data agree. As for off-diagonal elements, they represent misclassifications. In particular, n_{12} gives the number of negative cases classified as positive, known as False Positives or False alarms, while n_{21} indicates the number of positive cases classified as negative, often referred to as False Negatives or misdetections.

Actual Class

		change	no-change	Total
Predicted Class	change	**n_{11}** *True Positives*	**n_{12}** *False Positives*	n_{+1}
	no-change	**n_{21}** *False Negatives*	**n_{22}** *True Negatives*	n_{+2}
	Total	n_{1+}	n_{2+}	N

Table 2-1: Confusion matrix for a change detection problem.

From the CM, different quality parameters can be derived, such as Overall Accuracy, Cohen's Kappa coefficient, user's and producer's accuracy. While the overall accuracy represent the percentage of test cases correctly allocated, producer's accuracy focuses on

the accuracy of each individual class. Indeed, it represent the percentage of test cases belonging to a specific class that the algorithm correctly recognize. The producer's accuracies for the positive and negative classes of a binary problem are known as Sensitivity (also called Recall or True Positive Rate) and Specificity (or True Negative Rate), respectively, and they can be computed from the CM entries, as follows:

$$Sensitivity = \frac{n_{11}}{n_{11} + n_{21}} = \frac{n_{11}}{n_{1+}}$$ (2.33)

$$Specificity = \frac{n_{22}}{n_{12} + n_{22}} = \frac{n_{22}}{n_{2+}}$$ (2.34)

User's accuracy instead is computed as number of corrected predictions relative to the overall number of times a specific class was predicted. For a binary classification problem, user's accuracies for the positive and negative classes are respectively called Positive Predictive Value (PPV) or Precision and Negative Predictive Value (NPV) and they can be derived from the CM, according the following equations:

$$Precision \ or \ PPV = \frac{n_{11}}{n_{11} + n_{12}} = \frac{n_{11}}{n_{+1}}$$ (2.35)

$$NPV = \frac{n_{22}}{n_{21} + n_{22}} = \frac{n_{22}}{n_{+2}}$$ (2.36)

A further measure of agreement or accuracy frequently used in remote sensing is Cohen's Kappa coefficient (K). Conceptually it is defined as:

$$K = \frac{p_0 - p_c}{1 - p_c}$$ (2.37)

where p_o is the proportion of cases in agreement (i.e., correctly allocated) and p_c is the proportion of agreement that is expected by chance.

Given a CM, the Cohen's Kappa, can be computed as:

$$K = \frac{N \sum_{i=1}^{K} n_{ii} - \sum_{i=1}^{K}(n_{i+} \cdot n_{+i})}{N^2 - \sum_{i=1}^{K}(n_{i+} \cdot n_{+i})}$$ (2.38)

where k is the number of classes, thus the number of rows and columns of the CM, n_{ii} is entry (i,i) of the CM, n_{i+} and n_{+i} are the marginal totals of row i and column i, respectively, and N is the total number of test samples.

3. Experiment Description

3.1 Introduction and objective

The research work reported in this thesis has been carried out within the framework of the APhoRISM (Advanced PRocedures for volcanIc and Seismic Monitoring) project. APhoRISM[3] is a collaborative project under the theme FP7-SPACE-2013-1 of the Seventh Framework Programme of the European Commission, started in December 2013 and completed three years later (December 2016). It involved partners from five European countries: Istituto Nazionale di Geofisica e Vulcanologia (INGV), project leader, Italy; Sapienza - University of Rome, Italy; Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Spain; Bureau de Recherches Géologiques et Minières (BRGM), France; Gamma Remote Sensing Research and Consulting AG, Switzerland; University of Oxford, United Kingdom and ALMA Sistemi s.a.s, Italy. APhoRISM was aimed at developing and testing methods to generate products useful for seismic and volcanic risk management, through the integration EO satellite data from different sensors and ground data.

As for the seismic theme, which is the subject of this work, APhoRISM has developed a methodology, called APE - A Priori information for Earthquake damage mapping, focused on the generation of products to address the detection and estimate of damage caused by a seism. The use of satellite imagery to survey earthquake damages has been investigated in the literature and exploited in many projects. Usually the approach is based on change detection techniques and classification algorithms applied to satellite images only. This presents many limitations because of the insufficient availability of images (limitations due to cloud cover and satellite revisit time) and the capability of detecting

---

[3] www.aphorism-project.eu

certain type of damage from satellite (e.g., damage of internal structures). The novelty of APE relies on the exploitation of a priori information and concurrent data derived from many other sources, that include geological data (such as soil type, landslide and liquefaction susceptibility), structural, geometric and constructive characteristics of buildings, shakemaps and InSAR time series to measure surface antecedent movements which indicate zone prone to terrain deformation. The main goal of APE is to investigate the potentiality of such a priori information for increasing the reliability and accuracy of earthquake damage maps to be delivered to those involved in rescue management and damage assessment. The outputs of APE method are likelihood index damage maps (LIDaM) at different spatial scale. Two products can be generated, one at district level, where groups of buildings represent the elementary unit, the other at the scale of single buildings. In the first case, MR images can be exploited while, in the second case, images at VHR are needed.

This thesis deals with the generation of damage maps at single buildings level. Procedures were developed considering the earthquake that hit L'Aquila (Italy) on April 6, 2009.

3.2 The A Priori information for Earthquake damage mapping (APE) approach

A prompt response after an earthquake is of primary importance to save lives and effectively manage the emergency. A rapid damage mapping can provide a valuable support to rescue teams and to all the stakeholders involved in the emergency management cycle. In such a context, satellite remote sensing has already proved its potential to contribute to post-earthquake damage assessment. When an earthquake occurs the presence of soil instability factors, such as liquefaction phenomena and seismically induced landslides, may severely damage the surroundings or worsen the damage in areas already affected by the ground shaking. During the interseismic period, soil instabilities, such as natural or anthropogenic subsidence or slow slope movements, may alter the dynamic behaviour of the soil and the overlying urban settlement. Slow or fast subsidence below may induce differential settlements of the building foundation, giving rise to additional stresses on the superstructure, that may increase the vulnerability and worsen the seismic response. When the subsidence rate is very fast, the rapid

deformation may produce the changes in the natural oscillation periods of weak soils. This change is caused by the thickness reduction and the stiffness increment due to the compaction. The sinking and stiffening of soil shorten the site oscillation period, which is the main parameter that controls the seismic response of the soil, thus producing a double resonance in case of structures with lower natural oscillation periods. The APE approach proposed in APhoRISM generates the LIDaM taking into account these factors, that are known to influence the damage level during an earthquake (Kramer, 1996), together with the intrinsic building vulnerability with the aim to reduce false positives and misdetections on change detection maps from satellite imagery.
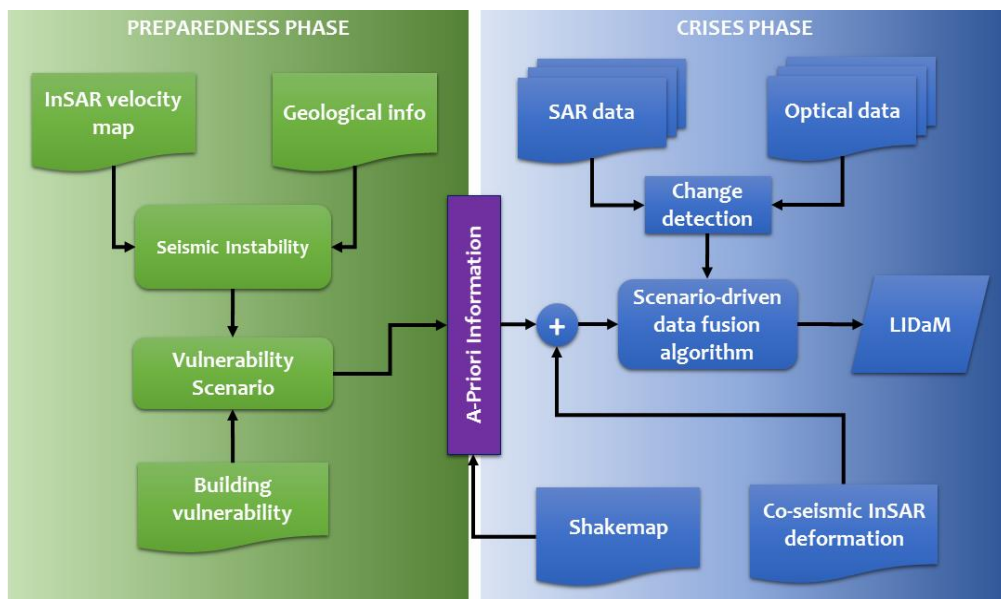
Figure 3-1: Flowchart of the APE method.

Figure 3-1 shows a flowchart of the method and the logical architecture used to generate the LIDaM (Devanthéry et al., 2016). Two main phases are foreseen: a *preparedness* phase (left side of Figure 3-1) and a *crises* phase (right side of Figure 3-1). During the preparedness phase, the APE procedure exploits a long sequence of pre-crisis SAR images in order to derive an InSAR velocity map by means of Persistent Scatterers Interferometry (PSI) (Crosetto et al., 2016). In particular, the Interferometric Point Target Analysis (IPTA) (Werner et al., 2003) method is adopted. The InSAR velocity map allows measuring possible slow surface deformations dealing with natural and/or anthropogenic activities. The outcomes of IPTA are coupled with geological information, such as soil natural oscillation period maps, landslide and liquefaction susceptibility maps, in order to obtain a parameter, called the seismic instability factor, which provide quantitative

information about factors able to increase the effects of an earthquake on buildings and infrastructures. The presence of a potential instability can be taken into account in the estimation of the building vulnerability together with information about building height, year of construction and structural typology, leading to the finalization of the so-called Vulnerability Scenario.

As soon as an earthquake occurs, the crises phase begins. The institutions in charge of the seismic networks generate the shakemaps that provide information about the soil velocity and acceleration in the epicentral region. Shakemaps are exploited to infer the macroseismic intensity throughout empirical models. The latter is used, together with the a priori information layers, to generate the first input to the data fusion algorithm, namely an estimation of the damage for single buildings and/or for a wider area. Furthermore, as soon as SAR data are available, a coseismic interferogram is generated to measure the displacements that, in principle, will help to restrict the damage evaluation analysis to the regions affected by surface deformation. In the crises phase, APE generates the change detection maps using both SAR and optical images. The change detection problem towards damage mapping is faced by segmenting the image before the seismic event, and identifying the objects of interest where focus the analysis. Objects could be single buildings, or groups of buildings, depending on the expected level of details and on the resolution of the available images. Alternatively, objects can be obtained from external sources, e.g., from cartography or GIS layers. The image segmentation step is done in background, so that, soon after the event, the first image available, SAR or optical, can be compared with the pre-event one on an object basis in order to derive information on the changes occurred. The LIDaM is finally generate by integrating a priori information and change detection product from EO data.

The work reported in this thesis contributed to the APE procedure implementation developing and testing object oriented change detection methodologies for mapping earthquake damages at single building scale, given a pair of VHR satellite images, both SAR and optical. Moreover, this thesis work has supported activity related to the implementation of the data fusion module by assessing different approaches to combine remote sensing change detection maps with a priori information layers generated according routines under the responsibility of different partners of the APhoRISM project.

3.3 Case study and datasets

After providing a summary of the selected seismic event, this section describes remotely sensed and reference data exploited during the experiments. Ground truth layers available from different data sources are also compared.

3.3.1 Event summary

On April 6, 2009 at 1:32 GMT(Greenwich Mean Time), an earthquake hit L'Aquila city, in Central Italy. The mainshock was rated 6.3 on the moment magnitude (M_w) scale; the epicentre was located near L'Aquila, at a depth of about 9 km, and was followed in the next week by 7 aftershocks with $M_w > 5$. L'Aquila and the surrounding villages suffered the highest damage. The earthquake was felt throughout central Italy; 309 people died, making this event the deadliest earthquake that hit Italy since the 1980 Irpinia one. The earthquake caused damage to thousands of buildings in the medieval city centre of L'Aquila. Several buildings also collapsed resulting in a sparse damage distribution within a high-density urban area.

3.3.2 Earth Observation datasets

The satellite dataset exploited for the L'Aquila case study consists of two pairs of optical and SAR images collected before and after the earthquake from the Quickbird (QB) and the COSMO-SkyMed (CSK) sensors, respectively (see Table 3-1).

The QB image taken before the earthquake is dated September 4, 2006, while the post-seismic one was collected on April 8, 2009, only 2 days after the catastrophic event. Each acquisition is composed of a panchromatic (PAN) and a multispectral (MS) image. The latter is collected in four spectral channel in the blue (450-520 nm), green (520-600 nm), red (630-690 nm) and near-infrared (760-900 nm) wavelength regions. Nominally, at nadir, the spatial resolution of the PAN image is 0.6 m, while the MS image has a 2.44 m resolution (DigitalGlobe, 2014).

The CSK system is a constellation of four satellites, developed in cooperation between ASI and the Italian Defense Ministry. The platforms host an X-band SAR, with right and left looking imaging capabilities, an incidence angle range of 20°- 60°, and a minimum revisit time of 12 h. The radar was designed to implement three different operation modes: 1) Spotlight mode, for metric resolutions over small images; 2) Stripmap mode for metric resolutions over tens of km images with single and dual polarization capability; and 3)

ScanSAR mode, for medium to coarse (100 m) resolutions over large swaths (Covello et al., 2010). The pair of images used in this work is in Spotlight mode. The pre-event image was taken on April 5, 2009 while the post-event one, two weeks after the seism (21/4/2009). Both observations were performed along ascending orbit with a right-side looking and 50.57° of incidence angle. The spatial resolution is about 1 m in both range and azimuth.

Optical Dataset (Quickbird)			
Date of Acquisition	Acquisition Mode	Looking Angles	Spatial resolution
04/09/2006	PAN + MS	-3.7° in- track -10.3° cross-track 10.9° off-nadir	0.6 m PAN 2.44 m MS
08/04/2009	PAN + MS	2.8° in- track 3.9° cross-track 4.8° off-nadir	0.6 m PAN 2.44 m MS

SAR Dataset (COSMO-SkyMed)					
Date of Acquisition	Acquisition Mode	Polarization	Look/ pass Direction	Incidence Angle	Spatial resolution
05/04/2009	SPOTLIGHT	HH	Right/ Ascending	50.57°	1 m
21/04/2009	SPOTLIGHT	HH	Right/ Ascending	50.57°	1 m

Table 3-1: Earth Observation dataset available for the L'Aquila test case.

3.3.3 Reference Data

The ground truth data available for the L'Aquila test site comes from two different sources. The first one is the survey performed by the INGV macroiseismic team (QUEST - QUick Earthquake Survey Team, http://quest.ingv.it), while the second one was produced by the Italian Department of Civil Protection (DPC).

After the catastrophic event, several groups of INGV researchers, in one week of fieldwork, collected information related to type of building and its vulnerability class, and the suffered damage. The macro-seismic survey was carried out according to the European Macroseismic Scale 1998 (EMS-98), which defines five damage grades: grade 1 for negligible to slight damage, grade 2 for moderate damage, grade 3 for substantial to heavy damage, grade 4 for very heavy damage, grade 5 for completely collapsed buildings

(Grünthal, 1998). Figure 3-2 shows the classification of damage for masonry and reinforced concrete buildings, as reported in Grünthal (1998).

Figure 3-2: EMS-98 scale: masonry (left) and reinforced concrete (right) building type.

It is worth noticing that the inventory data were collected by a visual inspection, looking from outside the buildings, because the INGV teams were not allowed to enter the edifices, for safety reasons. More than 1600 buildings were surveyed in the central area of L'Aquila, and a georeferenced vector file was produced to map the collected data in a GIS (Tertulliani et al., 2011, Tertulliani et al., 2012). The resulting map is shown in Figure 3-3.

The DPC survey was carried out during the six months following the seismic event and includes a detailed review of the interior parts of the buildings. The final goal of this survey was classifying the building usability and assess the damage affecting the individual structural elements of each edifice. The surveyed buildings for this dataset are both for private and public use. As for private buildings, the form (Agibilità e Danno sugli Edifici pubblici e privati: AeDES) filled in by the field operators contains more than 250 attributes for each building. The most relevant for our analysis are the geometrical characteristics (e.g. number of floor, area, volume, and regularity), the structural typology and the sustained damage level. The latter is provided for each structural element (vertical

structures, floor, stairs, roof, infills and partitions) as a damage grade following the EMS-98 scale. In addition, a second dataset from the DPC survey is specifically related to the monumental and historical buildings in L'Aquila. The survey was performed in the frame of the Safeguard of Cultural Heritage from Natural Disasters action, independently on the acquisitions of the AeDES forms, collecting similar but not identical information.

The digitized AeDES forms were made available to the APhoRISM project by DPC as Microsoft Excel© files (indeed, DPC is one of the potential user of APhoRISM). A geocoded version of the AeDES forms was provided as GIS layer by Istituto per le Tecnologie della Costruzione (ITC) of the Italian National Research Council (CNR). CNR-ITS partially contributed to perform the survey, developed the georeferenced database using the Carta Tecnica Regionale (CTR) of Regione Abruzzo as base cartography and also performed some evaluations. More specifically, a damage indicator and a vulnerability class were calculated following the EMS-98 scale. Figure 3-4 shows an overview of the EMS-98 damage grade classification in the downtown area of L'Aquila derived from the AeDES forms merged with the Cultural Heritage survey, as processed by CNR-ITC.
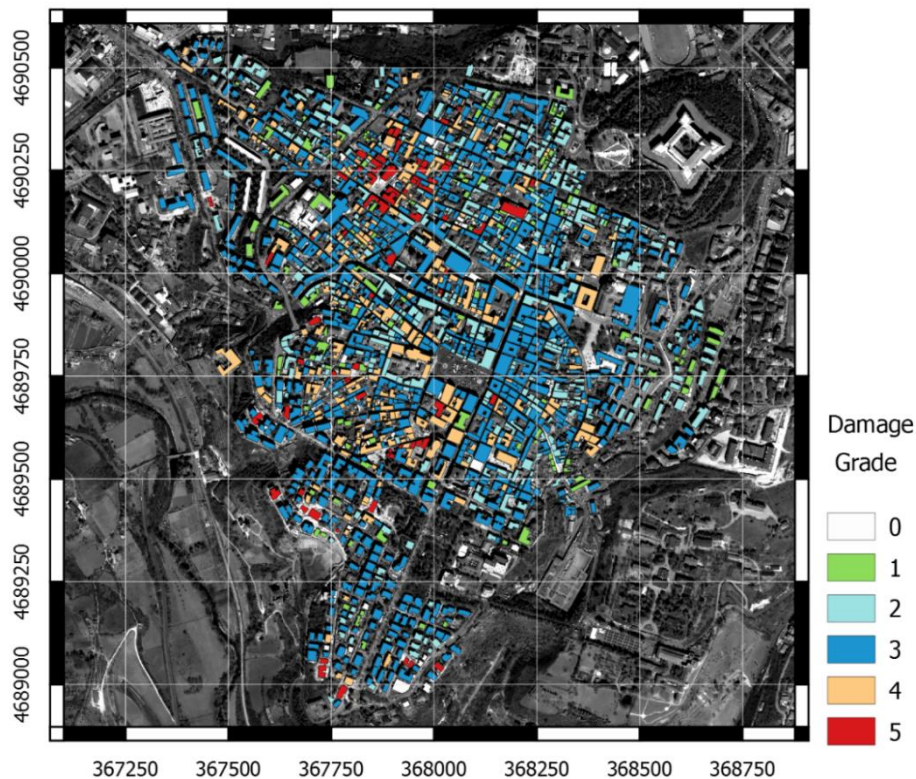
Figure 3-3: Damage distribution of L'Aquila city centre according to INGV ground survey. The polygons of surveyed buildings, with colour representing the damage grade, are superimposed on a VHR panchromatic image acquired by the Quickbird satellite after the earthquake.
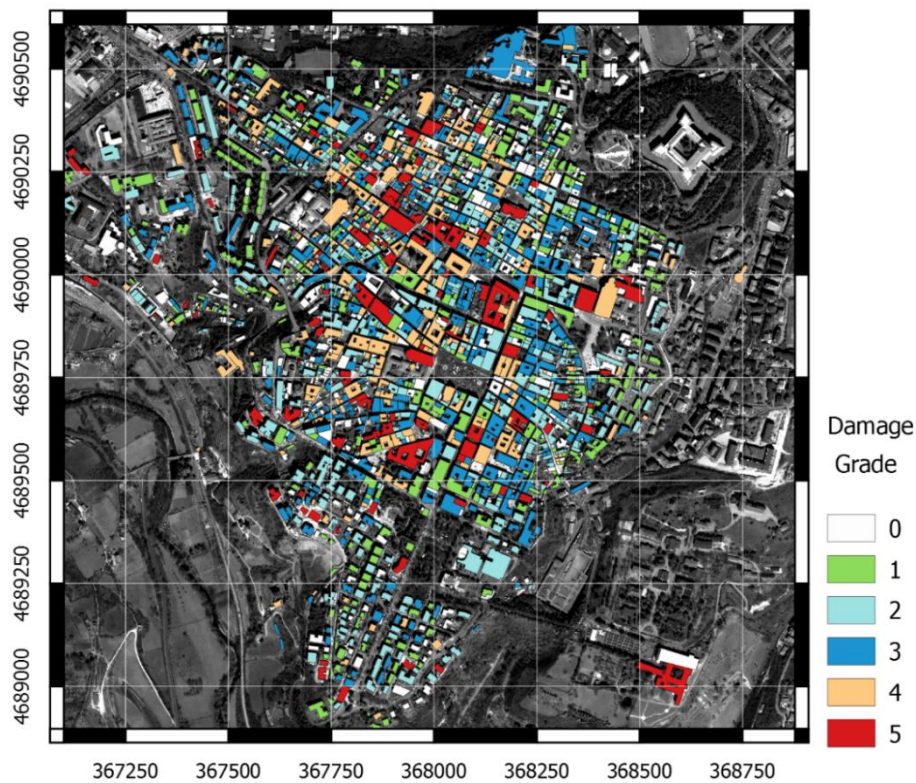
Figure 3-4: Damage distribution of L'Aquila city centre from the DPC-AeDES dataset processed by CNR-ITC. The polygons of surveyed buildings, with colour representing the damage grade, are superimposed on a VHR panchromatic image acquired by the Quickbird satellite after the earthquake.

Figure 3-5 provides a statistical summary, i.e., the number of occurrence of different damage grades, comparing the two surveys. We can notice the different distribution of damage, which is somehow related to the different purposes and the different conditions the surveying team encountered during their field work, but also to the method adopted to interpret and summarise the information collected in the field. More specifically, the distribution of the INGV damages is mostly concentrated on damage grade D=3 and reports only 74 collapses (D=5). The DPC-AEDES distribution is more uniform, exhibiting both many low (D=0) and high (D=5) damage grades.

