MECHANICAL AND AEROSPACE ENGINEERING DEPARTMENT

DOCTORATE OF PHYLOSOPHY

in AERONAUTICAL AND SPACE TECHNOLOGY

Earthquake Damage Analysis and Mapping

with the Use of

Satellite Remote Sensing

Ph.D. CANDIDATE: ADVISOR:

Tanya Scalia   Prof. Nazzareno PIERDICCA

CO-ADVISOR:

Ing. Fabrizio NOTO

Academic Year 2015-2016

XXVIII Ciclo

This page is intentionally left blank

To Ginevra

Acknowledgments

It is always very difficult for me to start such a page… so many people I'd like to thank, hopefully without forgetting anybody.

First of all, I would like to take this opportunity to thank my excellent PhD Advisor Prof. Nazzareno Pierdicca who patiently and outstandingly supported the work of such an "old" student who, having being away from real books for too long, really needed to be brought back to an academic framework. His guidance helped me throughout the research and writing of this thesis.

I also would like to thank Prof. Bernardo Favini for his foresight in understanding the direction I should take, long before I did myself. Had I not followed his suggestion I would not have tried to move into the "new" world of Earth Observation, which instead has greatly inspired me.

I wish to give special thanks to Ing. Fabrizio Noto for giving me the possibility of working with such a distinguished expert. This represented a valuable opportunity to open my mind to new perspectives and approaches which I consider an invaluable gift for my professional growth.

I would also like to thank Roberta Anniballe who has shared with me the same doubts and issues and supported me with precious information throughout the whole project.

There's a long list of people I would also like to thank because they had to cope with my anxiety - my close family (Gabriele, Vito and Mum), my extended family (Chiara, Michela, Poletta, Mirko), together with my wonderful colleagues and ex-colleague (Mauro) for the possibility of sharing ideas and insights.

A special thanks to my dearest friends who, I know, are always there!!! (Susanna, Rosaria, Gabriella, Eleonora, Pera, etc…) and the people who operatively support my daily life (Anna, Adriana).

Last but not least, I wish to send a very special thought to Prof. Maurizio Di Giacinto who, in a far from easy period of his life, readily supported me with the enthusiasm that I will always carry in my heart. I will keep his example of approach to life as a model to pursue.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

ACC	Accuracy
AeDES	Agibilità per edifici ordinari nell'emergenza post-sismica
ANN	Artificial Neural Network
APhoRISM	Advanced Procedures for Volcanic and Seismic Monitoring
APE	A Priori information for Earthquake damage mapping
BN	Bayesian Network
BT	Beni Tutelati
CDF	Cumulative Distribution Function
CEOS	Committee on Earth Observation Satellites
CN	Condition Negative
CNR	Consiglio Nazionale delle Ricerche
CP	Condition Positive
CPD	Conditional Probability Distribution
CPT	Conditional Probability Table
CTR	Carta Tecnica Regionale
DAG	Directed Acyclic Graph
DBELA	Displacement-Based Earthquake Loss Assessment
DBO-LO	Double Bounce LayOver
DEM	Digital Elevation Model
DoD	Department of Defence
DPC	Department of Civil Protection
DPM	Damage Probability Matrix
DRM	Disaster Risk Management
DRR	Disaster Risk Reduction
DSM	Digital Surface Model
EEAS	European External Action Service
EGNOS	European Geostationary Navigation Overlay System
EM	Expectation-Maximization
EMS	European Macroseismic Scale
EO	Earth Observation
EQuOS	EarthQUake Observation System
ERD	Earthquake Resistant Design
ESA	European Space Agency
EU	European Union
EVOSS	European Volcano Observatory Space Services
FAST	Foreshore Assessment using Space Technology
FEMA	Federal Emergency Management Agency
FN	False Negative
FP	False Positive
FP7	Seventh Framework Programme

GEO	Geostationary Earth Orbit
GIS	Geographic Information System
GMES	Global Monitoring for Environment and Security
GNDT	Gruppo Nazionale per la Difesa dai Terremoti
GEOSS	Global Earth Observation System of Systems
GM	Graphical Model
GPS	Global Positioning System
HCI	Human Computer Interface
HR	High Resolution
IASPEI	International Association of Seismology and Physics of the Earth's Interior
IGW	Internal Gravity Waves
INGV	National Geophysical and Volcanic Institute
InSAR	Interferometric Synthetic Aperture Radar
JDL	Joint Directors of Laboratories
JPDA	Joint Probabilistic Data Association
KL	Kullback-Leibler
LAMPRE	Landslide Modelling and Tools for Vulnerability Assessment Preparedness and Recovery Management
LEO	Low Earth Orbit
LF	Low Frequency
LiDAR	Light Detection And Ranging
LPI	Liquefaction Potential Index
MI	Mutual Information
MOVE	Methods for the Improvement of Vulnerability Assessment in Europe
MSK	Medvedev-Sponheuer-Karnik Scale
NaN	Not a Number
OPT	Optical
PAN	Panchromatic
PDA	Probabilistic Data Association
PDF	Probability Density Function
PGA	Peak Ground Acceleration
PGD	Peak Ground Displacement
PGV	Peak Ground Velocity
PPV	Positive Predictive Value
PSH	Pan Sharpened
PSI	Persistent Scatterers Interferometry
RASOR	Rapid Analysis and Spatialisation Of Risk
RC	Reinforced Concrete
RGR	Rapid Geospatial Reporting
ROI	Region Of Interest
RPC	Rational Polynomial Coefficient
SAFER	Services and Applications For Emergency Response
SAR	Synthetic Aperture Radar

SAW	Situation Awareness
SDOF	Single Degree of Freedom
SAE	Support to External Action
SEMEP	Search for electro-magnetic earthquake precursors combining satellite and ground-based facilities
SIGE	Sistema Informativo per la Gestione dell'Emergenza
SM	Seismic Microzonation
SP-BELA	Simplified Pushover-Based Earthquake Loss Assessment
SRTM	Shuttle Radar Topography Mission
STRUCT	Structural
TN	True Negative
TON	Test Outcome Negative
TOP	Test Outcome Positive
TP	True Positive
TPR	True Positive Rate
UNISDR	United Nations Office for Disaster Risk Reduction
USGS	United States Geological Survey
VHR	Very High Resolution
VLF	Very Low Frequency
VULSAR	Sistema per la generazione di mappe di VULnerabilità sismica mediante dati SAR

1   INTRODUCTION

Following the United Nations Resolution [1] deep concern is expressed at International level towards "... *the recent increase in the frequency and intensity of extreme weather events and associated natural disasters in some regions of the world and their substantial economic, social and environmental impacts, in particular upon developing countries...".*

Since 1989, the United Nations, following the adoption of Resolution 44/236 (22 December 1989) has launched the International Decade for Natural Disaster Reduction, with the objective of reducing loss of life, property damage and social and economic disruption caused by natural disasters, especially in developing countries.

The UNISDR resolution 63/217 clearly emphasizes the relevance of a concerted International action "...*addressing vulnerability and integrating risk reduction into all phases of disaster management, post-disaster recovery and development planning*".

Coherently with this broader perspective aiming at a higher reduction of vulnerability to all natural hazards and associated natural disasters, a closer and more systematic International cooperation, and information-sharing on disaster preparedness among the scientific and academic communities and disaster managers at all levels, should be strengthened.

Many initiatives have been implemented in the last years towards this direction. Within this framework, more than 40 nations [3] invested in EO satellites, "*…amounting to government investment of approximately 7-8 billion US$ per annum, with further and increasing investment coming from the commercial sector and through public-private partnerships*".

Many of the biggest programmes, such as Copernicus [4], are moving to fully free and open data policies, in order to apply EO generated data across a whole range of societal benefit areas, without restrictions. In fact, satellite observations can supply regular, detailed updates on the status of hazards on a global, regional, or national basis.

Summarising the outcomes of the CEOS Handbook on DRM [3], we can state that the contribution of remote sensing to the hazard and exposure components of Disaster Risk Reduction (DRR) is primarily related to the following functions in particular:

- Exposure mapping services to support preparedness/mitigation, early warning & response of all the humanitarian aids players;

- Hazard mapping and risk assessment (geo-hazards, hydro-meteorological hazards, climatological hazards, technological hazards). For example, in our work we focus on supporting earthquake risk managers. Different types of information can be extracted, such as the classified distribution of the land cover and socio-economic units in areas at risk or hazard damage information based on previous events and on the geological data of the area. Again, the satellite archives can provide also historical data showing damage extents for major events over a given area;

- Critical infrastructure monitoring: Up-to-date, synoptic, and objective infrastructure information helps maintaining and monitoring the status of the relevant assets at risk. Such data can be used to provide improved knowledge of the potential impact of natural hazards in areas at risk. For example, geophysical monitoring of subsidence and structural stability can be monitored by satellite radar images that can support the

measurement of sub-centimetre-scale changes in deformation over spans of days to years;

- Early warning/alert and tracking of a range of natural hazards, such as earthquakes, tropical cyclones, landslides, floods and volcanoes;

- Disaster response following natural and man-made hazards and support for Crisis Mapping/Damage Assessment: this function also includes rapid assessment and identification of the location, scale and severity of the disaster impact;

- Support to recovery/reconstruction/ rehabilitation: satellite EO supports precise post-disaster needs assessment and evaluation of early and long-term recovery needs and priorities. This function is mostly related to the identification of safe and unsafe areas in post-events and supports the prioritisation of infrastructure repair. EO can support the monitoring of risks from repeated hazards and focus the disaster management players on specific reconstruction needs.

In this work, coherently with the framework described above, we make use of EO data to assess the post-event earthquake damage. This activity was carried out in the wider context of an FP7 funded project, APhoRISM that aims 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. In particular, in this PhD work we focus on the seismic product, which consists of a damage map addressing the detection, analysis and estimate of changes to buildings and infrastructures occurring during an earthquake in the epicentral regions. This damage map is obtained by integrating *a priori* information and change detection images.

The *a-priori* information refers to structural data related to the built environment but may include, as well, information derived by processing pre-crisis interferometric Synthetic Aperture Radar data to measure possible surface movements. Such kind of analysis can provide evidence on the stability of the terrain. An integration algorithm merges the *a-priori* information with the detection of changes observed by the available Earth Observation sensors, (namely SAR and optical sensors) and ground shakemap from seismological data.

The goal of this PhD work is to design, develop and test the just mentioned integration algorithm to provide a Likelihood Index Damage Map, which represents a map of group of buildings or single buildings and infrastructures that are likely collapsed or strongly damaged.

Following the scheme above, we are going to present the results of this work in the next chapters, according to the following structure:

- Chapter 2: In these paragraphs, we are going to examine the post-earthquake damage assessment in urban environment. This chapter aims at clarifying which approaches are currently employed by the experts of post-earthquake assessments and defining the most relevant parameters on which those experts focus on. Therefore, definitions of risk, hazard, exposure and vulnerability in the earthquake damage scenario are reported to allow the reader in better understanding what kind of indicators and parameters we need to manage for the correct development of a detailed integration module.

- Chapter 3: in this chapter, we are going to review the current state-of-the-art in the use of Remote Sensing for damage assessment in general and for post-earthquake

damage assessment, in particular. As already mentioned in the previous paragraphs, many initiatives have been implemented in the last years on this topic. However, the approach adopted in our work, and in the APhoRISM project, brings an innovative perspective since it aims at addressing individual buildings damage by using a highly modular approach where, if only partial data is accessible, the probability of damage at single building level can still be assessed making use of the available data. Also in this chapter, we aim at understanding on what kind of indicators EO data is able to support civilian engineers in assessing post-earthquake buildings damage.

- Chapter 4: after having analysed, in the previous chapters, the different types of information and data that can be provided by structural engineers and remote sensing (from radar and/or optical sensors) experts, we need to understand how to fuse the entire or partial information available to give an assessment on the individual building collapse probability. In order to do this, we first review the general approach and methods to data fusion and we identify what is the level of information that is better to fuse referring to our goals. In this chapter, we also examine how the structural information (mainly coming from in-situ data, i.e. structural features of the building) is evaluated and we then focus on the description of Bayesian approaches and, more specifically, of Bayesian networks. Such type of graphical approach for our data fusion tool is implemented to assess post-earthquake building damage.

- Chapter 5: summarising all the information gathered in the previous chapters, we substantiate here our A Priori information for Earthquake damage mapping using a real test case. Our test case is based on L' Aquila earthquake which reached its peak in the event of April 6th 2009. In this case, we have a set of data available to build the Ground Truth validation test set. For what concerns remote sensing data, for this event, both COSMO-Skymed Radar and Quickbird VHR optical sensors are available thus allowing a complete remote sensing dataset. The in-situ information, though fragmentary, was built using data coming from different sources, mainly from INGV and the Italian Civil Protection Department. The results of different Bayesian networks are presented showing the step-by-step approach adopted which aims at generalising the methodology in order to further implement the network in future cases.

- Chapter 6: this chapter summarises the results of the whole work and establishes some main actions that can be implemented in the next future to further improve the proposed methodology and identify possible generalised approaches that can support a rapid disaster response and contribute to the post-event recovery and reconstruction phases in support to all humanitarian and emergency actors.

2 POST-EARTHQUAKE DAMAGE ASSESSMENT IN URBAN ENVIRONMENT

In the next paragraphs, we are going to examine the approaches, currently employed, to make post-earthquake damage assessments. On this regard, a detailed review of the main definitions and parameters related to earthquake risk is performed. These include namely risk, hazard, exposure and vulnerability in the earthquake damage scenario.

2.1 SEISMIC EVENT SCENARIO OVERVIEW

The use of Remote Sensing [6] within the domain of natural hazards and disasters has become increasingly common, due in part to increased awareness of environmental issues such as climate change, but also to the increase in geospatial technologies and the ability to provide up-to-date imagery to the public through the media and internet. As technology is enhanced, demand and expectations increase for near-real-time monitoring and visual images to be relayed to emergency services and the public in the event of a natural disaster.

For over ten years, also the European Space Agency has been supporting activities to increase the utilisation of satellite based EO in Disaster Risk Management (DRM) through different application development mechanisms. There are two fundamental DRM activities developed by ESA with other space agencies and organisations worldwide. Firstly the International Charter Space & Major Disasters (the Charter [7]), an International collaboration for the immediate disaster response phase, created in 2000, after UNISPACE III, and currently operating globally with 15 members – mainly with EO mission owners or operators - and providing basic operational services to Authorized Users. Secondly, the newly started Committee on Earth Observation Satellites [8] (CEOS) initiative on DRM looking at different phases of application development such as research, demonstrations of service capabilities, and pilot projects with strategic users to achieve acceptance of EO in the operational environment of end users. Under CEOS DRM, ESA has taken a lead role and coordinates the Study Team. ESA is involved in the definition of propositions for four (4) coordinated pilots started in 2014 in the themes of flooding, volcanoes and seismic hazards and a multi-thematic pilot that is a recovery observatory.

Practically, the CEOS DRM comprises a broad portfolio of different activities, some are operational EO capabilities, such as for instance the Charter for disaster response, some are dedicated to science users, such as the Geohazard Supersites and some are in the early stage of the EO application development cycle. The main emphasis of the CEOS DRM is on disaster risk reduction where EO applications are less developed than in other facets of the risk management cycle, such as disaster response [9].

After a seismic event [10], the primary information for a prompt response is related to dimension, extension and localization of the damages. For this reason, some evaluation tools built on the simulations of damage scenarios allow the implementation of an accurate emergency plan almost real-time and before the in-situ ground surveys are carried out. Jointly to this, further activities for the damage assessment also need to be implemented to assess the preliminary results and predictions based on the instrumental data available from the seismic monitoring network.

During the first hours after an earthquake, it is fundamental to know as better and as fast as possible the dimensions of the event and its potential impact on the territory and on the population in order to organise and dimensioning the aids correctly. On this regard, the particularly negative experience gained in Italy after the Irpinia earthquake in 1980 allowed preparing the background for using advanced Geographic Information Systems (GIS). Nowadays the Italian Civil Protection Department uses a system able to generate semi-real time simulation damage scenarios. In the event of a significant earthquake, the Italian Institute of Geophysics and Volcanology (INGV) communicates to the Civil Protection Department the core data (magnitude and coordinates) of the event. An automatic procedure is then activated to generate a report that is expected to be available within 10 minutes [11] after the event, as specified in the so called SIGE [12], [13] - Sistema Informativo per la Gestione dell'Emergenza (Informative System for Emergency Management). Such report contains data, maps and information on all the municipalities included in 100 km from the epicentre and, more specifically:

- Description of the territory (anthropic, physical and administrative aspects; infrastructures and buildings features; seismic monitoring networks);
- Hazardousness (seismic-genesis areas, historical earthquakes, isoseismals and contour plans, mitigation of ground movements);
- Vulnerability (private and public building assets, including schools, hospitals, roads and transport connections);
- Exposure (features and distribution of the population residing in the affected areas);
- Preliminary evaluation of damages and losses (damaged and unfit for use buildings, casualties and injured estimate, economic damage estimate).

After the event, usability surveys are characterised by a very large number of calls for inspections after the seismic event, by the need of providing answers in a short time and, more in general, by the emergency situation. Such inspections need to be carried out very quickly, therefore the usability judgment is only based on expert judgment; it is not definitive, but temporary; it does not have, finally, a precise objective in terms of risk. As reported in [14], usability can be defined as:

"The evaluation of usability in the post earthquake emergency is a temporary and rough evaluation – i.e. based on an expert judgment and carried out in a short time, on the basis of a simple visual inspection and of data which can be easily collected – aiming at determining whether, in case of a seismic event, buildings affected by the earthquake can still be used, with a reasonable level of life safety."

The contribution of data coming from remote sensing is intended to speed up the process of post-event damage assessment and to provide reliable and quick assessment on the individual building damage.

In order to better identify the problem related to vulnerability and damage assessment of the buildings before and after a seismic event, we are going to summarise in the following paragraphs some main concepts related to damage assessment for buildings.

2.2   SEISMIC RISK AND VULNERABILITY

Many definitions can be found in literature for the word "Risk" associated to the different causes. More in general, the UNISDR [15] defines the word risk simply as "*The combination of the probability of an event and its negative consequences*", following closely the ISO/IEC Guide 73. The word "risk" is usually associated to the chance or possibility that an event occurs in the popular use of the term, whereas in technical settings the emphasis is usually placed on the consequences, in terms of "potential losses" for some particular cause, place and period. Also the Italian GNDT [16] used a similar definition referring specifically to the seismic risk as the "*probability of occurrence and level of severity of possible effects due to an earthquake in a given period of time*". Therefore, also in this case the emphasis is given to the expected damages due to the earthquake. Whichever is the preferred definition, the risk is to be intended as a combination of the following factors:

$$Risk = Hazard \; x \; Vulnerability \; x \; Exposure \qquad \text{Eq. 2-1}$$

In MIT – IST Group [17] risk is associated only to hazard and vulnerability while loss also includes exposure, however as the distinction between risk and loss is quite loose, these terms can be used almost interchangeably.

Referring to earthquakes we can describe the parameters contributing to the risk evaluation as in the following:

- **Seismic Hazard**: probable level of ground shaking associated with the recurrence of earthquakes [18]. It is worth noticing that this parameter depends on the earthquake, on its magnitude, on the distance of the location under examination from the epicentre, and basically on the geomorphological conditions of the ground, thus it does not depend on what is built by humans and it is an intrinsic feature of the area under examination [19] .

- **Vulnerability**: The characteristics and circumstances of a community, system or asset that make it susceptible to the damaging effects of a hazard. Vulnerability refers to different aspects arising from various physical, social, economic, and environmental factors, for example poor design and construction of buildings. It is worth noticing that vulnerability is a characteristic of the element of interest (community, system or asset) which is independent of its exposure and location. In this sense, the concept of seismic vulnerability is an intrinsic feature of the element (building, etc...) and does not depend on any external factor.

- **Exposure**: it refers to the overall quantitative value of the assets (population, environmental and historical goods, etc...) that may suffer from the earthquake. More specifically, it is "*the extension, the quantity, the quality of the anthropic elements that are part of the territory (population, buildings, infrastructures, etc...) whose conditions and / or working conditions are damaged or altered by a seismic event*". [19]

In the following paragraphs, we are going to explore in more detail the concepts of hazard and vulnerability as on these two parameters we are specifically going to explore the possibility to better exploit remote sensing data.

2.3   HAZARD MAPS

Seismic hazard [20], intended in a probabilistic way, refers to the expected earthquake ground motions at the earth's surface and the estimate of the probability of exceeding a certain amount of ground shaking, or ground motion, in a given time interval, i.e. the probability that a certain shaking motion happens in a given period of time.

This type of estimation is based on the definition of some input elements (such as earthquake catalogues, epicentres regions, mitigation relation of ground motions, etc…) and of some reference parameters (for example: peak ground accelerations and displacements, type of ground, time interval, etc…).
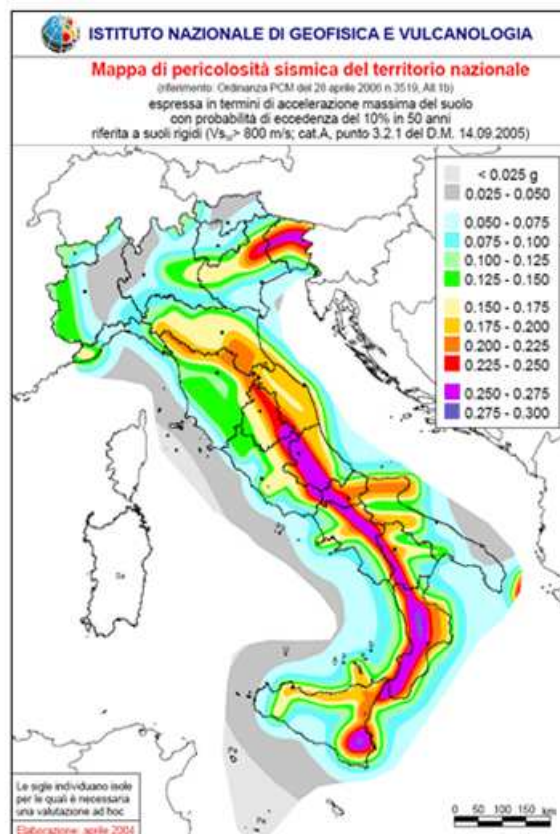
Figure 2-1: Seismic hazard Maps for Italy *(Source: INGV [21])*

The Italian municipalities have been attributed seismic hazard values in decreasing order, according to the following four zones:

- Zone 1 – The most dangerous zone, where heavy earthquakes can occur
- Zone 2 – Municipalities belonging to this zone may suffer from significant earthquakes
- Zone 3 – Moderate earthquakes can occur in the municipalities belonging to this group
- Zone 4 – This is the less dangerous zone

Each zone has been assigned a value for the seismic hazard useful for designers, as this is expressed in terms of Rock Peak Ground Acceleration (zone 1=0.35 g, zone 2=0.25 g, zone 3=0.15 g, zone =0.05 g) [22].

As from 2006, a new hazard study has provided the Italian Regions with an updated tool for the classification of the territories, introducing some acceleration ranges (ag) to be assigned to the four seismic zones with 10% probability of exceeding them in 50 years. Such classification corresponds to the one reported in the following table:

Table 2-1: Classification of seismic hazard

Seismic Zone	Acceleration that has a 10% chance of being exceeded in 50 years
1	$a_g > 0.25$
2	$0.15 < a_g \leq 0.25$
3	$0.05 < a_g \leq 0.15$
4	$a_g \leq 0.05$

Whichever is the regional choice regarding the zonation of its territory, each zone and sub-zone is associated to a basic hazard value, expressed in terms of peak ground acceleration on rigid ground (ag). Such basic hazard value has no influence on the design phase.

In the past years, seismic classification used to provide design parameters such as peak ground acceleration and the elastic response spectrum to be used for the calculation of the seismic actions. Following the standards approved in 2008, for each building there should be a reference to its own reference acceleration identified on the basis of its geographical coordinates and on its nominal life. A reference hazard value may therefore be defined for each point in the Italian territory on a 5 km square mesh, regardless of the municipalities' administrative borders.

Some of the main variables to examine ground seismic hazard are reported in the following table:

Table 2-2: Reference variables for seismic hazard assessment

Variables	Definitions
PGA – Peak Ground Acceleration	The peak ground acceleration is the maximum amplitude of the ground acceleration time-history. In terms of structural response, it corresponds to the peak value of the absolute acceleration of a Single Degree Of Freedom (SDOF) system with infinite stiffness, that is, with a natural period of vibration equal to zero. The meaning of this variable is similar to ag, but it also takes into account for the influence of amplifier effects due to underground or topographic features.
Peak Ground Displacement (PGD)	The highest absolute ground displacement recorded during an earthquake event, necessary to characterise the seismic motion amplitude at low frequencies.
Peak Ground Velocity (PGV)	Maximum highest ground velocity characterising the seismic motion amplitude at in-between frequencies. This variable specifies the hazard potential of structures sensitive to intermediate frequencies, such as high buildings or bridges.
Elastic Response Spectrum	The elastic response spectrum Se(T) is a diagram that provides, for different oscillation periods T, the maximum value of the pseudo-acceleration of the elementary oscillator (ground or structure) subject to seismic action, given a fixed damping ratio, which is a physical characteristic of the structure/ground. The spectrum provides us with the vibration frequencies where the stress on the structure is maximum and thus resonance phenomena are possible, increasing the collapse risk.
Nominal Life	This variable indicates the number of years during which a structure can be used for its original design. It is a parameter requested by the Italian rules on buildings. For ordinary buildings, the nominal life shall be at least ≥ 50 years.
Basic seismic hazard	Component of the seismic hazard due to the area seismologic features (type, dimensions and depth of the seismic source, energy and earthquake frequency). This parameter is calculated (generally on probabilistic terms), for a certain region and a given period of time, through the values of correspondent parameters at fixed exceeding probabilities. Such parameters (velocity, acceleration, intensity, spectral intercept) describe the shaking produced by the earthquake in rigid ground conditions and without morphological irregularities (reference earthquake). The reference scale is usually at regional level. One of the goals for future actions is to extend seismic classification on large scale areas. Nowadays this parameter is at the basis for the definition of a reference earthquake for seismic microzonation studies.
Local seismic hazard	This component of seismic hazard is due to local features (lithostratigraphic and morphological, vs. Local effects). The study of local seismic hazard is carried out starting from basic seismic hazard studies (reference earthquake) and analysing the geological, geomorphological, geotechnical and geophysical features of the site under examination; thus allowing identifying local amplitudes and ground instability phenomena. The most important outputs for this kind of analysis are the seismic microzonation maps.

2.4 SHAKEMAPS & MICROZONATION

A ShakeMap is a representation of ground shaking produced by an earthquake [23] . The information it presents is different from the earthquake magnitude and epicentre that are released after an earthquake because ShakeMap focuses on the ground shaking produced by the earthquake, rather than the parameters describing the earthquake source. Therefore, while an earthquake has one magnitude and one epicentre, it produces a range of ground shaking levels at sites throughout the region depending on:

- distance from the earthquake
- the rock type
- soil conditions at sites
- variations in the propagation of seismic waves from the earthquake due to complexities in the structure of the Earth's crust.

A "ShakeMaps" [25] Atlas has been developed for more than 1000 recent earthquakes. The methodologies followed are those described in [26]. The Atlas maps were built using constraints from macroseismic intensity data, instrumental ground motions and regional topographically-based site amplifications. The Atlas of ShakeMaps provides a consistent and quantitative description of the distribution of shaking intensity for global earthquakes (January 1973 – September 2007 for Version 1.0). The Atlas is regularly updated with more data constraints for historical events and the addition of future significant events as time progresses. The main goal of such an atlas is to be able to plan a prompt response after earthquakes and be able to calibrate and standardise worldwide methodological approaches.

The following table summarises the main types of available Shakemaps:

Table 2-3: Summary of ShakeMaps

Peak Acceleration Maps	Peak horizontal acceleration at each station is contoured in units of percent-g (where g = acceleration due to the force of gravity = 981 cm/s^2). The peak values of the vertical components are not used in the construction of the maps because they are, on average, lower than the horizontal amplitudes and ground motion prediction equations used to fill in data gaps between stations are based on peak horizontal amplitudes. The contour interval varies greatly and is based on the maximum recorded value over the network for each event. For moderate to large events, the pattern of peak ground acceleration is typically quite complicated, with extreme variability over distances of a few km. This is attributed to the small scale geological differences near the sites that can significantly change the high-frequency acceleration amplitude and waveform character. Although distance to the causative fault clearly dominates the pattern, there are often exceptions, due to local focussing and amplification. This makes interpolation of ground motions at one site to a nearby neighbour somewhat risky. Peak acceleration pattern usually reflects what is felt from low levels of shaking up to moderate levels of damage.
Peak Velocity Maps	Peak velocity values are contoured for the maximum horizontal velocity (in cm/sec) at each station. As with the acceleration maps, the vertical component amplitudes are disregarded for consistency with the regression relationships used to estimate values in gaps in the station distribution. Typically, for moderate to large events, the pattern of peak ground velocity reflects the pattern of the earthquake faulting geometry, with largest amplitudes in the near-source region, and in the direction of rupture (directivity). Differences between rock and soil sites are apparent, but the overall pattern is normally simpler than the peak acceleration pattern. Severe damage and damage to flexible structures is best related to ground velocity. For reference, the largest recorded ground velocity from the 1994 Northridge (Magnitude 6.7) earthquake made at the Rinaldi Receiving station, reached 183 cm/sec.
Spectral Response Maps	Following earthquakes larger than magnitude 5.5, spectral response maps are made. Response spectra portray the response of a damped, single-degree-of-freedom oscillator to the recorded ground motions. This data representation is useful for engineers determining how a structure will react to ground motions. The response is calculated for a range of periods. Within that range, the Uniform Building Code (UBC) refers to particular reference periods that help define the shape of the "design spectra" that reflects the building code. ShakeMap spectral response maps are made for the response at three UBC reference periods: 0.3, 1.0, and 3.0 seconds. For each station, the value used is the peak horizontal value of 5% critically damped pseudo-acceleration.
Rapid Instrumental Intensity Maps	As an effort to simplify and maximize the flow of information to the public, USGS developed a means of generating estimated Modified Mercalli Intensity maps based on instrumental ground motion recordings. These "Instrumental Intensities" are based on a combined regression of peak acceleration and velocity amplitudes vs. observed intensity for eight significant California earthquakes. From the comparison with observed intensity maps, it was found that a regression based on peak velocity for intensity > VII and on peak acceleration for intensity < VII is most suitable. This is consistent with the notion that low intensities are determined by felt accounts (sensitive to acceleration). Moderate damage, at intensity VI-VII,

typically occurs in rigid structures (masonry walls, chimneys, etc.) which also are sensitive to high-frequency (acceleration) ground motions. As damage levels increase, damage also occurs in flexible structures, for which damage is proportional to the ground velocity, not acceleration. By relating recorded ground motions to Modified Mercalli intensities, it is possible to estimate shaking intensities within a few minutes of the event based on the recorded peak motions made at seismic stations. A descriptive table of Modified Mercalli Intensity is available from ABAG (Association of Bay Area Governments). ShakeMap uses PGA to estimate intensities lower than V, it linearly combines PGA & PGV for intensities greater than V and less than VII, and it uses PGV for intensities greater than VII. [26]

After an earthquake, the observation of damages on constructions and infrastructures often highlights substantial differences in different built-up areas, even at short distance among them. In some cases, noteworthy collapses and damages occurred in places lying at great distances from the epicentre [27].

During L'Aquila earthquake of 6 April 2009, these type of situations were found both in the municipal territory of L'Aquila and in some distant municipalities about 30 km SE of the epicentre. The quality of buildings definitely plays a major role in the damage entity, but often causes are linked to the local seismic hazard, determined also by the different earthquake propagation, or by the instability of the soil.

Such considerations are part of the studies of Seismic Microzonation (SM), through which it is possible to pick out and characterize stable areas, stable areas susceptible to local amplification and areas subject to instability, such as landslides, surface fractures and soil liquefaction.

2.5   VULNERABILITY

2.5.1   Definition of Vulnerability

The concept of vulnerability is linked to the attitude of a certain environmental component (human population, buildings, services, infrastructures, etc.) to withstand the effects of an event, depending on its intensity. Therefore, the concept of vulnerability is related to the level of losses caused by an event of a determined power. It is expressed in a range between zero and one where zero stands for no damage occurrence and one corresponds to total collapse.

The concept of Physical vulnerability to which we refer when we talk about the effects of earthquakes on the built environment is actually linked to the susceptibility of a building to withstand the effects of an earthquake of a given intensity level. Three types of vulnerabilities can be identified, namely [19]:

- **Direct Vulnerability**: attitude of a single element, single or complex (e.g. building, roads, etc..), to withstand damages caused by an earthquake
- **Indirect Vulnerability**: this vulnerability is linked to the effects that the crisis generates on the organisation of the territory (e.g. crisis of the mobility system due to impossibility of passing through certain roads)

- **Deferred Vulnerability**: effects of the earthquake that modify or strongly affect the normal behaviours of the population, for example the need to locate large camps and containers in the territory, or the effect due to the impracticability of schools, hospitals or other public services.

Another important aspect, worth mentioning here, is linked to the extension of the concept of building vulnerability to the whole territory or to a part of it. It is quite clear, in fact, that the global vulnerability of the system is not a simple addition of the vulnerabilities of the single elements, but it is a broader evaluation that depends on the vulnerability of the elements that are part of the territory and on their interrelations and organisation. Therefore, quite often a hierarchical order is assigned depending on the weights attributed to the vulnerabilities of the single elements.

The building "vulnerability" expresses differences on the way how buildings respond to earthquake shaking. If two groups of buildings are subject to exactly the same earthquake shaking, and one group performs better than the other, then the buildings that were less damaged had lower earthquake vulnerability than the ones that were more damaged, or it can be stated that the buildings that were less damaged are more earthquake resistant, and vice versa [28].

2.5.2 Evaluation of Vulnerability

Among the various classifications that exist on the methods of evaluation of physical vulnerability, we can focus on the following three methods:

- **Mechanical/Analytical Methods**: usually this is a non-linear analysis of the structure. The damage is associated to reaching limit states. Such two limit states may be limit rotation or the collapse acceleration.

- **Empirical Methods**: these methods are based on the statistical analysis of the damages produced by an earthquake. These methods are particularly suitable for the evaluation of the vulnerability for a group of buildings associating similar buildings to the same functional classes depending on the materials used, on their geometry, morphology, age, etc.

- **Experts-Based Methods**: these methods are based on the possibility to associate to each building a vulnerability index, which is evaluated usually by qualitative or typology/diagnostic evaluation of the building seismic behaviour.

Being a detailed description of each of the above mentioned methods out of scope of this thesis, a summary of the most important features for those methods is given below with the aim of understanding which are the most important data needed to evaluate vulnerability.

Such step is necessary to fully understand how remote sensing data can operatively contribute to the evaluation of vulnerability.

2.5.2.1   Mechanical/analytical Methods

Analytical and mechanical methods allow numerical quantification of the level of risk. They are based on structural modelling of the building and evaluate its aptitude to be damaged by an earthquake.

A vulnerability function describes the probability distribution of loss ratio conditioned on an intensity measure level [29].

In summary, we can represent the seismic risk function for a building (or a class of building) as the probability for the structure not to withstand the limit states, such as the Ultimate Limit State [30].

$$P_f = P[Z(\bar{X}) \leq 0] = P[C(\bar{X}) \leq D(\bar{X})]$$ Eq. 2-2

Where

Z →represents the limit state function. If it is ≤0 it represents a failure state

$C(\bar{X})$ →represents the Capacity Curve and it is a random variable

$D(\bar{X})$ →represents the Seismic Demand and it is a random variable

This representation clearly shows the need to separate the analysis of the structural response from the probabilistic characterisation of the seismic demand.

The evaluation of the Capacity Curve for a building (or a class of buildings) can be performed in different ways and many methods have been proposed in literature. In general, some simplified methods can be used such as those suggested by Calvi [31], Borzi [32] (SP-BELA), Crowley [33] (DBELA) or by using pushover analysis ([34] , [35], [36]).

Once the capacity curve is known using any of the above mentioned methodologies, the evaluation of the seismic demand can be carried out through the Capacity Spectrum Method, either by using over-damping as suggested by Freeman [37] (this method is used in HAZUS [38]) or by using inelastic spectra (N2) as suggested by Fajfar [39] .

At the end of this process, the evaluation of the failure probability for each seismic intensity level can be performed by evaluating the function Z is ≤0, which can be done through Montecarlo methods, for instance. By repeating this kind of calculation for different earthquake intensities (in terms of PGA for example) and for different limit states of the structure, the fragility curves can be obtained for the class of buildings (or single building) under examination.

The following picture shows the HAZUS curves as an example of the output:
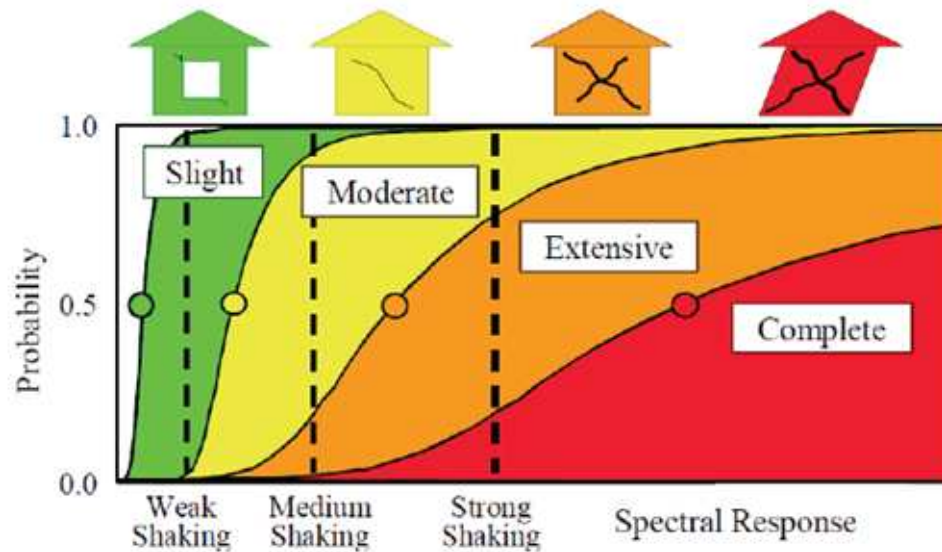
Figure 2-2: Fragility Curves in the HAZUS® method (FEMA)

2.5.2.2   Empirical Methods

The assessment on seismic vulnerability of buildings is a fundamental step in seismic hazard analysis and in the definition of earthquake damage scenarios, according to different earthquake intensities [40]. Vulnerability functions provide relations among the typology of the damaged element – the intensity of the dangerous event and the damage levels. The vulnerability functions refer to building classes that share similar behaviours in response to a dangerous event.

Already the previously used MSK (Medvedev-Sponheuer-Karnik Scale) scale could define the vulnerability classes with the aim of facilitating and improving macroseismic mapping for the evaluation of local intensities of an earthquake. Such definition has improved through the introduction of the EMS (European Macroseismic Scale) scale where building vulnerability classes have increased with respect to the previous classification. Six classes were identified in decreasing order of vulnerability starting from class A to class F [28].

The first three classes (A-B-C) include structures with resistances typical for masonry and reinforced concrete buildings. Such classification makes the EMS scale fully compatible with the previous MKS.

Classes D and E approximately represent a linear decrease in vulnerability thanks to the increasing level of anti-seismic design degree (Earthquake Resistant Design) and thus include structures that typically withstand earthquakes, such as well built wooden structures, confined masonry and steel structures. Class F includes structures with a very low level of vulnerability since they were built following all relevant anti-seismic criteria.

The following picture shows the differences among the structures according to the EMS classification.

Figure 2-3: EMS Scale Vulnerability Classes

The use of data based on observed damages, properly processed and organised in the form of Damage Probability Matrices has been introduced in Italy for the analysis of vulnerability and the prediction of the expected damage since the Irpinia earthquake in 1980 [42]. DPM provide the probability of occurrence of different damage classes – defined on the basis of the observed consequences on structural and non-structural elements – due to seismic impact given in terms of macroseismic intensity and for different vulnerability classes (classes containing buildings with homogeneous behaviour).

Starting from the intensity maps of different seismic events [41], and using Damage Probability Matrix (DPM) it is possible to define, on a probabilistic basis, for each vulnerability class (A, B, C, D) and for different seismic levels, the percentages of damage levels on the

buildings in a determined district, where the damage levels, as defined by MSK and EMS are the following:

0. No damage
1. Negligible to slight damage
2. Moderate damage
3. Substantial to heavy damage
4. Very heavy damage
5. Destruction

The following picture shows the classification of damage for masonry and reinforced concrete buildings, as reported in Grunthal, 1998 [28].

Figure 2-4: EMS 98 scale: masonry (left) and reinforced concrete (right) building type [28]

CLASS A

Intensity	Damage Level					
	0	1	2	3	4	5
VI	0.188	0.373	0.296	0.117	0.023	0.002
VII	0.064	0.234	0.344	0.252	0.092	0.014
VIII	0.002	0.020	0.108	0.287	0.381	0.202
IX	0.0	0.001	0.017	0.111	0.372	0.498
X	0.0	0.0	0.002	0.030	0.234	0.734

CLASS B

Intensity	Damage Level					
	0	1	2	3	4	5
VI	0.36	0.408	0.185	0.042	0.005	0.00
VII	0.188	0.373	0.296	0.117	0.023	0.002
VIII	0.031	0.155	0.312	0.313	0.157	0.032
IX	0.002	0.022	0.114	0.293	0.376	0.193
X	0.0	0.001	0.017	0.111	0.372	0.498

CLASS C

Intensity	Damage Level					
	0	1	2	3	4	5
VI	0.715	0.248	0.035	0.002	0.0	0.00
VII	0.401	0.402	0.161	0.032	0.003	0.00
VIII	0.131	0.329	0.330	0.165	0.041	0.004
IX	0.050	0.206	0.337	0.276	0.113	0.018
X	0.005	0.049	0.181	0.336	0.312	0.116

Figure 2-5: Example of Damage Probability Matrix (DPM) *(Source: Bernardini et al. [42])*

2.5.2.3   Expert based Methods

The Italian Civil Protection has prepared, supported by INGV (Italian Geophysical and Volcanic Institute) a procedure to assess post-earthquake damages and building usability [14]. Groups of experts perform the in-situ analysis filling in some forms answering to specific questions.

The availability of first and second level forms [44] allows performing an accurate analysis on buildings vulnerability. The use of first level forms enables the identification of vulnerability classes in which each building may be classified [40]. The use of second level forms allows evaluating punctual vulnerability on each building, by using a score between 0 and 100, for masonry and between –25 and 100 for concrete buildings.

As anticipated, two types of forms can be identified:

- First Level Forms: these forms are more general and usually used for statistical evaluation of groups of buildings or for accessibility evaluations.
  - First Level Forms for the evaluation of exposure and vulnerability of buildings (masonry or RC) and for particular buildings (e.g. churches, industrial facilities, etc…)
  - First Level Forms for the evaluation of damage, first aid, and accessibility for ordinary buildings to manage the emergency
- Second Level Forms: usually these forms are similar to the first level ones but require a higher level of detail, therefore they are usually required for the analysis of specific buildings
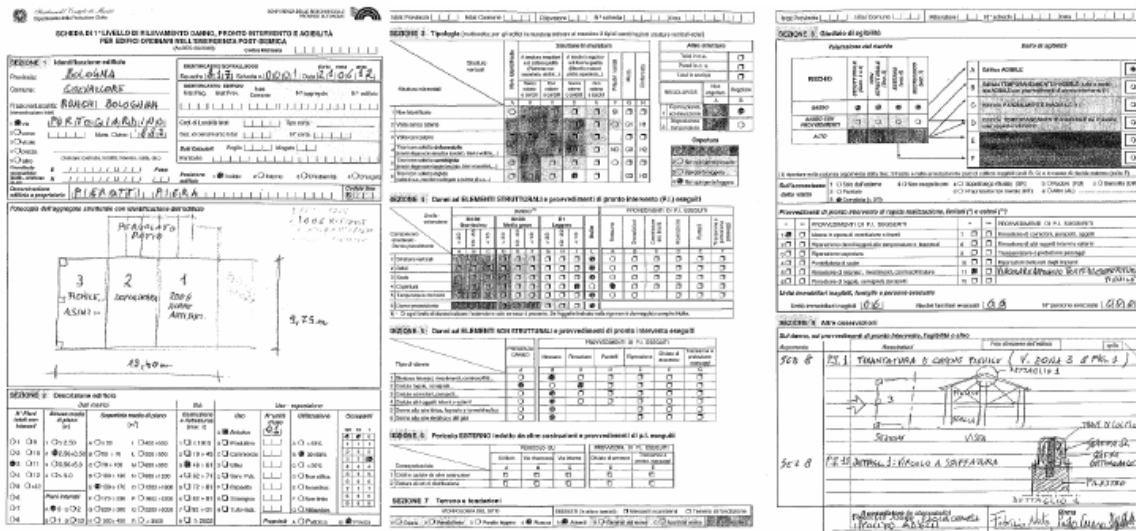
Figure 2-6: Examples of forms for the collection of vulnerability information

A summary of the most significant data required in the forms are detailed in the following table, particularly referring to those needed for the assessment on the conformity of the building to standards:

Table 2-4: Summary of the main Data Requirements in I/II level forms

Summary of Forms Data Requirement	Type of Data	Comment
Building Identification	Address	
	Cadastral Documents	
	Location	Isolated / Internal/ On the corner / On the far end
Building Description	Metrical Data	Average covered area
		Total n° of storeys including basements
		Average Storey Height
		Minimum and Maximum height above ground
		Street width on the main facade
	Age	of the building and /or of the renovations
	Usage/Exposure	N° of units and their predominant typ e of use (e.g. private, public, strategic, etc.)
		% of utilisation
		Residents
Building Type	Masonry	Horizontal (type of vaults, beams, etc..)
		Vertical Structures (regular or irregular textures, pillars, etc…)
	Reinforced Concrete	Regularity, Plant and cladding type
	Coverings/ Roof	Heavy push / no-push type/ Light push / no-push type
Damages to structural elements	Structural Component	Type of structural component and extension of the damage
Damages to non-structural elements	Type of damage	Finishing status / Plaster detachment / shingle, ledge fall / damages to the water, electric or sewage pipes
External danger (generated by other buildings)	Type of external Danger	on the building / on its entries / internal roads
	Potential cause	Collapse of buildings nearby / interruption of distribution networks / etc…
Land and Foundations	Morphology of the site	Peak / slope % / level ground
	Instability	Falling slopes / Foundations

2.5.3   Indicators of Vulnerability

In general, non analytical methods lead to the need of performing a statistical analysis of the data observed; therefore the vulnerability of a building can be evaluated from its "symptoms" [45] (resistance of its components, global resistance, level of use, etc...). The basic idea is to assign to each building a vulnerability indicator Iv that can be interpreted as its suitability in withstanding earthquakes. The Vulnerability index Iv is a conventional measure and is related to the building behaviour versus damage, it can be immediately used only for heuristic analyses, for example for practical comparisons among buildings. The vulnerability index does not provide an evaluation of the expected damage for a given level of severity of the earthquake. It can be calculated as shown below:

$$I_v = \sum w_i p_i$$

Eq. 2-3

where

w_i → weight

p_i → parameter that has an influence on the global building behaviour against earthquakes

An example of choice of parameters and relative assignment of the weights is given in the following table:

Table 2-5: Example of Weights and Parameters for Vulnerability index evaluation *(Source: Zuccaro [30])*

PARAMETER (p_i)	WEIGHT (w_i)
Type and organisation of the resisting system	1,00
Quality of the resisting system	0,25
Conventional resistance	1,50
Building location and foundations	0,75
Horizontal structures	Variable
Planimetric configuration	0,50
Height configuration	Variable
Maximum distance between the walls	0,25
Roof and coverings	Variable
Non structural elements	0,25
Overall conditions	1,00

The vulnerability indicator is a non-dimensional number between 0 and 1, going from the lowest to the highest level of vulnerability.