The matching between the two available ground surveys was analysed in detail since they were used as a reference to assess the accuracy of the classification of the satellite images. To properly compare the two ground surveys, same processing of the corresponding GIS layer was necessary to reduce as much as possible co-registration errors. First of all, INGV damage grades, originally mapped on polygons manually drawn on a satellite image, were associated to polygons from the CTR which represents the geographical base map of the DPC survey as provided by the ITC. When necessary CTR polygons were

aggregated or disaggregated, in order to match the original INGV ground truth layer. Depending on which buildings were really surveyed and how a certain built up area was (or was not) split into single elements (i.e., polygons), when intersecting the layers, '*one-to-one*', '*one-to-many*', '*many-to-one*' or '*many-to-many*' correspondences can be found. In order to evaluate a CM between the two surveys, we have considered only polygons that are in common between the two datasets (i.e., that have an intersecting area above a threshold of ~15%).
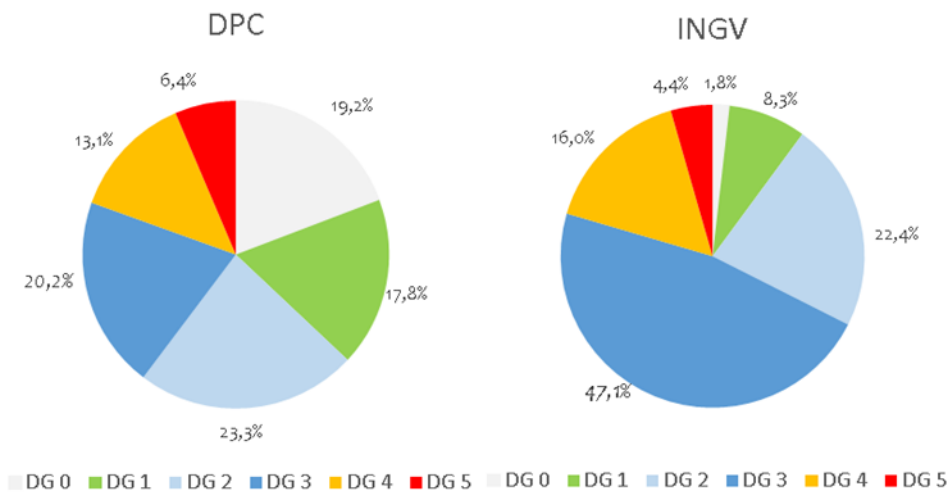
Figure 3-5: Damage frequencies in the INGV ground survey (left) and from the DPC-AEDES dataset processed by CNR-ITC (right).

	DPC							
	D0	**D1**	**D2**	**D3**	**D4**	**D5**	**n.a.**	
D0	**14**	7	2	2	0	0	5	25
D1	40	**35**	21	12	3	0	27	111
D2	52	84	**92**	63	22	9	51	322
D3	41	96	232	**196**	105	38	77	708
D4	10	13	21	71	**75**	28	49	218
D5	2	3	8	7	10	**31**	13	61
	159	238	376	351	215	106	222	1445

(INGV labels the rows D0–D5 on the left side)

Percentage of coincident outcomes	30.7%	**Cohen's Kappa**	11.4%

Table 3-2: INGV vs DPC ground truth confusion matrix. The n.a. column reports, for each damage grade, the number of buildings for which the damage assessment from the DPC is not available.

From such analysis, we can identify a high dispersion mainly due to the different human evaluations of damage classes, even though coming from expert teams working on the

field. Out of 74 buildings that have sustained a damage of grade 5 according the INGV survey, 61 AeDES forms were filled and only 31 of them confirm the damage grade 5 (see Table 3-2). The percentage of equally classified objects is 30.7%, while the Cohen's kappa is equal to 11.4%. When the binary classification problem to distinguish between damage grade 5 and damage grade less than 5 is considered, a Cohen's kappa of 33.6% and an overall accuracy equals to 92.7% were found.

4. Data analysis and results

4.1   Introduction

The new generation of spaceborne SAR and optical sensors provides metric or submetric resolution imagery, thus allowing, in principle, the detection of single building damaged after an earthquake. However, the complexity of the image forming mechanisms within urban settlements, especially of radar images, makes the automatic detection of damaged buildings still a challenging task. To cope with time constraints and accuracy requirements of emergency response activities, the visual interpretation of VHR optical imagery is still the most widely used technique to assess urban damage from EO data (Chini et al., 2009a; Ehrlich et al., 2009).

Considering the earthquake that hit L'Aquila city (Italy) on April 6, 2009, this chapter presents semiautomatic procedures exploiting VHR images acquired before and after the seismic event from both SAR and optical sensors for providing damage assessment products at single building scale. In order to make the delivered product suitable for the user community (e.g. civil protection), it is important to rely on a damage scale which is recognized as a standard. To this aim we considered the EMS-98 scale and assessed the possibility to discriminate between collapsed or heavy damaged (D = 5 in the EMS-98 scale) buildings and less damaged or undamaged buildings (D < 5 in the EMS-98 scale). The proposed methodologies were developed and tested using optical images from Quickbird satellite and Spotlight CSK SAR imagery (see section 3.3.2). In our experiments, we considered the whole L'Aquila historical centre comparing classification results with ground based damage maps. This kind of assessment can be hardly found in the literature, especially when the main focus is on the development of sophisticated and advanced algorithms. Most of the works provide a change map as final product and report only a qualitative validation, based on the comparison with visual interpretation results,

often related to a limited area.

In the wider context of the APE approach, the OBIA techniques, described in sections 4.3 and 4.4 of this chapter, are exploited, soon after an earthquake occurs, to compare the first available image, either SAR or optical, with the pre-event one in order to derive information on the changes occurred. Change detection products from EO data are then integrated with a priori information layers and concurrent data for the generation of the LiDAM at single building level. A data fusion module accomplish this task processing information that each module of the APE approach is able to provide, according to different strategies that will be described in section 4.5 of this chapter.

4.2   Visual interpretation benchmark

In order to have a reference when assessing the performance of an automatic algorithm in detecting damage, a photointerpretation of pre- and post- event Quickbird images was performed building by building. The comparison between photointerpretation results and INGV ground truth is reported in Table 4-1 that provides a sort of benchmark against which automatic algorithms can be judged.

		Ground Truth data (INGV)				
		D = 5	D = 4	D = 3	D <=2	total
Visual inspection	damaged	45	46	4	0	95
	undamaged	29	221	781	541	1572
	total	74	267	785	541	1667

Table 4-1: Confusion matrix of the photointerpretation of Quickbird data.

A missed detection error of 39.2% and 82.8% was found for buildings that, according the INGV ground survey, sustained a damage of grade 5 and 4, respectively. Just 4 out of 785 buildings classified as damage grade 3 were recognized as damaged by visual inspection. This result demonstrates that only the detection of the damage grade 5 is feasible using remote sensing data.

A few samples of the appearance of buildings and the problems that can be encountered when an automatic algorithm is used for damage detection are shown in Figure 4-1. This

figure compares pre- and post- event Quickbird images for some buildings classified as damage grade 5 by the INGV ground survey. Provided that the ground survey is telling the truth, the building collapse in Figure 4-1 (a) and (b) is easily detectable by visual inspection, whereas the same is not true for the building in Figure 4-1 (c) that appears unchanged.

(a)   **(b)**   **(c)**
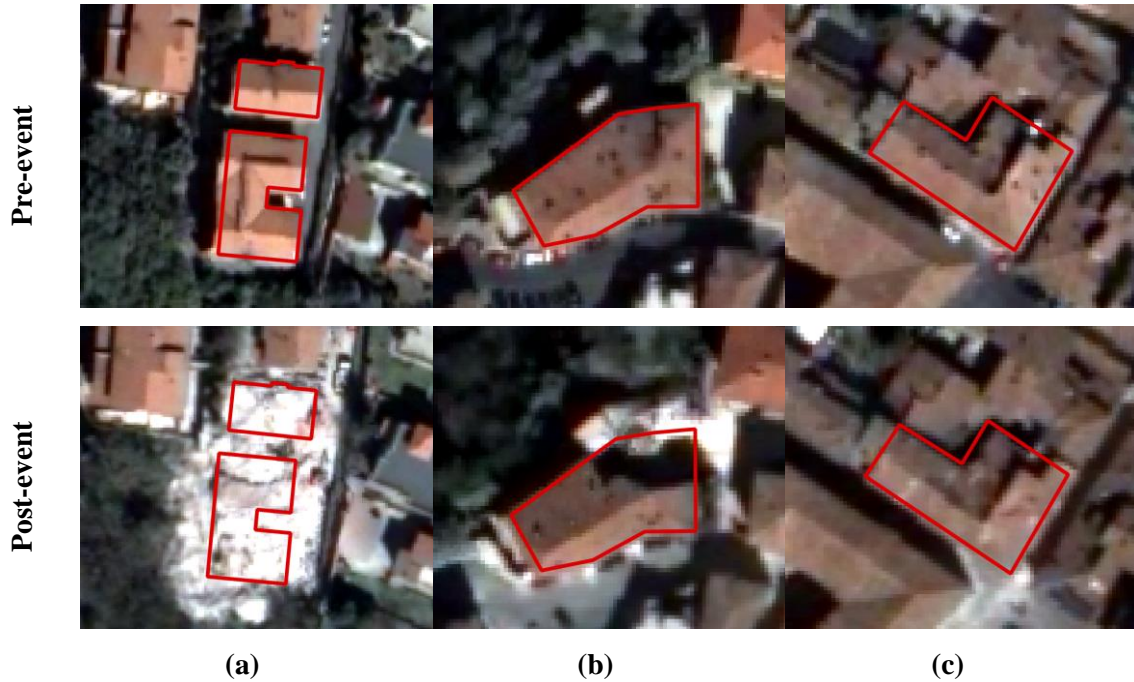
Figure 4-1: Comparison of the pre- (upper panels) and post- event (lower panels) Quickbird images for some buildings classified as damage grade 5 by the INGV ground survey.

4.3   Damage assessment from VHR optical imagery

To assess the feasibility of a damage product at level of individual building from a pair of VHR optical data, we implemented a change detection approach that works at object scale. In the contest of the earthquake damage assessment from optical data, segmenting the pre-event image into objects corresponding to a building allows the change analysis to be focused on the objects of interest, avoiding false alarms due for example to vegetation changes and temporary objects. In this work, the segmentation of the scene is carried out using a pre-existing building map provided as GIS layer, which is generally available in most of the cases. When a map of the urban area is not available, objects corresponding to buildings can be identified exploiting image segmentation algorithms, as described for example in Chini et al. (2009a). Within each building footprint (i.e., an

image object) a number of features, potentially capable of detecting changes associated with the building collapse, were computed from pre- and post-seismic images, both panchromatic and multispectral. Different classification approaches were tested and compared.

4.3.1 Pre-processing

This section describes the pre-processing carried out on Quickbird images. Panchromatic images were orthorectified exploiting the Rational Polynomial Coefficients (RPC) provided with the data and a Digital Terrain Model (DTM) obtained from a LiDAR overflight of the study area. Pan-sharpened (PSH) pre- and post-event images were produced from PAN and MS imagery using the Gram-Schmidt Pan Sharpening method (Brower and Laben, 2000) implemented into IDL/ENVI©. Furthermore, these data were orthorectified exploiting the RPCs of PAN images and the LiDAR DTM according to the same procedure used for PAN data. A further registration between orthorectified images was required in order to achieve a better alignment of the images at street level and with the GIS map. Namely, a rigid shift of 0.6 meters in the North-South direction and of -4.8 meters in West-East direction was applied to the pre-event image. As for the post-event image, a shift of -2.4 meters in the North-South direction and a shift of 7.8 meters toward the East were necessary.

As the orthorectification was done with respect to the surface level (i.e., street level), in order to better superimpose the building footprints reported in the GIS layer to the actual image pixels associated to the roof we compensated the parallax error, i.e. the apparent displacement of building roofs with respect to their bases. The parallax displacement occurs along the sensor Line Of Sight, over a distance that depends on building height h and off-Nadir view angle θ, according to the equation:

$$\Delta s = h \tan \theta$$

Fixed a Cartesian reference frame where the x-axis is oriented towards the geographic East, the y-axis towards the geographic North and the z-axis is oriented along the vertical upward direction, the displacement components along x- and y-axes can be found as follows:

$$\Delta s_X = \Delta s \sin \varphi$$
$$\Delta s_Y = \Delta s \cos \varphi,$$

where φ is the sensor azimuth measured clockwise from the North. Based on the building height reported in the CTR of Regione Abruzzo, building polygons were shifted, generating two layers matching pre-and post-event images, respectively.

The results of this fine registration step can be appreciated in Figure 4-2. Figure 4-2(a) and Figure 4-2(b) show respectively pre- and post-event PAN images orthorectified and co-registered. Buildings outlines shifted for compensating the parallax error are reported in yellow, while red dashed polygons represent buildings footprints in their original position. The improvement of the position of the polygons with respect to the roofs is particularly evident in the pre-event image (Figure 4-2(a)) that was collected with a greater off-nadir angle (see Table 3-1).

(a)   (b)
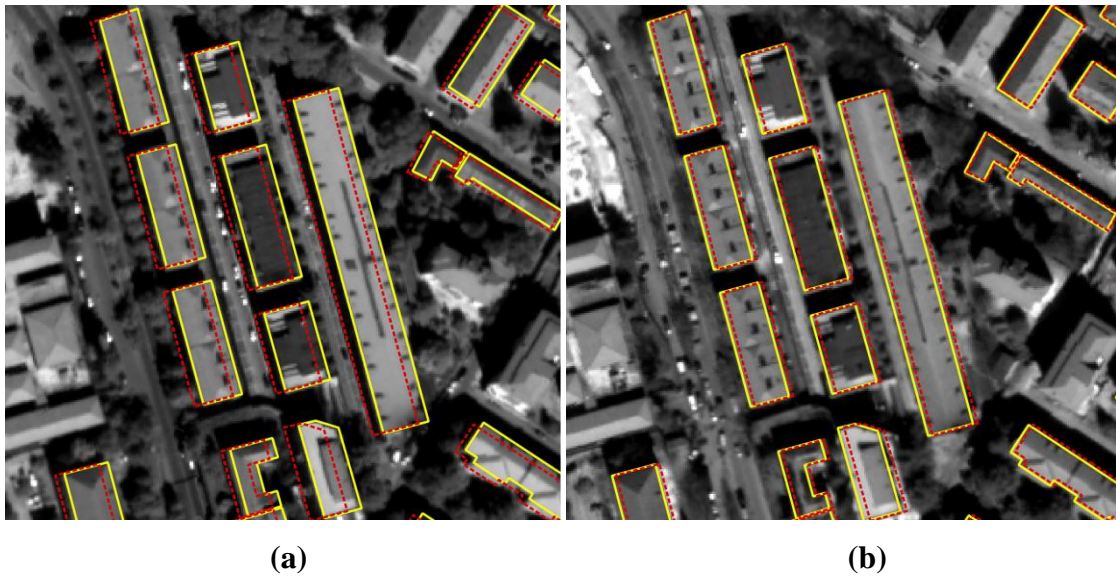
Figure 4-2: Pre- (a) and post- (b) event PAN images superimposed with buildings footprints in their original position (red dashed polygons) and shifted for compensating the parallax error (yellow polygons).

Before extracting change detection features, a histogram matching was performed to radiometrically compensate for the different season and illumination conditions of the pre- and post-seismic acquisitions. A linear transformation was applied to each band of the post-event image in order to get a data distribution with the same median and interquartile range of the corresponding pre-event data. We carried out this step considering only urban pixels that have been identified using the building polygons shifted for compensating parallax error, as described before.

4.3.2 Feature extraction and classification

In this work, we performed an extensive investigation of many change parameters with the objective of identifying the most suitable set for damage detection purpose. In particular, we have considered a set of 13 features that can be grouped in four categories: standard change detection metrics, change indicators from information theory, features describing texture and colour changes, as shown in Table 4-2. According to the OBIA approach, features were locally evaluated considering only pixels pertaining to image object corresponding to buildings.

Group	Name	Description
Information theory features	MI_{pan}	Mutual Information from PAN images
	KLD_{pan}	Kullback-Leibler Divergence from PAN images
	MI_{psh}	Mutual Information from PSH images
	KLD_{psh}	Kullback-Leibler Divergence from PSH images
Texture features (from PAN images)	Δcon	Change in Contrast ($\Delta con=conpost - conpre$)
	Δcor	Change in Correlation ($\Delta cor=corpost - corpre$)
	Δene	Change in Energy ($\Delta ene=enepost - enepre$)
	Δhom	Change in Homogeneity ($\Delta hom=hompost - hompre$)
	Δent	Change in Entropy ($\Delta ene=entpost - entpre$)
Colour features (from PSH images)	Δhue	Change in Hue ($\Delta hue=huepost - huepre$)
	Δsat	Change in saturation ($\Delta sat=satpost - satpre$)
	Δlum	Value component difference ($\Delta lum=lumpost - lumpre$)
standard change detection features	Δint_pan	Change in the intensity from PAN images ($\Delta int\_pan=int\_panpost - int\_panpre$)

Table 4-2: Change detection features extracted from VHR optical data.

Statistical similarity measures, such as KLD e MI, were extracted from PAN images, using equations (2.15) and (2.16), and from PSH images, according to equations (2.13) and (2.14).

As a difference in the spatial arrangement of pixel intensity is expected as a consequence of the building collapse, changes in the object texture were considered in our analysis. Texture changes were assessed by means of five parameters: Contrast, Correlation, Energy, Homogeneity and Entropy. They were derived from GLCMs computed at object scale. Considering four angular directions (0°, 45°, 90°, 135°) and a fixed distance d=1, four symmetrical GLCMs were derived for each object and afterward summed for obtaining the rotational invariant GLCM from which texture descriptors were extracted. Features were computed from pre- and post- event images according to equations (2.4), (2.5), (2.6), (2.7) and (2.8), and their difference was considered to detect possible changes. As additional source of information for discriminating between damaged and undamaged buildings, changes in the Hue (H), Saturation (S), Value (V) colour space were taken into account, considering for instance that rubbles are expected to have low saturation as opposed to changes due to building restoration. The RGB value of a pixel was first transformed into the HSV space using a method suggested in Smith (1978). For each HSV colour channel, the mean value within the building footprint was calculated. The difference between the mean values of each HSV component computed respectively from the post- and pre-event image were then evaluated in order to obtain the change metric.

Let $\Theta = \{\theta_1, \theta_2, ..., \theta_d\}$ be the change parameters vector extracted from the pre- and post-event optical images and associated to a building. Let $\Omega = \{\omega_1, \omega_2\}$ be the set of classes, identifying a building with damage grade $D = 5$ or $D < 5$. We carried out the classification task considering two supervised learning algorithms, namely, the Bayesian Maximum A Posteriori criterion and Support Vector Machine. Training and test sets in our analysis were provided by the INGV ground survey. Since the dataset is very unbalanced (the number of instances for the damaged buildings class is much lower than the number of samples belonging to the class of buildings with damage grade less than 5) and the number of collapsed building is low, it was not possible to split data into training and test samples. Thus, we exploited a k-fold CV procedure to assess the classification performances. The dataset was split into k=10 disjoint subsets of approximately equal size, preserving the original class proportions: in turn, each of the k subset was used for testing the classifier trained on the remaining k-1 subsets. For each iteration, we derived the CM referred to the current test set and finally, by summing up all the k CMs, we generated a global CM to assess the classification performances on the whole dataset. Classification accuracy

measures have to be considered with care in case of unbalanced sets and just one parameter could be not enough for assessing the classification results. Here, several parameters were derived: overall accuracy, sensitivity, precision, and Cohen's Kappa with and without normalization with respect to the number of samples belonging to each class. The normalized Cohen's Kappa, is computed from a normalized form of the confusion matrix, where each element represent the percentage of occurrence with respect to the total number of samples falling into each class.

Following the implementations of the Bayesian MAP criterion, given the observed feature vector Θ, we assign the building to the class ω_k (k=1,2) with the highest a-posteriori probability $P(\omega_k|\Theta)$.

We estimate the posterior probabilities assuming that both classes are equally likely a priori. As for the class-conditional pdfs, , also known as class likelihood functions, they are estimated from the training set through the non-parametric approach known as Parzen window method, assuming, moreover, the class-conditional independence of all the features, i.e., the Naïve hypothesis. In detail, given a set of n_k training samples from class k, and considering a Gaussian kernel, the Parzen window method estimates the distribution $p(\Theta|\omega_k)$ by the following (Parzen, 1962):

$$p(\Theta|\omega_k) = \prod_{i=1}^{d} p(\theta_i|\omega_k) = \prod_{i=1}^{d} \frac{1}{n_k\sqrt{2\pi}h_i} \sum_{j=1}^{n_k} exp\left(-\frac{1}{2}\left(\frac{\theta_i - \theta_{ij}^{(k)}}{h_i}\right)^2\right) \qquad (4.1)$$

where $\theta_{ij}^{(k)}$ is the jth observation of the ith feature from the kth class, and h is the so called bandwidth parameter or kernel width. Equation (4.1) shows that, for each feature, the pdf of the data given the class, $p(\theta_i|\omega_k)$, is estimated as sum of Gaussian kernel functions placed on each training data point.

As discussed in section 2.2.4.1, when a SVM is used to solve a binary classification problem, new input data are mapped in a higher dimensional feature space where they are labelled based on which side of the optimal separating hyperplane, constructed during the training phase, they fall. We recall that the non linear mapping of the input data into a high dimensional feature space is implicitly provided by a kernel function, i.e. a function returning the inner product $\Phi(x)\cdot\Phi(y)$ between the images of two data points x, y in the feature space. For this work we chosen the Radial Basis Function (RBF) kernel defined in section 2.2.4.2. Moreover, dealing with an unbalanced dataset, we exploited the cost-sensitive SVM algorithm proposed by Veropoulos et al. (1999) where two different

misclassification cost, C^+ and C^-, are considered for positive and negative classes (see section 2.2.4.2 for more details).

Both the selected classification algorithms have parameters that need to be tuned. In order to estimate the class-conditional pdf, the kernel widths of the Parzen window method have to be specified. In this study, we used a single bandwidth h for each of the d dimensions ($h_i=h$, for all i). To make input data more spherical and justify the use of a common bandwidth for each dimension, data were preliminarily standardized to have mean of zero and standard deviation of one. Many different techniques are available in the literature for choosing the optimal bandwidth from the data. A review of these methods can be found for example in Chiu (1996). Most of these bandwidth selection methods minimize the mean integrated square error (MISE) between the true pdf and its kernel based estimate. However, the bandwidth that is optimal for the MISE of a density estimator may not always be good for discriminant analysis, where the main emphasis is on the minimization of misclassification rate (Ghosh et al., 2006). For example, sometimes the bandwidth that minimizes the misclassification rate might be much larger than the bandwidth minimizing MISE (Ghosh and Chanduri, 2004). One of the most popular methods for selecting the kernel width is the Silverman's rule (Zang et al., 2006) which assume that the true distribution is normal. In this work, we used such rule for initializing our algorithm that empirically search for the optimal kernel width, as detailed in the following. For the SVM algorithm, the hyperparameters to tune are the regularization parameters of the positive and negative classes (C^+ and C^- respectively) and the parameter γ of the RBF kernel which define its width. We expressed the misclassification cost C^+ of the positive class, which is the minority class we are interesting in, as a function of the misclassification cost C^- of the negative class, through the relation $C^+=wC^-$, were w is a weight greater than 1. Then, the parameters we tuned were C^-, w and γ.

In order to optimize the parameters $\boldsymbol{\varphi}$ of the classifiers, i.e. the kernel width h for the MAP classifier ($\boldsymbol{\varphi}_{MAP} = \{h\}$) and the parameters C^-, w and γ for SVM ($\boldsymbol{\varphi}_{SVM} = \{C^-, w, \gamma\}$) we ran a repeated k-fold CV, with k = 10. For each repetition of the CV a different set of classifier parameters was tested and chosen according to simulated annealing algorithm, implemented through a built-in Matlab® function, which solves minimization problems. We initialised the kernel width of the Parzen window method to

a value h_0 computed according the Silveman's rule (Zang et al., 2006). Then we searched for its optimal value in the range [h_0/10, 10h_0]. As for the hyperpameters of SVM, $\boldsymbol{\varphi}_{SVM} = \{C^-, w, \gamma\}$ they were inizìtialised to $\boldsymbol{\varphi}_{SVM_0} = \{10, 1, 1/d\}$ where d is the dimension of the feature vector used as input. The regularization parameter C^- and the weight w were varied in the ranges [1, 10] and [2, 50], respectively The RBF kernel width γ was tuned in the range [10^{-3}, 2].

The additive inverse of Cohen's Kappa was chosen as objective function of the optimization algorithm. We preferred the Cohen's Kappa to other classification performance measures, such as the overall accuracy, because the latter is inappropriate when dealing with unbalanced datasets. If the dataset is highly unbalanced, the classification algorithm can achieve high overall accuracy simply assigning all the samples to the majority class. As a consequence, searching for the classifier parameters giving the best cross validated performances in terms of overall accuracy can lead to a classifier strongly biased towards the majority class. The Cohen's Kappa, considers also the off-diagonal elements of the CM, thus providing a more reliable estimate of the capability of a classifier in discriminating between two classes.

In detail, for each given set of change features, our approach for searching the classifiers parameters vector $\boldsymbol{\varphi}$ that provide the best cross-validated classification performances, works as follows.

For a fixed classifier parameters vector $\boldsymbol{\varphi}$, the training and the validation are repeated k times until all data are used. At the end of kth iteration of the CV, the Cohen's kappa, K= -f($\boldsymbol{\varphi}$), obtained with the current parameters vector, $\boldsymbol{\varphi}$, is derived from the global CM. Then, following a simulated annealing algorithm, (Dekkers and Aarts, 1991; Goffe et al., 1994; Bouleimen and Lecocq, 2003), the classifier parameters are updated.

A pseudo-code description of the algorithm is reported Figure 4-3.

Starting from the current point $\boldsymbol{\varphi}$, a new trial solution $\boldsymbol{\varphi}$' is selected from a neighbourhood of $\boldsymbol{\varphi}$ according to the equation: $\boldsymbol{\varphi}$'= $\boldsymbol{\varphi}$ + $\boldsymbol{\theta_n}$ Δr_n, where $\theta_n \in R^d$, with $\|\boldsymbol{\theta_n}\| = 1$, is a direction vector randomly generated, and Δr_n, is the step length defining the distance $\boldsymbol{\varphi}$' from $\boldsymbol{\varphi}$ for the nth iteration of the optimization algorithm. Then, a CV is performed in order to evaluate the objective function for the trial solution $\boldsymbol{\varphi}$' which is compared with value of the objective function achieved using $\boldsymbol{\varphi}$. If the difference in the objective function, Δ= f($\boldsymbol{\varphi}$')-f($\boldsymbol{\varphi}$) is less or equal to zero, $\boldsymbol{\varphi}$' is chosen to replace the current solution $\boldsymbol{\varphi}$. If in

addition $f(\boldsymbol{\varphi'})$ is the lower value of the objective function found so far, $\boldsymbol{\varphi'}$ becomes the best solution $\boldsymbol{\varphi_{opt}}$.

The algorithm avoids being trapped in local minima, allowing a neighborhood of a worse solution $\boldsymbol{\varphi'}$ to be explored, if specific conditions are met. Whenever $f(\boldsymbol{\varphi'})$ is greater than $f(\boldsymbol{\varphi})$, an acceptance probability $p=\exp(-\Delta/T_n)$ is computed and compared to y, a random number generated from a uniform distribution on $[0, 1]$. If p is greater than y, $\boldsymbol{\varphi'}$ is accepted to replace the current solution $\boldsymbol{\varphi}$, otherwise another trial solution $\boldsymbol{\varphi'}$ is selected from the neighbourhood of $\boldsymbol{\varphi}$.

The acceptance probability depends on the difference in the objective function values, Δ, and on the value of the parameter T_n at the current iteration of the algorithm. At each iteration, such parameter, called temperature, is lowered according an exponential cooling scheme $T_n = T_0 \times 0.95^n$, where T_0 is the initial temperature. Besides controlling the acceptance probability, the temperature determines the distance of the trial point from the current solution. Thus, as the algorithm progresses the probability of accepting worse solution decreases and increasingly smaller neighbourhood of the current solution are explored. The algorithm stops when a specified number of iteration is reached or when the average change in the objective function observed in a given number of iterations is smaller than a given tolerance, returning the best achieved solution $\boldsymbol{\varphi_{opt}}$.

```
Procedure Simulated Annealing with Cross Validation (CV)
φ ← φ₀                        // Initialization
f(φ) ← perform a CV to evaluate objective function for the initial solution φ₀
φ_opt ← φ
n ← 0
while (stopping criterion is not meet) do
        Tₙ ← T₀×0.95ⁿ  // Cooling scheme
        // select a new point from the neighbourhood of the current solution
        φ'← φ + θₙ Δrₙ, where θₙ∈Rᵈ, with ‖θₙ‖ = 1
        f(φ') ← perform a CV to evaluate objective function for the trial solution φ'
        Δ ← f(φ') - f(φ)
        if (Δ ≤ 0) then
            φ ← φ'
            φ_opt ← φ'
        else if
            p ← exp(-Δ / Tₙ) // Acceptance probability
            p ← generate a random number from a uniform distribution on [0, 1]
            if (p > y) then φ ← φ'
        endif
        n← n+1
endwhile
```

Figure 4-3: Pseudo-code of the simulated annealing algorithm exploited for optimizing the MAP classifier and SVM parameters.

4.3.3 Feature selection and classification results

A careful selection of the considered change detection features (see Table 4-2) was carried out. We searched for the most relevant subsets of features according to a wrapper approach which consists in ranking subset of features based on the classification results that they achieve (Guyon and Elisseeff, 2003).

We evaluated the effect on the classifiers performances by varying the number and the combination of the features used as input. We considered classification results obtained after tuning classifiers parameters and, for each set of d features, where d ranges from 1 to 13, we selected the combinations that provided the best results in terms of the Cohen's Kappa. In particular, due to the uncertainty in the Cohen's kappa estimation, for each group of d features, we selected those combinations giving a Cohen's Kappa within the range [maxKappa(d), maxKappa(d) - 0.02], where maxKappa(d) is the maximum Cohen's kappa achieved with d features. We performed this analysis independently for MAP classifier and SVM. Results, for groups of features of size from 1 to 5, are summarized in Table 4-3 and Table 4-4, which refer respectively to MAP classifier and SVM. These tables consist of five columns indicating the number of features, and 13 rows, one for each of the considered change parameters. Coloured cells in each sub-column represent selected features. The first sub-column of each group, reports the combination of features giving the best performance, i.e. the combination of d features for which the Cohen's Kappa was equal to maxKappa(d).

When a single feature was considered as input of the change detection algorithm, change in the contrast provided the best results, either with SVM or MAP classifier. Looking at Table 4-3, it can be seen that the best results of the MAP classifier were all achieved with feature sets containing change in the contrast, which is the most selected feature also in the case of SVM (see Table 4-4). When a building collapses the contrast generally increases due to presence within the building footprint of pixels with very different grey levels close each other. Consistent performances were also achieved using feature subset including changes in Homogeneity and Entropy, which are frequently included in the optimal subsets showing that changes in textural properties are the most reliable features for detecting changes due to the building damage. Being inversely but strongly correlated, Homogeneity and Entropy are interchangeable. In fact, either using SVM or MAP classifier, these two features are selected alternatively.

Table 4-3 (MAP classifier) — columns: number of features, 1, 2, 3, 4, 5, Number of occurrences in the optimal subsets

number of features	Number of occurrences in the optimal subsets
Mlpan	1
KLDpan	0
Mlpsh	5
KLDpan	4
Δcon	23
Δcor	8
Δene	9
Δhom	5
Δent	8
Δhue	6
Δsat	13
Δlum	0
Δint_pan	0

Table 4-3: Exhaustive feature selection based on Cohen's Kappa for different feature vector size using MAP classifier.

Table 4-4 (SVM) — columns: number of features, 1, 2, 3, 4, 5, Number of occurrences in the optimal subsets

number of features	Number of occurrences in the optimal subsets
Mlpan	8
KLDpan	0
Mlpsh	12
KLDpan	0
Δcon	16
Δcor	2
Δene	5
Δhom	4
Δent	8
Δhue	5
Δsat	0
Δlum	3
Δint_pan	5

Table 4-4: Exhaustive feature selection based on Cohen's Kappa for different feature vector size using SVM.

Comparing Table 4-3 and Table 4-4 it can be seen that there are some features that demonstrated their usefulness depending on the classification approach. The Mutual Information, for example, being very sensitive to residual registration errors between the two images, generally determines a lot of false alarms when used as input to the MAP classifier, while it provides satisfactory results if exploited, in combination with other features, as input to SVMs. In combination with other features, high values of such

parameter can help avoiding false positive due to changes not attributable to the earthquake but associated, for example, to a building restoration intervened in the timeframe between the two image acquisitions. It can be noted that, in the case of the MAP classifier, feature combinations providing the best results in terms of Cohen's kappa often include change in the colour saturation which conversely does not appear within the features subset for which SVM provided the best results. This is likely because the Mutual Information from PSH images (i.e., multispectral data), which is one of the most frequently selected features in the case of SVM, brings colour signature as well.

Figure 4-4(a) reports the Cohen's Kappa as a function of the features number for both SVM and MAP classifier, considering the best subset for each group of d features. In Figure 4-4(b) and Figure 4-4(c) classification approaches are compared in term of sensitivity and precision, respectively. In the context of damage detection, the sensitivity, computed from the global CM using equation (2.33), represents the percentage of building to which a damage grade 5 was assigned from both ground survey and classification algorithm. As for the precision, defined by equation (2.35), it represents the percentage of buildings classified as damage grade 5 from satellite data that have actually sustained a damage of grade 5, according the ground survey. Looking at Figure 4-4(a), it can be seen that increasing the number of features improves the classification Cohen's Kappa, until reaching the highest value with a subset of five features, for both the considered classification approaches. Then, the performances in term of Cohen's Kappa tend to decrease as the input space dimensionality increases. From Figure 4-4(b) and Figure 4-4(c), it can be noted that increasing the number of features from 1 to 5 does not improve the sensitivity of the algorithms to the damage detection, but it increases its precision, i.e. the rate of buildings classified as damage grade 5 that are real damages (i.e., decrease of false alarms). The results we obtained using SVM outperform those achieved by the MAP classifier. With respect to the MAP classification approach, SVMs are able to provide classifications characterized by a lower rate of false alarms. This could be due to the complexity of the input data that a SVM, working in a higher dimensional feature space, is more able to recover. Besides providing better performance in term of precision, in most of case, SVM correctly recognize a greater number of building with damage grade 5. However, performances in term of sensitivity are not so markedly distinct and both algorithms often fail in detecting a building with damage grade 5 according to the ground

survey, mainly because, as pointed out, in section 4.1, about 40% of these damages are not detectable even by visual inspection.
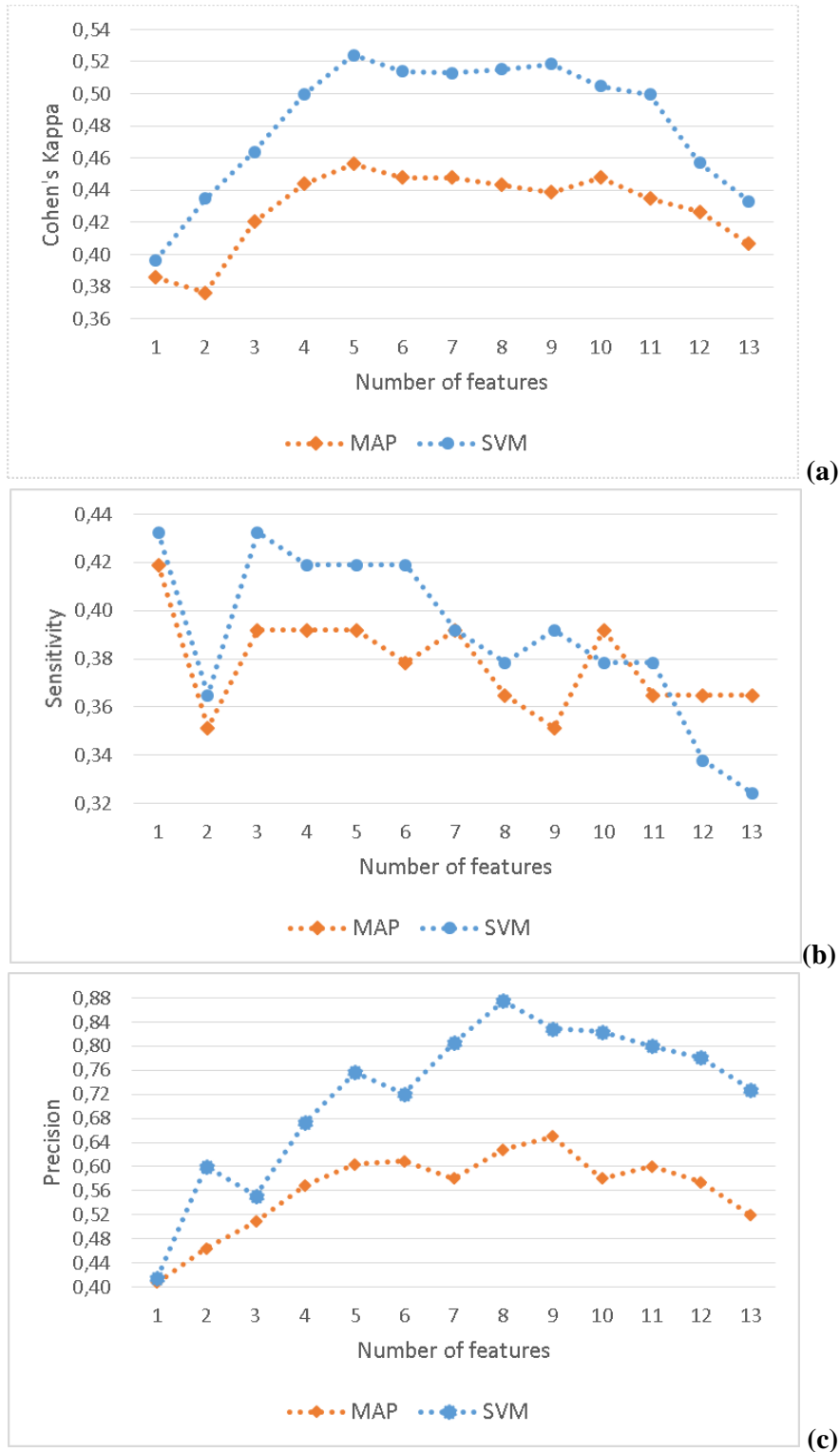
Figure 4-4: Classification performances of SVM and MAP classifier as a function of the number of features used as input: (a) Cohen's Kappa (b); Sensitivity; (c) Precision.

The confusion matrix shown in Table 4-5 summarizes the results of the MAP classification approach with the subset of five features that achieved the highest Cohen's Kappa. Using changes in Contrast, Energy, Entropy, Hue and Saturation the MAP classification algorithm exhibits an overall accuracy of 96.2% and a Cohen's Kappa of 45.6%. We correctly detect 39.2% of buildings with damage grade 5, generating 19 false positives.

		INGV						
		(D=5)	(D<5)					
classifier	(D=5)	**29**	19	48		60.4%	*precision*	
	(D<5)	45	**1574**	1619		97.2%	*NPV*	
		74	1593	1667		*overall*	*Cohen's*	*normalized*
		sensitivity	*specificity*			*accuracy*	*kappa*	*kappa*
		39.2%	98.8%			96.2%	45.6%	38.00%

Table 4-5: Confusion Matrix obtained using the MAP classifier with the optimal subset of features (changes in Contrast, Energy, Entropy, Hue and Saturation).

The result in terms of false alarms can be considered satisfactory, especially if it is considered that 5 false positives are determined by buildings that suffered a damage $D < 5$ that is apparent according to the photointerpretation, as shown in Figure 4-5(a) and Figure 4-5(b), where two buildings classified respectively as damage grade 4 and 3 in the INGV survey, are reported.

As for the low sensitivity to the damage, it is worth pointing out that the buildings correctly recognized as damaged from our algorithm represent about 60% of the buildings with damage level 5 that we were able to identify by visual inspection of the Quickbird images. In other word, most of missed detection of damage 5 according to INGV were not detectable even by visual interpretation of the images. As discussed in section 4.1, the class of the buildings to which a damage grade 5 was assigned is very heterogeneous. In some cases, the damage sustained by the building roof, although detectable by a human image interpreter, was not so extensive, thus preventing its identification without generating a large amount of false alarms. An example is reported in Figure 4-6(a). There are, in addition, same buildings suffering floor collapse while retaining an intact roof (pancake type collapse), such as that shown in Figure 4-6(b). In these cases, buildings were wrongly classified as belonging to the class of damage $D < 5$.
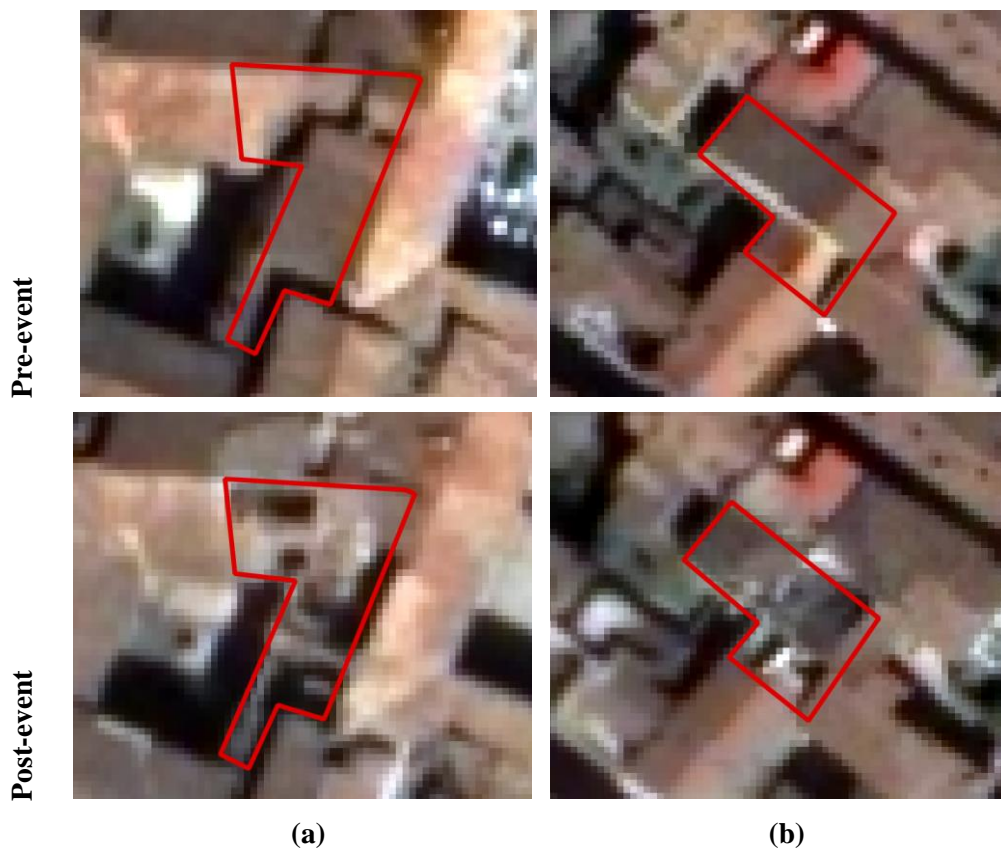
Figure 4-5: Examples of false positives generated by the MAP classifier. Upper panels: pre-event Quickbird images. Lower panels: post-event Quickbird images. Polygons delineate buildings that have sustained a damage of grade 4 (figure (a)) and 3 (figure (b)) according the INGV ground truth. Because the roof of both buildings is evidently damaged, they were wrongly classified as damage of grade 5.
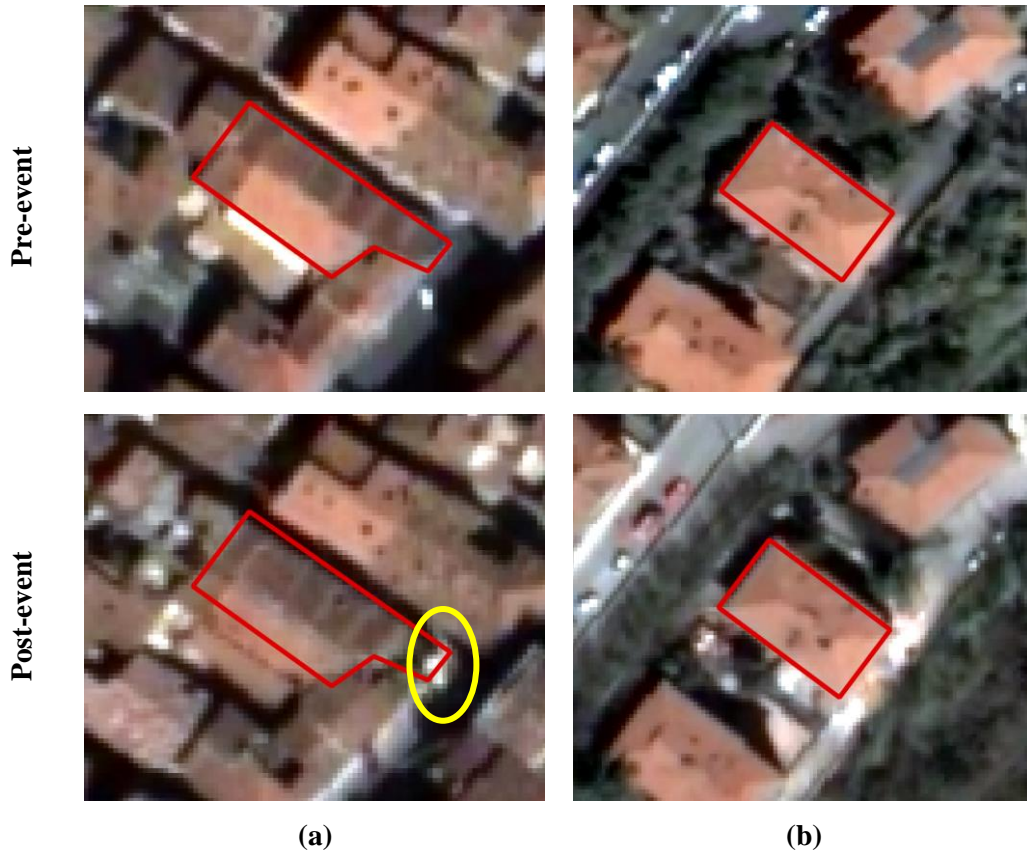
Figure 4-6: Examples of false negatives generated by the MAP classifier. Upper panels: pre-event Quickbird images. Lower panels: post-event Quickbird images. (a) The roof is almost completely intact, except for the portion within the yellow ellipse. (b) Due to the pancake type collapse, the only change is the presence of dust and debris around the building footprint.

Using SVM we obtained better result, as shown in the confusion matrix reported in Table 4-6 which refers to the SVM outcomes when Mutual Information from PSH images, changes in Energy, Entropy, Hue and Intensity from PAN images were used as input.

		INGV					
		(D=5)	(D<5)				
classifier	(D=5)	**31**	10	41	75.6%	*precision*	
	(D<5)	43	**1583**	1619	97.4%	*NPV*	
		74	1593	1667	*overall*	*Cohen's*	*normalized*
		sensitivity	*specificity*		*accuracy*	*kappa*	*kappa*
		41.9%	99.4%		96.8%	52.4%	41.26%

Table 4-6: Confusion Matrix obtained using SVM with the optimal subset of features (Mutual Information from PSH images, changes in Energy, Entropy and Hue, change in Intensity from PAN images).

Sensitivity and Precision increase to 41.9% and 75.6%, respectively. We correctly classify 31 out of 74 buildings with damage grade 5, identifying about 68% of those detected by visual inspection. Even in this case some of the false alarms (3 out of 10) are due to buildings recognized as damaged by photointerpretation.

The classification performances against the INGV ground survey demonstrates the challenge of detecting earthquake damage at single building scale from satellite data. However, it is very important to notice that the classification results assessed against the INGV ground truth (overall accuracy=96.2%, Cohen's kappa=45.6% using the MAP classifier; overall accuracy=96.8%, Cohen's kappa=52.4% using SVM) are quite similar and even better than the direct comparison between the INGV and DPC ground surveys, for which, as reported in section 3.3.3, an overall agreement of 92.7% and a Cohen's kappa of 33.6% were found. In other words, the damage map obtained using satellite data over the entire historical town of L'Aquila shows a mismatch with respect to the INGV ground truth comparable to the mismatch of the two ground surveys. It demonstrates the undoubted value of remote sensing for a rapid response, when compared to the uncertainty, highly costly and time consuming ground survey. This aspect will be further detailed in the chapter 5 of this thesis.