It is worth noticing here that, in literature, some methods have been suggested to evaluate vulnerability using remote sensing data. For example, Taubenbock et al. [46] suggested considering a wider definition of vulnerability, which also includes exposure and coping capacity. In such analysis, the concept of vulnerability is to be intended as:

$$Vu\ln erability = \frac{Exposure \times Susceptibility}{CopingCapacity}$$

Eq. 2-4

Where the Coping capacity is the ability to cope with or adapt to hazard stress. It is the product of planned preparation, spontaneous adjustments and relief and reconstruction made in response to the hazard. The Susceptibility reflects the capacity of individuals, groups or the physical or socio-economic system to withstand the impact of the hazard. The exposure is to be intended as already presented in Par. 2.2.

The following table shows the overall concept, as proposed by Taubenbock et al. [46]. This is a holistic concept that includes in the analysis also joint indicators related to hazard, exposure and vulnerability

Figure 2-7: Holistic approach to Vulnerability *(Source: Taubenbock et al.[46])*

The methodology followed by [46] follows the same procedure presented above for the evaluation of a vulnerability indicator as in Eq. 2-3. In this case, however, a wider set of parameters has been considered due to the holistic definition of vulnerability proposed by the authors. The final calculated I_v allowed comparing four different locations in the same urban area identifying those that were more vulnerable than others. The results presented by the

authors clearly show the capabilities of remote sensing to assess aspects of physical and demographic vulnerability as well as aspects defining the spatial distribution of after-effects in case of an earthquake.

We have reviewed in this chapter the main concepts related to earthquake risk, vulnerability, exposure and hazard. For each of these concepts we have examined the main characterising parameters. We understood that the hazard concept is related to the geographical positioning of the building and its main indicators (e.g. PGA, shakemaps, etc.) are completely independent from the type of building, while the concept of vulnerability is strictly related to the building type. In summary, the building vulnerability is mainly depending on three indicators, which are a) height, b) age, c) type of building, roughly if masonry or reinforced concrete. This information is very important to understand our modular approach for the building damage assessment that is presented in the next chapter.

3 STATE-OF-THE-ART IN THE USE OF REMOTE SENSING FOR DAMAGE ASSESSMENT

In this chapter, we are going to review the current state-of-the-art in the use of remote sensing for damage assessment and, in particular, for post-earthquake damage assessment. Our goal is to understand which are the indicators that can be further exploited by the use of EO data. After reviewing the state-of-the-art, this chapter introduces also the specific requirements of the data fusion integrated module that we are going to develop afterwards.

3.1 USE OF REMOTE SENSING FOR RISK MANAGEMENT & EARTHQUAKE ASSESSMENT

As already mentioned in Par.2, the increased interest in the use of remote sensing data for risk prevention and management has been increasing significantly in the last years, in particular for earthquake assessment:

"*The observation of a strong ground motion and aftershock sequences as well as the identification of the destruction from these earthquakes provide seismology and earthquake engineers with informative and valuable data, experiences and lessons and raise some important scientific problems. Application of new technologies such as GIS, GPS and remote sensing, among others, has been shown to be of significant help in the study of seismic hazard and risk*" [47].

Application of satellite images to loss quantification and classify building stock are needed for a real-time loss estimation and can still advance significantly, as highlighted in the IASPEI Commission on earthquake Hazard, Risk and ground strong motion .

The topic of emergency management is grabbing attention worldwide and many resources were put towards gaining an International structured methodology to perform quick, coordinated and useful emergency management at International level. The example of the Haiti 2010 earthquake and the quick availability of data and maps, also provided by public tools such as the Google Crisis Response, show how such effort is shared by the International Community and how many governmental and private players are involved in such work.

Many initiatives have been going through in the recent years to optimise the use of remote sensing for earthquake and, more generally, also for emergency management, both in Europe and worldwide. It is worth mentioning here that most of these initiatives do not approach the evaluation of post-earthquake damage assessment at individual building level, as done in our work. For example, the project EVOSS [48], which is dedicated to the creation of European Volcano Observatory Space services, aims at increasing the capacity of volcano observatories by providing remote monitoring services using space-based systems. The technology proposed is intended to enhance the speed and efficiency of responses to multi-hazard events. Volcano monitoring techniques are currently, as for earthquakes, mainly focused on predicting volcanic eruptions in order to save lives and property. In both cases, the majority of techniques are ground-based, such as measuring ground deformations, seismicity and local gases (e.g. sulphur dioxide (SO_2). Recent technological advances enable remote sensing using satellite-based observations and Global Positioning System (GPS) measurements. European scientists sought to exploit space-borne technological support of volcanic monitoring systems to enhance capabilities in the EU and Africa

significantly. The system that was developed relies on geostationary (stationary with respect to a given point above the Earth) and Low Earth Orbit (LEO) satellites together with advanced signal processing techniques. The EVOSS virtual volcano observatory infrastructure is also aiming at producing an automated alert procedure that combines data from 12 orbiting platforms hosting seven different payloads acquired at four different places and processed by eight different scientific content providers.

Another study (*SUBCOAST* [49]) approached a similar problem on coastal areas by developing a GMES service for monitoring and forecasting subsidence hazards in coastal areas around Europe. SubCoast developed a GMES-downstream service for assessing and monitoring subsidence hazards in coastal lowland areas around Europe, based on satellite data, in-situ measurements and geo-scientific models. In this work the location of radar scatterers (InSAR), GPS, levelling and gravimetry observations were used to produce the deformation model of the Netherlands.

The topic of using both space and in-situ data is also addressed by DORIS ([50] Ground Deformations Risk Scenarios) which aims at an advanced downstream service for the detection, mapping, monitoring and forecasting of ground deformations, that integrates traditional and innovative Earth Observation (EO) and ground based (non-EO) data and technologies. The service delivers innovative products tailored for Civil Defence authorities. DORIS uses ERS-1/2 and ENVISAT C-band SAR archives to provide very long time-series of ground deformations. It evaluates new SAR sensors, including ALOS, COSMO-SkyMed and TERRASAR-X, exploiting the different bands (L/X), the significantly reduced revisiting time, and the higher spatial resolution offered by these sensors. DORIS developed the integration of satellite and ground-based SAR interferometry, coupled with GPS measurements and geophysical probing. Novel satellite imaging and advanced sensors are penetrating the Earth's surface to reveal danger zones, forecast weak points and alert communities of specific natural hazards. This study proved that the use of powerful satellite technology yields to much more detailed ground information than ever before, since valuable satellite information reveals the mechanical behaviour of ground deformation phenomena. This, combined with meteorological data, radically improves European ability to predict the long-term behaviour of slow-moving landslides. Importantly, the technology is expected to upgrade and update ground deformation maps, facilitating the evaluation of associated hazards and risk.

Some studies on earthquakes also addressed the problem of combining geophysical observations and parameters with refined data analysis methods to improve earthquake prediction (*PRE-EARTHQUAKES* [51] - Processing Russian and European Earth Observations for Earthquake Precursors Studies). To this purposes, observations from more than 20 different satellite systems (particularly those from European, Russia and US Space Agencies) and about 100 ground-based stations were systematically collected, integrated and analysed. The main goal of this kind of study was to reduce the number of false alarms and improve reliability in earthquake prediction. To achieve this, the project incorporated information from satellite and ground-based systems. Researchers used the invaluable resources provided by the European Space Agency (ESA) and the Russian Federal Space Agency, ROSKOSMOS to study their archived satellite data to amass data relevant to the preparatory phase of earthquakes. Another major goal was to create a global earthquake observation system (EQUOS) as part of the Global Earth Observation System of Systems (GEOSS). Altogether, observations from 20 satellite systems were used to study both surface and atmospheric, up to the ionosphere, anomalies that are regarded as possible precursors to earthquakes. The prime test areas were in Italy, eastern Russia and Turkey.

During the first year of PRE-EARTHQUAKES the scientists collected a massive amount of data. A total of 6165 kinds of observations were performed (different for measured parameter, observation technology, analysis method, geographic area or test period). 2011 is instead the number of different integrations/comparisons experiments performed. The data can be viewed and compared at the common integration platform, PRE-EARTHQUAKES Geo-Portal (PEG). However the initiative PRE-EARTHQUAKES mainly tackles the preparatory phase of earthquakes, not really the damage assessment post-earthquake. The main course of prevention against earthquake damage is to build safe houses. As that is not economically possible at present, better earthquake prediction promises to offer the best protection for European citizens.

The problem of addressing highly vulnerable areas is also addressed by the LAMPRE study ([52]- LAndslide Modelling and tools for vulnerability assessment Preparedness and REcovery management) which aimed at reaching technological developments to increase GMES limited operational capacity to cope with triggered landslide events and their consequences, in Europe and elsewhere. LAMPRE enhances landslide risk mitigation/preparedness efforts and post-event-landslide recovery and reconstruction activities, in highly vulnerable geographic and geologic regions. The project tackled the ability to detect/map landslides, assess/forecast the impact of triggered landslide events on vulnerable elements, and model landscape changes caused by slope failures. These goals were achieved by

(i)   researching and developing new techniques and products to dynamically integrate satellite/airborne imagery,
(ii)   designing and using intelligent image processing techniques,
(iii)   modelling landslide-infrastructure interactions using advanced numerical modelling and ground based thematic information,
(iv)   proposing standards for landslide mapping, susceptibility zonation and image processing.

The results of LAMPRE are relevant in the framework of the more general EU strategy for the prevention, preparedness and response to natural hazards, and the protection of people, property, infrastructures and the environment, to implement the EU Soil Thematic Strategy, and to the design of novel GMES landslide services based on images taken by the ESA Sentinel-2 satellites.

In recent years, researchers coordinated by the University of Sheffield in UK have also found a link between changes in electromagnetic phenomena (e.g. radio signals and plasma waves) in the ionosphere that occur before a large earthquake. Through advanced satellite and ground monitoring of radio signals that change a few days before a major earthquake, researchers can radically improve the prediction of large seismic events. The EU-funded project 'SEMEP - Search for electro-magnetic earthquake precursors combining satellite and ground-based facilities' [53] sought to build on this potential. The project studied anomalies in the electromagnetic environment and outlined the types of measurements required to improve forecasting. Interestingly, the concept can also be applied to volcanic eruptions, which also instigate a change in electromagnetic behaviour several days or hours before a natural event occurs. To achieve its aims, the project team analysed data from French and Russian microsatellites, looking as well into ground-based measurements that could also help predict earthquakes. In addition, the team investigated data from Very-Low–Frequency (VLF) and Low-Frequency (LF) receiver stations in Europe, Japan and Russia. It studied plasma waves in the vicinity of earthquake epicentres and sought to correlate the waves'

behaviour with seismic activity. SEMEP also undertook an in-depth study of the coupling between the Earth's lithosphere, atmosphere and ionosphere, investigating Internal Gravity Waves (IGWs) caused by changes within the crust from increased seismic activity. Moreover, the project team noted that conducting a joint satellite and ground-based analysis in two different seismic zones increases the chance of finding events related to different geophysical conditions.

Another initiative with the objective of setting up a space-based end-to-end information services, based on satellite remote sensing data, to support prevention/preparedness and recovery phases of the Forest Fires emergency cycle in the European Mediterranean Region is the project *PREFER* [54] (Space-based information support for prevention and recovery of forest fires emergency in the Mediterranean area). The key-drivers of this approach can be summarised in the four main topics below:

- Systematic Fuel Estimates
- Systematic Burn Scar Mapping
- Systematic Analysis of fire Effects on Slope Stability
- Systematic Vegetation Recovery Analysis

Multi-hazard applications such as the prediction/early detection of emergencies, population alerting, environmental monitoring, crisis management and risk assessment are also proposed to be managed by a single, user friendly interface, targeting several users (crisis managers, operators of critical infrastructures, insurance companies, scientific/academic). Such approach was the one proposed by DLR in PHAROS [55] - Project On A Multi-Hazard Open Platform For Satellite Based Downstream Services. This study focused on developing a sustainable, pre-operational open service platform, which integrates space-based observation, communications and navigation technologies to provide innovative services for a wide variety of users and application domains using:

- Satellite EO data access and in-situ sensors and processing;
- Decision support services based on data fusion and situation assessment techniques;
- A simulation platform for hazard modelling and risk assessment;
- An alert message gateway to alert the population through a variety of communications means.

A novel energy-efficient satellite uplink optimised for long messages is added to in-situ sensors improving their availability and deployment flexibility. Location information from remote responders is integrated into the common operational picture and GNSS services are applied to disseminate alert messages to the population. A free-space optical downlink can be used to download data from a real EO satellite, achieving throughput in the order of Gbps. This methodology was also field tested with a pre-operational demonstration in a realistic forest fire exercise, organised and controlled by fire brigades, during which the entire platform was challenged and all services were validated.

The Integration of GMES Emergency Services with satellite navigation and communication for establishing a flood information service (FLOODIS [56]) was also examined and developed by using accurate location based applications for portable devices. The proposed solution closes a critical gap for disaster management teams, civil protection, field/emergency response units to better address and mitigate crises arising before, during, and after heavy flooding. The project provides an open-source; location based smart phone

application for the general public to enable the capacity for individuals to take pre-cautionary actions, therefore vastly reducing the likelihood of human and economic loss. The project also considered rescuers relying on professional terminals and legacy communication channels. In this project, Earth Observation and GNSS (GPS, Galileo, EGNOS/EDAS) technologies were combined to deliver alerts and interactive maps on flooding risk/events to users in the geographical area at risk. A key and innovative capability of the system is the ability of the user to send back actual information about the situation "*on field*" to a remote gateway for ingestion and subsequent dissemination to all other users. As such, it serves to provide up-to-date, local information to disaster management centres, civil protection agencies, emergency response units, as well as affected citizens.

Another study, recently carried out, aimed at adapting the newly developed 12m resolution TanDEM-X Digital Elevation Model (DEM) to risk management applications, using it as a base layer to interrogate data sets and develop specific disaster scenarios. The Rapid Analysis and Spatialisation Of Risk (RASOR [57]) project overlays archived and near-real time very-high resolution optical and radar satellite data, combined with *in-situ* data for both global and local applications. A scenario-driven query system allows users to project situations into the future and model multi-hazard risk both before and during an event. Managers can determine the extent of flooding in a given area and determine, for example, the risk pending on Critical Infrastructure Systems in terms of their residual functionality as a basis for a systemic vulnerability analysis. Public authorities are able to determine the impact of coastal subsidence on flood defence over several years given various sea surge scenarios and based on actual, accurate subsidence information. The goal of this study was to provide disaster managers to use real scenarios when determining new mitigation or prevention measures, and integrate real-time data into their operational system when organising response activities.

Since also river floodplains are only managed to maximize river discharge capacity, a dedicated study was performed for the assessment of foreshore using Space Technology (*FAST* [60]). Marine foreshores are currently not included in water safety assessments and in levee design. However, foreshores and floodplains deliver several services, such as increasing sedimentation, reducing erosion and attenuating waves that mitigate flood risk by improving levee stability and lifetime. Including foreshores and floodplains in levee design and safety, assessments can result in considerable cost reductions for flood risk management. The FAST (Foreshore Assessment using Space Technology) project aimed to develop a new GMES/Copernicus downstream service by developing products based on Sentinel data to gain spatial information on foreshore and floodplain characteristics, such as morphology, sediment characteristics and vegetation properties. Necessary ground referencing in combination with measurements on wave attenuation and erosion/deposition regimes were executed at eight characteristic case-study sites across Europe (Spain, Romania, United Kingdom and the Netherlands). From the collected data, general relationships between foreshore and floodplain characteristics and flood risk mitigation properties were derived and implemented in a GIS-based software package. Software calculates effects of foreshores and floodplains on hydraulic conditions either derived from satellite images or engineering manuals. Effects of foreshores and floodplains on wave impact and bed stability are translated to potential reduction in levee width and crest height.

Some recent studies have also been focusing more on the improvement of earthquake damage assessment (and prevention). Some of them are also recalled here in the following.

A group of companies and research centres, coordinated by the Helmholtz Centre Potsdam German Research Centre for Geosciences (GFZ), performed a very important analysis on the integration of Space-based and in-situ data. ([61] - Framework to integrate Space-based and in-situ – *SENSUM* - SENSing for dynamic vUlnerability and recovery Monitoring). The SENSUM project is a response to the fact that human society is increasingly exposed to natural hazards through urbanisation, greater dependency on technological infrastructures and environmental change. It is therefore essential to understand the changes in society's vulnerability, and to integrate this into robust estimates of risk and of losses that follow an extreme natural event. This is especially important in countries such as those in Central Asia, where area-wide knowledge of the existing building stock is lacking, and the urban environment is rapidly changing. It is this lack of information about the extension and conformation of an urban area's vulnerability, which is one of the fundamental problems addressed by this study. It is for this reason that innovative methods of integrating observations and data from Earth Observing satellites and ground-based methods such as omni-directional camera surveys were a primary focus. However, the results are highly dependent on the availability of the necessary data and information, which is not uniformly distributed worldwide.

All these studies prove that the assessment of vulnerability is a fundamental step in order to reduce it. Another example of study, which is more focused on the vulnerability aspect, is the project *MOVE* [62] - Method for the Improvement of Vulnerability Assessment in Europe. This Consortium, coordinated by Università Degli Studi Di Firenze, studied natural hazards such as floods, temperature extremes, droughts, landslides, earthquakes, wildfires and storms. The vulnerability assessment carried out by the researchers aimed at reducing disaster risks by assessing elements of climate change from both theoretical and practical standpoints. The vulnerability assessment included all major pillars of vulnerability covering physical, environmental, economic, social, cultural and institutional criteria that are evaluated by quantitative and qualitative parameters. Scientists evaluated existing methodology and modified or developed tools and indicators for the clear measurement of vulnerability for all scales and hazard types. Application of the generic framework and data analysis methods to seven case studies in different European regions enabled the development of a standard approach described in the project's 'Manual of Vulnerability Assessment in Europe'. An accompanying handbook [58] presents the seven case studies. MOVE adopted a holistic approach focusing on both vulnerability and hazard concepts. Scientists included the human factor in vulnerability and risk both before and after an event. Such work therefore made an important attempt to include also the intricate coupling between emergencies and human behaviour.

The pre-operational GMES emergency service provided through the FP7 project SAFER ([64] - Crisis management – prevention, response and recovery), which ended in April 2012, also addressed the following situations:

a) Natural disasters (e.g. storms, fires, floods earthquakes, tsunamis, volcanic eruptions),
b) Man-made disasters, either deliberate or accidental (e.g. urban fires, chemical incidents on industrial sites),
c) Humanitarian disasters.

The project focused on rapid mapping during the response phase, with a response time objective of six hours after the emergency situation for the provision of reference mapping and of 24 hours for the provision of assessment maps. The services of the SAFER project rely on a full time on-duty staff available 24 hours/day, 365 days/year. This initiative was coordinated by Spot Image (Infoterra) Astrium Geo-information services.

The problem of gaining a rapid response in the event of emergency was also addressed by the G-MOSAIC Project ([65] - GMES services for Management of Operations, Situation Awareness and Intelligence for regional Crises), where the Rapid Geospatial Reporting (RGR) Service provides users with relevant information in the context of a crisis. The RGR's products include both basic cartography information and the analysis of layers. In order to improve the timeliness of responses, the service products are predefined with the intention of covering possible needs in the context of crisis activation. Nevertheless, as seen in the experiences of the earthquakes of Haiti (January 12th, 2010) and Chile (February 27th, 2010), the products can also be tailored during the requested time to meet users' requirements once the crisis is on-going, focusing on new necessities that might arise. This initiative, which ended in 2012, provided a system that can be applied to early warning and crisis prevention as well as to crisis management and rapid interventions in hot spots around the world. However, its main goal was oriented towards the identification and development of products, methodologies and pilot services for the provision of geo-spatial information in support to EU external relations policies and at contributing to the definition and demonstration of the sustainability of GMES global security services. Therefore, also in this case the use if remote sensing was not specifically oriented towards the damage assessment at individual buildings level.

As complementary activity to this one, a Geospatial Intelligence in support of EU external action, called *G-SEXTANT* [66] was also initiated. Its main goals were:

a) The preparation and delivery of pre-operational services, developed in the context of user-driven Support to External Action (SEA) scenarios;
b) The enhancement of products and services, as requested by users;
c) The development of a standardised portfolio of products and services.

Another extension of the above mentioned studies is a project the further enhanced the geo-spatial crisis management services in both rush (event-driven response) and non-rush (periodic or non-event-driven) modes by addressing the gap between research-based and operational services in the programmatic context of Copernicus services for Security applications. In particular, G-NEXT ([66]) supplies information and intelligence data to support the operations of the European External Action Service (EEAS), including mapping and geo-information products ready for use in emergency and crisis situations.

Already in 2011, a study lead by DLR ended showing the possibility to build a truly international Global Earth Observation System of Systems (GEOSS) to the benefit of global

societies. Such preliminary action on EUropean GEoss NEtwork (EUGENE - [67]) brought together European experts working in this area to improve collaboration and maximise their input into the international system. The two-year project aimed at specifically strengthening the coordination among national and regional programmes on *in-situ* and space-based earth observation and modelling techniques. Focusing on the specific areas of climate, water and disaster management, including earthquakes, the project looked at a number of questions. These included how to improve the systems architecture when pooling research infrastructure; how to promote data sharing to ensure informed decision-making in a timely manner; and how to ensure the integration of technological innovations across a network. The APhoRISM initiative is fully in line with such framework, since it aims at implementing such combination in both earthquake and volcanic disaster management.

The use cases selected also by the *INCREO* [68] - Increasing Resilience through Earth Observation - project include dam failure, storm surge and wave height, flood, earthquake and landslides and the transfer of solutions to a specifically multi-risk prone test site. From a technical point of view, the IncREO solutions are based on state-of-the-art methodologies. These were implemented by means of up-to-date mapping and modelling / procedures and, finally, disseminated to the relevant stakeholder groups

The overall objective of this kind of initiative shows that the ability to provide actors responsible for disaster management, risk prevention, civil protection and also spatial planning with EO-based solutions can significantly contribute to an improved preparedness and mitigation planning for areas that are highly vulnerable to natural disasters and may also support already noticeable climate change trends. These solutions were adjusted to the users and end-users needs and reflect on short-term climate change scenarios and related legislature both on national, supranational and European level. However, in this kind of study a multi-risk oriented concept *per-se* addressed any type of natural disaster, including earthquakes, thus not specifically addressing the problem of fast post-earthquake damage assessment at buildings level that we intend to address in this work.

The problem of providing near-real–time image communications and satellite earth observations for situational maps during emergencies and crisis management was addressed in the *GEOPICTURES* [69] project as well, which is related to GMES and Earth observation with position-based image and sensor communications technology for universal rescue, emergency and surveillance management. Successful field validations prove global coverage via both terrestrial and satellite communications, for fast, effective response when disaster strikes. In emergency situations, with associated danger to civilians and the environment, authorities require accurate and real-time information from the field in order to quickly take effective measures. Global emergency management must rely heavily on mobile satellite communications to secure operations anywhere and anytime, but SatCom has limited bandwidth and is slow and costly. The GEO-Pictures project incorporated geo-tagged photo, video, audio and sensor data for Situational Awareness, and improved interpretation of Global Monitoring for Environment and Security (GMES) Earth Observation (EO) information. The integrated system provides near-real–time situational maps with a fusion of required data establishes a missing link in larger-scale emergency management support. The satellite time- and location-referenced multi-input and multi-output communications system enables efficient distribution of large image files like high-resolution photos, video and satellite Earth Observation (GMES/Copernicus) data over unknown and capacity-limited

satellite or mobile channels without broadband infrastructure. GEO-Pictures was also used during several real events worldwide, including the Youth Olympic Games in 2012, Thailand/Bangkok flooding in 2011 and several UN operations.

Ground instability can be dangerous and costly, yet information on these phenomena has, to date, been difficult to obtain. PanGeo ([70] - Protecting EU cities from ground instability hazards) provides free access to ground instability geohazard information for many of Europe's largest cities. Users of the PanGeo service include local authorities, civil protection agencies, geological surveys, insurers and businesses providing environmental and land reporting services and of course the general public. This work aimed at providing a 'ground stability layer', which describes the spatial location and extent of geohazards for all the mapped towns.   Each polygon, within the ground stability layer, is linked to a full interpretation made by that country's National Geological Survey. The combination of data performed in this work is mainly given by the following information:

- Satellite measurements of ground and building movement
- Geological information already held by National Geological Surveys

Also included in the PanGeo product is the European Commission's Urban Atlas land use data; this provides insights into what types of urban land use are affected by geohazards described in the ground stability layer. During the study, a free online geohazard information for 52 of the largest towns in Europe was provided, mapping geohazards that could potentially affect up to 13% of the EU population.

Galileian Plus in Italy, also carried out a specific use of radar data from COSMO-SKYMed constellation,, in a project that was funded by the Italian Space Agency. In this project, called *VULSAR* [71] Seismic VULnerability assessment via SAR data, the main goal was the development of an informatic platform based on GIS able to create large-scale semi-automatic seismic hazard maps, including an evaluation of the vulnerability of the built facilities through the SP-BELA model. Following this line of research, the Italian company Galileian Plus S.r.l (GPLUS) developed an innovative adaptive filtering and classification algorithm, which was patented in 2009 with European Patent Number EP2146315 [67]. It was designed following a major scope: to optimize and "opticize" the filtered image, by taking into account the specificity of radar remote sensing and facilitating the automatic detection of features. The patented methodology aims to reduce noise primarily in areas affected by speckle, keeping the original spatial resolution where noise is not present.

The focus of another study, funded by the DG-ECHO (Humanitarian Aid and Civil Protection European Directorate), was also to provide Civil Protection authorities with accurate and up-to-date maps of seismic hazard, urban vulnerability, risk of exposed people and elements at local scale,. The state-of-the-art population evacuation models for two European pilot sites, heavily struck by earthquakes in Western Athens (GR) and Abruzzo Region (IT) were examined. This study, *MASSIVE* [73] - Mapping Seismic Vulnerability and Risk of Cities – is an interesting attempt to provide Civil [74] with evaluation on seismic hazard considering the three main parameters:

- the seismic sources that produce the hazard
- the ground motion that is attenuated away from the earthquake epicentre
- the local soil conditions defined from a geological map

In this project, coordinated by the National Observatory of Athens, a quantitative expression of earthquake risk, R, for a region was adopted as the convolution of seismic hazard (PGA) and building vulnerability. In many cases, especially in metropolitan cities, the hazard might be low but the risk is high due to the high average density of buildings. This recognition was verified with the earthquake of 7 September 1999 in Athens which was of moderate physical size but caused the most costly impact ever reported from earthquakes in Greece.

It is interesting to note that, even though many studies and initiatives have been carried out in the years, still the evaluation at building level of the post-earthquake damage is a complex task in which the use of advanced and highly sophisticated information from satellites has not been standardised and still need to be fully explored. In this work, we will examine how the use of in-situ data for individual buildings structures can improve the evaluation of post-earthquake damage assessment.

3.2 REMOTE SENSING FOR EARTHQUAKE DAMAGE ASSESSMENT

Besides the above mentioned projects, there are lots of other significant works in literature which clearly show the interest towards the topic of earthquake management and, more specifically, towards the identification of indicators to assess earthquake risk.

Among these works, it is worth mentioning the project Armagedom [74] that is fully dedicated to large-scale seismic risk analyses. Many works concentrate their focus on evaluating risk through the convolution of hazard and vulnerability. Armagedom, validated on different urban seismic contexts, offers different levels of precision depending on the surface areas examined. In order to meet differing project targets, three levels of seismic risk assessment were defined based on the macroseismic and mechanical approaches for vulnerability and damage estimation presenting different levels of precision, namely:

- Level N0: estimates seismic risk on a regional territorial scale based on the macroseismic approach and existing statistical data;
- Level N: yields the seismic risk at a district level based on the macroseismic approach and on visual evaluation of the vulnerability of structures over an itinerary in the area to be analysed;
- Level N2: also establishes the seismic risk at a district level, but the hazard description is represented by a spectrum and vulnerability is estimated based on mechanical models.

The software modules outputs provided by Armagedom are illustrated by maps derived from the seismic risk analyses performed.

A thorough review of the products available or developed in the past years on the use of remote sensing for vulnerability and damage assessment on urban context is reported also

by Panagiota [76] et al, considering both the two conditions of pre- and post-earthquake assessment. From such review of the state of the art in damage and vulnerability assessment in urban areas by using remote sensing, we can state:

- In the years, many attempts of data fusion, at different levels (pixel, object or decision) have been adopted stimulating the development of different algorithms now that VHR images are more easily available

- The need for different ancillary data, such as DEM (Digital Elevation Model), GIS, in-situ, is increasing to optimise data fusion with SAR and VHR images

- Nowadays there is a lack of a clear damage assessment methodology especially for urban areas when the 3D characteristics of the buildings may be difficult to assess using only remote sensing information and also because of the different morphology of urban areas.

The contribution of Remote Sensing to the evaluation of vulnerability and damage pre- or post-earthquake events may come from different parameters. A list of attributes derived from remote sensing, as summarised by Panagiota et al. is reported in Table 3-1.

Table 3-1: List of attributes of Vulnerability from Remote Sensing *(Source: Panagiota et al. [76])*

Buildings	Homogeneous areas
Footprint: (boundaries of target building)	Boundaries of target area
Roof Shape: flat or gable	Number of buildings included in the area
Shape of the building:(square, rectangular, round)	Concentration of buildings within the target area: (dense or sparse)
Size of the building(Small, Big)	Identification of homogeneous areas (Downtown, Residential, Commercial, Recreational areas)
Average height	Average height of buildings within the target area
Land use (Buildings, road, river, sport park, bridge)	
Land Cover(Street, Avenue, tree, water, shadow)	
Adjacent of a building:(Building, street, tree)	
Distance of adjacent buildings	
Characterization of well shaped buildings: (Yes or No)	
Total square footage	

3.2.1   The APhoRISM Project Goals

The present PhD Thesis is performed in the framework of the APhoRISM project [77] (Advanced Procedures for Volcanic and Seismic Monitoring) - Grant Agreement No. 606738. The Seismic product tool expected to be released in the APhoRISM project is namely the APE tool, i.e. the 'A Priori information for Earthquake damage mapping'. The proposed approach aims at generating maps of damage caused by a seism using satellite remote sensing data and ground data.

Soon after a seismic event a fast and, as much as possible, accurate damage map of the hit urban area is extremely important to guide and support rescue team operations. In this respect, Earth Observation data demonstrated their capability to provide useful information, especially if the earthquake occurs in remote regions or communication infrastructures fail.

As briefly shown in the previous paragraphs, we can find a wide literature concerning the use of EO data to investigate earthquake damages, but usually the approaches are only based on change detection techniques and classification algorithms. The currently damage mapping procedures usually do not account for information about soil vulnerability, which can play a very important role on local amplification effects. Moreover, the seismological information is not accounted as well, such as acceleration data that are strongly linked to collapsed building or heavy damage.

The novelty of APE relies on the exploitation of a-priori information derived by InSAR time series to measure surface movements, shakemaps obtained from seismological data, and other vulnerability information available a-priori. Such a-priori information is then integrated with change detection maps to improve accuracy and limit false alarms.

More details on the APE procedure developed in the framework of the APhoRISM project are provided in the following Par. 3.3.

However, we can anticipate that the seismic product proposed by APhoRISM is a Likelihood Index Damage Map (LIDaM). The LIDaM will be composed of polygons representing groups of buildings or single buildings and infrastructures. Each polygon will have an attribute describing the likelihood of the considered object being collapsed or strongly damaged.[78]

The product will be delivered as a shape file with attributes and/or as a raster image colour, coded according to the considered attribute. Metadata will include time tag, cartographic projection, quality indices and input data used in the processing. [79]

The product is expected to have a spatial scale related to the input data available for processing. According to whether the elementary unit is a single building or a cluster of buildings, the attribute may be the probability of being heavily damaged or the collapse ratio, according to the standard definition, with associated confidence intervals.

3.3 DEFINITION OF THE DATA FUSION NEEED FOR APE

As anticipated in the previous paragraph, the APE method is expected to provide an earthquake damage product obtained by integrating a-priori information and change detection images.

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). In particular, the IPTA - Interferometric Point Target Analysis - method is adopted. The InSAR velocity map allows us measuring possible slow surface deformation related to natural and/or anthropogenic activities. The outcomes of IPTA coupled with geological information and data concerning the stratigraphy of the studied area, are used to obtain the soil vulnerability of this area and to take into account any site effect locally amplifying the seismic waves. In addition to soil vulnerability, the APE method takes into account the building vulnerability, in order to finalise the Vulnerability Scenario. The latter contributes to the preparation of the a-priori information layers.

The a-priori information is then completed as soon as an earthquake occurs and the seismological Institutes, in charge of the seismic accelerometer networks, generate the shakemaps that provide information about the soil velocity and acceleration in the epicentral region. At this step, the a-priori information layers are ready and constitute the first input to the data fusion algorithm.

Furthermore, as soon as SAR data are available, a coseismic interferogram is also generated to measure the displacements and restrict the damage evaluation analysis to the regions affected by surface deformation.

During the crises phase, the APE generates the change detection maps using Optical or SAR or both kinds of images. The changes highlight building damages, at a scale of groups of buildings or single buildings, depending on the spatial resolution of the available satellite data.

The a-priori information is then integrated with the change detection products using a data fusion algorithm that is defined through the work performed in this PhD thesis.

The APE method, by integrating a-priori information and change detection images, provides a Likelihood Index Damage Map (LIDaM), which represents a map of groups of buildings or single buildings and infrastructures that are likely collapsed or strongly damaged.

The following picture shows the overall flowchart of the APE procedure within the APhoRISM project.

In the following picture, the yellow boxes stand for pre-event available data blocks. The red boxes are instead related to post-event available data. The green boxes represent information that is calculated by the cyan modules, which represent different routines.

For the sake of clarity, in the framework of the APhoRISM project, the different routines are under different partners' responsibilities. In this work, we focus only on the Fusion Module and on its input and output data, regardless of who has the responsibility to provide the input information.
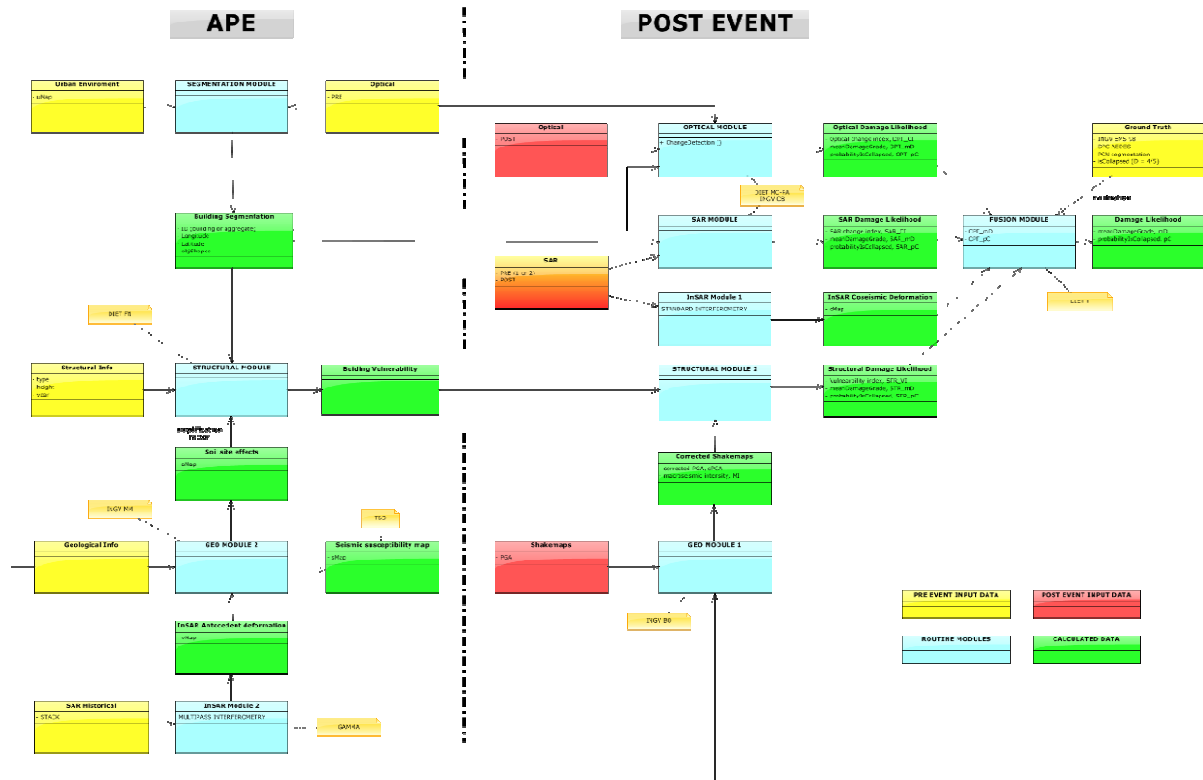
Figure 3-1: APE Flowchart

As briefly anticipated in this paragraph, the goal of the APhoRISM data fusion module that was developed in this work, is to fuse the information coming from the available remote sensing data (e.g. SAR, Optical, etc…) with in-situ information (e.g. structural analysis building information) and geohazard related data, when available.

Before entering the details of the selected procedure, we are going to review in the following chapter the main data fusion methods.

In this chapter, we reviewed the current state-of-the-art in the use of remote sensing for damage assessment and, in particular, for post-earthquake damage assessment. We were able to confirm in our review that, following the already mentioned UNISDR approach, international, national and regional efforts have been particularly strengthened in the recent years in the use of remote sensing to improve emergency response procedures. Our project can be positioned in the broader perspective of this worldwide scenario as one of the possible attempts to make use of remote sensing to improve the preparedness and reaction in post-earthquake emergencies. We introduced at the end of this chapter the general flowchart of the APE procedure that was designed by the project team. We can clearly see how the fusion module represents an integration of data coming from the structural (*a-priori* data), optical, and SAR sensors. The SAR sensor may contribute to the damage assessment through both a change detection index (as for the optical module) and through interferometric data to measure possible surface movements, which affect the hazard geological evaluation.

4 DATA FUSION APPROACH FOR EARTHQUAKE VULNERABILITY AND DAMAGE ASSESSMENT

As briefly anticipated in the previous chapter, the different types of information and data that structural engineers can provide and remote sensing (from radar and/or optical sensors) experts can identify, is expected to be integrated in order to provide a more detailed and accurate evaluation of each building damage likelihood. At this point, we need to better understand how to fuse the entire or partial information available. In fact, in view of a future generalisation of this problem, we have to note that we do not always have the possibility to access all EO data or even gather information on the built environment, especially in earthquake events happening in developing countries.

However, before extending our problem, in this chapter we go through a general review of data fusion methodologies, since one of the aspects we have to deal with includes also the level of fusion that we intend to pursue in the APE procedure.

For example, all the different modules used as input for the data fusion one are able to pre-elaborate their individual likelihoods; therefore, we can elaborate a fusion module that can process the information at different levels, either processing individual damage likelihoods or using the features that each module brings in. We exclude the case for which the fusion module takes raw data as input, since we want to maintain a highly modular approach for our algorithm.

After providing a broad review of the main data fusion methods, we start entering the Bayesian approach and, more specifically, the Bayesian network that will be further implemented in our APE algorithm.

However, before entering more in detail on how the fusion module is conceived in this work, we provide here in the following some general information on data fusion techniques and on the selection of the most suitable one for our purposes.

The most accepted definition of data fusion was provided by the Joint Directors of Laboratories (JDL) of the Department of Defence (DoD) [80]:

"*A multi-level process dealing with the association, correlation, combination of data and information from single and multiple sources to achieve refined position, identify estimates and complete and timely assessments of situations, threats and their significance*".

In general, data and information fusion can provide enhancement to the outcomes of processes for solving various application problems. Some advantages of carrying out DIF (Data & Information Fusion) include [81]:

- improvement in the accuracy of data, as well as reduction in uncertainty and ambiguity within data,
- improvement in Situation Awareness (SAW) and inference that lead to better decision making.

4.1   REVIEW OF DATA FUSION METHODS

Many classification models for the Data fusion process have been provided in the years; however, one of the most common and internationally shared classification methods for data fusion is the JDL Data Fusion Classification.

This classification, which is the most popular in the Data fusion community, has also been updated in the years, but in general the overall revised scheme can be presented as in the following picture:
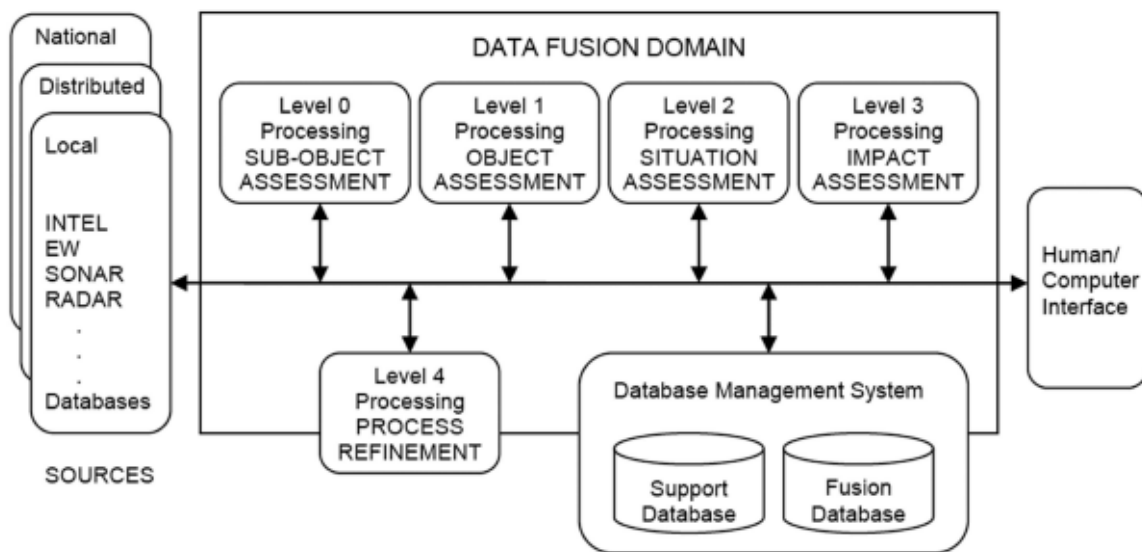
Figure 4-1: Revised JDL DF model *(Source: Steinberg et al. [82])*

Despite all the suggested revisions existing in literature and the different approaches, the general definitions of the levels of fusion are reported in the following table:

Table 4-1: Data Level

FUSION LEVEL	FUSION TYPE	DESCRIPTION
Level 0	Sub-Object Data Assessment	Estimation and prediction of signal/object observable states on the basis of pixel/signal level data association and characterization
Level 1	Object Assessment	Estimation and prediction of entity states on the basis of observation-to-track association, continuous state estimation (e.g. kinematics) and discrete state estimation (e.g. target type and ID)

FUSION LEVEL	FUSION TYPE	DESCRIPTION
Level 2	Situation Assessment	Estimation and prediction of relations among entities, to include force structure and cross force relations, communications and perceptual influences, physical context, etc.
Level 3	Impact Assessment	Estimation and prediction of effects on situations of planned or estimated/predicted actions by the participants; to include interactions between action plans of multiple players (e.g. assessing susceptibilities and vulnerabilities to estimated/predicted threat actions given one's own planned actions)
Level 4	Process Refinement	It is an element of Resource Management. These are adaptive data acquisition and processing to support mission objectives

Different types of classifications of data fusion methods are available depending on their main viewpoint. As summarised in Castanedo, 2013 [80], the following table shows the main ones:

Table 4-2: Data Fusion Methods Classifications *(Source: Castanedo, 2013 [80])*

Type of Data Fusion Classification	Description
Classification Based on the Relations between the Data Sources	***Complementary***: input sources show different parts of the same scene, thus completing the information provided
	Redundant: two or more input sources provide information of the same target, thus enabling the increase of confidence
	Cooperative: combination of input information into a more complex than the original ones
Dasarathy's Classification	***Data in-data out*** (DAI-DAO): The algorithms employed at this level are based on signal and image processing algorithms
	Data in-feature out (DAI-FEO): raw data from the sources employed to extract features or characteristics that describe an entity in the environment
	Feature in-feature out (FEI-FEO): also known as feature fusion, symbolic fusion, information fusion or intermediate level fusion
	Feature in-decision out (FEI-DEO): set of features as input and provides a set of decisions as output.
	Decision In-Decision Out (DEI-DEO): It fuses input decisions to obtain better or new decisions.