4.4 Damage detection from VHR SAR images

In a medium resolution SAR image, build-up areas appear almost uniformly bright, although with a complex texture, due to the predominance of the double reflections from the ground to the buildings wall facing the radar. Thus, the decrease of the backscattering is a key feature for detecting severely damaged urban settlement from a pair of SAR images with a resolution in the order of tens of meters (Chini, 2014). In metric resolution SAR images, such as those provided by CSK or TerraSAR-X, a building appears as a complex of image structures associated to different scattering mechanisms (Chini et al., 2015b), as discussed in section 2.1.1. We can observe bright linear features with different thicknesses as well as dark areas. The former are related to double-bounce or, as the line thickness increases, to layover areas, where the roof or the facade scattering may be dominant depending on the building characteristics. The latter are due to the shadow effect (Ferro et al., 2013).

When a building collapses the double bounce and the layover areas drastically decrease

their brightness. At the same time, an increase of backscattering it is expect due to the scattering from debris laying within the building footprint and to the return coming from the ground previously occluded by the building itself (Anniballe et al. 2014, 2015, 2016; Marin et al., 2015), as sketched in Figure 4-7.

Taking advantage from the theoretical model of building SAR response and its expected changes after a collapse, we developed an object-oriented change detection approach for mapping earthquake damage at single building level.
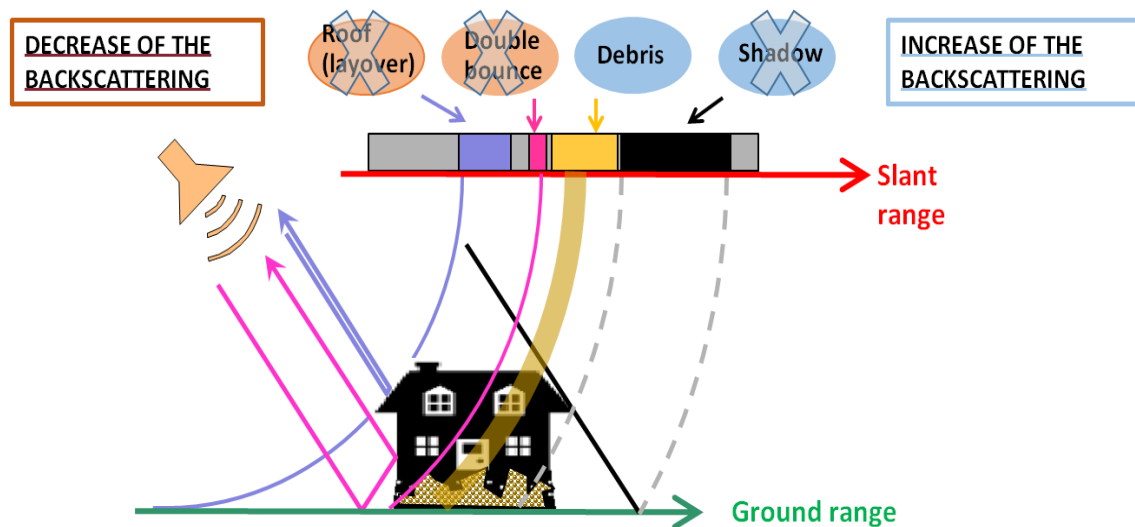
Figure 4-7: Decrease and increase of the backscattering intensity after the building collapse. Double bounce and layover effects are no longer observed. The scattering contributions from debris determine the increase of the backscattering within the building footprint. The region previously occluded by the shadow becomes visible.

4.4.1 Methodology

The present study has been conducted using a pair of Spotlight CSK images collected before and after L'Aquila earthquake (see section 3.3.2).

The proposed procedure for identifying collapsed buildings goes through several steps.

After some preprocessing, where pre- and post- event images are co-registered and radiometrically calibrated, the pre-event image is segmented to extract layover areas and bright stripes due to the double bounce (DB/LO objects), that are the regions where a decrease of the backscattering in the post-event SAR image is expected when a building collapses. The identification of DB/LO is performed using open and close morphological filters and assuming linear SEs with different orientation and length (Chini et al., 2015b).

In the case of partially collapsed buildings it may happen that the radar illuminates an undamaged building wall, and then the damage cannot be identified looking only at changes in the DB/LO area. Taking into account this aspect, the second stage of our methodology aims at identifying changes occurred within the building footprint and in the shadow area, generally characterized by an increase of backscattering in case of collapse (INC objects). For this purpose a thresholding and region growing based approach exploiting the KLD, computed between pre- and post- event SAR image after masking DB/LO regions, is used. Other scientists of the APhoRISM project team developed the segmentation procedures for singling out DB/LO and INC regions. They have been applied in the frame of the present thesis work to the L'Aquila images taken from CSK. In the following section a summary of those algorithms is provided. Once extracted from the images, DB/LO and INC objects are associated to polygons representing individual buildings. The association of DB/LO objects to a building is performed after correction of range displacement of the building itself to compensate for layover which depends on building height and radar incidence angle, as detailed in section 2.1.1. At the end of this step, a building may be associated to only one or both types of objects. Finally, exploiting change features extracted using only pixel belonging to DB/LO and INC objects, a classification step based on the Bayesian MAP criterion is performed. Each stage of the proposed methodology is explained in detail in the next subsections.

4.4.1.1 Image segmentation

Segmentation aims at identifying structures rather than making statistical analysis at single pixel bases (Blaschke and Strobl, 2001). It implicitly assumes that the actual scenario imaged by a SAR consists of regions that are homogeneous, i.e., considers a uniform backscattering within segments (Caves et al., 1998), with abrupt changes in the edges; it is the case of urban areas when we work with VHR SAR data, where a mixture of very bright and dark backscattering values is present.

Double-bounce and layover objects are the image features that mostly enable the identification of damages due to an earthquake. In order to extract these features from a single pre-event SAR image, a methodology, made up of two steps in cascade was implemented (Chini et al., 2015b).

In a first step, to highlight geometric characteristics of objects we are interested in, such

as bright thin lines and ridge representative of the double bounce and layover area, we extracted a multiscale morphological profile using the opening and closing operators described in section 2.2.2. Since structural classes from mathematical morphology depend on the size of the SE it is important to define the most suitable shape and dimension for classifying a certain type of target. The dimension of SE is directly related to pixel resolution and to the objects in the scene (Chini et al., 2009b). Since our target are the DB/LO objects which are basically linear, we implemented an anisotropic SE with linear shape, where anisotropic means that the filtering window privileges some specific directions. In particular, we have applied the opening and closing operators to the original pre-event CSK intensity image, using a linear SE with eight different directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) and three different lengths (5, 9, 13 pixels). In this way, we obtained additional features for describing each image pixel besides the pre-event SAR image itself.

In the second step, we used the morphological profiles and the pre-event SAR image itself as input to a K-means clustering algorithm (Richards and Jia, 2006) to obtain the DB/LO objects. The number of clusters in the K-means was fixed to 10 after a trial and error step, where different numbers were tested, and the one that maximizes the Jeffries Matusita (J-M) distance (Richards and Jia, 2006) between clusters was selected. The procedure for extracting the DB/LO objects was performed in the SAR slant range geometry, in order to preserve as much as possible the geometric characteristics of the linear structures we are looking for. Once identified, DB/LO were projected onto ground, using a LiDAR DTM, in order to assign the geographic coordinates, which allow their association to the buildings provided by a digital cartography of the area.

The second stage of our approach to damage detection aims to highlight changes occurred within the building footprint and in the shadow area. Given two SAR images collected on the same geographical areas at two different times, the most widely used approach to automatically detect regions of change is applying a thresholding algorithm to a further image generated from the original multitemporal images according to some comparison operator, such for example the ratio. In this work, for the automatic extraction of changes, we adopted a statistical-based approach, which exploits the KLD change feature and assumes it has a bimodal distribution locally. The KLD was estimated from pre- and post-event SAR intensity images using a moving window of 7x7 pixels, after masking DB/LO

objects. One major constrain for the statistical modelling-based algorithms is that they need a sufficient number of changed pixels to estimate a reliable threshold. Indeed, statistical modelling-based automated algorithms require that the changed class is identifiable, so that a pdf can be parameterized for this class. This is not the case when changes represent only a small portion of the area of interest. The Split-Based Approach (SBA) has been proposed in the literature to deal with images characterized by small percentages of changed pixels with respect to the total number of pixels (Bovolo and Bruzzone, 2007). The idea behind this approach is to split the image resulting from the comparison between two acquisitions, to which we refer as change image, into sub-images and to extract a threshold based on the statistics of each obtained tile. An alternative is to consider all tiles containing a sufficient number of changed pixels and to estimate one global threshold value that is applied to all tiles. SBA has also been used to map flood extent by considering only backscatter statistics inferred from a single flood image as a way to separate the 'flood' class from the rest (Pulvirenti et al., 2014). Usually SBA is applied with tiles of fixed size, defined, by the expert, based on the sensor resolution, the image size and the extension of the expected areas of change.

Here we considered a Hierarchical Split-Based Approach (HSBA) (Chini et al., under review), where the size of tiles, is not fixed a priori. Instead, a hierarchical tiling of the change image, i.e. the KLD image, is carried out in order to automatically select tiles (of potentially different sizes) where the statistical distribution function attributed to changes of backscatter values can be parameterized. Subsequently, based on the statistics of the selected tiles, we automatically delineate changed regions, making use of a methodology where, based on the fitted distribution functions, an optimization procedure is run in order to extract the optimal thresholds for generating the seed data set of a subsequent region growing step.

In the HSBA a hierarchical tiling of the change image is realized by starting with 4^0 tiles (i.e. the entire image) on the first level and iteratively subdividing the image into 4^L sub-images, where L is the hierarchical level of splitting. In other words, at L = 1 we are splitting the image into quarters, with L = 2 the image is subdivided into sixteenths, etc. Depending on L, tiles will thus be characterized by different sizes. At each level, descending from the upper level to the lower one, only tiles fulfilling the following criteria are retained:

a)   The histogram computed with pixels of the change image located in the considered tile must be bimodal.

b)   The smallest population in the change image pixels histogram must represent at least 10% of the considered tile.

Regardless of their size (i.e. the hierarchical level to which they belong) all tiles fulfilling the previous criteria are used to create a bimodal mask, which is binary. This mask is then applied to change image to extract areas where High and Low KLD values, corresponding to "Change" and "Non-change" regions, are present with a sufficient amount of pixels and the related distribution function is identifiable. In other words, the histograms of the masked areas in the change image are assured to be bimodal.

To calibrate the parameters of the two Gaussian distributions composing the histogram of each tile, we exploited the Levenberg-Marquardt algorithm (Marquardt, 1963) which is a standard technique used to solve nonlinear least squares problems, combining the steepest descent and inverse-Hessian function fitting methods. Once all distributions are fitted, the Ashman D (AD) coefficient (Ashman et al., 1994) was used for evaluating the bimodality. This coefficient quantifies how well two Gaussian distributions are separated, by considering the distance between mean values and their dispersions, i.e. standard deviations. For a mixture of two Gaussian distributions AD > 2 is required for a clean separation of the distributions.

Once tiles fulfilling conditions a) and b) have been identified, they were exploited to identify seed pixels for the subsequent region growing step through which regions of change (INC objects) were extracted. Starting from seed pixels, the region growing algorithm search for pixels within the whole change image that are connected neighbours of the seeds and that lie within a tolerance of the KLD value (Haralick and Shapiro, 1985). The choice of the threshold value for generating seed regions and the identification of the tolerance KLD value represent critical aspects. Our strategy is to select as seeds of the region growing step those pixels that have KLD values with a particularly high likelihood of belonging to the change class. To this aim, the histogram related to the pixels located in all the selected tiles has been fitted to a distribution given by the sum of two Gaussian functions, representing populations with high and low KLD values. Then, pixels with KLD value greater than the mean value of high KLD distribution were selected as seeds of the region growing algorithm. As for the tolerance we choose a KLD value that is

higher than the seeds threshold and that minimizes the RMSE between the theoretical high KLD distribution and the empirical distribution, which is the histogram of detected changed pixels resulting from the region growing.

4.4.1.2 **Feature extraction and classification**

We carried out the change analysis considering four change indicators, such as: intensity ratio, interferometric coherence (equation (2.17)), intensity correlation equation (2.18)), and Kullback-Leibler divergence. The latter was computed according its cumulant-based approximation, proposed by Inglada et al. (2007). We evaluated these parameters for each DB/LO and INC object and two independent classifications were performed: the one based on features extracted from DB/LO objects, the other based on features evaluated on the INC objects. The classification step has been carried out according the Bayesian MAP criterion considering two classes, discriminating between collapses (D = 5 in the EMS-98) and less severe damages (D < 5 in the EMS-98). For those building having associated both DB/LO and INC objects, a fusion of INC and DB/LO MAP classifications was performed. Given the features vectors $\Theta_{DB/LO}$ and Θ_{INC} associated to DB/LO and INC objects, respectively, we assign the building to the class ω_k (k = 1, 2) with the highest a-posteriori probability conditioned to the observed DB/LO and INC features, $P(\omega_k \mid \Theta_{DB/LO})$ and $P(\omega_k \mid \Theta_{INC})$, i.e. to the class:

$$\omega^* = argmax_{\omega_k \epsilon \Omega}\{P(\omega_k \mid \Theta_{DB/LO}), P(\omega_k \mid \Theta_{INC})\} \qquad (4.2)$$

As in the case of optical data, we retrieved the class conditioned pdfs from training data according to the Parzen window method considering again a Gaussian kernel. The simulated annealing algorithm, described in section 4.3.2, was used to tune the kernel width. The performances were assessed against the INGV ground truth using a Leave One Out CV (LOOCV) approach. LOOCV requires N iterations, where N is the number of sample in the data set. At each iteration, the training is performed using all the available data except for one sample, which is used for filling the classification CM.

4.4.2 Results

Figure 4-8(a) shows a multi-temporal SAR image for a portion of L'Aquila city generated by assigning to the Red channel the pre-event image and to the Green and Blue channels the post-event image. In such an image, red colour indicates the decrease of the backscattering in the post-seismic image due to the buildings collapse, while cyan pixels

single out the increase in the backscattering values due to the presence of debris and to the vanishing of the building SAR shadow. In Figure 4-8(c) and DB/LO and INC maps for the same area are reported, respectively. Polygons superimposed to each image delineate buildings that have sustained a damage of grade 5, according to the INGV ground survey. As expected, DB/LO regions associated to damaged buildings (DB/LO objects within the red dashed circles in Figure 4-8(b)) appear red in Figure 4-8(a) while INC objects are those regions where, in most of cases, a decrease of the backscattering can be observed in Figure 4-8(a).
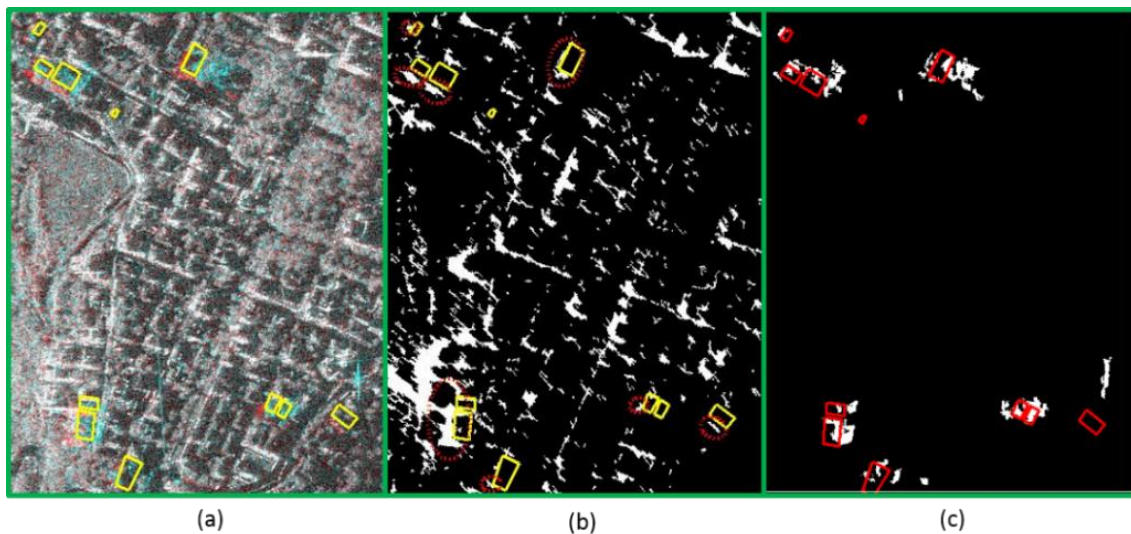
Figure 4-8: (a) RGB composite of pre-event (Red and Green) and post-event (Blue) CSK intensity images. Red and cyan areas single out respectively the decrease and the increase of backscattering in the post-seismic image. Polygons delineate buildings that have sustained a damage of grade 5. (b) DB/LO objects extracted from the pre-event image. Red dashed circles highlight DB/LO objects associated with damaged buildings. (c) INC objects extracted from the KLD map.

Figure 4-9 shows the four change detection features calculated for each DB/LO object. All parameters exhibit a certain degree of sensitivity to the damage. DB/LO objects associated to damaged buildings, generally are characterized by high values of intensity ratio and KLD, and low values of intensity correlation and interferometric coherence.

Drawing the attention on the intensity ratio (Figure 4-9(a)) it is possible to see how the brighter DB/LO objects (where the change is higher) are those associated to damaged buildings except for the buildings in the lower left corner of the figure. This happens because the layover effect due to the topography prevents the correct separation of objects, so, more objects, corresponding to damaged and undamaged buildings have been

classified as unique one. However, it can be noted that interferometric coherence (Figure 4-9(d)) and KLD (Figure 4-9(b)) single out this damage.

(a)   **(b)**

(c)   **(d)**
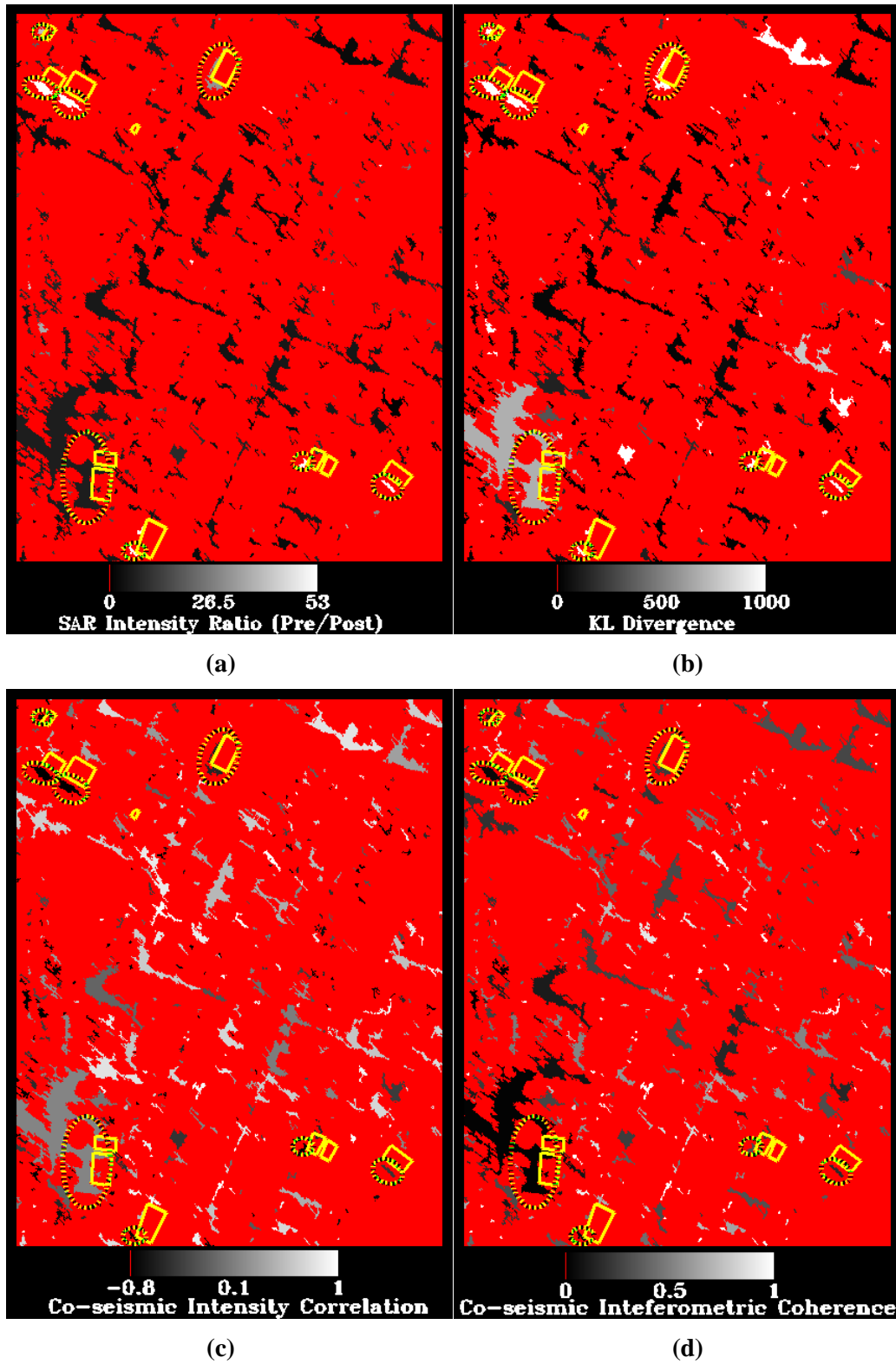
Figure 4-9: Change features calculated for each DB/LO object. Polygons delineate buildings that have suffered a damage of grade 5. Circles highlight the associated DB/LO objects.

In order to evaluate the usefulness of the investigated features and assess their contributions to the classification performances, we varied the number and the combination of features used as input space of the MAP classification algorithm.

KLD, coherence and intensity ratio turned out to be the most significant features for discriminating between collapsed and not collapsed buildings from changes in DB/LO regions. Combining DB/LO and INC classification results, the best performance in term of Cohen's kappa, was obtained classifying INC object based on all the available features. When the classification was performed based on features extracted only from DB/LO objects, 29.7% of the building with damage grade 5 were correctly identified. An increase of the sensitivity (40.5%) to the damage grade 5 was observed when also the INC objects were considered. All the results are summarized in the CMs reported in Table 4-7 and Table 4-8.

		INGV						
		(D=5)	(D<5)					
classifier	(D=5)	**22**	78	100		22.0%	*precision*	
	(D<5)	52	**1515**	1567		96.7%	*NPV*	
		74	1593	1667		*overall*	*Cohen's*	*normalized*
		sensitivity	*specificity*			*accuracy*	*kappa*	*kappa*
		29.7%	95.1%			92.2%	21.3%	24.83%

Table 4-7: Classification confusion matrix (features from DB/LO objects).

		INGV						
		(D=5)	(D<5)					
classifier	(D=5)	**30**	94	124		24.2%	*precision*	
	(D<5)	44	**1499**	1543		97.1%	*NPV*	
		74	1593	1667		*overall*	*Cohen's*	*normalized*
		sensitivity	*specificity*			*accuracy*	*kappa*	*kappa*
		40.5%	94.1%			91.7%	26.2%	34.64%

Table 4-8: Classification confusion matrix (features from DB/LO and INC objects).

The high rate of misdetection can be ascribed to the fact that some buildings classified as damage grade 5 by the ground survey are not easily detectable even by visual inspection

of optical imagery; indeed, the complexity of L'Aquila urban fabric represents a challenge for damage detection from space. Most of the buildings not recognized as damaged are located in the city historical centre. Instead, considering the new build-up area shown in Figure 4-8(a) better classification performances in terms of sensitivity to the damage were achieved. Being this area less densely settled with respect to the historical centre, the SAR response of the buildings is more similar to the canonical one, so that we are able to detect all the damaged buildings except one, probably due to its small size. Comparing Table 4-8 and Table 4-5 it can be seen that SAR data allow comparable performance to that obtained classifying optical images according to the MAP criterion in terms of sensitivity to the damage grade 5. (40.5% and 39.2% for SAR and optical data, respectively). However, optical data provide better performance in term of false alarm rate (5.9% and 1.2% for SAR and optical data, respectively) and thus precision (24.2% and 60.4% for SAR and optical data, respectively).