Type of Data Fusion Classification	Description
Classification Based on the Abstraction Levels.	*Signal Level*: directly addresses the signals that are acquired from the sensors
	Pixel Level: operates at the image level and could be used to improve image processing tasks
	Characteristic: employs features that are extracted from the images or signals (i.e., shape or velocity)
	Symbol: at this level, information is represented as symbols; this level is also known as the decision level
JDL Data Fusion Classification The five levels of data processing are shown in Table 2-1	*Sources*: Different types of sources can be employed, such as sensors, a priori information (references or geographic data), databases, and human inputs;
	Human-Computer Interaction (HCI): inputs from the operators and outputs to the operators. HCI includes queries, commands, and information on the obtained results and alarms
	Database Management System: it stores the provided information and the fused results
Classification Based on the Type of Architecture	*Centralized Architecture*: all of the fusion processes are executed in a central processor that uses the provided raw measurements from the sources.
	Decentralized Architecture: a network of nodes in which each node has its own processing capabilities and there is no single point of data fusion. Therefore, each node fuses its local information with the information that is received from its peers
	Distributed Architecture: measurements from each source node are processed independently before the information is sent to the fusion node; the fusion node accounts for the information that is received from the other nodes.
	Hierarchical Architecture: combination of decentralized and distributed nodes, generating hierarchical schemes in which the data fusion process is performed at different levels in the hierarchy

Depending on the level at which the fusion occurs, the available data fusion techniques can be classified into three non-exclusive categories:

  I.   Data association
  II.   State estimation
  III. Decision fusion

The most common techniques are those reported in the following table:

Table 4-3: Common techniques for data fusion at different levels

LEVEL	TECHNIQUE	REFERENCE
DATA ASSOCIATION	Nearest neighbour & K-means	[83][84]
	Probabilistic Data Association (PDA)	[85]
	Joint Probabilistic Data Association (JPDA)	[86]
	Multiple Hypothesis Test (MHT)	[87][88]
	Distributed Joint Probabilistic Data Association	[89]
	Distributed Multiple Hypothesis Test	[90][91]
	Graphical Models • Direct GM : Bayesian • Indirect GM: Markov Random Field	[92][93]
STATE ESTIMATION METHODS	Maximum Likelihood & maximum posterior	[94]
	Kalman Filter	[95][96]
	Particle Filter	[97][98]
	Distributed Kalman filter	[99][100]
	Distributed Particle Filter	[101][102][103]
	Covariance Consistency Methods: • Covariance Intersection • Covariance Union	[104][105]
DECISION FUSION METHODS	Bayesian Methods	[106][107]
	Dempster-Shafer Inference	[108][109]
	Abductive reasoning: • Fuzzy logic • Neural Network	[110][111]
	Semantic Methods	[112][113]

The concept of fusion can also be extended to image fusion, which also can be performed at three different processing levels according to the stage at which the fusion takes place (namely Pixel, feature or Decision level). Image fusion at pixel level means fusion at the lowest processing level referring to the merging of measured physical parameters.

Image fusion at feature level requires the extraction of objects recognised in the various data sources, e.g., using segmentation procedures. Features correspond to characteristics extracted from the initial images, which are depending on their environment such as extent, shape and neighbourhood. These similar objects (e.g., regions), from multiple sources, are assigned to each other and then fused for further assessment, using statistical approaches or Artificial Neural Networks (ANN).

Image fusion at Decision - or interpretation level fusion represents a method that uses value-added data (e.g. ancillary data) where the input images are processed individually for

information extraction. The obtained information is then combined applying decision rules to reinforce common interpretation, resolve differences and furnish a better understanding of the observed objects [117].

4.2 BAYESIAN DATA FUSION APPROACH

The approach to the fusion module for the APhoRISM project builds on the use of Bayesian networks for representing statistical dependencies. A Bayesian network (see Par. 4.2.3) is a graph-based model of joint multivariate probability distributions that captures properties of conditional independence between variables. This kind of model is attractive for its ability to describe complex stochastic processes and because it provides a clear methodology for learning from noisy observations.

In the following paragraphs, we first review the Bayes' theorem, then exploit why the sequential inference can be helpful for our approach and finally examine Bayesian Networks.

4.2.1 The Bayesian statistical framework

Bayes' theorem is a result in probability theory that relates conditional probabilities. If A and B denote two events, P(A|B) denotes the conditional probability of A occurring, given that B occurs.

The two conditional probabilities P(A|B) and P(B|A) are in general different. Bayes theorem gives a relation between P(A|B) and P(B|A). An important application of Bayes' theorem is that it gives a rule how to update or revise the strengths of evidence-based beliefs in light of new a-posteriori evidence. Bayes' theorem relates the conditional and marginal probabilities of stochastic events A and B:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Eq. 4-1

Each term in Bayes' theorem has a conventional name:

- P(A) is the **prior probability** or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B;
- P(A|B) is the conditional probability of A, given B. It is also called the **posterior probability** because it is derived from or depends upon the specified value of B;
- P(B|A) is the conditional probability of B given A, also known as **likelihood**;
- P(B) is the prior or marginal probability of B, and acts as a normalising constant.

We can also represent the theorem by using the following expression, by use of the total probability law, which introduces the concept of "*complementary event*" (¬A: not A), then we can write the equation above as:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)}$$

Eq. 4-2

Inference, or model evaluation, is the process of updating probabilities of outcomes based upon the relationships in the model and the evidence known about the situation at hand [122]. When a Bayesian model is actually used, the end user applies evidence about recent events or observations. This information is applied to the model by "instantiating" or "clamping" a variable to a state that is consistent with the observation. Then the mathematical mechanics are performed to update the probabilities of all the other variables that are connected to the variable representing the new evidence.

After inference, the updated probabilities reflect the new levels of belief in (or probabilities of) all possible outcomes coded in the model. These beliefs are mediated by the original assessment of belief performed by the author of the model.

The beliefs originally encoded in the model are known as prior probabilities, because they are entered before any evidence is known about the situation. The beliefs, computed after evidence, are known as posterior probabilities, because they reflect the levels of belief computed in light of the new evidence.

The Naive Bayes model assumes that instances fall into one of a number of mutually exclusive and exhaustive classes [93]. Thus, we have a class variable C that takes on values in some set.

The model also includes some number of features $(X_1 \ldots X_n)$ whose values are typically observed. The naive Bayes assumption is that the features are conditionally independent given the instance's class. In other words, within each class of instances, the different properties can be determined independently. More formally, we have that

$$(X_i \perp X_{-i}|C) \quad \text{for all i}$$

Eq. 4-3

Where $X_{-i} = \{X_1, \ldots, X_n\} - \{X_i\}$

This model can be graphically represented by using the Bayesian network of Figure 4-2. Based on these independence assumptions, the model factorises as:

$$P(C|X_1, \ldots, X_n) = P(C) \prod_{i=1}^{n} P(X_i|C)$$

Eq. 4-4

In this model, we can represent the joint distribution using a small set of factors: a prior distribution P(C), specifying how likely an instance is to belong to different classes a priori, and a set of Conditional Probability Distributions (CPDs) $P(X_j | C)$, one for each of the *n* finding variables. These factors can be encoded using a very small number of parameters. For example, if all of the variables are binary, the number of independent parameters required to specify the distribution is 2n+1.
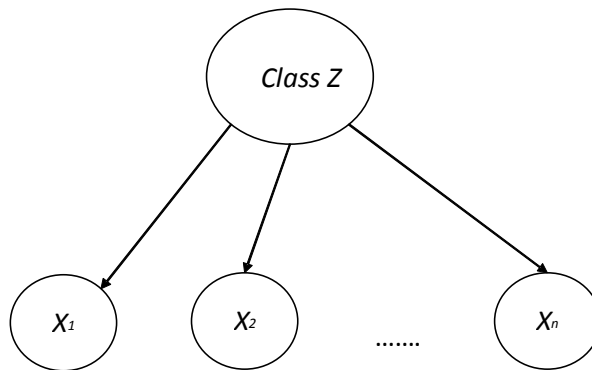
Figure 4-2: Example of Bayesian Network

4.2.2   Sequential Bayesian Inference

In this section, we briefly explain, following the definitions given above how the posterior [123], $p(C|X_1, X_n)$ evolves as we observe increasingly more data points X_i. Actually, we note that we can compute the posterior incrementally since here the data are assumed independent (conditioned on C), e.g. for $X = (X_1; X_2; X_3)$:

$$\lambda_0 = p(C)$$
Eq. 4-5

$$\lambda_i = p(C|x_i) \quad \text{for all } i$$
Eq. 4-6

$$P(C|X_1, \ldots., X_n) = \frac{P(C|X_1, \ldots., X_{n-1})\lambda_n}{P(C|X_1, \ldots., X_{n-1})\lambda_n - \left(1 - P(C|X_1, \ldots., X_{n-1})\right)(1 - \lambda_n)}$$
Eq. 4-7

So, more generally, we can treat the posterior having observed $(X_1,\ldots, X_k)$ as the `*prior'* for the remaining data $(X_{k+1},\ldots, X_N)$ and obtain the equivalent result to seeing all the data at once.

Applying the above mentioned concept to the context presented in Par. 3.2.1 of the APhoRISM project, we can consider our specific case, i.e., trying to merge the information derived from the Optical (OPT), SAR, Geological (GEO) and Structural (STRUCT) modules of the APE process. Referring to Par. 4.1, we can perform the fusion at different levels of information:

a) Using the classifications as already carried out in the respective modules of the APE algorithm (SAR, optical and structural modules) - *Decision Fusion*
b) Using the probabilities of damage as evaluated by the relevant modules of the APE algorithm – *Rule Fusion*
c) Using a more complex network that establishes relations among the different features coming from the different modules, thus mixing optical, SAR and structural features – *Feature Fusion*

If we assume to operate as in case a), the classifications performed by the SAR, OPT and STRUCT modules can be summarised as simply binary nodes:

- 0 = Not collapsed
- 1 = Collapsed

It is possible to combine all these outcomes from different modules, for instance by adopting a majority rule. Alternatively, we can combine the probability of damage D provided by the individual modules (being D a discrete or continuous random variable), or even combine the observations (or features) computed for each sample (i.e., each building in case of the high resolution product).

Let **X** be the observation vector representing the state for the optical module observations X_{OPT} and that of the SAR module observations X_{SAR}. The observations of the Vulnerability module, related to the structural evaluations, are indicated as X_{STRUCT}. Assuming the class variable D_i is the damage grade *i* (with i = 0, ..,5) we want to infer, the Bayes' theorem can be written as:

$$P(D_i|X_{SAR}, X_{OPT}, X_{STRUCT}) = \frac{P(X_{SAR}, X_{OPT}, X_{STRUCT} |D_i) \cdot P(D_i)}{P(X_{OPT}, X_{SAR}, X_{STRUCT})} \qquad \text{Eq. 4-8}$$

In summary, we can consider our problem as a **Bayesian Inference** problem. This means that we want to answer the following question: "*How do we update our subjective belief (e.g. in the building being damaged) after we observe some data (e.g. SAR or OPT) or account for some other features (e.g., geological, structural)*? ". In practice this means that we want to estimate the PDF (probability density function) of unknown states given some known observations and an initial (prior) probability distribution.

In the following, we consider that the probability of changes in the images (SAR and Optical) depends only on the damage grade which affects the imaging process (i.e., the appearance of the building after the earthquake) and not on the prior vulnerability of the building, the actual seismic intensity or the geological hazard (although the visibility of the damage can be affected by factor like for example terrain slopes that may determine geometric distortions or other effects, or density of the urban settlements). It means that the variables X_{SAR}, X_{OPT} are conditionally independent on X_{STRUCT} given damage grade *D*. This hypothesis allows formulating the posterior PDF in two different ways suitable for the integration between modules. Using the Bayes' theorem, it is straightforward to write:

$$P(D_i|X_{SAR}, X_{OPT}, X_{STRUCT}) = \frac{P(X_{SAR}, X_{OPT}|D_i) \cdot P(X_{STRUCT}|D_i) \cdot P(D_i)}{P(X_{OPT}, X_{SAR}, X_{STRUCT})} \qquad \text{Eq. 4-9}$$

Using the Bayes' theorem again to introduce the posterior probability coming from the structural module (the latter was modelled through the Beta function, as explained in Par. 5.2.1), and adding the hypothesis of independence of SAR and Optical features associated to the same class of damage, we can express n observations generally as a vector $\mathbf{X}=(X_1, X_2, \dots X_n)$, we can find how the posterior $P(D|X_1, \dots X_n)$ evolves as we observe increasingly more features X_i [123] Actually, we note that we can compute the posterior incrementally since here the data are assumed independent again (conditioned on *D*). The incremental formulas are Eq. 4-6 and Eq. 4-7. More generally, we can treat the posterior, having observed $(X_1,...X_k)$ as the 'prior' for the remaining data $(X_{k+1}, ..., X_n)$ and obtain the equivalent result to seeing all data at once. We can translate this in our APhoRISM case, as in the following.

Let us assume D is the damage grade and start with only two observations X and Y to see how the posterior probabilities combine. We then assume:

$$P(D|X) = \frac{P(X|D) \cdot P(D)}{P(X|D)P(D) + P(X|\neg D)P(\neg D)} \qquad \text{Eq. 4-10}$$

$$P(D|Y) = \frac{P(Y|D) \cdot P(D)}{P(Y|D)P(D) + P(Y|\neg D)P(\neg D)} \qquad \text{Eq. 4-11}$$

The posterior, which accounts for both observations, can be evaluated making the assumption of conditional independence, and using Bayes' theorem iteratively as shown in the following formulation:

$$P(D|XY) = \frac{P(X,Y|D)P(D)}{P(X,Y)} = \frac{P(X|D)P(Y|D)P(D)}{P(X,Y)} = \frac{P(D|X)P(X)P(D|Y)P(D)}{P(D)P(D)P(X,Y)}$$
$$= \frac{P(D|X)P(X)P(D|Y)P(Y)}{P(D)P(X,Y)} \qquad \text{Eq. 4-12}$$

Then the joint probability, applying the conditional independence and the Bayes' theorem again, is given by:

$$P(X,Y|D) \propto P(X,Y,D) + P(X,Y,\neg D) = P(X,Y|D)P(D) + P(X,Y|\neg D)P(\neg D) =$$
$$= P(X|D)P(Y|D)P(D) + P(X|\neg D)P(Y|\neg D)P(\neg D)$$

Eq. 4-13

$$P(X,Y|D) = \frac{P(D|X)P(X)}{P(D)}\frac{P(D|Y)P(Y)}{P(D)}P(D) + \frac{P(\neg D|X)P(X)}{P(\neg D)}\frac{P(\neg D|Y)P(Y)}{P(\neg D)}P(\neg D)$$
$$= \frac{P(D|X)P(X)P(D|Y)P(Y)}{P(D)} + \frac{P(\neg D|X)P(X)P(\neg D|Y)P(Y)}{P(\neg D)}$$

Eq. 4-14

Considering that $P(\neg D|X)$=1-$P(D|X)$ and $P(\neg D)$ = 1- $P(D)$, we find that Eq. 4-14 can written as:

$$P(D|XY) = \frac{P(D|X)P(D|Y)}{P(D|X)P(D|Y) + \frac{P(D)[1 - P(D|X)][1 - P(D|Y)]}{1 - P(D)}}$$

Eq. 4-15

which represents the posterior probability combination in the case of a 2-class (binary) problem with equal prior probability, i.e., P(Di)=0.5 with i =1,2.

In the following, we will use also such approach, since it merges homogeneous quantities produced by the different modules to be combined as posterior probabilities of damage. We exploit this approach in Par. 5.5.3, where we illustrate how the posterior distribution updates with increasing amounts of data.

4.2.3 The Bayesian Networks

Probabilistic graphical models use a graph-based representation as the basis for compactly encoding a complex distribution over a high-dimensional space. The nodes (or ovals) correspond to the variables in our domain, and the edges correspond to direct probabilistic interactions between them.

"*Bayesian networks build on the same intuitions as the naive Bayes model by exploiting conditional independence properties of the distribution in order to allow a compact and natural representation. However, they are not restricted to representing distributions satisfying the strong independence assumptions implicit in the naive Bayes model. They allow us the flexibility to tailor our representation of the distribution to the independence properties that appear reasonable in the current setting*". [93]

A Bayesian network [124] is a specific type of graphical model that is represented as a Directed Acyclic Graph (DAG). Each node of the DAG represents an object or event that exists in the real world and is called variable (or state). Causal relations between nodes are

represented by the edges linking them. For any given edge between the nodes (variables), if a causal relationship exists, the edge will be directional, leading from the cause variable (parent) to the effect variable (child). In Figure 4-2, a simple network is shown where variable class Z represents the parent. Variables X_1, X_2 and X_n are conditionally independent, given variable class Z. This is not the same as saying that X_1, X_2 and X_n are totally independent, since their independence is conditioned to the state of node Z.

The framework of a graphical model allows experts [125] to concentrate on building up the qualitative structure of a problem, before even beginning to address issues of quantitative specification. As emphasised by Pearl (1998) [92], such models are intended to encode natural judgments of relevance and irrelevance and can be formed prior to any probabilistic consideration. Nodes in the graph represent variables; missing links in the graph represent the irrelevance properties. A directed edge is put between two variables to represent a direct influence. To avoid inconsistencies, a sequence of directed edges is not allowed to return to its starting node: the graph is thus a Direct Acyclic Graph (DAG).

As mentioned above, the core of the Bayesian network representation is a directed acyclic graph (DAG) $\mathcal{G}$, whose nodes are the random variables in our domain and whose edges correspond, intuitively, to direct influence of one node on another. This kind of graph $\mathcal{G}$ can be viewed either:

- as a **data structure** that provides the skeleton for representing a joint distribution compactly in a factorized way or;
- as a **compact representation** for a set of conditional independence assumptions about a distribution.

This type of network allows using the so called "Explaining away" which is an instance of a general reasoning pattern called intercausal reasoning, where different intercausal reasoning causes of the same effect can interact.

Although based on the same statistical framework, Bayesian Networks (BN) are different from Bayesian Neural Networks (BNN). A BN is a graphic model, whereas a BNN is a neural network that uses the Bayesian inference to maximise the posterior probability distribution of the parameters (weights and biases) of the network in the training phase to avoid the problem of overfitting and generalization.

Referring to the problem faced in the APhoRISM project, we can start from a simplified Bayesian network as in the following Figure 4-3, where each node represents an available set of features and each arc connects a parent (influencing) node to a child (influenced) node. In case of a high-resolution damage assessment product, this network could be applied to features associated to each individual building.
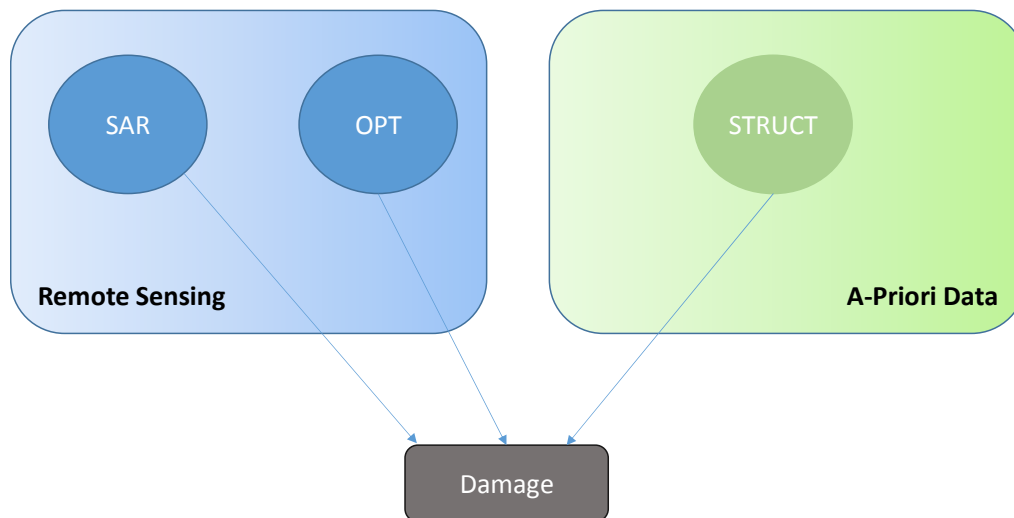
Figure 4-3: Simplified Bayesian Network (BN)

In this simplified topology of our network, we omit the relations among the parents. However, this step is interesting to start evaluating the different behaviours when considering a binary classification for each of the parents' nodes (i.e. a decision fusion approach) or when each parent node is described in terms of probability of damage (i.e. rule fusion) produced by different modules of APhoRISM to be merged. In the above shown examples, our SAR, OPT and STRUCT modules take discrete values. Such values are, in our hypothesis mutually exclusive and exhaustive, which means that the variable must take on exactly one of these values at a time.

Once the topology of the BN is specified, the next step is to quantify the relationships between the connected nodes. It consists in specifying the conditional probability distribution for each node. This is done in the learning phase which may take statistical relationships already learned and created manually by an expert, or based on theoretical knowledge of the problem at hand (perhaps expressed by an equation). Alternatively, it makes use of a number of "cases" (an experimental training set) provided to the network all together or one-by-one.

In our work, the software NETICA [126], one of the world's mostly widespread Bayesian network development software, was used. When using discrete variables NETICA creates the relationships between nodes in the form of a Conditional Probability Table (CPT) [127]. Each variable is assigned a CPT of the variable given its parents. For variables without parents, this is an unconditional (also called a marginal) distribution. Root nodes also have an associated CPT, containing only one row representing its prior probabilities. For Boolean networks a variable with n parents requires a CPT with 2^{n+1} probability entries. The conditional probability table specifies the probabilities, and must sum to one over all possible states.

Note that we can also define hidden variables (or latent variables), which is a way to easily describe correlation between child variables, even if there is no experimental evidence on such variables. In this case there are even more analogies with Neural Networks, being latent variable similar to hidden nodes of a NN.

The Learning algorithms used by NETICA to learn CPTs are: [129]

- Counting,
- Expectation-Maximization (EM)
- Gradient Descent.

The Counting algorithm is usually the one that performs better whenever there are no latent variables, and not much missing data or uncertain findings for the learning nodes or their parents. When learning the CPT of a node by counting, Netica only uses those cases that supply a definite value for the node and all of its parents. When learning the CPT of a node by counting, Netica only uses those cases that supply a definite value for the node and all of its parents. For all three algorithms, the order of the cases does not affect the results.

During Bayes net learning [128], we are trying to find the maximum likelihood Bayes net, which is the net that is the most likely given the data. If N is the net and D is the data, we are looking for the N which gives the highest P(N|D). Using Bayes rule, P(N|D) = P(D|N) P(N) / P(D). Since P(D) will be the same for all the candidate nets, we are trying to maximize P(D|N) P(N), which is the same as maximizing its logarithm:

$$\log(P(D|N)) + \log(P(N)) \qquad \text{Eq. 4-16}$$

The more data you have, the more important the first term will be compared to the second. There are different approaches to dealing with the second term log(P(N)), which is the prior probability of each net (i.e. how likely you think each net is before seeing any data). One approach is to say that each net is equally likely, in which case the term can simply be ignored, since it will contribute the same amount for each candidate net. Another is to penalize complex nets by saying they are less likely (which is of more value when doing structure learning). NETICA bases the prior probability of each net on the experience and probability tables that exist in the net before learning starts. If the net has not been given any such tables, then NETICA considers all candidate nets equally likely before seeing any data.

The first term log(P(D|N)) is known as the net's log likelihood , If the data D consists of the *n* independent cases d_1, d_2, … d_n, then the log likelihood is:

$$\log(P(D|N)) = \log(P(d_1|N)\,P(d_2|N)\,\ldots\,P(d_n|N)) = \log(P(d_1|N)) + \log(P(d_2|N)) + \ldots + \log(P(d_n|N)) \qquad \text{Eq. 4-17}$$

Each of the $\log(P(d_i|N))$ terms is easy to calculate, since the case is simply entered into the net as findings, and NETICA's regular inference is used to determine the probability of the findings.

In this chapter, we reviewed the data fusion methods. We then examined Bayesian networks in more detail, since BNs have been selected for our work. In fact, they have many capabilities that make them well suited for the proposed application. Most important among

all is their ability to update the probabilistic states of the variables upon receiving relevant information. Inference can also be implemented in Bayesian networks, which is a very important feature if we want to extend the results of our methodology to other cases. In fact, when evidence on one or more variables, e.g., increase in contrast or in building vulnerability, or observed states of components, can be entered into the BN and this information propagates throughout the network to provide up-to-date probabilistic characterizations of the building damage probability. This can be done in near-real time and under the uncertain and evolving state of information that is characteristic of the post-event period.

Such type of graphical approach for our data fusion model is implemented to assess post-earthquake building damage. Since our network – as will be better detailed in the next paragraphs – is not too complex or large, we also manage in this way to avoid the highly demanding computational requirements. Therefore, the selection of this method is also due to the relatively easy possibility to extend the results of our approach to future similar events without making the computational analysis too heavy.

5 TEST CASE: L'AQUILA RESULTS

In this chapter, we are going to test our BN on a real case based on L' Aquila earthquake which reached its peak in the event of April 6th 2009.

In this case, we have a set of data available to build the Ground Truth validation test set. The identification of this ground truth is the result of an integration of data coming from different sources, mainly from INGV and the Italian Civil Protection Department (DPC). Then, as previously anticipated, we provide a brief description of what the individual modules of the APE procedure (see Figure 3-1) calculate in terms of structural, optical and SAR modules. For what concerns remote sensing data, for this event, both COSMO-Skymed Radar and Quickbird VHR optical sensors are available thus allowing a complete remote sensing dataset. A simplified model that is described in Par. 5.2 provides the structural information. In addition, a brief summary of the information derived from the geotechnical module is considered here. Such details are provided as in the following paragraphs:

- Optical Damage Likelihood→OPT (see Par. 5.3)
- SAR Damage Likelihood → SAR (see Par. 5.3)
- Structural Damage Likelihood → STRUCT (see Par. 5.2)
- InSAR – Interferometric SAR → GEOTEC (see Par. 0)

The fusion module is implemented, using Bayesian networks. The results of different Bayesian networks are presented showing the step-by-step approach adopted which aims at generalising the methodology in order to further implement the network in future cases.

5.1 IDENTIFICATION OF THE GROUND TRUTH SURVEY

In this paragraph, we briefly summarise all the activities carried out for the reconstruction of a reference ground Truth dataset against which to validate our methodology. The test set refers to L'Aquila earthquake that took place on April 6th 2009.

The cartographic data available for L'Aquila test site are:
- GIS layers including building footprints and heights from:
  - Regione Abruzzo (Carta Tecnica Regionale)
  - Ministry of Environment
  - DEM SRTM
  - DEM Lidar
- A GIS layer from INGV consisting of building footprints manually drawn on a VHR optical image taken by the Quickbird satellite a day after the seism.

The ground truth data available for the selected test sites come from two different sources. The first one concerns on the survey performed by INGV Macroseismic team, while the second one was produced by the DPC (Italian Department of Civil Protection). The ground

truth coming from the DPC is divided into three different coverages that partly overlap namely AEDES, BT and COLLAPSES.

Here in the following a brief summary of the main results of the two sources:

- **INGV dataset**: INGV researchers, after the event collected information on the types of building, on their vulnerability class, and damage occurred, according to the European Macroseismic Scale 1998 (EMS-98) (Grünthal, 1998) [28]. As reported in Par. 2.5, the damage grade ranges from zero (0) to five (5) going from no damage to full collapse (see Figure 2-4). It is worth noticing that the inventory data was collected only by visual inspection, building-by-building, since INGV teams are not allowed to enter the buildings for safety reasons. More than 1700 buildings were surveyed in L'Aquila city, and a georeferenced vector file was produced to map the collected data on a GIS platform [130], [131]. The resulting map is shown in Figure 5-1 that shows also the statistics of the damage grade presented in terms of number of buildings labelled with different damage grades;

- **DPC dataset**: the overall DPC dataset includes three types of different sources, namely :
  - *DPC-AEDES dataset*: This survey was carried out during the six months following the seismic event by the Civil Protection Department (DPC) and includes a detailed review of the interior parts of the buildings. The final goal of this survey was to classify the building usability and assess the post-earthquake damage. The surveyed buildings are both for private and public use, and only monumental and historical buildings are excluded, since a specific survey is dedicated to them (see DPC-BT dataset). As explained previously in Par. 2.5, the complete form (Agibilità e Danno sugli Edifici pubblici e privati: AeDES), filled in for each building, includes more than 250 fields (columns) - see Figure 2-6. The most relevant for our analysis are the following:
    - Geometry identification (~5 columns) including information such as number of stories, height, volume, etc…
    - Structural typology (~40 columns)
    - Damage (60 columns)
    - Building regularity or additional modifications (~9 columns)

    A damage indicator and a vulnerability class were also calculated following the EMS'98 scale. CNR-ITC also geo-referenced all the buildings based on the Technical maps of Regione Abruzzo (CTR - Carta Tecnica Regionale). Figure 5-2a) shows an overview of the EMS'98 damage grade classification in the downtown area of L'Aquila derived from the AeDES forms as processed by CNR-ITC. Figure 5-2b) provides a statistical summary, i.e., the number of occurrence of different damage grades.

  - *DPC-BT dataset*: The DPC survey named BT (Beni Tutelati) is methodologically similar to the AeDES survey described above, with the only difference that it was carried out only on buildings with cultural heritage relevance, either public or private. 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. As a matter of fact, there is the possibility of duplication of the surveys of some of these buildings. The dataset available

for the APhoRISM project was provided by CNR-ITC in a geocoded format as for the AeDES form described above. This GIS layer does not include the full information as for the AeDES survey, but only refers to an average and a maximum damage level, as identified in localised areas of the building. An overview of the DPC-BT survey is shown in Figure 5-3 a) and b) and Figure 5-4 a) and b), reporting the mean and maximum damage, respectively. And summarising the statistics of the two (average and local maximum) damage indicators.

o *DPC-COLLAPSES dataset*: The AeDES survey was conceived for assessing the usability of the buildings after the seism. Therefore, it may be possible that the operational teams did not consider completely collapsed buildings. Then this dataset was derived from the GIS AeDES layer, by performing a second inspection on the polygons, identifying those that were associated to fully collapsed buildings. However, this dataset contains also some objects that are not included in the AeDES layer and does not contain any other information useful for the APhoRISM project for further vulnerability assessment. Furthermore, some buildings, here labelled as collapsed, can be marked with a damage level lower (or equal to) than 4 in the AeDES sheets, a situation that has to be properly managed. A survey of this data set is shown in Figure 5-5.

As shown in Figure 5-6, the frequencies of the damage classes are quite different among the various datasets. More specifically, the distribution of the INGV damages is mostly concentrated on average values (3) and only has few collapses (damage 5), namely seventy-five (75). The DPC-AEDES distribution is flatter, if compared to the previous one, since it has both many low level (0) and high level (5) damages. The DPC BT average damages look more like INGV distribution, however without the same percentage of collapses (damage 5). However, for the DPC BT dataset, the maximum damage feature is not useful for comparison purposes, since more than 50% of the buildings are marked with level 5 damage and these refer to localised damages only.

For the sake of the APhoRISM project purposes, considering that the geographical base maps are different, since the building polygons are either derived from the regional CTR (i.e., AeDES) or from satellite images (i.e., INGV), a rule for geometric correction was introduced to manage possible geolocation errors, and associate the different information to the right polygons. As a matter of fact, the polygons of the layers do not always have a one-to-one matching since this depends on which buildings were really surveyed and how a certain built up area was (or was not) split into single elements (i.e., polygons).

As already observed, some polygons can be found in different GIS layers even with different labels, including for the damage grade (an example is the possible superimposition among DPC AEDES and DPC BT). Therefore, when intersecting the layers different combinations are possible. There could be a '*one-to-one*' correspondences of polygons (generally a minimum intersection area of 15% is considered to assume that two polygons corresponds to the same building), but also '*one-to-many*', '*many-to-one*' or '*many-to-many*' correspondences can be found. Also in these cases, we compared only the polygons that are in common between the two datasets (i.e., that have an intersection area above a given threshold of ~15%).

The results of the analysis of the comparison among the datasets following this superimposition rule show a high dispersion mainly due to the different human evaluations on damage classes, even though coming from expert teams working on the field. It is worth pointing out that different datasets may provide different ground truths in a remote sensing image classification testing procedure. The significant variability in damage assessment (the ground truth) is confirmed by the comparison between DPC-COLLAPSES and all the other datasets. Among the buildings labelled as "*Collapsed*", only half of the AEDES and about two-third (2/3) of INGV are assigned level 5 damage.

In order to identify a proper number of collapses actually visible using remote sensing data, the APhoRISM project approach was to keep the seventy-five (75) damages coming from INGV layer which was actually based on QB images and associate to each building the information coming from the other layers which were not available in the INGV one.
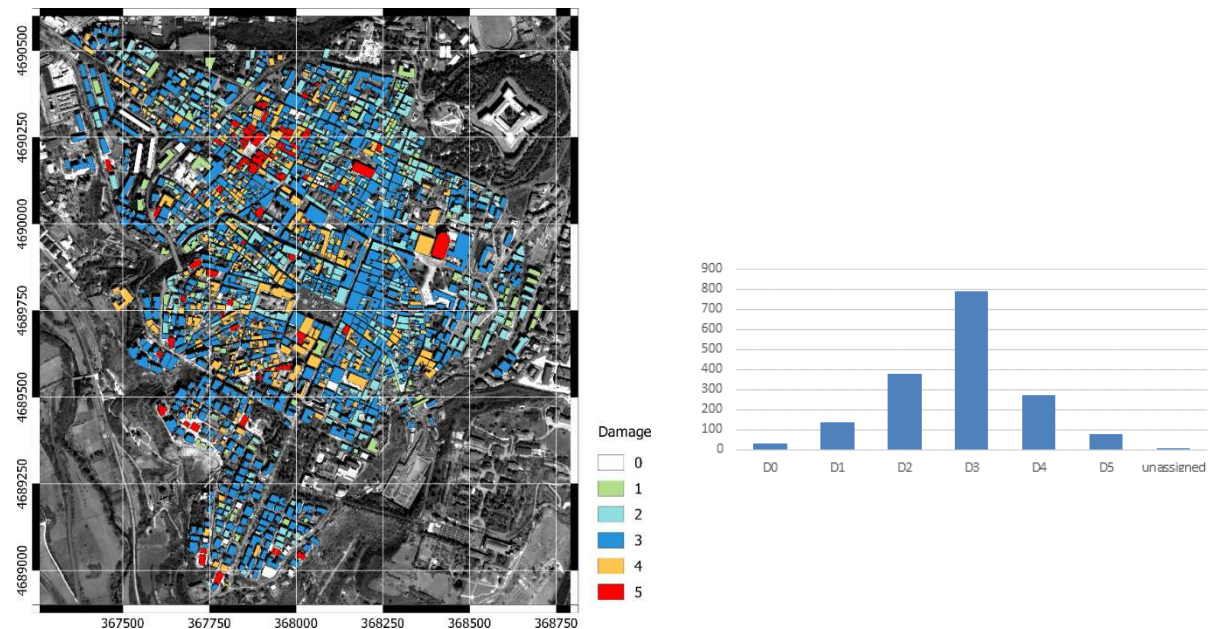
a) The polygons of surveyed buildings are superimposed on a VHR image (QuickBird)

b) Damage frequencies INGV dataset

Figure 5-1: Damage distribution of L'Aquila city centre (INGV Dataset)

a) GIS damage distribution DPC-AEDES dataset processed by CNR-ITC

b) Damage frequencies DPC-AEDES dataset processed by CNR-ITC
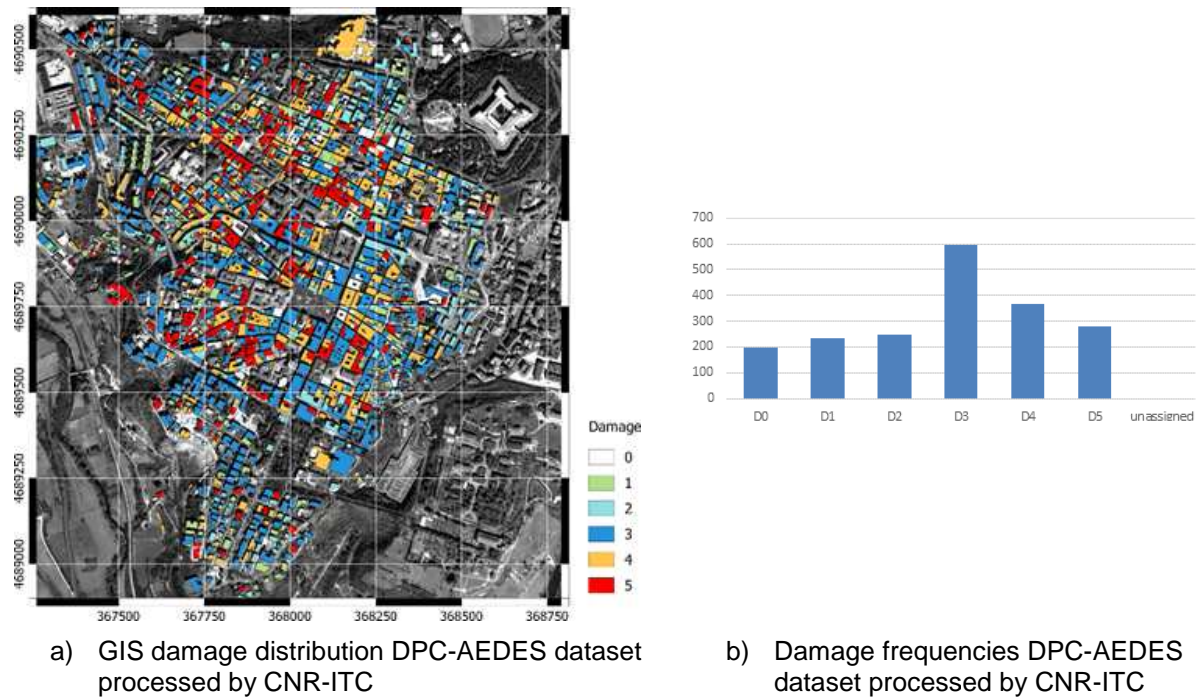
Figure 5-2: Damage distribution of L'Aquila city centre (DPC Dataset – CNR-ITC)

a) GIS damage distribution DPC-BT dataset (average damage)

b) Damage frequencies DPC-BT dataset (average damage)
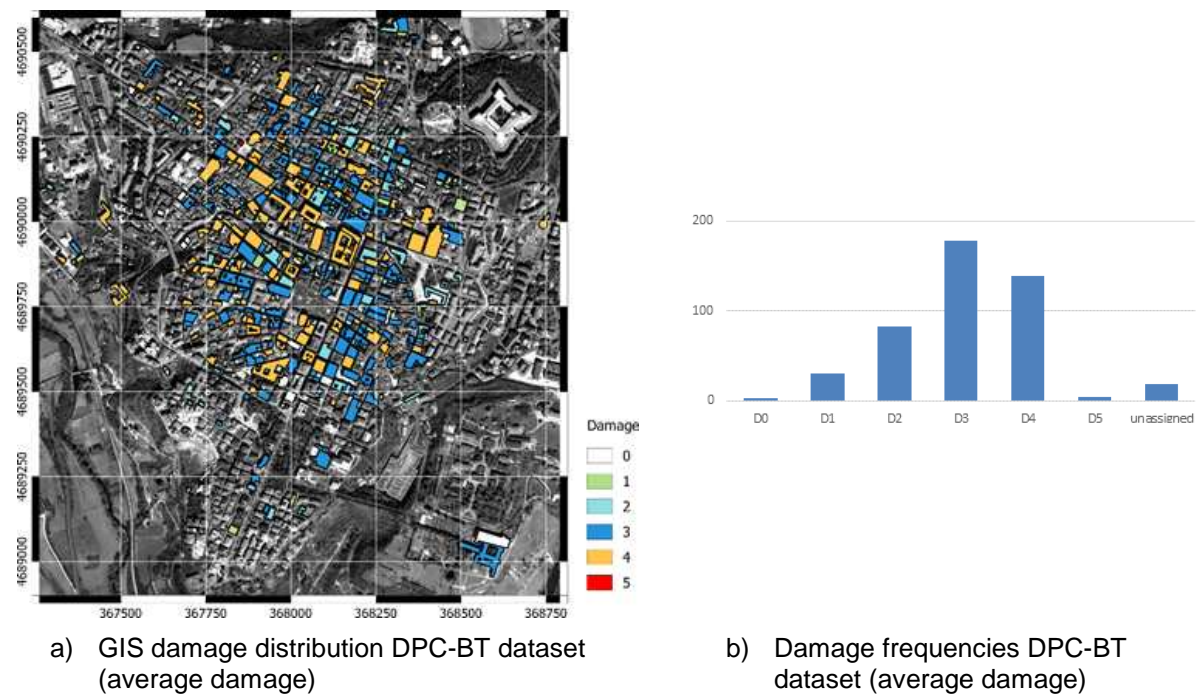
Figure 5-3: DPC-BT dataset - Average damage distribution

a) GIS damage distribution DPC-BT dataset (maximum local damage)

b) Damage frequencies DPC-BT dataset (maximum local damage)
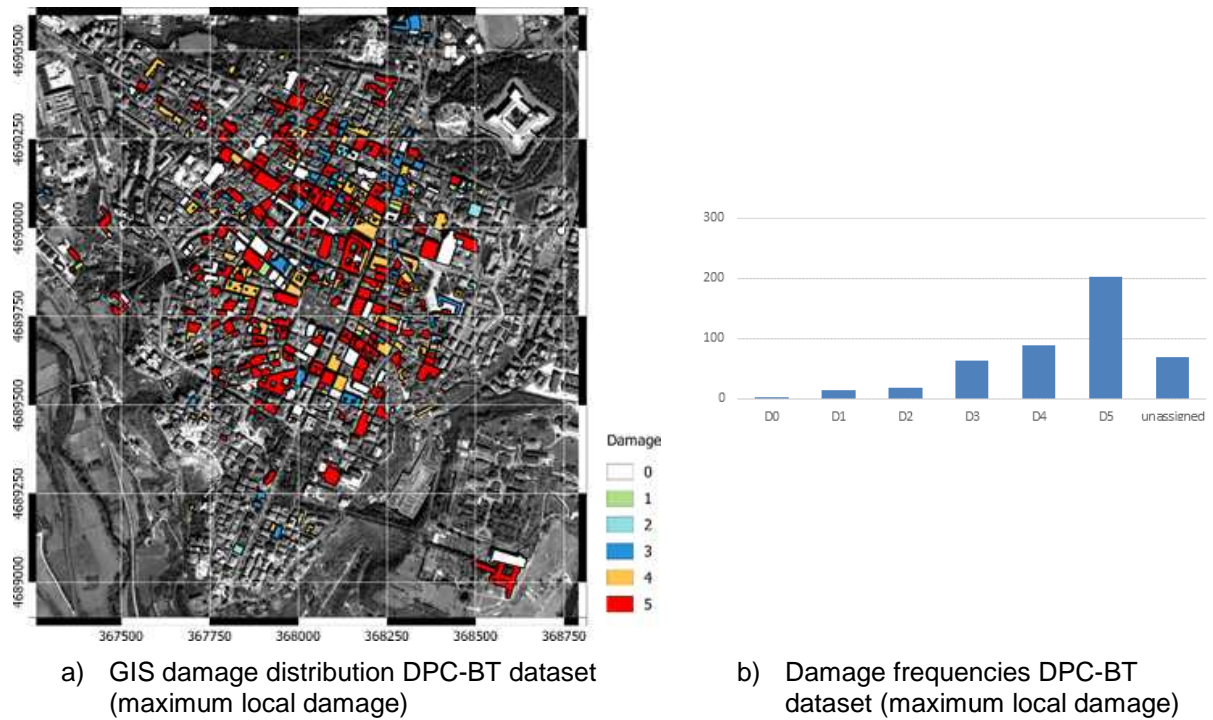
Figure 5-4: DPC-BT dataset - Maximum local damage distribution
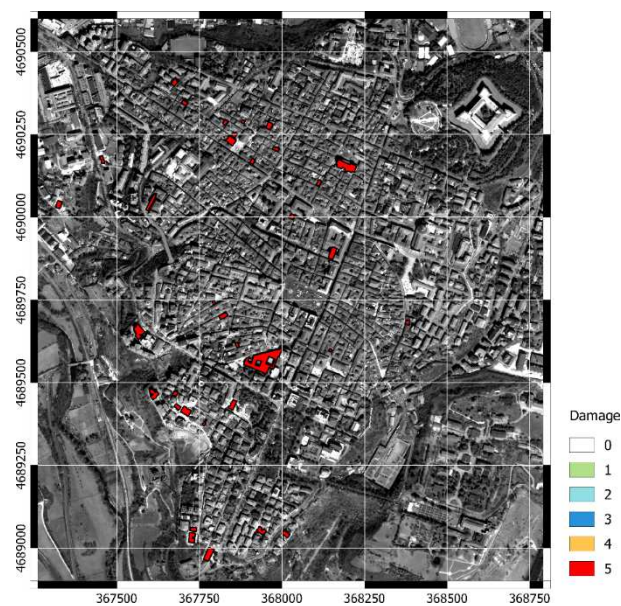
Figure 5-5: Damage distribution DPC-COLLAPSES dataset (38 objects)
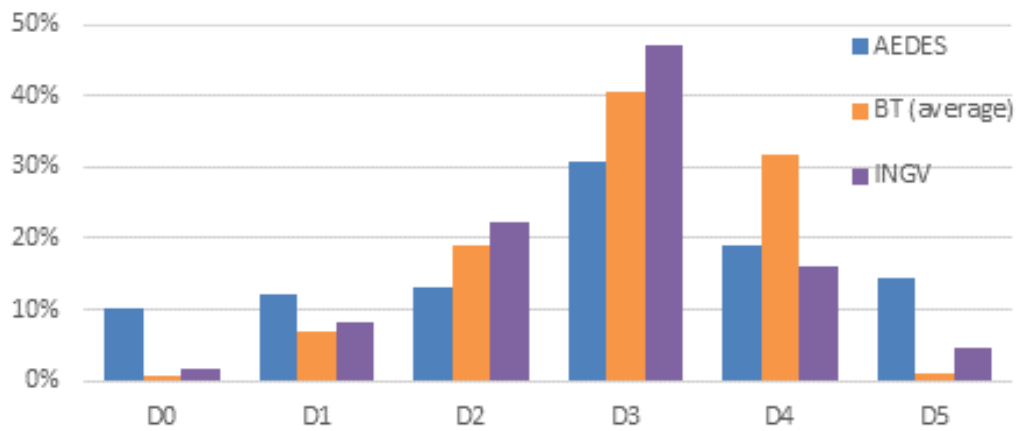
Figure 5-6: Damage distribution dataset comparison

The following table summarises all the available data for L'Aquila Test Case for all the remote sensing and in-situ modules. Par. 5.5 presents the results of the Bayesian Networks based on these data.

Table 5-1: Summary of L'Aquila test case Dataset

	OPT	SAR	GEOTEC	STRUCT
# Features	13	4	1	1
# Obj	1676	1677	1646	1359
# NaN	1	0	31	318

5.2   STRUCTURAL MODULE

The Italian Civil Protection (DPC) designed, supported by the Italian Geophysical and Volcanic Institute (INGV), a procedure to assess post-earthquake damages and building habitation safety (usability). Groups of experts performed the *in-situ* analysis filling in some forms (AeDES forms) answering to specific questions for each building, as better detailed in Par. 2.5.2. The availability of first and second level forms [44], in L'Aquila earthquake test case, allows the identification of vulnerability classes in which each building may be classified [28], and a complete assessment of the damage at urban scale can be performed. In order to do this, we also need an estimate of the intensity of the earthquake. This estimate may be inferred from seismic hazard maps, in the case in which the scenario is developed in a pre-event situation, or from a correlation of the instrumental measurements of the event in the case, the scenario is developed post-event. In the latter case, literature correlations between, for example, the a_g recorded at ground level and Macroseismic intensity can be used. These correlations have usually the form shown in Eq. 5-2.