4.5   Data Fusion

This section presents data fusion strategies implemented for integrating damage detection products from EO data with a priori information as foreseen by the APE approach proposed in APhoRISM and described in section 3.2. For L'Aquila earthquake case study it was possible to assess data fusion results using a significant set of heterogeneous data, namely features extracted from optical and SAR images, data related to the structural vulnerability of buildings and to the soil instability. The interseismic deformation from SAR data and ground shaking from in situ measurements were also analysed, as foreseen by the APE approach, however no significant dynamic of this parameters was found in L'Aquila central area to assume a positive contribution to the final damage assessment. Two data fusion approaches were attempted, the one based on the Bayesian probability theory, the other exploiting SVMs.

The Bayesian data fusion approach implemented in this work merges damage probabilities from four independent modules: optical, SAR, geotechnical and structural. Under the assumption of conditional independence, the outputs of the modules are combined sequentially according to their availability. By expressing n observations coming from different modules as a vector $\mathbf{X} = (X_1, X_2, \dots X_n)$, we determine how the posterior probability $P(D/X_1, \dots X_n)$ evolves as we observe increasingly more features X_i

(i.e., we gather additional information), according to the following incremental formula:

$$\lambda_0 = p(D)$$

$$\lambda_i = p(D|X_i) \quad for\ all\ i$$

(4.3)

$$P(D|X_1, \dots, X_n)$$

$$= \frac{P(D|X_1, \dots, X_{n-1})\lambda_n}{P(D|X_1, \dots, X_{n-1})\lambda_n - \big(1 - P(D|X_1, \dots, X_{n-1})\big)(1 - \lambda_n)}$$

So, we can treat the posterior having observed $(X_1, \dots, X_k)$ as the "prior" for the remaining data $(X_{k+1}, \dots, X_n)$, and obtain the equivalent result to seeing all the data at once.

The posterior probabilities $P(D|X_{EO})$ coming from the EO modules, both optical or SAR, are evaluated according the non-parametric approach described in sections 4.3 and 4.4 of this work. The same approach is used to estimate the posterior probability $P(D|X_{GEO})$ from the soil resonant period, provided by the geotechnical module. As far as the posterior probability coming from the structural module $P(D|X_{STR})$ is concerned, we have adopted an approach available from the literature (Giovinazzi, 2005). It is a macroseismic method that leads to a probabilistic framework consistent with the Damage Probabilities Matrices (DPM) implicitly contained in the EMS-98 scale. It is parametric, in the sense that it assume a Beta pdf to represent the damage distribution around the mean value of the expected damage grade μ_D, as shown in Figure 4-10.
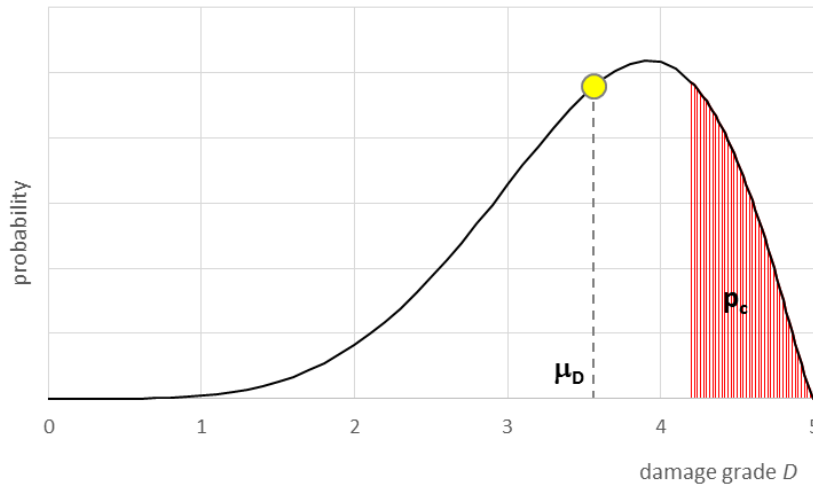
Figure 4-10: Structural module – Building damage pdf and collapse probability p_c. μ_D is the mean value of the expected damage grade.

The mean damage grade μ_D is a function of the macroseismic intensity and depends on the building vulnerability index, V. The latter is a parameter that quantifies the disposition of a building to be damaged by an earthquake. It has been defined taking into account the

EMS-98 vulnerability table (Table 4-9) which establishes a correspondence between vulnerability classes and building typologies categorized based on the construction material (masonry, reinforced concrete, wood and steel). Six vulnerability classes are defined, ranging from class A, the most vulnerable, to class F, the less vulnerable.

Table 4-9: EMS-98 building typologies and identification of their seismic behavior by vulnerability classes.

As it can be seen from Table 4-9, the correspondence between vulnerability classes and building typologies is not deterministic because the seismic behaviour of a building does not only depends on the behaviour of its structural system but it is affected by many other factors such as, for example, the height, the plan and vertical irregularity and the condition of maintenance. This aspect suggested the following definition for the vulnerability index:

$$V = V_0 + \Delta V \qquad\qquad (4.4)$$

where V_0 is the score that depends on the building typology and ΔV represents a contribution to take into account the presence of seismic behaviour modifiers.

For L'Aquila earthquake, the AeDES forms provided the information reported in Table 4-9, necessary to evaluate the vulnerability index V, namely the type of structure and the year of construction. The latter defines the parameter ERD (Earthquake Resistant

Design), according to the date when modern codes (i.e., the rules that specify the minimum standards for constructed aseismic objects) are actually used in the country under examination. In addition, AeDES forms provided information about the building height, exploited to compute the term ΔV of equation (4.4).

To evaluate the benefits of the Bayesian data fusion procedure, Table 4-10 first reports confusion matrices and performances of each independent module, assessed against the INGV ground truth. For the sake of clarity we point out that, for same buildings surveyed by INGV, information from structural and geotechnical module was missing This is due to the lack of same AeDES forms and to the fact that the available soil resonant period map does not cover the complete test area.

Looking at Table 4-10 it can be noted that, the optical module, which uses the 13 features listed in Table 4-2, produces the best classification performances. The Cohen's Kappa measures the performance of the classifier, with and without normalization with respect to the number of samples within each class (cardinality of the test set). In this case, the optical module achieve a Cohen's Kappa of 29%. The second remote sensing module, the SAR one, which exploits the four features described in section 4.3, provides worse performance with respect to the optical module (Cohen's Kappa, K=21.3%), since it is able to detect almost the same number of collapsed buildings at the expense of a higher number of false positives. The geotechnical module, in the L'Aquila test case, has just one feature related to soil resonant period. It was quite expected that the performance of this module was poor (K=3.6%) because soil oscillation period was found slightly correlated to the damage, but very much scattered when observed at the scale of a single building. Finally, the structural module, that also uses just one feature (the parameter μ_D), is not able to detect collapsed buildings, due to the fact that the posterior collapse probability conditioned to the vulnerability class never raises over the value of 0.5, at least according to the approach adopted in this work. It is worth noticing that confusion matrices from EO data, reported in Table 4-10, were derived without performing any optimization of the classifier parameters. If a calibration is possible using a training set, satellite false alarms are very much reduced and the data integration, which allows only the decrease of the collapse probability, as will be shown in the following, does not provide a significant contribution to the damage assessment.

		Optical	SAR	Geotechnical	Structural
CM	collapsed	36	35	42	0
	false alarms	107	157	617	0
	misdetections	38	39	32	46
	not collapsed	1486	1436	945	1312
	total	1667	1667	1636	1358
performances	sensitivity	48,6%	47,3%	56,8%	
	precision	25,2%	18,2%	6,4%	
	overall accuarcy	91,3%	88,2%	60,3%	
	Cohen's kappa	29,0%	21,3%	3,6%	
	normalized kappa	41,9%	37,4%	17,3%	

Table 4-10: Independent module classifications.

Table 4-11 shows the results of the incremental classification procedure, starting from the optical module, and combining one at time the posterior probabilities of the other modules. The first iteration combines optical and SAR module. The overall effect is quite negligible, likely because the optical dataset holds a similar and in the average richer information with respect to the SAR dataset. The second iteration combines damage probabilities estimated from the geotechnical module with the results of the data fusion obtained at the first iteration (optical+SAR). Since posterior probabilities from the geotechnical module are not very different between classes, the result of the classification remains almost unchanged, with a small reduction of false alarms. The third iteration is the most interesting, because the posterior probabilities of the structural module, despite of the very poor classification when used alone, are suitable to change the belief of previous modules. The building class vulnerability, i.e., the contribution coming from the structural module, is capable to supersede the belief of remote sensing mainly removing elements previously leading to the false positive collapses in the confusion matrix. Knowing that a building falls into a class of "low vulnerability" (i.e. steel, new reinforced concrete frame) imply a very low probability that the building has collapsed, regardless of what is observed from remote sensing data, optical or SAR.

At the expense of a reduction of 5 damage detections (going from 36 to 31), which represents 13% of the initial number of buildings correctly detected as collapsed, the data fusion reduces the number of false positives of 44 units. This result suggest that, when remotely sensed data processing cannot take advantage from a suitable training set for

optimizing the classification algorithm, a priori information can improve the final damage assessment product.

		Optical	Optical + SAR	Optical + SAR + Geotechnical	Optical + SAR + Geotechnical + Structural
CM	collapsed	36	37	36	31
	false alarms	107	99	96	63
	misdetections	38	37	38	43
	not collapsed	1486	1494	1497	1540
	total	1667	1667	1667	1677
performances	sensitivity	48,6%	50,0%	48,6%	41,9%
	precision	25,2%	27,2%	27,3%	33,0%
	overall accuarcy	91,3%	91,8%	92,0%	93,7%
	Cohen's kappa	29,0%	31,3%	31,0%	33,6%
	normalized kappa	41,9%	43,8%	42,6%	38,0%

Table 4-11: Incremental module classifications.

Finally, it is important to note that, as a consequence of the conditional independence hypothesis at the base of the data fusion procedure, the final result does not depend on the sequence of combination of outcomes of individual modules. Reversing or mixing the order of posterior probability of individual modules, one would get the same result.

In parallel to the experiment using the Bayesian approach, it was implemented a data fusion approach exploiting SVMs.

In order to integrate information coming from different data sources using SVMs, two approaches can be used. The first method combines features from different data sources in a unique feature vector, which is used as input to a unique SVM classifier. The second strategy uses a SVM for independently classifying each dataset. Then there are two options: the resulting rule-images, proportional to the distance of the sample from the hyperplane, can be combined using an additional SVM, or it can be kept the final decision based on the SVM that provides the output function with the maximum absolute value (Waske and Benediktsoon, 2007).

We exploited SVMs to integrate information coming from the optical module with that provided by structural and geotechnical modules. As a first approach, a unique SVM is used for classifying buildings described by a set of 17 attributes: the 13 optical features listed in section 4.3.2; the soil resonant period at the building site from the geotechnical module; building height, ERD and EMS-98 vulnerability class provided by the structural

module. The second strategy integrates data at level of decision function. Such approach exploits two SVMs for independently classifying optical data and information from both structural and geotechnical modules. An additional SVM is used for integrating the resulting rule images $f_{OPT}(x)$ and $f_{STRUCT-GEO}(x)$, as sketched in

Figure 4-11. Buildings for which attributes from structural and/or geotechnical modules were missing have been classified using a SVM trained with only optical features.

A k-fold cross validation approach with k=10 was used for evaluating the classification performances.
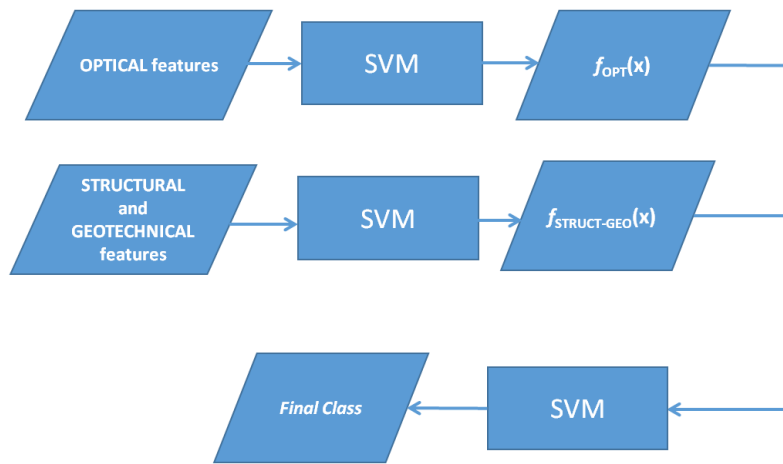
Figure 4-11: Schematic diagram of the SVM- based rule image fusion.

In Table 4-12 data fusion results are compared with classification performances achieved using a SVM exploiting only optical features. As in the case of the Bayesian data fusion approach, confusion matrices were derived without performing any optimization of the classifier parameters.

| | | Optical features SVM | Data fusion | |
			Multisource features vector SVM	Rule fusion SVM
CM	collapsed	30	28	33
	false alarms	57	35	56
	misdetections	44	46	41
	not collapsed	1536	1558	1567
	total	1667	1667	1697
performances	sensitivity	40,5%	37,8%	44,6%
	precision	34,5%	44,4%	37,1%
	overall accuarcy	93,9%	95,1%	94,3%
	Cohen's kappa	34,1%	38,4%	37,5%
	normalized kappa	37,0%	35,6%	41,1%

Table 4-12: SVM Data Fusion Results.

The results we obtained show that integrating optical data with information from different data sources can improve the results of the damage detection but not so significantly. When different data sets are integrated at features level, a considerable decrease of the false alarm rate is achieved at the expense of a slight reduction of the sensitivity to damage. Overall the results are slightly better, as shown by the increase of the Cohen's Kappa (driven by the decrease of false alarms), and only a small decrease of its normalized version (due to the decrease of correct detection). When the integration is performed at decision level, the results are not very much different with respect to those obtained using only optical features. In this case, however, an improvement of both Cohen's Kappa and its normalized version is observed, driven by a better sensitivity to damage (more detections) and a slight improved accuracy (less false alarms).

5. Triple Collocation to assess Classification Accuracy

5.1 Introduction

Assessing the accuracy of an image classification product is a fundamental step when remote sensing is used to produce or update thematic maps. There are many works in the literature addressing this problem (Congalton and Green, 1999; Foody 2002). Speaking about thematic accuracy, a first step is the acquisition of a set of reference samples belonging to well-known classes and the computation of the CM, which counts the occurrences of each category in the classification and what is considered the ground truth. From the confusion matrix different quality parameters can be derived, such as the Overall Accuracy (percentage of cases correctly allocated), the Cohen's Kappa coefficient, and the user's and producer's accuracy (see section 2.2.5).

In many cases, getting samples of the ground truth is not feasible; they can be affected by errors and thus do not represent the real situation (Foody 2002; Baraldi et al., 2005). Indeed, errors can affect test sets acquired from aerial images and visual interpretation used as reference to validate an automatic image classification algorithm. Moreover, even a ground survey can be affected by errors, especially when it is carried out in a difficult situation or is conceived for a purpose different from providing a reference for satellite image classification assessment. This is the case, for instance, of ground surveys performed after a disastrous event like an earthquake. The survey teams act in unsafe conditions (with a limited access to the affected area) with the purpose of giving a rapid assessment of the seismic intensity or checking the structural integrity of the buildings after the earthquake and their usability.

In Foody (2002) the impact of imperfect reference data on the assessment of classification performances was simulated considering errors either statistically independent or correlated to a certain degree. The case of a change map was considered, where the objective is discriminating between changed ($\Delta = 1$) or not changed ($\Delta = 0$) samples. It was found that in order to cope with the problem of imperfect reference and recover the real accuracy of the map, one should know exactly the quality of the reference set (which is not always the case), or to rely on latent class analysis when this information is not available at all (Foody, 2010). In addition, in Carlotto (2009) it was shown that when the reference data are affected by errors we could still compare the accuracy of two classifications relatively, but only if the number of test samples is large, although the retrieved accuracies remain biased.

A similar problem is faced in the field of medical diagnostic, when an accurate diagnostic test (denoted as gold standard) is not feasible, being too expensive or invasive. Therefore, the prevalence of a disease in a population has to be inferred by tests that have some unknown errors (Rutjies et al. 2007; Pepe and Janes, 2007). It can be note that in medical diagnostic generally a 2-classes problem is considered (positive or negative test result in a patient), whereas land cover classification implies a general number of N categories.

In a quite different field of remote sensing, i.e. the retrieval of geophysical parameters, a similar problem exists. Validation of remotely sensed products become questionable when the true values of the target parameter are not surely known, as the source of validation data is known to have its own errors. The Triple Collocation (TC) technique is successfully applied to cope with this problem. TC is able to retrieve the error variances of three systems observing the same target parameter, assuming their errors are independent. The technique has been originally conceived to assess the accuracy of wind speed retrieval over sea (Stoffelen, 1998) and successively largely adopted for testing soil moisture retrieval from space sensors or hydrological models (Dorigo et al., 2010; Pierdicca et al., 2015a), even combining retrievals from four systems (Pierdicca et al., 2015b). In the TC the unknown target parameter as well as the errors are continuous random variables, so that the method is not directly applicable to a classification problem. In this chapter, we propose a method to assess the accuracy of thematic maps without ground truth, (i.e., a gold standard) which is a common situation in post-disaster damage mapping. We exploit the TC idea to test the classification accuracy, which in our case

consists in comparing three different classifications of the same test samples to infer their CMs with respect to the unknown class the samples actually belong to. The main assumption is the conditional independence of the errors committed by the three systems. The proposed novel TC for Classification Assessment (TCCA) method represents a manageable way to solve a latent class model (Foody, 2010). We formulate it for a general number N of classes, and then solve it numerically for the 2-class problem. We first demonstrate the reliability of the proposed approach considering a simulated scenario, then we apply the TCCA to the problem of detecting collapsed building from a pair of VHR optical images collected before and after L'Aquila earthquake. The damage classification map was compared with two ground based damage maps resulting from the surveys carried out after the earthquake by INGV and DPC with different purposes and protocols.

Section 5.2 presents the TCCA approach, with all the mathematical details included in the Appendices. Section 5.3 analyses the proposed approach considering a simulated scenario. Section 5.4 presents the results of the TCCA, and finally Section 5.5 draws the conclusions.

5.2 TCCA theory

TC is used for validating geophysical parameter retrievals (e.g., wind speed, soil moisture) assuming that three systems, X, Y, Z, measuring the same target parameter θ, are affected by systematic calibration and additive random errors. Let us consider one system as the reference (i.e. unitary gain and zero bias) and assume the additive random errors are statistically independent and independent from the true parameter θ. Three variance-covariance matrices can be computed from the three sets of observations x, y, z of the same target parameter. It can be demonstrated that the unknown gains and variances of the random errors affecting each system can be derived. Although slight differences in the solution may exist, as discussed in Pierdicca et al. (2015a), the fundamental hypothesis of uncorrelated errors is common to most papers on this topic.

In a problem of image classification, three classification systems, X, Y, and Z, associate to each sample a label x, y, z, respectively, which is an integer between 1 and N, for an N-classes problem. Since we do not have a gold standard, all of them are imperfect indicators of the unobserved (latent) status θ of the samples (Foody, 2010).

We can therefore build three different CMs from each pair of classification systems. This can be translated into a probabilistic formulation as follows. If we normalize the three CMs with respect to the number of samples in the test set (getting the three normalized confusion matrices NCMs, from now on denoted as XY, XZ and YZ), we obtain an estimate of the joint probabilities $P_{XY}(x,y)$, $P_{XZ}(x,z)$, $P_{YZ}(y,z)$. With regard to XY, and similarly for the others, one can write:

$$P_{XY}(x,y) \leftrightarrow XY = \begin{bmatrix} p_{11}^{XY} & p_{12}^{XY} & \cdots & p_{1N}^{XY} \\ p_{21}^{XY} & p_{21}^{XY} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{N1}^{XY} & p_{N2}^{XY} & \cdots & p_{NN}^{XY} \end{bmatrix} \qquad (5.1)$$

Then, XY contains the joint probabilities arranged in a matricial form.

Actually, one wants to know the joint probabilities $P_{X\Theta}(x,\theta)$, $P_{Y\Theta}(y,\theta)$ and $P_{Z\Theta}(z,\theta)$ of the three system outcomes with respect to the latent variable, i.e. the NCM with respect to the class the samples really belong to. With reference to system X, we should retrieve $X\Theta$, which corresponds to the joint probability $P_{X\Theta}(x,\theta)$:

$$P_{X\Theta}(x,\theta) \leftrightarrow X\Theta = \begin{bmatrix} p_{11}^{X} & p_{12}^{X} & \cdots & p_{1N}^{X} \\ p_{21}^{X} & p_{22}^{X} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{N1}^{X} & p_{N2}^{X} & \cdots & p_{NN}^{X} \end{bmatrix} \qquad (5.2)$$

where rows indicate the outcomes of the classifier and columns the true classes, so that p_{ij}^{X} is the joint probability of getting class i out of the classifier, being j the true class. The diagonal terms indicate the probability of correct classification. Similarly, we refer to $Y\Theta$ and $Z\Theta$ as the NCMs of the other systems with elements p_{ij}^{Y} and p_{ij}^{Z}.

By adapting the hypothesis of TC, we assume that the errors of each classification system are conditionally independent to the errors of the other systems (e.g. the outcome of one classifier cannot be deduced from the outcome of the others) (see also Foody (2010) for a discussion on this hypothesis). It is demonstrated in the Appendix A that, under this assumption, we can derive the following equations for $X\Theta$, $Y\Theta$ and $Z\Theta$, which are expressed as function of the known matrices XY, XZ, YZ derived from the test set:

$$\begin{aligned} X\Theta \cdot P \cdot X\Theta^T &= XZ \quad \cdot YZ^{-1} \cdot XY^T \\ Y\Theta \cdot P \cdot Y\Theta^T &= YZ \quad \cdot XZ^{-1} \cdot XY \\ Z\Theta \cdot P \cdot Z\Theta^T &= YZ^T \cdot XY^{-1} \cdot XZ \end{aligned} \qquad (5.3)$$

where superscripts "T" and "-1" indicate transposition and matrix inversion, respectively. P is a diagonal matrix containing the inverse of N unknown probabilities p_j ($j=1, \dots N$) of

the classes (sometime denoted as prevalence of the classes), which can be computed from any of the target NCMs, so that the following $3N$ constraints also apply:

$$p_j = \sum_{i=1}^{N} p_{ij}^X = \sum_{i=1}^{N} p_{ij}^Y = \sum_{i=1}^{N} p_{ij}^Z \qquad (5.4)$$

They correspond to the marginalization of any of the joint probabilities, i.e. $P_{X\Theta}(x,\theta), P_{Y\Theta}(y,\theta)$ or $P_{Z\Theta}(z,\theta)$, to get the prevalence of the classes.

Note that matrices in equations (5.3) (e.g. $X\Theta \cdot P \cdot X\Theta^T$) are symmetric, so each matricial equation corresponds to $N(N+1)/2$ polynomial equations in the N^2 unknown $p_{ij}^{X,Y,Z}$, plus 3 constraints requiring that by marginalizing the joint probability we get one, i.e., $\sum\sum p_{ij}^X = \sum\sum p_{ij}^Y = \sum\sum p_{ij}^Z = 1$.

In TTCA we can also build a 3-dimensional NCM, denoted as XYZ, representing the joint probability $P_{X,Y,Z}(x,y,z)$, with elements p_{ijk} being the probability of encountering a sample labeled as classes i, j, and k by the three classifiers, respectively. This leads to other constraints for the target NCMs (see Appendix A):

$$p_{i,j,k} = \sum_{m=1}^{N} \frac{p_{im}^X p_{jm}^Y p_{km}^Z}{p_m^2} \qquad (5.5)$$

which are not all independent, as by marginalizing with respect to the classes one obtains $\sum_k p_{i,j,k} = p_{i,j}^{XY}, \sum_j p_{i,j,k} = p_{i,j}^{XZ}, \sum_i p_{i,j,k} = p_{i,j}^{YZ}$.