Data reported in Table 5-2 and Table 5-3 are directly available from the forms, and are typically:

- the structure type (masonry, reinforced concrete both walls and frames)
- the height of the building
- the age of the building
- the floor system (in case of masonry buildings)

These kind of data are very similar to that present in any cadastre archive, so that they can be assumed available when running the APE procedure operationally in the event of an earthquake.

In L'Aquila test case, the AeDES forms and the ground truth were available in the central area of the town (mainly the historical town and part of the neighbourhood), so the ground shake was uniform (highly uniform amplification factor, as better explained in Par. 0. The macroseismic intensity in L'Aquila is assumed equal to 8.5, as indicated by experts' evaluation post-event.

For what concerns the reporting of the actual damage of the buildings observed on ground we have the results of different surveys. This information can be considered as the ground truth when testing the damage classification network. The order of magnitude of the mismatches identified in the ground truth was comparable to those of the image classification with each individual survey. Therefore, although the image classification does not provide very good matching with the "ground truth", the uncertainty of the satellite product is comparable to the uncertainty inherent to the ground surveys themselves.

In summary, the number of objects (i.e., buildings) available with associated features from optical images, SAR images, geotechnical and structural modules are reported in Table 5-1. The smaller number of objects with structural data (only one feature, i.e., the parameter μ_D defined in Par. 5.2.1) is due to the lack of AeDES forms associated to some of the buildings surveyed by INGV. Moreover, the available soil frequency maps do not cover the complete test area; therefore, we lose 31 buildings that are then labelled as Not A Number (NaN) in our structural vector of buildings.

5.2.1 **Structural Vulnerability Assessment**

Referring to the methods presented in Par. 2.5.2 for vulnerability assessment, in the APhoRISM project, in order to infer the structural vulnerability, an empirical method was adopted. This is also the mostly used in the assessment of damage scenario at urban scale. The mechanical method, to which the scientific community has been paying much attention in recent years, requires good and detailed data for each building that are not frequently available for an entire urban area. The mechanical method also requires be accurately calibrating, comparing to literature data available from the empirical method, otherwise it leads to inconsistent results with respect to the evidence collected on damage of past earthquakes. The expert method is not feasible for the committed requirements in terms of resources, and it is more appropriate for the evaluation of individual buildings or confined areas affected by an earthquake.

The method adopted is a Macroseismic method [118], which leads to a probabilistic framework consistent with the damage matrices implicitly contained in the EMS'98 scale. This is accomplished by de-fuzzification of linguistic terms for the percentage of buildings in damage matrices and percentage of each structural type in vulnerability classes, present in the EMS'98 [119], [120].

Given the vulnerability class of a building or, equivalently, its type of structure, Table 5-2 provides the mean value of the base vulnerability index V_0. The year of construction 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.

Some behaviour modifiers can adjust the base vulnerability index V_o, allowing a better prediction of the response of the building, as shown in Table 5-3. Among the many modifiers available, the most relevant is the one related to the height of the building. Other potential modifiers refer to the following structural features:

- state of preservation
- plan and vertical irregularity
- roof system type
- presence of retrofitting interventions
- aggregate position and composition
- foundations

Table 5-2: Vulnerability index values for building classes and structural typologies

	Type of Structure	A	B	C	D	E	F	Vulnerability Index
	V_{min}	0.78	0.62	0.46	0.3	0.14	-1.02	■ most likely / ◀▶ probable / ◁▷ less probable
	V_o	0.9	0.74	0.58	0.42	0.26	0.1	
	V_{max}	1.02	0.86	0.7	0.54	0.38	0.22	
M1	rubble stone	■						0.85
M2	adobe (earth brick)	■	▶					0.82
M3	simple stone	◁	■					0.75
M4	massive stone		◀	■	▷			0.60
M5	unreinforced masonry, with manufactured stone units	◁	■	▷				0.74
M6	unreinforced masonry, with RC floors		◀	■	▷			0.60
M7	reinforced or confined masonry			◁	■	▶		0.40
RC1	frame without eartquake-resistant design (ERD)	◁	◀	■	▷			0.63
RC2	frame with moderate level of ERD		◁	◀	■	▶		0.45
RC3	frame with high level of ERD			◁	◀	■	▶	0.29
RC4	walls without ERD		◁	■	▶			0.56
RC5	walls with moderate level of ERD			◁	■	▶		0.40
RC6	walls with high level of ERD				◁	■	▶	0.24
S	steel structures			◁	◀	■	▶	0.29
W	timber structures		◁	◀	■	▶		0.45

MASONRY: M1–M7 · REINFORCED CONCRETE (RC): RC1–RC6 · STEEL: S · WOOD: W

Vulnerability Index scale: 1 0.8 0.6 0.4 0.2 0

Table 5-3: Scores for height behaviour modifier factors for Masonry and RC buildings
(Source: Giovinazzi [118])

Behaviour modifier	Masonry		Reinforced Concrete	
	Height	ΔV_m	Height	ΔV_m
Number of floors	Low (1÷2)	-0.08	Low (1÷3)	-0.02
	Medium (3÷5)	0	Medium (4÷7)	0
	High (≥6)	0.08	High (≥8)	0.04

Then, the total vulnerability index, to be used in the damage scenario estimation, will be evaluated as:

$$V = V_0 + \Delta V$$

<div align="right">Eq. 5-1</div>

To complete the assessment of the damage at urban scale, we also need an estimate of the intensity of the earthquake. This estimate may be inferred from seismic hazard maps, in the case in which the scenario is developed pre-event, or from a correlation of the instrumental measurements of the event in the case in which the scenario is developed post-event. In the latter case, literature correlations between, for example, the a_g recorded at ground level and macroseismic intensity could be used. These correlations have usually the form reported in the following equation

$$a_g = c_1 c_2^{(I-5)}$$

<div align="right">Eq. 5-2</div>

Following the assumptions of Giovinazzi [118], the mean value of the expected damage μ_D, given the vulnerability index V of a building and the macroseismic intensity *I* is given by:

$$\mu_D = 2.5 + 3 tanh \left(\frac{I + 6.25V - 12.7}{3} \right)$$

<div align="right">Eq. 5-3</div>

and the full distribution of the continuous random variable *D*, ranging from zero to five, is given by:

$$D \sim Beta(\alpha, \beta, a, b)$$

<div align="right">Eq. 5-4</div>

Where, more specifically, as reported in Giovinazzi, 2005 [118], the use of the beta distribution is proposed for the damage, since it does not need to be truncated and suits the specific requirement of allowing varying the scatter around the mean value. The beta probability density function and the beta cumulative density function are respectively reported in (3.4) and (3.5).

$$PDF\!: p_\beta(x) = \frac{\Gamma(t)}{\Gamma(r)\Gamma(t-r)} \frac{(x-a)^{r-1}(b-x)^{t-r-1}}{(b-a)^{t-1}}$$

<div align="right">Eq. 5-5</div>

$$CDF: P_\beta(X) = \int_a^x p_\beta(y) dy$$

Eq. 5-6

With a ≤ x ≤ b and where a, b, t and r are the parameters of the distribution and Γ is the gamma function.

Parameters t and r (or equivalently the mean and the variance) control the shape of the distribution. In general, low values of t give broad distributions (in fact t=2 and r=1 give a uniform distribution) and high values of t (greater than 8) give narrow distributions. According to McGuire (2004) [121] t =3÷6 may result in reasonable building damage distribution. The values attributed to the distribution parameters are the following:

a = 0, b = 5, t ≅ 8, r≅ t μ_D / 5, α = r, β = t-r

In the following picture (see Figure 5-7), we define the probability of collapse of a building p_C (i.e. that the damage corresponds to a discrete damage grade equal to 5):

$$p_C = \int_{4.16}^5 B(\alpha, \beta, a, b) \, dD$$

Eq. 5-7
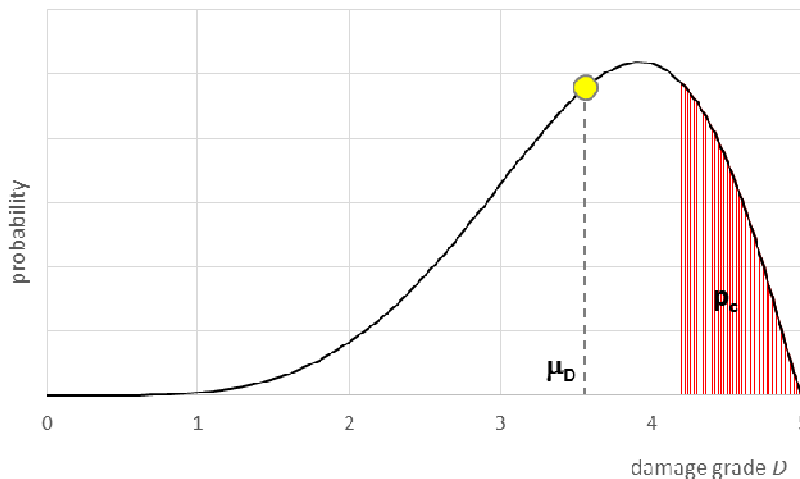
Figure 5-7: Building Damage Probability

5.3   REMOTE SENSING: SAR & OPTICAL MODULES

This section provides a summary of the Earth Observation (EO) data at Very High Resolution (VHR) available for L' Aquila test case and summarises the methodological approach adopted to generate the remote sensing input data set for the data integration.

Two pairs of Optical and Synthetic Aperture Radar (SAR) images of the L'Aquila earthquake were used, as reported in Table 5-4.

Optical data are from Quickbird (QB). The pre- event image is dated September 4, 2006 while the post-event one was collected on April 8, 2009. Both panchromatic (PAN) and

multispectral (MS) images are available. The multispectral images consist of four bands in the blue, green, red and near-infrared wavelength regions with resolution of 2.44 m at nadir. The resolution of panchromatic images is 0.6 m.

The SAR data set is composed by two SPOTLIGHT COSMO-SkyMed (CSK) images. The pre-event image was taken on April 5, 2009 while the post-event one, two weeks after the seism (April 21, 2009). They have a spatial resolution of about 1 m, both in range and azimuth, and were acquired along ascending orbit with a right-side looking angle of 50.57°.

Table 5-4: Earth Observation Dataset for the L'Aquila test case

VHR Optical Dataset (QuickBird)			
Date of Acquisition	*Acquisition Mode*	*Looking Angle*	*Spatial Resolution*
04/09/2006	Panchromatic (PAN)+ Multi spectral (MS)	-3.7° in-track -3.10° cross-track 1.09° off-nadir	0.6 m (PAN) 2.4 m (MS)
08/04/2009	Panchromatic (PAN)+ Multi spectral (MS)	2.8° in-track 3.9° cross-track 4.08° off-nadir	0.6 m (PAN) 2.4 m (MS)

VHR 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

The pre-processing steps made on VHR optical data consist of a series of manipulations of the pre-event and post-event data to obtain a co-registered dataset, ready to be overlapped to the ground truth information layer. Panchromatic images were orthorectified using a Digital Surface Model (DSM) from LiDAR data, and exploiting the Rational Polynomial Coefficient (RPC) provided with the QB dataset. The two resulting images were compared with the polygons related to the building footprint (potentially coming from any available GIS layers, but in our case from the ground truth survey, as reported in Par. 5.1), and then manually shifted of few meters towards East and North coordinates to remove residual bias and to match the footprints of the ground truth.

Finally, a histogram matching was performed to radiometrically compensate the different season and illumination conditions of the pre-seismic and post-seismic acquisitions. It was based on al linear stretching to get the same digital counts of the 25th and 75th percentiles in the two images.

From the PAN and MS images, a Pan Sharpened (PSH) dataset was also derived. First, the Gram-Schmidt Pan Sharpening method was applied, and then the orthorectification and the final bias removal was performed as for the PAN image.

As for the pre-processing of CSK-SAR data, pre- and post-event images were first co-registered in the complex domain, assuming the pre-seismic image as the reference

geometry, and then radiometrically calibrated. For this purpose the Sarscape© software in an IDL/ENVI© environment was used.

For segmenting optical images in objects corresponding to buildings the GIS layer reporting building footprints was used. For what concerns SAR data, because of the complexity and peculiarity of building appearance in radar images, an *ad-hoc* segmentation technique of the pre-event image was developed. It is based on the use of morphological profiles to extract bright stripes and ridges caused by double bounces and/or layover mechanisms, the most distinctive features of the SAR building response. Looking at changes in these regions, heavily damaged buildings could be identified. When a building collapses, changes are also expected in the building footprint and in its shadow area. Typically, an increase of the backscattering is observed in these regions due to the scattering contribution from debris and to the return coming from the ground previously occluded by the building shadow. In order to single out such changes, a segmentation approach exploiting the Kullback-Leibler (KL) Divergence computed from the pre-and post-event SAR intensity images was used. The task of extracting objects corresponding to double bounce/layover areas (DB/LO objects), as well as regions of change where an increase of the backscattering in the post-event image is expected (INC objects) was performed in the SAR slant range geometry, in order to preserve the geometrical characteristics of these objects as much as possible. Once the DB/LO and INC objects are identified, they are projected on the ground, using the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM), in order to assign the geographic coordinates which permit their association to the buildings in the available map. The association of DB/LO objects to buildings was performed after correction of the range displacement of the building itself to compensate for layover which depends on the building height and on the radar incidence angle.

Once the objects are identified, a vector consisting of change features derived from the pre- and post-event optical and SAR images was associated to each building in the available GIS map.

Regarding optical data a set of thirteen (13) features was considered as shown in Table 5-5, consisting of the following parameters:

a) Change metrics derived from the Information theory, i.e. the KL Divergence and the Mutual Information, computed from both PAN and PSH images;
b) Changes in the textural parameters retrieved using the grey level co-occurrence matrix from PAN data. Energy, Correlation, Homogeneity, Contrast and Entropy were extracted, and the difference between their post-seismic and pre-seismic values evaluated;
c) Changes in the colour space, i.e. differences in the Hue, Saturation and Intensity from PSH imagery
d) Difference between post-event and pre-event PAN images

As for SAR data, the ratio between the pre- and post-event backscatter intensity images, the interferometric coherence, the intensity correlation and the KL Divergence were computed for both DB/LO and INC objects, as reported in Table 5-6. The following tables summarise all the remote sensing features taken into consideration.

Table 5-5: Optical VHR Features

GROUP	FEATURE	DESCRIPTION
Information theory features	MI_{sb}	Mutual Information from panchromatic
	KLD_{sb}	Kullback-Leibler divergence from panchromatic
	MI_{mb}	Mutual Information from pansharpened
	KLD_{mb}	Kullback-Leibler divergence from pansharpened
Texture features (from panchromatic)	dcon	Contrast difference
	dcor	Correlation difference
	dene	Energy difference
	dhom	Homogeneity difference
	dent	Entropy difference
Colour features (from pan sharpened)	dhue	Hue difference
	dsat	Saturation difference
	dlum	Value component difference
Other	diff	Mean difference from panchromatic

Table 5-6: VHR SAR Features

GROUP	FEATURE	DESCRIPTION
Double Bounce/LayOver areas	dbo cbkld	Double bounce KL Divergence
	dbo coh	Double Bounce interferometric coherence
	dbo corr $_b$	Double bounce intensity correlation
	dbo i ratio	Double Bounce ratio between the pre- and post-event backscatter intensity images
Increase of the backscattering in the post-event image (INC Objects)	inc cbkld	KL Divergence
	inc coh	Interferometric coherence
	inc corr	Intensity correlation
	inc i ratio	Ratio between the pre- and post-event backscatter intensity images

We should take into account that, the optical and SAR features used are continuous variables. However, the Bayesian network makes use of nodes to which we want to associate different states (*Collapsed* or *Not Collapsed* in our case). Therefore, the Bayesian networks will be built considering these as continuous variables node. In this way, the network can accept continuous values when learning from cases, or generate them when simulating cases. In addition, this provides better documentation of the node, and makes it easier if, at a later time, we want to discretize the node in another way (i.e. with a different number of states, or different thresholds for the state intervals). However, such approach leads us to the difficulty of identifying the right threshold (or cutoff) where the interval related to the "*Not Collapsed*" status ends and the "*Collapsed*" status begins.

In order to preliminarily understand the feature behaviour and identify the right cutoff range, each feature was divided into N intervals and, for each interval, the occurrence of buildings that belong to the "*Not Collapsed*" class (*orange colour*) and those belonging to class "*Collapsed*" (*blue colour*) was counted. In this way, we can identify the intervals for which the probability of occurrence of a damage is higher. The following pictures show the identified bins. This activity is useful not only to better understand the feature behaviour but, as we will see later in Par. 5.5, it is relevant to support the identification of the appropriate thresholds (or cutoffs) to discretize a continuous variable in our network nodes.