If we expand the matricial relations (5.3) we get, together with equations (5.4) and (5.5) a polynomial system of equations in the unknown probabilities $p_{ij}^{X,Y,Z}$. The analysis of its solution (i.e. existence and unicity) for the general case of N classes is beyond the scope of this work and is left to future studies. Here, we face the case of a 2-class problem, which is typical of any change detection problem, such as the detection of collapsed building after an earthquake. The solution is found in Appendix B. It requires to solve the equation (B.5) for the prevalence of class $i = 1$ (i.e. p_1), where $p_{11}^{X,Y,Z}, p_{12}^{X,Y,Z}, p_{21}^{X,Y,Z}, p_{22}^{X,Y,Z}$ as function of p_1 are given by equations (B.4). This method will be used in section 5.3 for a simulated scenario and in section 5.4 to compare damage classifications provided by satellite remote sensing and two ground surveys.

The hypothesis of conditional independence of the errors of the three systems is often assumed when exploiting the latent class analysis, as discussed in Foody (2010). Such paper investigated the presence of correlated errors by a simulated experiment and concluded that the satisfaction of the model assumption is critical. Indeed, in case a

correlation between system errors exists, with only three systems there is no way to solve for the unknown correlation, which should be known a priori. Solutions to this issue were proposed for TC, for instance using four systems instead of three (Gruber et al., 2016). This issue for TCCA is left for future studies.

The test data should be properly sampled in order to estimate the CMs and any other quality parameter. This is common to any classification assessment based on a truth (a gold standard). Then, the TCCA method is not adding any additional assumption in this respect.

5.3   TCCA applied to a simulated scenario

We tested the proposed approach on a simulated data set to verify the correctness of the solution, the effect of number of training samples, and the behavior of the TCCA when the main assumption, i.e. the conditional independence of the errors of the three system, does not hold. The approach is similar to what done in Foody (2010).

We considered a binary classification problem, where two classes, 0 and 1, have to be distinguished. Samples of the latent variable (i.e. the true class) were generated randomly according to a given prevalence of classes. Then, a random generation of errors was run, according to given conditional probabilities of errors $p_{XYZ}(1|0)$ and $p_{XYZ}(0|1)$ to produce the outcome of the three independent classifiers X, Y and Z. Finally, three confusion matrices were obtained pairwise comparing the outcomes of the three systems. The number of available test samples is a critical aspect when evaluating the classification performance, and could be even more critical in our TCCA. Therefore, the experiment was repeated by considering test sets of different size. The first experiment was carried out considering a balanced data set of 2000 samples, with equal prior probability (or prevalence) of the classes, i.e. $p_0=p_1=0.5$. The errors of the classifiers X, Y, and Z were generated assuming false alarm rate $p_{XYZ}(1|0)$ and misdetection $p_{XYZ}(0|1)$ of (8%, 12%), (10%, 30%) and (20%, 40%) for X, Y and Z, respectively. Consequently, among the 2000 samples, the number of samples of class 0 (1) correctly classified were 920 (880), 900 (700) and 600 (800), respectively. The resulting CMs for each pair of classifiers are reported in Table 5-1, whereas the 3-D CM with all combinations of classification outcomes is shown in Table 5-2. By applying the solution described in Appendix B, the CMs with respect to the class each simulated sample really belongs to (i.e. the gold

standard we know in the simulated exercise) are finally retrieved and reported in Table 5-3. The ones retrieved from the TCCA are in the first row, whilst those imposed in the simulation are in the second row of each cell.

X		Y	
		0	*1*
	0	854	186
	1	346	614

Y		Z	
		0	1
	0	599	601
	1	201	599

X		Z	
		0	1
	0	568	472
	1	232	728

Table 5-1: 2×2 confusion matrices computed from the three pairs of simulated classifications X, Y and Z.

	y=0		*y=1*	
	z=0	*z=1*	*z=0*	*z=1*
x=0	496	358	72	114
x=1	103	243	129	485

Table 5-2: 3-D confusion matrix for the three classifications as reported in Table 5-1.

X		Truth	
		0	*1*
	0	*896* / *920*	*144* / *120*
	1	*90* / *80*	*870* / *880*

Y		Truth	
		0	*1*
	0	*891* / *900*	*309* / *300*
	1	*94* / *100*	*706* / *700*

Z		Truth	
		0	*1*
	0	*593* / *600*	*207* / *200*
	1	*393* / *400*	*807* / *800*

Table 5-3: Retrieved confusion matrices of the three classifications X, Y and Z with respect to the "true" (first row in black), and actual CM according to the imposed error probabilities in the simulation (second row, in red). The former represents the number of correct classifications, misdetection and false alarms with reference to the unobserved true status of the samples.

We can appreciate that our technique was capable to retrieve fairly well the "true" CMs we look forward. There is an uncertainty that can be ascribed to the limited size of the test set in the simulated exercise. Such an error is inherent to each classification assessment, so it is interesting to see at what extent our method is reliable compared to a theoretical and sometime infrequent event in which a gold standard is available to test the classification. For this purpose, we generated a huge number of samples, from which smaller subsets were randomly selected. In order to avoid a tedious comparison of all CMs, we limited our comparison to an overall parameter commonly used to measure the classification performances, which is the Cohen's Kappa coefficient (Congalton and

Green, 1999). In Figure 5-1(a) we show the Cohen's Kappa retrieved from a finite amount of samples by the proposed TCCA for the simulated classification systems (i.e., same error rates) presented in Table 5-3. In Figure 5-1(b) we report the Cohen's Kappa obtained comparing the classification outcomes of the finite size test set with the actual class each sample belongs to (i.e. the gold standard). We can see that in both cases the uncertainty of the estimates decreases with increasing number of test samples, as expected. The TCCA has larger fluctuations, even though not much worse respect to the comparison with the truth.

| (a) | (b) |
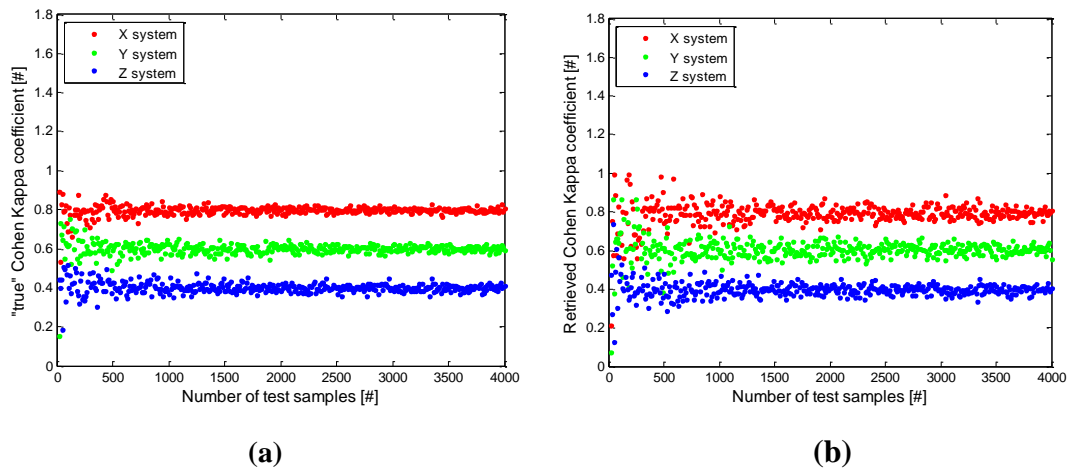
Figure 5-1: Cohen's Kappa coefficient as function of number of samples of the test set. (a): retrieved by TCCA. (b): evaluated with respect to the "true". Case of balanced test set (i.e., prevalence p_1 of damage equal to 0.5) and systems X, Y, Z with "true" Cohen's Kappa=0.8, 0.6, 0.4, respectively.

To better quantify this, we extracted from our simulated database many different test sets with same number of samples and computed the standard deviation of the Cohen's Kappa, with respect to both the gold standard and the class retrieved using TCCA. The result is reported in Figure 5-2 as function of the number of samples. We can see that the uncertainty of Cohen's Kappa retrieval reduces with the number of samples, and is larger for TCCA, as expected. It is also dependent on its expected value (i.e. the "true" Cohen's Kappa of the three simulated systems). The best retrievals of the Cohen's Kappa (i.e., lower standard deviation) when a gold standard is available is obtained for system X, i.e., the one with better performances, whereas using TCCA they are found for system Z with lower performances (this was already apparent in Figure 5-1(b)).

Roughly speaking, we need somehow more than double amount of test samples using TCCA to get the same precision in the Cohen's Kappa estimate when compared to the

availability (quite unlike indeed) of a gold standard. Despite this drawback, the TCCA offers the chance to estimate each classification accuracy without needing the "true" reference.
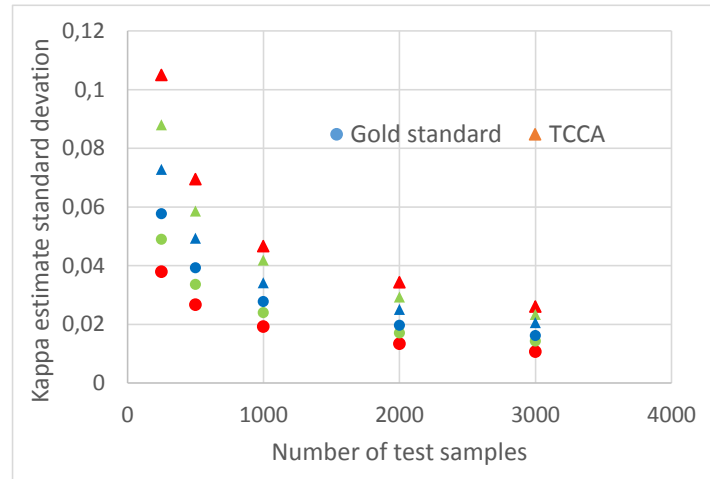
Figure 5-2: Standard deviation of Cohen's Kappa estimates as function of number of samples of the test set, when considering the "true" class of simulated samples (i.e., the gold standard) (circles) and using TCCA (triangles). Colour refers to the simulated classification systems X (red), Y (green) and Z (blue) as in Figure 5-1.

Another point worth to be investigate is the impact of an unbalanced test set. Assuming a prevalence of class 1 equal to 0.2, the precision of the Cohen's Kappa estimates worsens using the TCCA more than when computing the CMs with respect the gold standard. The standard deviation of the Cohen's Kappa estimates for the three systems X, Y and Z are compared in Figure 5-3 as a function of the number of test samples, as done in Figure 5-2.
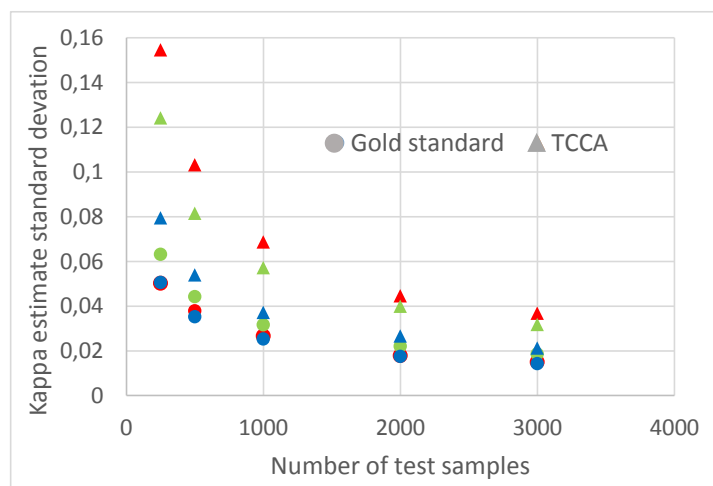
Figure 5-3: As in Figure 5-2, but for unbalanced test set (i.e., prevalence of damage equal to 0.2) and systems X, Y, Z with "true" Cohen's Kappa=0.74, 0.58, 0.26, respectively.

We can see that the TCCA uncertainty strongly increases for system X, which is the one with better classification performances, whereas the worsening is less for system Z, i.e. the one with poorer performance. This is not a problem uniquely of TCCA. Similarly, when the gold standard is available, the estimate of the Cohen's Kappa are not equally as good for the system with higher performance as in the case of a balanced test set. This is because with very unbalanced test sets and accurate classifiers, there are very few samples of error events in the smaller class, which makes the computation of the Cohen's Kappa less reliable.

Finally, we wanted to investigate the main assumption of the approach, which is the conditional independence of the outcomes of one classifier with respect to the others. In case there is a correlation among the classification results, the assumption does not hold anymore and the TCCA algorithm is expected to fail. To quantify the amount of this error, in the simulated exercise considered in Figure 5-1 (i.e. 2000 test samples, error probability leading to Cohen's Kappa=0.8, 0.6, 0.4 for X, Y and Z, respectively) we have imposed a percentage of the X classification outcomes in common with the Y classification. Figure 5-4 shows the retrieved Cohen's Kappa as function of such percentage.
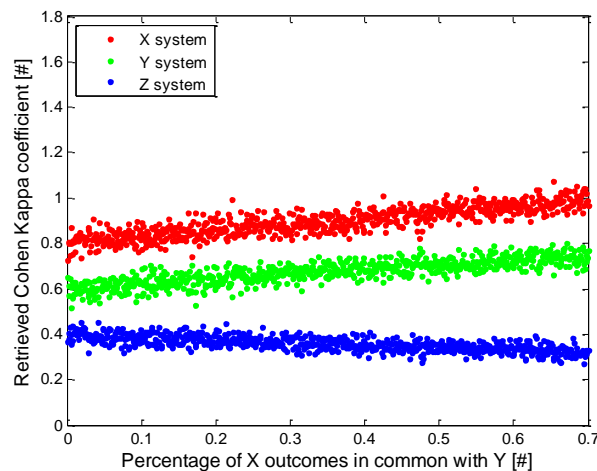
Figure 5-4: Cohen's Kappa retrieved by TCCA for increasing number of X outcomes in common with Y. Case of errors as in Figure 5-1 5-1, i.e. balanced test set and systems X, Y, Z with "true" Cohen's Kappa=0.8, 0.6, 0.4, respectively.

We can see that, besides fluctuations due to the finite sample size, the retrievals are increasingly biased. TCCA trusts the more on the systems that have similar outcomes, thus overestimating their performances and underestimating that of the third uncorrelated system. We remark that the validity of the assumption of uncorrelated errors must be

carefully verified before applying TCCA. In any case, the effect may be limited, as for instance in the example of Figure 5-4 where the bias overcomes the random uncertainty due to the test set finite size only when the number of common outcomes is well above 10%.

5.4 TCCA applied to the L'Aquila damage classification

We considered L'Aquila earthquake case study and applied the TCCA method to the three damage maps from DPC, INGV and from Quickbird images classification assuming the three systems provide independent classification maps.

The damage map from Quickbird data has been generated using the object oriented approach present in section 4 of this thesis. We considered a set of five change features for each object, which were selected by an exhaustive test of the classification performances among 13 candidate features (see Table 4-2). Buildings were classified according the Bayesian MAP criterion, using the Parzen window methods for estimating class conditional pdfs. For training the classifier we relied on a data set extracted by visual inspection of pre- and post- event Quickbird images. The fact that the two surveys were carried out by different teams and the use of an independent supervised image processing approach guarantee that the hypothesis of independent errors of the systems holds.

Since the two ground surveys do not refer exactly to the same building GIS layer, in order to apply the TCCA approach, it was first necessary to intersect the two layers, as detailed in Section 3.3.3. This leaded to a common set of 1445 buildings with associated labels, i.e. *collapsed* (D = 5) or *non-collapsed* (D < 5), according to the two surveys. The MAP classification was then applied to the same set of image objects.

The CMs resulting from the comparison of the MAP algorithm result and the classification provided by the ground surveys are reported Table 5-4 (2-D *standard* CMs), and Table 5-5 (3-D CM providing all combinations of classification outcomes). Table 5-4 shows the standard CM considering all combinations of two classifications. Table 5-5 refers to the triple collocation approach, as it counts the occurrences of all possible combinations of the outcomes of the three classifiers. Specifically, it reports the number of samples labeled as D = 5 by the three classifiers, or commonly labeled as D < 5, or labeled as D = 5 by two classifiers and D < 5 by the third one, and so on so forth.

		INGV	
		D<5	*D=5*
DPC	*D<5*	1309	30
	D=5	75	31

		EO	
		D<5	*D=5*
INGV	*D<5*	1341	43
	D=5	37	24

		EO	
		D<5	*D=5*
DPC	*D<5*	1297	42
	D=5	81	25

Table 5-4: 2×2 confusion matrices for the three pairs of systems, assuming X as the DPC survey, Y the INGV survey and Z the Earth Observation (EO) image classification.

	INGV D<5		INGV D=5	
	EO D<5	EO D=5	EO D<5	EO D=5
DPC D<5	1276	33	21	9
DPC D=5	65	10	16	15

Table 5-5: 3-D confusion matrix for the three systems as defined in Table 5-4.

We can notice that the performances of the image classification (denoted as EO in the tables) compared to the two ground surveys, supposing they could be considered as the *ground truth*, are not very good. Only 24 collapses were detected out of 61 according to INGV, or 25 out of 106 according to DPC, with overall accuracy of 94.5% (Cohen's Kappa, K=0.35) and 91.5% (K=0.25), respectively. However, the matching between the two ground surveys, which we remind were carried out according to different purposes, was not good as well, being 92.7% the percentage of buildings for which the two ground surveys agree (K=0.34). Then the question we want to answer by using the TCCA is: which is the actual accuracy of the three systems with respect to the *truth*?

Using the TCCA solution for the 2-class problem, after normalization of the CMs in Table 5-4, we derived the three NCMs XΘ, YΘ, ZΘ as function of the probability of collapsed building p_2 using equation (B.4). Then from the NCMs we computed the Overall Accuracy and Cohen's Kappa coefficients. In Figure 5-5 it is shown that the INGV classification has a Cohen's Kappa always greater than the others, despite of the true prevalence of *damaged* class. Concerning the other two classification results, EO is slightly better than the DPC classification when the probability of damage is low, whilst DPC becomes a bit better respect the EO classification if the probability of damage is higher (see Figure 5-5). This result can be justified if one considers that according to the CMs in Table 5-4 the INGV classification has a fair matching with both DPC and EO

(92.7% and 94.5%, respectively), whereas this is not true for the others that exhibit a pronounced disagreement (91.5% between EO and DPC). This experimental evidence makes the TCCA solution trust more on the INGV classification, as it can be intuitively understood. If the prevalence of damaged class were high, the TCCA would have trusted more on the DPC survey that has a higher occurrence of D = 5 class respect to EO. Figure 5-6 compares the Overall Accuracy of the systems. It is noticeable that according to this quality score the satellite classification has a quality intermediate when compared to the ground surveys.

Figure 5-5: Cohen's Kappa coefficient of systems DPC, INGV and image classification results (EO) as function of prior probability of collapsed building.
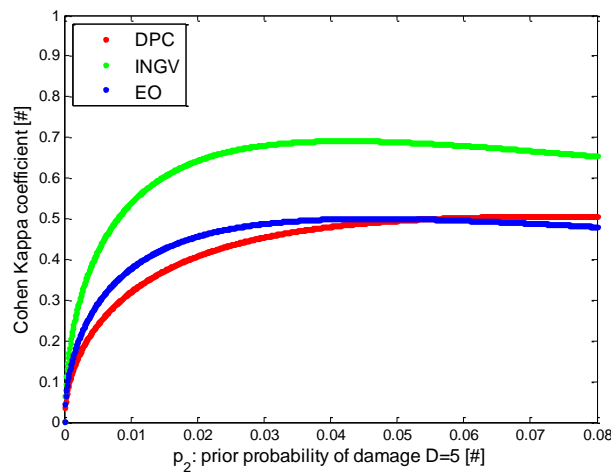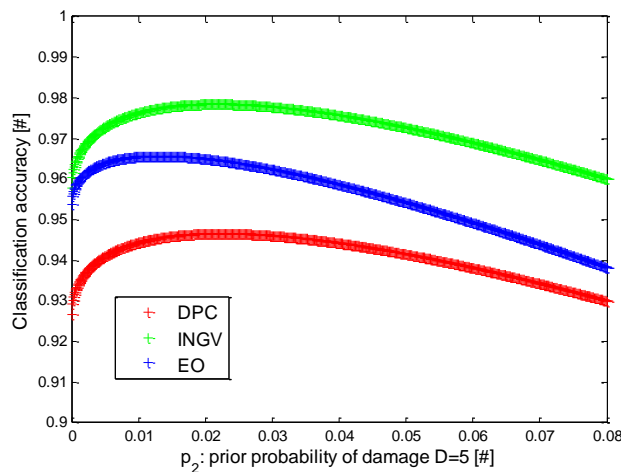

Figure 5-6: **Overall Accuracy of systems DPC, INGV and image classification results (EO) as function of prior probability of collapsed building.**

To retrieve the final solution, the probability of collapse p_2 has to be derived by solving equation (B.5). From the CM in Table 5-5 it comes out that the number of samples

classified as *undamaged* by the three systems is 1276, whereas the number of samples accordantly considered as *damaged* is 15. Then, dividing by 1445, it comes out that $p_{111}=$ 0.8830 and $p_{222}=0.0104$. A numerical solution of equation (B.5) was found by using the Matlab® Symbolic Tool. Here for sake of better clarity we depicts a graphical solution in Figure 5-7, where $p_{21}^X p_{21}^Y p_{21}^Z / p_1^2 + p_{22}^X p_{22}^Y p_{22}^Z / (1 - p_1)^2$ derived by using (B.4) is plotted as function of p_2, so that according to (B.5) the solution is found when the value $p_{222} = 0.0104$ is reached. Both methods provided $p_2 = 0.0529$, from which XΘ, YΘ and ZΘ and associated classification accuracy scores were finally retrieved using equation (B.4).
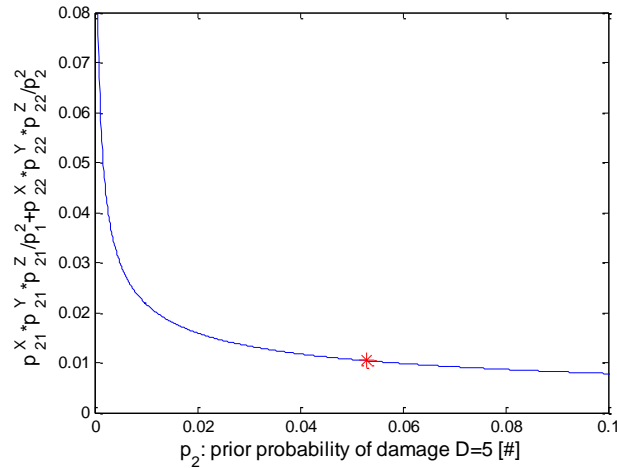
Figure 5-7: $p_{222} = p_{21}^X p_{21}^Y p_{21}^Z / p_1^2 + p_{22}^X p_{22}^Y p_{22}^Z / (1 - p_1)^2$ **as function of** p_2**. Red dot corresponds to the solution of eq. (B.5) with** p_{222}**=15/1445=0.0104.**

Multiplying XΘ, YΘ and ZΘ by the number of samples, we computed the CMs with respect to the *truth*, which are reported in Table 5-6. They provide the number of correct classifications, misdetections ad false alarms with respect to the unobserved true status of the samples. Note that column totals are the estimated number of *damaged* (D = 5) and *non-damaged* (D < 5) buildings.

DPC		Truth	
		D<5	D=5
	D<5	1311	28
	D=5	58	48

INGV		Truth	
		D<5	D=5
	D<5	1356	28
	D=5	13	48

EO		Truth	
		D<5	D=5
	D<5	1339	38
	D=5	30	38

Table 5-6: Confusion matrices of the three damage classifications with respect to the "true". X refers to the DPC survey, Y to the INGV survey and Z to the image classification (EO). They represent the number of correct classifications, misdetection and false alarms with reference to the unobserved true status of the samples. Column totals is the estimated number of really *damaged* (D = 5) and *non-damaged* (D < 5) buildings.

From these CMs, we can derive the Overall Accuracy scores and the Cohen's Kappa coefficients of the system. We observe again that the INGV survey is the most accurate (97.2 % accuracy, 0.687 Cohen's Kappa) with few false positives but a bit more misdetections of collapse. The image classification has worse performances (accuracy of 95.3 %, Cohen's Kappa of 0.500), with a much larger number of false alarms. However, its real accuracy is better than that found by comparing directly with the two ground surveys. The DPC survey, with an accuracy of 94.1 % and a Cohen's Kappa of 0.499, is a bit worse. This result can be justified by the different purposes of the surveys. In particular, the DPC survey provides a great amount of information which aims at assessing the structural condition of the buildings and their fitness for use, despite of their exterior appearance. For instance, a heavy damage of the vertical structures (e.g., pillars) may have been considered as a predominant damage of grade 5, even if not consistent with the general definition of EMS-98 when looking at the building from the exterior. Conversely, the image-based classification suffers from situations we were able to identify in some cases by a careful analysis of the satellite images but also from an aerial survey and ground based photographs. Heavy damage may be not visible looking down upon, for instance in case of the so-called *pancake* effect, which we recognized occurred in some buildings, and is characterized by the downfall of the ground floor only. A picture of a building taken from ground and the same building imaged by Quickbird after the earthquake are shown in Figure 5-8 to give an example of this effect.
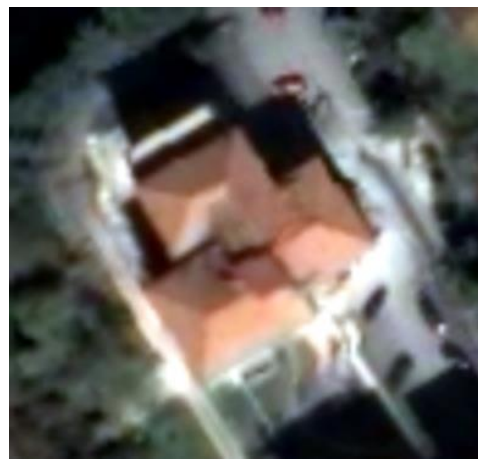
(a) (b)

Figure 5-8: (a) Photograph of a building in L'Aquila affected by the *pancake* type of damage and (b) the same building in the post-event Quickbird image. With respect to the pre-event image (not shown here) the only change is the presence of dust on the street due to the damage that was classified as EMS-98 grade *D*=5 from both ground surveys.

Additionally, false positives occurred when the building underwent a restoration just after the acquisition of the pre-event image, which was wrongly classified as a change due to the earthquake. Despite of these problems and the difficulty of change detection in a very dense historical town, the satellite classification exhibited a damage detection performance comparable or even better of at least one ground survey.

6. Conclusions and future work

In this work, change detection methodologies exploiting VHR SAR and optical data were developed to assess earthquake damage in urban area. The investigation was focused on the estimation of damage of individual buildings, and in particular the detection of fully collapsed buildings. Different approaches to combine heterogeneous data for a final assessment of the damage have been attempted.

Procedures were developed and tested considering the 2009 L'Aquila earthquake where a significant set of data from different sources was available. For such case study it was possible to exploit features from optical and SAR change detection analysis for each building of the central part of the town, together with data related to structural vulnerability and the soil frequency.

Considering the possible situation of missed data, likely occurring in other cases, an approach to data fusion based on the Bayesian theory, exploiting the hypothesis of conditional independence between remote sensing observations and other data sources, has been designed. The approach allows one to accumulate the evidences of damage as soon as new data become available by combining posterior probabilities.

Another way to solve the problem was assessed using SVMs, and the idea to apply different SVMs in cascade, taking the output of the first SVM stage in the sequence as input of the others and eventually keeping the decision of the first if a sample (i.e. a building) do not have associated information. This combination is done in the space of the distances of each sample from the hyperplane separating the two classes (i.e., "collapsed" and "not collapsed") in the higher dimension feature space.

It has been found that, with respect to the damage assessment generated by processing the remotely sensed data, the structural data in particular, notwithstanding their poor capability to predict the damage of buildings when used alone, can be able to improve the final damage assessment product, especially reducing the false alarms of the image change detection algorithm. This is surely true when the remotely sensed data processing

cannot take advantage from a suitable training set for optimizing the classification algorithm. Conversely, if a calibration is possible using the training set (e.g., optimization of the non-parametric likelihood estimator), the satellite false alarms are very much reduced and the vulnerability does not provide a positive contribution. As far as soil instability is concerned, the soil oscillation period was found slightly correlated to the damage, but very much scattered when observed at the scale of a single building, so that its contribution to the damage classification was very small (although perceptible).

The performances of damage classification from EO only, or after integration of additional information, have been quantitatively assessed by comparison mainly with INGV ground survey.

A cross-comparison of the EO-based classification and the two different surveys (INGV and DPC) has been performed by means of the TC approach, which, in this work, was extended to deal with classification problem. The TCCA was successfully verified on a simulated dataset and its robustness with respect to the scarcity of test samples and to the unsuitability of the main assumption (i.e. the independence among classifications errors) was evaluated. Then the TCCA was applied to the damage classification after L'Aquila earthquake. Although one can expect a ground survey would represent a reference for validating satellite image classifications, the TCCA investigation showed that assuming a ground survey as the reference for testing the satellite damage classification can be questionable. The concept of damaged or non-damaged can be controversial on its own, as it may be related to the purpose of the survey and to the conditions in which the teams were operating. The TCCA results demonstrated that EO is able to provide damage maps with accuracy comparable to at least one of the ground survey. This is an encouraging result for what concerns the applicability of space technology in this field, as the EO based classification is surely less expensive and rapidly available once a suitable constellation of satellites is in orbit, as it becomes true with the full deployment of the European Copernicus program.

Appendix

A. TCCA Analytical Formulation

The hypothesis of independent errors requires that the probability of outcome *x, y, z* of each system is not dependent on the result of the others, e.g. $P_X(x|y,\theta) = P_X(x|\theta)$. Then by using the Bayesian theorem we can write for the pair of systems X and Y (in the sequel we omit the subscript in the pdf symbol $P(\ )$ for sake of brevity):

$$P(x,y,\theta) = P(x|y,\theta)P(y,\theta) = P(x|\theta)P(y,\theta) = \frac{P(x,\theta)P(y,\theta)}{P(\theta)} \qquad \text{(A.1)}$$

By marginalizing with respect to θ we get the joint probability $P(x,y)$, i.e. the observed confusion matrix XY whose elements can therefore be expressed as:

$$p_{i,j}^{XY} = \sum_{k=1}^{N} \frac{p_{ik}^X p_{jk}^Y}{p_k} \qquad \text{(A.2)}$$

Similarly for the other pairs. It is straightforward to verify that this corresponds to write the following three matricial relations relating the target normalized confusion matrices XΘ, YΘ and ZΘ we want to retrieve to the observed ones, i.e., XY, XZ, YZ:

$$\begin{aligned} X\Theta \cdot P \cdot Y\Theta^T &= XY \\ X\Theta \cdot P \cdot Z\Theta^T &= XZ \\ Y\Theta \cdot P \cdot Z\Theta^T &= YZ \end{aligned} \qquad \text{(A.3)}$$

P is a N×N diagonal matrix given by:

$$P = \begin{bmatrix} 1/p_1 & 0 & \dots & 0 \\ 0 & 1/p_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/p_N \end{bmatrix} \qquad \text{(A.4)}$$

Where p_j (j=1, …, *N*) are the *N* unknown probabilities of the classes. By deriving XΘ from the second equation in (A.3) and YΘ from the third equation, substituting into the first equation and using well-known matrix operation rules, one obtains:

$$XΘ \cdot P \cdot YΘ^T = XZ \cdot [P \cdot ZΘ^T]^{-1} \cdot P \cdot [YZ \cdot [P \cdot ZΘ^T]^{-1}]^T = XY$$
$$XZ \cdot [ZΘ^T]^{-1} \cdot P^{-1} \cdot P \cdot [YZ \cdot [ZΘ^T]^{-1}P^{-1}]^T = XY$$
$$XZ \cdot [ZΘ^T]^{-1} \cdot I \cdot P^{-1} \cdot ZΘ^{-1} \cdot YZ^T = XY$$

Where I is the identity matrix and we considered that P is diagonal so that $P^{-1} \cdot P = I$. Then multiplying by XZ^{-1} on the left and by $[YZ^T]^{-1}$ on the right:

$$[ZΘ^T]^{-1} \cdot P^{-1} \cdot ZΘ^{-1} = XZ^{-1} \cdot XY \cdot [YZ^T]^{-1}$$
$$ZΘ \cdot P \cdot ZΘ^T = YZ^T \cdot XY^{-1} \cdot XZ \tag{A.5}$$

Which is the third of equation (5.3) in the main text we wanted to demonstrate.

Similarly for the joint probability $P(x,y,z,θ)$, and using (A.1) one can write:

$$P(x, y, z, θ) = P(x, y|z, θ)P(z, θ) = P(x, y|θ)P(z, θ) = \frac{P(x, y, θ)P(z, θ)}{P(θ)}$$
$$= \frac{P(x, θ)P(y, θ)P(z, θ)}{P^2(θ)} \tag{A.6}$$

Which leads to equation (5.5) in the main text once we marginalize with respect to $θ$.

B. TCCA solution for N=2 classes

Equations (5.3), (5.4) and (5.5) in the main text can be easily solved in case of $N = 2$ classes. We show the solution for XΘ only, considering it is identical for the other CMs. Each right term in equation (5.3) is a symmetric 2×2 known matrix (it is a function of the three CMs derived from the test set), whose elements are hereafter denoted as *a, b, c*. Then, equation (5.3) becomes (we omits superscript "X" for sake of simplicity):

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \dfrac{1}{p_1} & 0 \\ 0 & \dfrac{1}{p_2} \end{bmatrix} \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{B.1}$$

Then, using equation (5.4) the probabilities of each class are $p_1=p_{11}+p_{21}$, $p_2=1-p_1=p_{12}+p_{22}$. Equation (B.1) corresponds to three scalar equations, which are dependent since from (B.1) it follows:

$$\begin{bmatrix} \dfrac{p_{11}^2}{p_1} + \dfrac{p_{12}^2}{1-p_1} & \dfrac{p_{11}p_{21}}{p_1} + \dfrac{p_{12}p_{22}}{1-p_1} \\ \dfrac{p_{11}p_{21}}{p_1} + \dfrac{p_{12}p_{22}}{1-p_1} & \dfrac{p_{21}^2}{p_1} + \dfrac{p_{22}^2}{1-p_1} \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Indeed, it can be easily found that $a+b=p_{11}+p_{12},\ b+c=p_{21}+p_{22}$, and then $a+2b+c=1$. Therefore, we solve the equation with respect to an unknown quantity, that we assume being p_1. As for the upper-left element of (B.1), one can write:

$$\frac{p_{11}^2}{p_1} + \frac{p_{12}^2}{1-p_1} = a$$

Then:

$$a p_1(1-p_1) = (1-p_1)p_{11}^2 + p_1 p_{12}^2 \tag{B.2}$$

Using $p_{12} = a+b-p_{11}$ and $c = 1-a-2b$, we found a second-degree algebraic equation in p_{11}, whose coefficients are functions of p_1:

$$p_{11}^2 - 2p_1(a+b)p_{11} + p_1(ap_1 + b^2 - ac) = 0 \tag{B.3}$$

That we can easily solve, and then we can derive all the elements of XΘ:

$$p_{11} = p_1(a+b) \pm \sqrt{p_1^2(a+b)^2 - p_1(ap_1 + b^2 - ac)}$$
$$p_{12} = a + b - p_{11} \tag{B.4}$$
$$p_{21} = p_1 - p_{11}$$
$$p_{22} = b + c - p_{21}$$

The sign shall be chosen considering the constraint of p_{ij} being real and between zero and one.

To found p_1 to be substituted into equation (B.4) we have to rely on the 2-D CM and to equation (5.5). It is enough to impose only one of the eight constraints of equation (5.5) since they are all dependent, so for instance we can consider:

$$p_{222} = \frac{p_{21}^X p_{21}^Y p_{21}^Z}{p_1^2} + \frac{p_{22}^X p_{22}^Y p_{22}^Z}{(1-p_1)^2} \tag{B.5}$$

which can be easily solved numerically, as done in the main text.

Bibliography

Akbani, R., Kwek, S., Japkowicz, N. Applying support vector machines to imbalanced datasets. In Proceedings of the 15th European Conference on Machine Learning, pp. 39–50, 2004.

Anniballe, R., Bignami, C., Chini, M., Pierdicca, N., Stramondo, S. Detecting sparse earthquake damages in high density urban settlements by VHR SAR data. Proceeding of SPIE Vol. 9243, SAR Image Analysis, Modelling, and Techniques XIV, 924307, 2014.

Anniballe, R., Chini, M., Pierdicca, N. Earthquake damage detection in urban areas from VHR SAR imagery. Atti della XXI Riunione Nazionale di Elettromagnetismo, 2016.

Anniballe, R., Chini, M., Pierdicca, N., Bignami, C., Stramondo, S., Noto, F., Scalia T., Martinelli, A., Mannella, A. Detecting earthquake damage in urban area: application to Cosmo-SkyMed imagery of L'Aquila earthquake. Proc. SPIE 9642, SAR Image Analysis, Modelling, and Techniques XV, 96420C, 2015.

Aoki, H., Matsuoka, M. and Yamazaki, F. Characteristics of satellite SAR imaged in the damaged areas due to the Hyogoken-Nanbu earthquake. In Proceedings of the 19th Asian Conference on Remote Sensing, C7, pp. 1–6, 1998.

Ashman, K. M., Bird, C. M., Zepf, S. E. Detecting bimodality in astronomical datasets, The Astronomical Journal, Vol. 108, pp. 2348–2361, 1994.

Baraldi A., L. Bruzzone, P. Blonda. Quality assessment of classification and cluster maps without ground truth knowledge. IEEE Transactions on Geoscience and Remote Sensing, 43(4), 857-873, April 2005.

Baraldi, A., Parmiggiani, F. An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. IEEE Transactions on

Geoscience and Remote Sensing, 33, 293−304, 1995.

Batuwita, R., Palade, V. Class imbalance learning methods for support vector machines. In Imbalanced learning: Foundations, algorithms, and applications, (pp. 83–99). John Wiley & Sons, Inc., 2013.

Benediktsson, J.A., Palmason, J.A., Sveinsson J. Classification of hyperspectral data from urban areas based on extended morphological profiles. IEEE Transactions on Geoscience and Remote Sensing, 43(3), 480–491, 2005.

Benediktsson, J.A., Pesaresi, M., Arnason, K. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. IEEE Transactions on Geoscience and Remote Sensing, 41 (9), 1940–1949, 2003.

Bignami, C., Chini, M., Pierdicca, N., Stramondo, S., Comparing and combining the capability of detecting earthquake damages in urban areas using SAR and optical data. IEEE Proc. IGARSS 1, 55–58, 2004

Bignami, C., Chini, M., Stramondo, S., Emery, W. J., Pierdicca, N. Objects textural features sensitivity for earthquake damage mapping. In Joint Urban Remote Sensing Event (JURSE) 2011, pp. 333-336, IEEE, April 2011.

Bitelli, G., Gusella, L. Remote sensing satellite imagery and risk management: image based information extraction. Risk Analysis VI: Simulation and Hazard Mitigation, 39, 149, 2008.

Blaschke, T. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 65, 2–16, 2010.

Blaschke, T., Strobl, J. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS, GIS Z. Geoinf. Syst. 14 12–17, 2001.

Borfecchia, F., De Cecco, L., Pollino, M., La Porta, L., Lugari, A., Martini, S., Pascale, C. Active and passive remote sensing for supporting the evaluation of the urban seismic vulnerability. Italian Journal of Remote Sensing, 42(3), 129-141, 2010.

Bouleimen, K., Lecocq, H. A new efficient simulated annealing algorithm for the resource-constrained project scheduling problem and its multiple mode version. European Journal of Operational Research, 149.2, 268-281, 2003.

Bovolo, F., Bruzzone, L. A split-based approach to unsupervised change detection in large-size multitemporal images: application to tsunami-damage assessment. IEEE Transactions on Geoscience and Remote Sensing, 45(6), 1658-1670, 2007.

Brett, P. T., Guida, R. Earthquake damage detection in urban areas using curvilinear features. IEEE Transactions on Geoscience and Remote Sensing, 51(9), 4877-4884, 2013.

Brower, B., Laben, C. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. US Patent 6011875, 2000.

Brunner, D., Lemoine, G., Bruzzone, L. Earthquake damage assessment of buildings using VHR optical and SAR imagery. IEEE Transactions on Geoscience and Remote Sensing, 48(5), 2403-2420, 2010b.

Brunner, D., Lemoine, G., Bruzzone, L., Greidanus, H. Building height retrieval from VHR SAR imagery based on an iterative simulation and matching technique. IEEE Transactions on Geoscience and Remote Sensing, 48(3), 1487-1504, 2010a.

Bruzzone, L., Bovolo, F. A novel framework for the design of change-detection systems for very-high-resolution remote sensing images. Proceedings of the IEEE, 101(3), 609-630, 2013.

Burges, C. J. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery, 2(2), 121-167, 1998.

Carlotto M. J. Effect of errors in ground truth on classification accuracy. International Journal of Remote Sensing, 30(18), 4831-4849, 2009.

Caves, R., Quegan, S., White, R. Quantitative comparison of the performance of SAR segmentation algorithms. IEEE Transactions on Image Processing, 7, 1534−1546, 1998.

CEOS (Committees on Earth Observation Satellites), Satellite Earth Observations in Support of Disaster Risk Reduction – Special 2015 WCDRR Edition, 2015. Available on line: www.eohandbook.com/eohb2015/files/CEOS_EOHB_2015_WCDRR.pdf.

CEOS (Committees on Earth Observation Satellites). The Earth Observation Handbook 2012/Special Edition for Rio+20. [ESA SP-1325 July 2012] (ESA Communications), 2012. http://www.eohandbook.com

Chapelle, O., Vapnik, V., Bousquet, O., Mukherjee, S. Choosing multiple parameters for support vector machines. Machine learning, 46(1-3), 131-159, 2002.

Chen, C. H. Signal and image processing for remote sensing. CRC press, 2007. ISBN 0-8493-5091-3

Chesnel, A. L., Binet, R., Wald, L. Object oriented assessment of damage due to natural disaster using very high resolution images. Proc. Of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3736-3739, 2007.

Chini, M. Building Damage from Multi-resolution, Object-Based, Classification Techniques, in Encyclopedia of Earthquake Engineering, pp. 1-11, Springer Berlin Heidelberg, 2014a. DOI: 10.1007/978-3-642-36197-5_225-2.

Chini, M., Pierdicca, N., Emery, W.J. Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake, IEEE Transactions on Geoscience and Remote Sensing ,47 (1), 145–152, 2009a.

Chini, M., Albano, M., Saroli, M., Pulvirenti, L., Moro, M., Bignami, C., Stramondo, S. Coseismic liquefaction phenomenon analysis by COSMO-SkyMed: 2012 Emilia (Italy) earthquake. International Journal of Applied Earth Observation and Geoinformation, 39, 65-78, 2015a.

Chini, M., Anniballe, R., Bignami, C., Pierdicca, N., Mori, S., Stramondo, S. Identification of building double-bounces feature in Very High Resolution SAR data for earthquake damage mapping, Proc. Of IEEE Geoscience and Remote Sensing Symposium (IGARSS) 2015, pp.1144-1147, 2015b.

Chini, M., Chiancone, A., Stramondo, S. Scale Object Selection (SOS) through a hierarchical segmentation by a multi-spectral per-pixel classification, Pattern Recognition Letters 49, 214–223, 2014b.

Chini, M., Hostache, R., Giustarini, L., Matgen P. A Hierarchical Split-Based Approach (HSBA) for automatically mapping changes using SAR images of variable size and resolution: flood inundation as a test case. IEEE Transactions on Geoscience and Remote Sensing (Submitted).

Chini, M., Pacifici, F., Emery, W. J. Morphological operators applied to X-band SAR for urban land use classification. Proc. of IEEE Geoscience and Remote Sensing Symposium (IGARSS), IV, pp. 506−509, 2009b.

Chini, M., Piscini, A., Cinti, F.R., Amici, S., Nappi, R., De Martini, P.M. The 2011 Tohoku (Japan) Tsunami inundation and liquefaction investigated through optical, thermal and SAR data. IEEE Transactions on Geoscience and Remote Sensing, 10, 347–351, 2013.

Chiu, S. T. A comparative review of bandwidth selection for kernel density estimation. Statistica Sinica, 6(1), 129-145, 1996.

Congalton, R.G., Green, K., Assessing the Accuracy of Remotely Sensed Data, Boca Raton, FL: Lewis, 1999.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E. R. Digital change detection methods in ecosystem monitoring: a review. International Journal of Remote Sensing, 25(9), 1565-1596, 2004.

Cortes, C., Vapnik, V. Support-vector networks. Machine learning, 20(3), 273-297, 1995.

Cossu, R. Segmentation by means of textural analysis. Pixel, 1, 21−24, 1988.

Cova, T. J. GIS in emergency management. Geographical information systems, 2, 845-858, 1999.

Covello, F., Battazza, F., Coletta, A., Lopinto, E., Fiorentino, C., Pietranera, L., Zoffoli,

S. COSMO-SkyMed an existing opportunity for observing the Earth. Journal of Geodynamics, 49(3), 171-180, 2010.

CRED (Centre for Research on the Epidemiology of Disasters). 2015 Disasters in Numbers, 2016. Available on line: www.emdat.be/publications.

Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., Crippa, B. Persistent scatterer interferometry: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 115, 78-89, 2016.

Deichmann., U., Ehrlich, D., Small, C., Zeug, G. Using high resolution satellite data for the identification of urban natural disaster risk. Global Facility for Disaster Reduction and Recovery: Washington, DC, USA, 2011.

Dekkers, A., Aarts, E. Global optimization and simulated annealing. Mathematical programming, 50.1-3, 367-393, 1991.

Devanthéry N., Luzi G., Stramondo S., Bignami C., Pierdicca N. , Wegmüller U., Romaniello V., Anniballe R., Piscini A., Albano M., Moro M., Crosetto M. The validation activities of the Aphorism EC 7FP project, aimed at post seismic damage mapping, through a combined use of EO and ground data. Proc. Of ESA Living Planet Symposium 2016, Prague, May 9-13, 2016.

DigitalGlobe Inc., Core Imagery Product Guide V2.0, December 2014, available on line: http://global.digitalglobe.com/resources/technical-information.

Dorigo, W. A., Scipal, K., Parinussa, R.M., Liu, Y.Y., Wagner, W., de Jeu, R. A. M., Naeimi, V. Error characterization of global active and passive microwave soil moisture datasets, Hydrol. Earth Syst. Sci, vol. 14, pp. 2605-2616, 2010.

Ehrlich, D., Guo, H. D., Molch, K., Ma, J. W., Pesaresi, M. Identifying damage caused by the 2008 Wenchuan earthquake from VHR remote sensing data. International Journal of Digital Earth, 2(4), 309-326, 2009.

Erten, E., Reigber, A., Ferro-Famil, L., Hellwich, O. A new coherent similarity measure for temporal multichannel scene characterization. IEEE Transactions on Geoscience and

Remote Sensing, 50(7), 2839-2851, 2012.

Ferro, A., Brunner, D., Bruzzone, L. Automatic detection and reconstruction of building radar footprints from single VHR SAR images. IEEE Transactions on Geoscience and Remote Sensing, 51(2), 935-952, 2013

Ferro, A., Brunner, D., Bruzzone, L., Lemoine, G. On the relationship between double bounce and the orientation of buildings in VHR SAR images. IEEE Geoscience and Remote Sensing Letters, 8(4), 612-616, 2011.

Flusser, J., Suk, T., Zitová, B. 2D and 3D Image Analysis by Moments, John Wiley & Sons, 2016.

Foody, G. M., Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sensing of Environment, Vol. 114, Issue 10, pp. 2271–2285, October 2010.

Foody, G.M. Status of land cover classification accuracy assessment. Remote Sensing of Environment, Vol. 80, Issue 1, pp. 185–201, April 2002.

Franceschetti, G., Iodice, A., Riccio, D. A canonical problem in electromagnetic backscattering from buildings. IEEE Transactions on Geoscience and Remote Sensing, 40, 1787–1801, 2002.

Franceschetti, G., Iodice, A., Riccio, D., Ruello, G. SAR raw signal simulation for urban structures. IEEE Transactions on Geoscience and Remote Sensing, 41, 1986–1995, 2003.

Geiß, C., Jilge, M., Lakes, T., Taubenböck, H. Estimation of Seismic Vulnerability Levels of urban structures with multisensor Remote Sensing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(5), 1913-1936, 2016.

GFDRR (Global Facility for Disaster Reduction and Recovery). The making of a riskier future: How our decisions are shaping future disaster risk. World Bank: Washington DC, 2016. Available on line: https://www.gfdrr.org/sites/default/files/publication/Riskier%20Future.pdf

Ghosh, A. K., Chaudhuri, P., Sengupta, D. Classification Using Kernel Density Estimates. Technometrics, 48(1), 2006.

Ghosh, A. K., Chaudhuri.P. Optimal smoothing in kernel discriminant analysis. Statistica Sinica, 14(2), 457-484, 2004.

Giovinazzi, S. The vulnerability assessment and the damage scenario in seismic risk analysis. Ph.D., Faculty of Engineering Department of Civil Engineering of the University of Florence, 2005.

Goffe, W. L., Ferrier, G.D., Rogers, J. Global optimization of statistical functions with simulated annealing. Journal of Econometrics, 60(1), 65-99, 1994.

Gruber, A., Su, C.-H., Crow, W. T., Zwieback, S., Dorigo, W. A., Wagner, W. Estimating error cross-correlations in soil moisture data sets using extended collocation analysis, Journal of Geophysical Research: Atmospheres, 121(3), 1208–1219, 2016.

Grünthal G. (ed.). European Macroseismic Scale 1998 (EMS-98), Cahiers du Centre Européen de Géodynamique et de Séismologie, Vol. 15, 99 pp., 1998.

Guida, R., Iodice, A., Riccio, D. Height retrieval of isolated buildings from single High-Resolution SAR Images. IEEE Transactions on Geoscience and Remote Sensing, 48, 2967-2979, 2010.

Gusella, L., Adams, B. J., Bitelli, G., Huyck, C. K., & Mognol, A. Object-oriented image understanding and post-earthquake damage assessment for the 2003 Bam, Iran, Earthquake. Earthquake Spectra, 21(S1), 225–238, 2005.

Guyon, I., Elisseeff, A. An introduction to variable and feature selection. Journal of machine learning research, 3(Mar), 1157-1182, 2003.

Haralick, R. M., Shanmugam, K., Dinstein, I. H. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics, 6, 610-621, 1973.

Haralick, R. M., Shapiro, L. G. Image segmentation techniques. Computer vision, graphics, and image processing, 29(1), 100-132, 1985.

Harb, M. M., De Vecchi, D., Dell'Acqua, F. Physical Vulnerability Proxies from Remotes Sensing: Reviewing, Implementing and Disseminating Selected Techniques. IEEE Geoscience and Remote Sensing Magazine, 3(1), 20-33, 2015.

Herold, M., Gardner, M. E., Roberts, D. A. Spectral resolution requirements for mapping urban areas. IEEE Transactions on Geoscience and Remote Sensing, 41(9), 1907-1919, 2003.

Herold, M., Roberts, D. A., Gardner, M. E., Dennison, P. E. Spectrometry for urban area remote sensing—Development and analysis of a spectral library from 350 to 2400 nm. Remote Sensing of Environment, 91(3), 304-319, 2004.

Hoffman, J., Roth, A., Voigt, S. Interferometric displacement and damage assessment for the 2003 Bam earthquake. In Proceedings of the 2004 Envisat and ERS Symposium, Salzburg, Austria, 6–10 September 2004, ESA SP-572, 2004.

Hsu, C. W., Chang, C. C., Lin, C. J. A practical guide to support vector classification. Available online: https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf, 2010

Inglada, J., Mercier, G. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. IEEE Transactions on Geoscience and Remote Sensing, 45(5), 1432-1445, 2007.

IPCC (Intergovernmental Panel on Climate Change), Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the IPCC. [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley, (Eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp., 2012.

Ito, Y., Hosokawa, M., Lee, H. Liu, J.G. Extraction of damaged regions using SAR data and neural networks. In Proceedings of the XIX ISPRS Congress, 16–22 July, Amsterdam, Vol. XXXIII, 2000.

Joyce, K. E., Wright, K. C., Samsonov, S. V., Ambrosia, V. G. Remote sensing and the

disaster management cycle. Advances in Geoscience and Remote Sensing, Gary Jedlovec (Ed.), ISBN: 978- 953-307-005-6, InTech, 2009.

Kramer, S.L. Geotechnical Earthquake Engineering, Prentice Hall, Inc., Upper Saddle River, New Jersey, 653 pp. 1996.

Kullback, S., Leibler, R. A. On information and sufficiency. The annals of mathematical statistics, 79-86, 1951.

Li, X., Guo, H., Zhang, L., Chen, X., Liang, L. A new approach to collapsed building extraction using RADARSAT-2 polarimetric SAR imagery. Geoscience and Remote Sensing Letters, IEEE, 9(4), 677-681, 2012.

Lindell, M. K. Recovery and reconstruction after disaster. In Encyclopedia of natural hazards (pp. 812-824). Springer Netherlands, 2013.

Lindell, M. K., Perry, R. W., Prater, C., Nicholson, W. C. Fundamentals of emergency management. Washington, DC: FEMA, 2006.

Marceau, D.J. Hay, G.J. Remote sensing contributions to the scale issue, Can. J. Remote Sens. 25 (4), 357–366, 1999.

Marin, C., Bovolo, F., Bruzzone, L. Building change detection in multitemporal Very High Resolution SAR images. IEEE Transactions on Geoscience and Remote Sensing, 53(5), 2664-2682, 2015.

Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters, Journal of Society for Industrial and Applied Mathematics, Vol. 11, No. 2, pp. 431-441, 1963.

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., Rabaute, T. The displacement field of the Landers earthquake mapped by radar interferometry. Nature, 364(6433), 138-142, 1993

Mathieu, R., Aryal, J. Object-based classification of Ikonos imagery for mapping large-scale vegetation communities in urban areas. Sensors, 7(11), 2860-2880, 2007.

Matsuoka, M., Yamazaki, F. Application of the damage detection method using SAR intensity images to recent earthquakes. Proc. of IEEE Geoscience and Remote Sensing Symposium (IGARSS), 4, pp. 2042–2044, 2002.

Matsuoka, M., Yamazaki, F. Application of the Damage Detection Method Using SAR Intensity Images to Recent Earthquakes. In Proceedings of the IEEE IGARSS, Toronto, ON, Canada, 24–28, pp. 2042–2044, June 2002.Grünthal G. (ed.), European Macroseismic Scale 1998 (EMS-98), Cahiers du Centre Européen de Géodynamique et de Séismologie", Vol. 15, 99 pp., 1998.

Matsuoka, M., Yamazaki, F. Characteristics of satellite images of damaged areas due to the 1995 Kobe earthquake. In Proceedings of the 2nd Conference on the Applications of Remote Sensing and GIS for Disaster Management, Honolulu, HI, USA, 24–28 July 1999, The George Washington University: Washington, DC, USA, 1999.

Matsuoka, M., Yamazaki, F. Characteristics of Satellite SAR Images in the Areas Damaged by Earthquakes. In Proc. of IEEE Geoscience and Remote Sensing Symposium (IGARSS), 24–28, pp. 2693–2696, 2000.

Matsuoka, M., Yamazaki, F. Use of satellite SAR intensity imagery for detecting building areas damaged due to earthquakes. Earthquake Spectra, 20, pp. 975–994, 2004.

Melgani, F., Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 42(8), 1778-1790, 2004.

Mercer, J. Functions of positive and negative type, and their connection with the theory of integral equations. Philosophical transactions of the royal society of London. Series A, containing papers of a mathematical or physical character, 415-446, 1909.

Mück, M., Taubenböck, H., Post, J., Wegscheider, S., Strunz, G., Sumaryono, S., Ismail, F. A. Assessing building vulnerability to earthquake and tsunami hazard using remotely sensed data. Natural hazards, 68(1), 97-114, 2013

Mueller, M., Segl, K., Heiden, U., Kaufmann, H., Potential of high-resolution satellite data in the context of vulnerability of buildings. Natural hazards, 38(1-2), 247-258, 2006.

Pacifici, F., Chini, M., Emery, W. J. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. Remote Sensing of Environment, 113(6), 1276-1292, 2009.

Pan, G., Tang, D. Damage information derived from multi-sensor data of the Wenchuan Earthquake of May 2008. International Journal of Remote Sensing, 31, 3509–3519, 2010.

Parzen, E. On estimation of a probability density function and mode. The annals of mathematical statistics, 1065-1076, 1962.

Pepe, M. S., Janes, H. Insights into latent class analysis of diagnostic test performance. Biostatistics, 8(2), 474-484, 2007

Pesaresi, M., Benediktsson, J.A. A new approach for the morphological segmentation of high resolution satellite imagery. IEEE Transactions on Geoscience and Remote Sensing, 39 (2), 309–320, 2001.

Pierdicca, N., Fascetti, F., Pulvirenti, L., Crapolicchio, R., Muñoz-Sabater, J. Analysis of Ascat, SMOS, In-situ and Land Model soil moisture as a regionalized variable over Europe and North Africa, Remote Sensing of Environment, 170, 28, 2015a.

Pierdicca, N., Fascetti, F., Pulvirenti, L., Crapolicchio, R., Muñoz-Sabater, J. Quadruple Collocation Analysis for Soil Moisture Product Assessment, IEEE Geoscience and Remote Sensing Letters, Vol 12, pp. 1595-1599, 2015b.

Pierdicca, N., Chini, M., Pelliccia, F. The contribution of SIASGE radar data integrated with optical images to support thematic mapping at regional scale. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(7), 2821-2833, 2014.

Pittore, M., Wieland, M. Toward a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing. Natural hazards, 68(1), 115-145, 2013.

Powers, D.M.W. Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. Journal of Machine Learning Technologies, 2(1), 37-63, 2011.

Pulvirenti, L., Chini, M., Pierdicca, N., Guerriero, L., Ferrazzoli, P. Flood monitoring using multi-temporal COSMO-SkyMed data: image segmentation and signature interpretation. Remote Sensing of Environment, 115 (4), 990–1002, 2011.

Pulvirenti, L., Marzano, F. S., Pierdicca, N., Mori, S., Chini, M. Discrimination of water surfaces, heavy rainfall, and wet snow using COSMO-SkyMed observations of severe weather events. IEEE Transactions on Geoscience and Remote Sensing, *52*(2), 858-869, 2014.

Rastiveis, H., Samadzadegan, F., Reinartz, P. A fuzzy decision making system for building damage map creation using high resolution satellite imagery. Natural Hazards and Earth System Sciences, 13.2, 455, 2013.

Richards, J. A., Jia, X., Remote sensing digital image analysis: An introduction, Heidelberg, Germany: Springer-Verlag, 2006.

Roberts, D. A., Herold, M. Imaging spectrometry of urban materials. Infrared Spectroscopy in Geochemistry, Exploration and Remote Sensing, Mineral Association of Canada, Short Course Series, 33, 155-181, 2004.

Rutjes A.W.S., Reitsma J.B., Coomarasamy A., Khan K.S. Bossuyt P.M.M. Evaluation of diagnostic tests when there is no gold standard. A review of methods, Health Technology Assessment, Vol. II, No. 50, 2007.

Salvi, S., Stramondo, S., Funning, G. J., Ferretti, A., Sarti, F., Mouratidis, A. The Sentinel-1 mission for the improvement of the scientific understanding and the operational monitoring of the seismic cycle. Remote sensing of environment, 120, 164-174, 2012.

Serpico, S. B., Dellepiane, S., Boni, G., Moser, G., Angiati, E., Rudari, R. Information extraction from remote sensing images for flood monitoring and damage evaluation. Proceedings of the IEEE, 100(10), 2946-2970, 2012.

Silverman, B. W. Density estimation for statistics and data analysis (Vol. 26). CRC press, 1986.

Small, C. High spatial resolution spectral mixture analysis of urban reflectance. Remote Sensing of Environment, 88(1), 170-186, 2003.

Smith, A. R. Color gamut transform pairs. ACM Siggraph Computer Graphics. Vol. 12. No. 3. ACM, 1978.

Soille, P. Morphological Image Analysis, Principles and Applications, second ed., Springer-Verlag, Berlin, Germany, 2003.

Srinivasan, G. N., Shobha, G., Statistical texture analysis. In Proceedings of world academy of science, engineering and technology, Vol. 36, pp. 1264-1269), 2008.

Stramondo, S. Bignami, C., Chini, M., Pierdicca, N., Tertulliani, A. Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies. International Journal of Remote Sensing, 27, 4433 - 4447, 2006.

Tabak, E. G., Turner, C. V. A family of nonparametric density estimation algorithms. Communications on Pure and Applied Mathematics, 66(2), 145-164, 2013.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S. Monitoring urbanization in mega cities from space. Remote sensing of Environment, 117, 162-176, 2012.

Taubenböck, H., Post, J., Roth, A., Zosseder, K., Strunz, G., Dech, S. A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing. Natural Hazards and Earth System Sciences, 8(3), 409-420, 2008.

Tertulliani A., Leschiutta I., Bordoni P., Milana G. Damage distribution in L'Aquila city (central Italy) during the 6 April 2009 earthquake. Bulletin of the Seismological Society of America 102, 1543–1553. 2012.

Tertulliani, A., Arcoraci, L., Berardi, M., Bernardini, F., Camassi, R., Castellano, C., Del Mese, S., Ercolani, E., Graziani, L., Leschiutta, I., Rossi, A., Vecchi, M. An application of EMS-98 in a medium-sized city: The case of L'Aquila (Central Italy) after the April 6, 2009 Mw 6.3 earthquake. Bulletin of Earthquake Engineering, 9, 67–80, 2011.

Tolomei, C., Salvi, S., Boncori, J. M., Pezzo, G. InSAR measurement of crustal deformation transients during the earthquake preparation processes: a review. Bollettino di Geofisica Teorica ed Applicata, 56(2), 151-166, 2015.

Tuceryan, M., Jain, A. K. Texture analysis. Handbook of pattern recognition and computer vision, 2, 207-248, 1993.

Turker, M., Sumer, E., Building-based damage detection due to earthquake using the watershed segmentation of the post-event aerial images. International Journal of Remote Sensing, 29 (11), 3073-3089, 2008.

UN (United Nations), World Population Prospects: The 2015 Revision, Key Findings and Advance Tables, Working Paper No. ESA/P/WP.241, UN Department of Economic and Social Affairs, Population Division, New York, 2015. Available on line: http://esa.un.org/unpd/wpp/publications/files/key_findings_wpp_2015.pdf

UNISDR (United Nations International Strategy for Disaster Reduction). Living with risk: a global review of disaster reduction initiatives, Geneva, Switzerland, UN Publications, 2004.

UNISDR (United Nations International Strategy for Disaster Reduction), Terminology on Disaster Risk Reduction, Geneva, Switzerland, 2009, Available on line: http://www.unisdr.org/files/7817_UNISDRTerminologyEnglish.pdf

UNISDR (United Nations International Strategy for Disaster Reduction), Launching UNISDR science and technology partnership and the science and technology road map to 2030. Short concept note: Work stream 2, working Group 2, 2016. Available on line: www.preventionweb.net/files/45270_unisdrcnws2wg2vulnerabilityandexpos.pdf.

Van Westen, C. J. Remote sensing and GIS for natural hazards assessment and disaster risk management. Treatise on Geomorphology, edited by: Shroder, J. and Bishop, MP, Academic Press, San Diego, CA, 3, 259-298, 2013.

Vapnik, V. N. An overview of statistical learning theory. IEEE Transactions on Neural Networks, 10(5), 988-999, 1999.

Veropoulos, K., Campbell, C., Cristianini, N. Controlling the sensitivity of support vector machines. In Proceedings of the International Joint Conference on Artificial Intelligence, pp. 55-60, 1999.

Waske, B., Benediktsson, J. A. Fusion of support vector machines for classification of multisensor data. IEEE Transactions on Geoscience and Remote Sensing, 45(12), 3858-3866, 2007.

Werner, C., Wegmuller, U., Strozzi, T., Wiesmann, A. Interferometric point target analysis for deformation mapping. In Proc. Of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 7, pp. 4362-4364, 2003.

Yager, R. R. An extension of the naive Bayesian classifier. Information Sciences, 176(5), 577-588, 2006.

Yang, X., Lo, C. P. Relative radiometric normalization performance for change detection from multi-date satellite images. Photogrammetric Engineering and Remote Sensing, 66(8), 967-980, 2000.

Yonezawa, C., Takeuchi, S. Decorrelation of SAR data by urban damage caused by the 1995 Hoyogoken-Nanbu earthquake. International Journal of Remote Sensing, 22, 1585–1600, 2001.

Zhang, X., King, M. L., Hyndman, R. J. A Bayesian approach to bandwidth selection for multivariate kernel density estimation. Computational Statistics & Data Analysis, 50(11), 3009-3031, 2006.

Zhao L, Yang J, Li P, Zhang L, Shi L, Lang F. Damage assessment in urban areas using post-earthquake airborne PolSAR imagery. International Journal of Remote Sensing, 34(24), 8952–8966, 2003.

List of Publications

Journals

1) Pierdicca, N., Anniballe, R., Noto, F., Bignami, C., Chini, M., Martinelli, A., Mannella A. Triple collocation to assess classification accuracy without a ground truth in case of earthquake damage assessment. IEEE Transaction on Geoscience and Remote Sensing, 2016. (under review)

2) Romaniello, V., Piscini, A., Bignami, C., Anniballe, R., Stramondo S. A multisensor approach for the 2016 Amatrice earthquake damage assessment. Annals of Geophysics, 59, 2016.

3) Bonafoni, S., Anniballe, R., Gioli, B., Toscano, P. Downscaling Landsat Land Surface Temperature over the urban area of Florence. European Journal of Remote Sensing, 49, 553-569, 2016. DOI: 10.5721/EuJRS20164929

4) Anniballe, R., Bonafoni, S. A stable Gaussian fitting procedure for the parameterization of remote sensed thermal image. Algorithms, 8, pp. 82-91, 2015. DOI: 10.3390/a8020082, ISSN: 1999-4893.

5) Baldinelli, G., Bonafoni, S., Anniballe, R., Presciutti, A., Gioli, B., Magliulo, V. Spaceborne detection of roof and impervious surface albedo: potentialities and comparison with airborne thermography measurements. Solar Energy, 113, pp. 281-294, 2015. DOI:10.1016/j.solener.2015.01.011, ISSN: 0038-092X.

6) Anniballe, R., Bonafoni, S., Pichierri, M. Spatial and temporal trends of the surface and air heat island over Milan using MODIS data. Remote Sensing of Environment, 150, pp. 163-170, 2014. DOI:10.1016/j.rse.2014.05.005, ISSN: 0034-4257; 1879-0704.

Proceedings

7) Anniballe, R., Chini, M., Pierdicca, N. Earthquake damage detection in urban areas from VHR SAR imagery. Atti della XXI Riunione Nazionale di Elettromagnetismo, 2016.

8) Devanthéry, N., Luzi, G., Stramondo, S., Bignami, C., Pierdicca, N., Wegmüller, U., Romaniello, V., Anniballe, R., Piscini, A., Albano, M., Moro, M., Crosetto, M. The validation activities of the Aphorism EC 7FP project, aimed at post seismic damage mapping, through a combined use of EO and ground data. Proc. Of ESA Living Planet Symposium 2016, Prague, May 9-13, 2016.

9) Anniballe, R., Chini, M., Pierdicca, N., Bignami, C., Stramondo, S., Noto, F., Scalia T., Martinelli, A., Mannella, A. Detecting earthquake damage in urban area: application to Cosmo-SkyMed imagery of L'Aquila earthquake. Proc. SPIE 9642, SAR Image Analysis, Modelling, and Techniques XV, 96420C, 2015. DOI:10.1117/12.2195723.

10) Chini, M., Anniballe, R., Bignami, C., Pierdicca, N., Mori, S., Stramondo, S., Identification of building double-bounces feature in Very High Resolution SAR data for earthquake damage mapping. In Proc. of IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2015, pp.1144-1147, 2015. DOI: 10.1109/IGARSS.2015.7326376.

11) Anniballe, R., Casa, R., Castaldi, F., Fascetti, F., Fusilli, L., Huang, W., Laneve, G., Marzialetti, P., Palombo, A., Pascucci, S., Pierdicca, N., Pignatti, S., Qyuniao, X., Santini, F., Silvestro, P.C., Yang, H., Yang, G. Synergistic use of radar and optical data for agricultural data products assimilation: a case study in central Italy. In Proc. of IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2015, pp.3381-3384, 2015. DOI: 10.1109/IGARSS.2015.7326544.

12) Bonafoni, S., Anniballe, R., Pierdicca, N. Downscaling of the land surface temperature over urban area using Landsat data. In Proc. of IEEE International

Geoscience and Remote Sensing Symposium (IGARSS) 2015, pp.1144-1147. DOI: 10.1109/IGARSS.2015.7325973

13) Bonafoni, S., Anniballe, R., Pichierri, M. Comparison between surface and canopy layer urban heat island using MODIS data. In Proc. of IEEE Joint Urban Remote Sensing Event (JURSE) 2015, pp.1-4. DOI: 10.1109/JURSE.2015.7120457.

14) Anniballe, R., Bignami, C., Chini, M., Pierdicca, N., Stramondo, S. Detecting sparse earthquake damages in high density urban settlements by VHR SAR data. Proceeding of SPIE, Vol. 9243, SAR Image Analysis, Modelling, and Techniques XIV, 924307, 2014. DOI:10.1117/12.2068738.

15) Anniballe, R., Bonafoni, S., Pierdicca, N. Surface and air heat island of Milan: spatial and temporal analysis from MODIS. Atti della XX Riunione Nazionale di Elettromagnetismo, 2014. ISBN: 978-88-907599-4-9.

Other contributions

16) Anniballe, R., Pierdicca, N., Noto, F., Bignami, C., Scalia, T. Rilevazione dei danni da terremoto in area urbana da immagini ottiche satellitari ad altissima risoluzione. VIII Convegno Nazionale dell'Associazione Italiana di Telerilevamento (AIT), Palermo, 22-24 June 2016. (Abstract and oral presentation).

17) Anniballe, R., Bignami, C., Chini, M., Martinelli, A., Mannella, A., Noto, F., Pierdicca, N., Scalia T., Stramondo, S. A data integration approach for earthquake damage assessment using VHR SAR and optical imagery. Living Planet Symposium 2016, Prague, 9-13 May 2016. (Abstract and poster presentation).

18) Anniballe, R., Chini, M., Pierdicca, N., Bignami, C., Stramondo, S., Noto, F., Scalia T., Martinelli, A., Mannella, A. An object oriented approach to detect earthquake damage in urban area from VHR optical imagery. Mapping Urban Areas from Space 2015 Conference, ESRIN, Frascati (Rome) Italy, 4-5 November 2015. (Abstract and poster presentation).

19) Bonafoni, S., Anniballe, R., Pierdicca, N. Mapping the Land Surface Temperature over Urban Areas from Space: a Downscaling Approach. Mapping Urban Areas from Space 2015 Conference, ESRIN, Frascati (Rome) Italy, 4-5 November 2015. (Abstract and poster presentation).

20) Bignami, C., Di Paolo, M., Pierdicca, N., Chini, M., Stramondo, S., Anniballe, R. VHR SAR images for detecting sparse damaged buildings in urban areas: the L'Aquila earthquake case study, IEEE International Geoscience and Remote Sensing Symposium, Québec, Canada, July 13-18, 2014. (Abstract and poster presentation).