Table 5-7: Optical Features
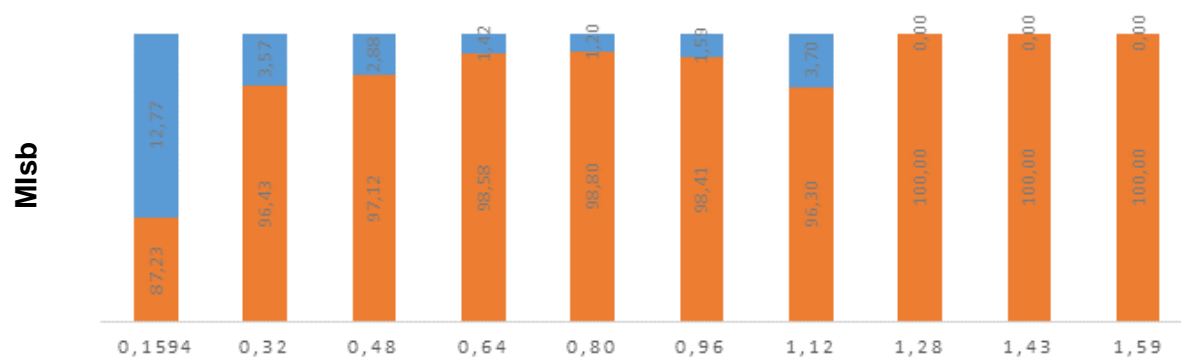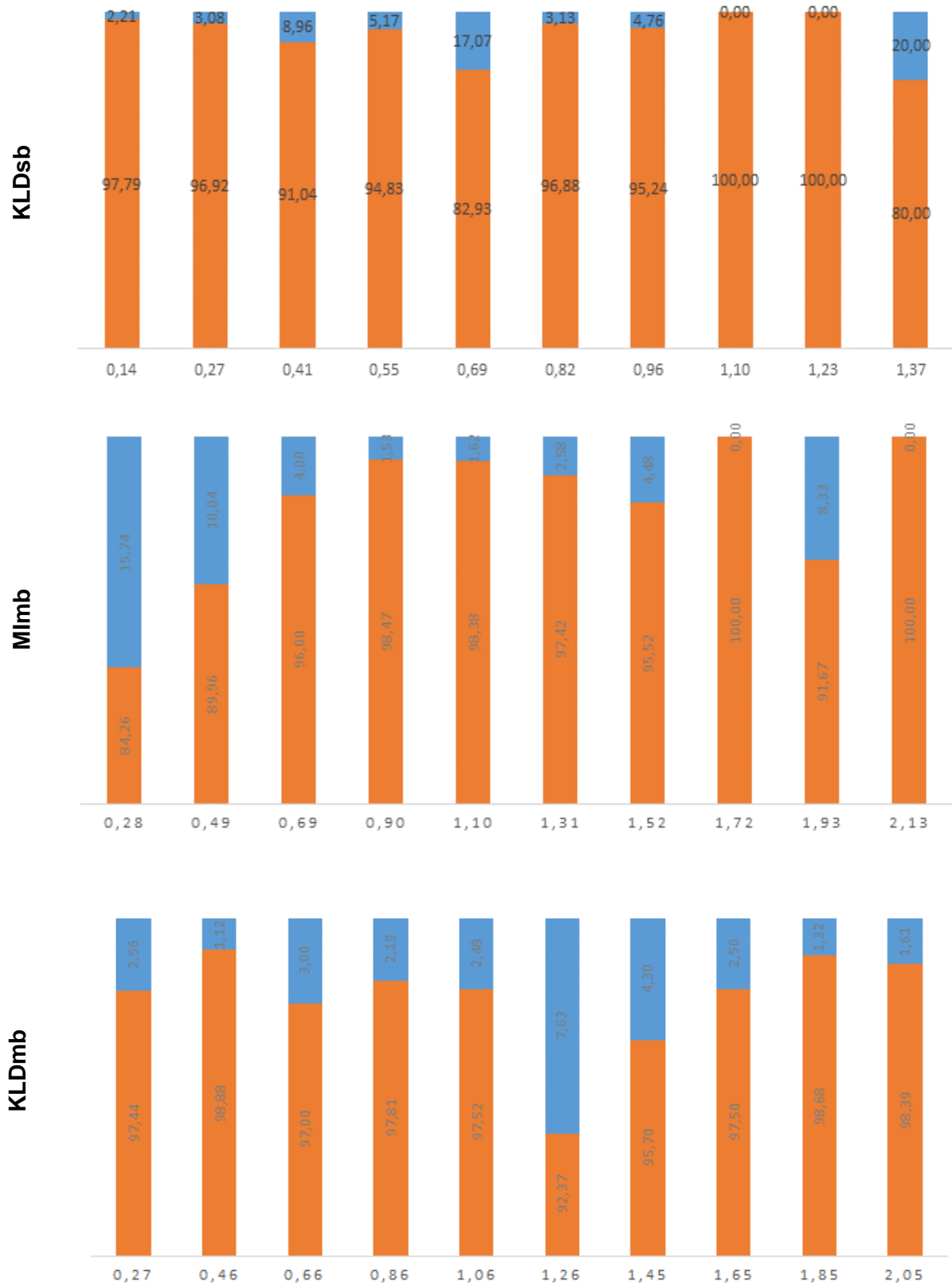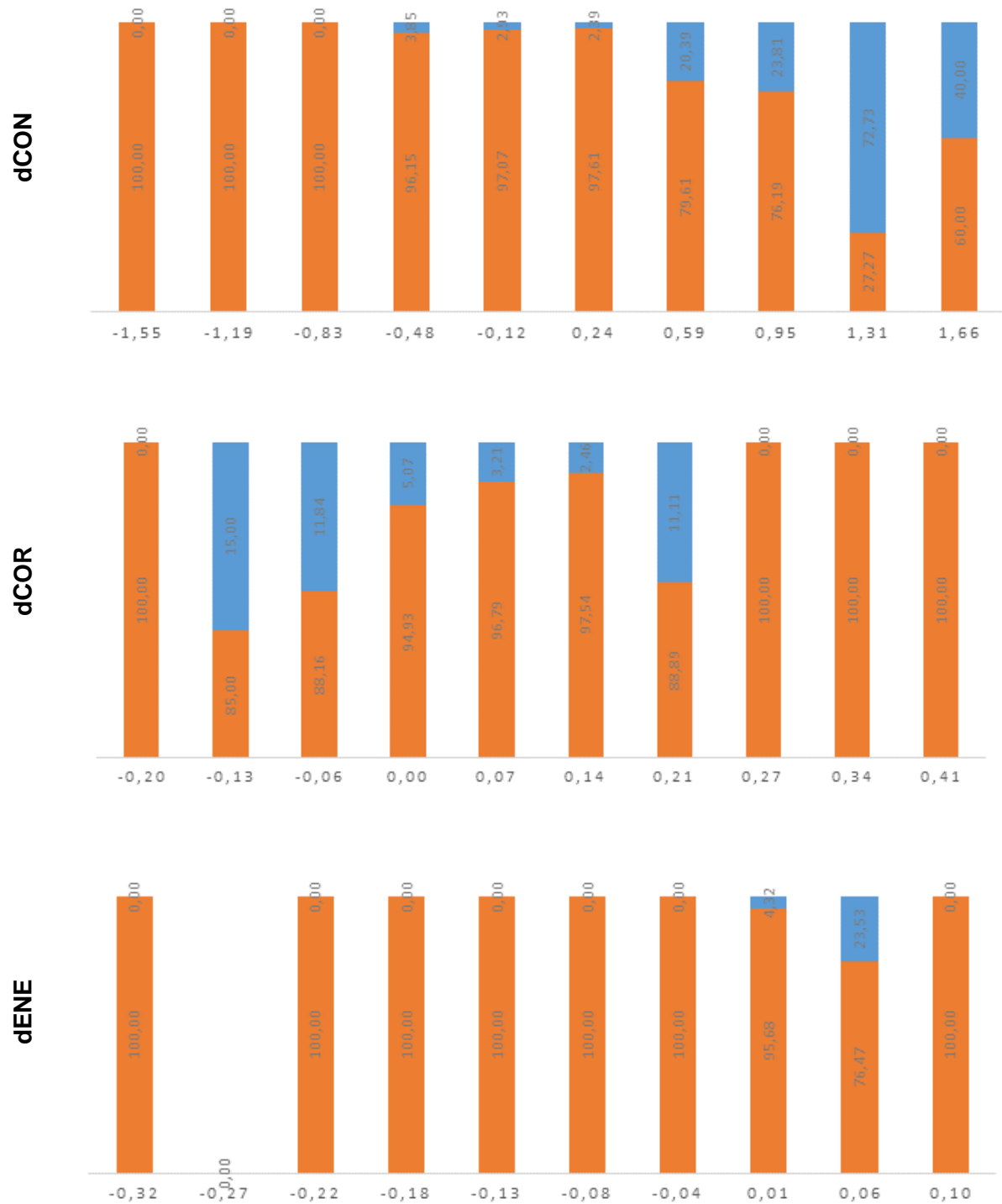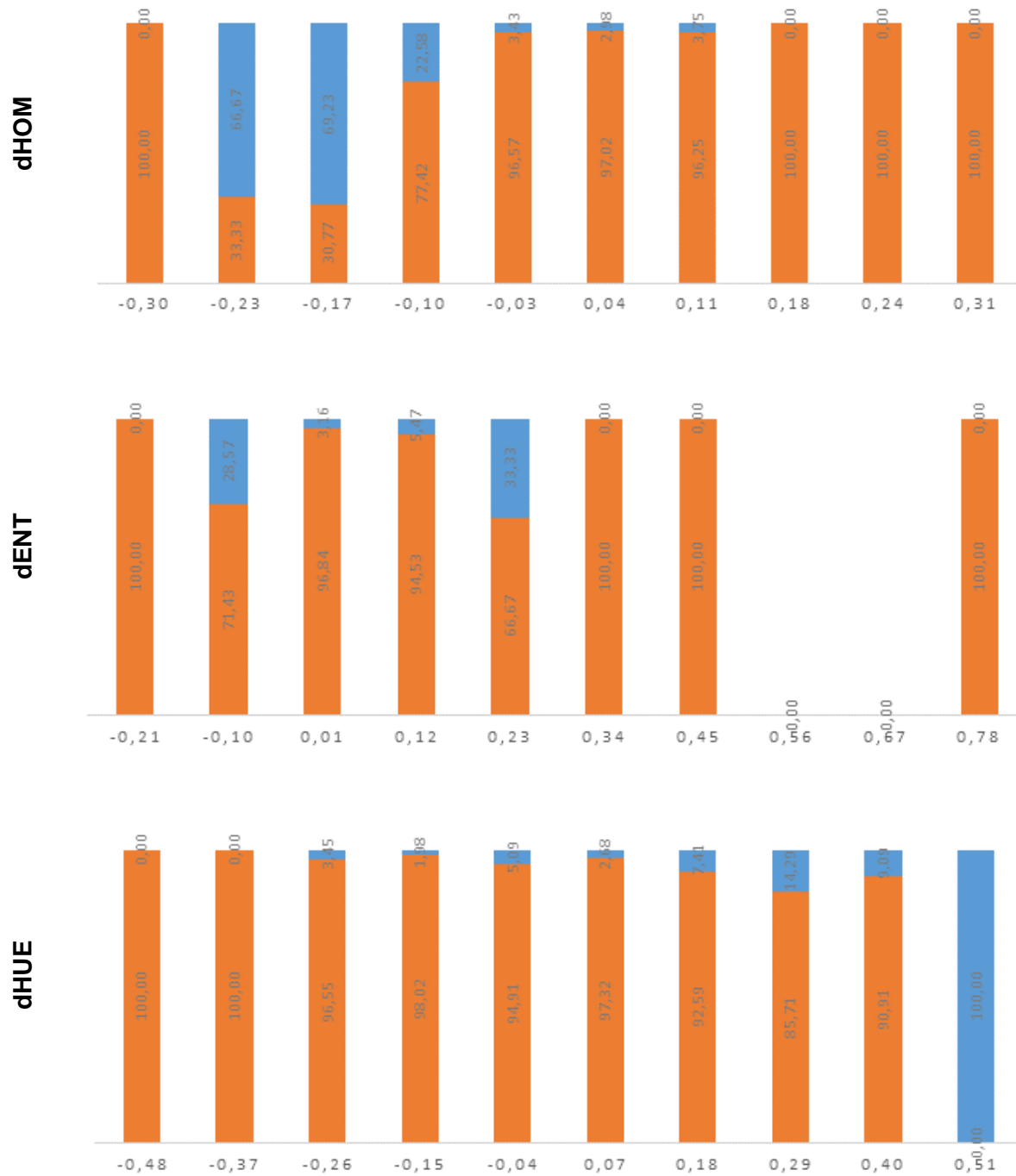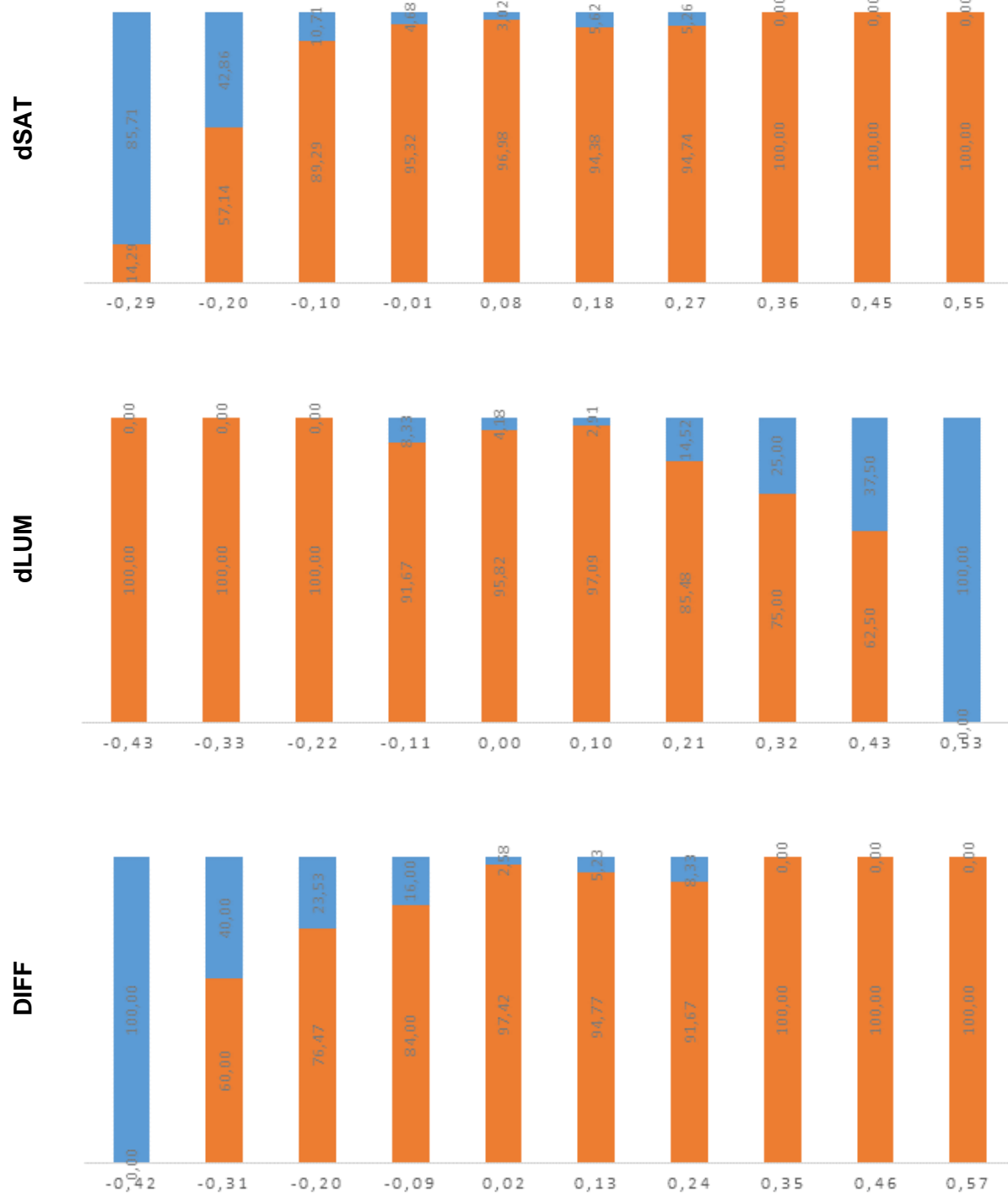
Table 5-8: SAR Features
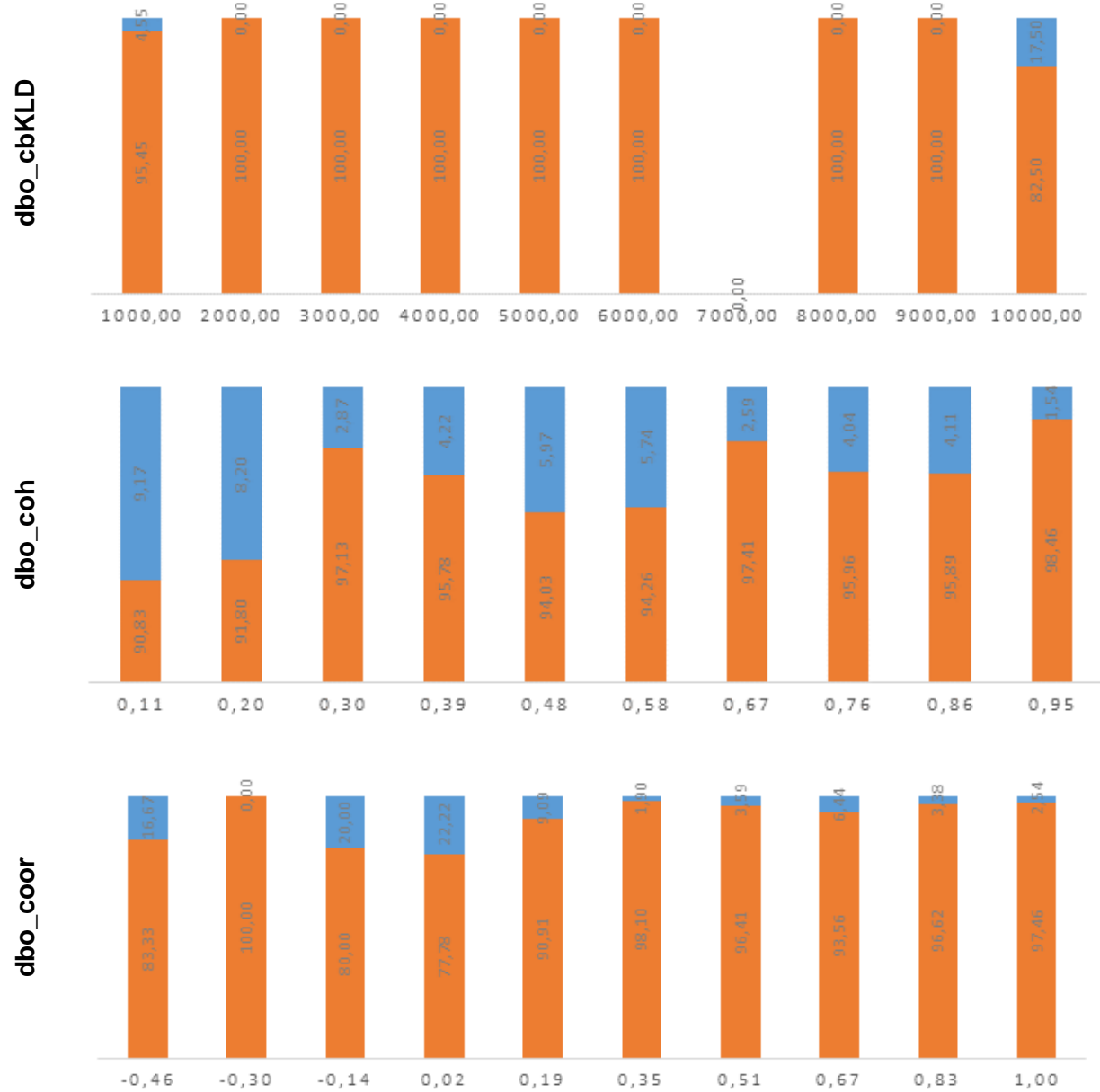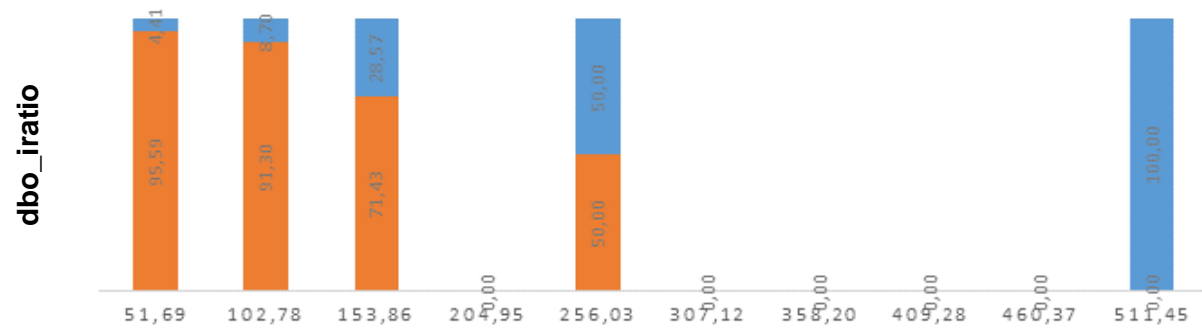
5.4   GEOTECHNICAL MODULE

The geotechnical module is a combination of two kinds of geological information (see Figure 3-1). The first one (Geomodule1) defines the ShakeMap strategies to improve the accuracy and resolution of the ground shaking ShakeMap (see Par. 2.4 for the detailed definition of a Shakemap). This module takes as inputs:

   a) Real data from seismological stations and intensity measurements
   b) Models of ground motion prediction equations and maps of the site amplifications

The output defined by this module consists of maps of the ground shaking in terms of Peak Ground Acceleration, Peak Ground Velocity, spectral acceleration at 0.3s, 1.0s and 3.0s, and macroseismic intensity, available in different formats as, for example, shapefiles (for GIS applications) or text files.

INGV calculated the ShakeMap for the main event of L'Aquila earthquake sequence and provided the output focusing on the area in which the microzonation was also available, which covers only the metropolitan area of L'Aquila.

The second module related to geological information (Geomodule2) is related to information on soil instability. Such information is given by different data that are quantified according to the available *a-priori* data regarding soil amplification, soil resonant periods, liquefaction, landslides and ground subsidence. Generally, such kind of data in Italy are provided by seismic microzonation studies. All this information is known to influence the damage level during an earthquake (Kramer, 1996) [132]. The basic concept is that the soil instability, in conjunction with the structural features (e.g., building height, age, typology, material), aims at reducing the false positive and negative alarms on change detection maps from satellite imagery, which may originate from different factors, such as shadows and sun illumination in optical images, changes in looking angle and speckle in SAR images. Such improvement is expected because if an infrastructure stands on an area more susceptible to seismic amplification, or prone to slide or liquefy, then such infrastructure is expected to behave badly during an earthquake and, presumably, it will be affected by higher damaged with respect to other similar infrastructures placed on stable areas.

This module provides, as output, a vector file and its associated table containing information on:

- I_A = soil amplification maps.
- I_T = soil natural oscillation period index, related to the available resonance period maps for the soils susceptible of seismic amplification
- I_V = InSAR velocity index, that is related to the deformation velocity measured by SAR multipass interferometry data during interseismic periods. This index depends on the resolution of the available dataset, on the density of Point Scatterers (PS) and on the gradient of the detected deformation
- I_L = liquefaction instability index, that is based on the available liquefaction susceptibility maps, in particular to the Liquefaction Potential Index (LPI) and on the coverage density of the liquefiable area with respect to the monitored infrastructures
- I_S = landslide instability index, related to the available information on the landslide distribution, susceptibility and hazard.

For L'Aquila test case, the soil instability is given only by the soil fundamental frequency of oscillation, which was found to be slightly correlated to the damage grade, as shown in Figure 5-8 a) and b) (Credits: INGV).
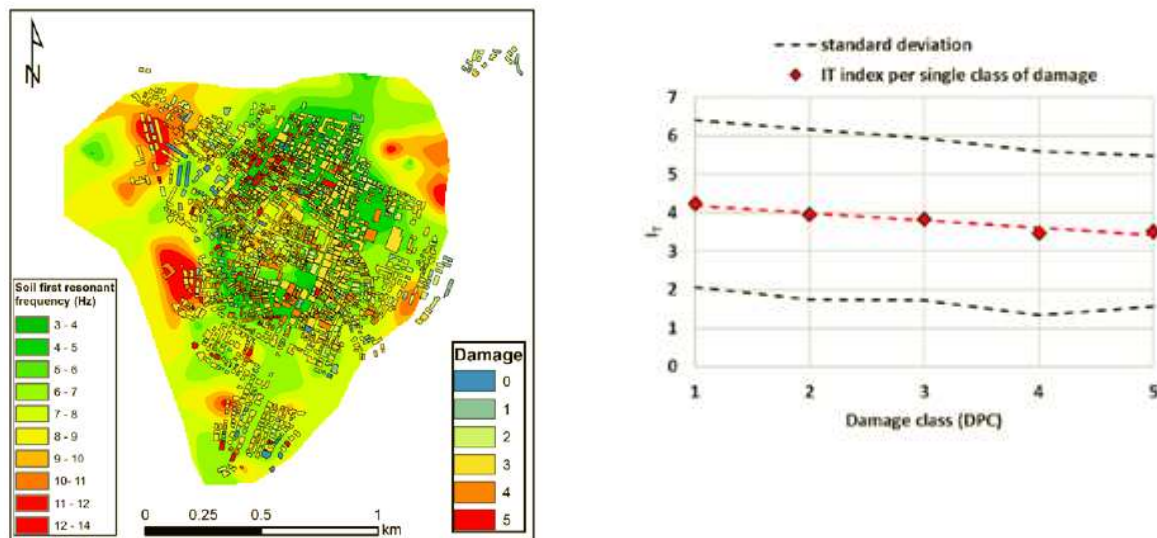
a) Damage grade (DPC dataset) superimposed on the soil resonant frequency map for L'Aquila city centre

b) Comparison between IT index and damage grade for L'Aquila city centre

Figure 5-8: Geotechnical evaluation of damage grade (DPC dataset) for L'Aquila city centre *(Source: INGV)*

5.5   DATA FUSION MODULE RESULTS

In the following paragraphs, we present the results of the Bayesian networks built for the L'Aquila earthquake test case. Before describing the network results in detail, we provide in the next paragraph, a summary on how the performances of the network were evaluated in this study.

5.5.1   Data Fusion Performances Evaluation

Since the results of the Data Fusion module were chosen to be provided by classifying each building into a two levels collapse probability conditions (low/high), in the following paragraphs, we are going to present the results in the form of Confusion Matrix and measuring the performances using two parameters, namely Cohen's Kappa and Cohen's Normalised Kappa. Hereinafter a brief summary of the meaning of such measures is reported.

A Confusion Matrix [133] contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. The following table shows the confusion matrix for a two class classifier.

Table 5-9: Confusion Matrix *[133]*

	Total Population	Ground Truth	
		Condition Positive (CP)	*Condition Negative (CN)*
Predicted Condition	*TestOutcomePositive - TOP*	a = TP (True Positive)	b = FP (False Positive)
	TestOutcomeNegative - TON	c = FN (False Negative)	d = TN (True Negative)

The entries in the confusion matrix have the following meaning in the context of our study:

- a is the number of correct predictions that an instance is positive
- b is the number of incorrect predictions that an instance is negative
- c is the number of incorrect of predictions that an instance is positive
- d is the number of correct predictions that an instance is negative

The Accuracy and Precision (or Positive Predictive Value – PPV) and the Sensitivity (True Positive Rate) are defined as:

$$ACC = \frac{\sum True\ Positive + \sum True\ Negative}{\sum Total\ Population}$$

Eq. 5-8

$$PPV = \frac{\sum True\ Positive}{\sum Test\ Outcome\ Positive}$$

Eq. 5-9

$$TPR = \frac{\sum True\ Positive}{\sum Condition\ Positive}$$

Eq. 5-10

The measurement of the performance is carried out using Cohen's Kappa [134] [135]. This coefficient springs from the notion that the observed cases of agreement include some cases for which the agreement was by chance alone. Cohen assumed that there were two raters, which rate *n* subjects into one of *m* mutually exclusive and exhaustive nominal categories. The raters operate independently; however, there is no restriction on the marginal distribution of the ratings for either rater.

Let p_i, be the proportion of subjects that were placed in the *i*, *j*th cell, i.e., assigned to the i_{th} category by the first rater and to the j_{th} category by the second rater (*i, j* = 1, . . . , *m*).

Also, let

$$p_i = \sum_{j=1}^{m} p_{i,j}$$

Eq. 5-11

denote the proportion of subjects placed in the i_{th} row (i.e., the i_{th} category by the first rater), and let

$$p_j = \sum_{i=1}^{m} p_{i,j}$$

Eq. 5-12

denote the proportion of subjects placed in the j_{th} column (i.e., the j_{th} category by the second rater). Then, the kappa coefficient proposed by Cohen is:

$$\hat{k} = \frac{p_0 - p_c}{1 - p_c}$$

Eq. 5-13

where

$$p_0 = \sum_{i=1}^{m} p_{ii}$$

Eq. 5-14

is the observed proportion of agreement and

$$p_c = \sum_{i=1}^{m} p_i \cdot p_i$$

Eq. 5-15

Often situations arise when certain disagreements between two raters are more serious than others. However, k makes no such distinction, implicitly treating all disagreements equally. Cohen [136] introduced an extension of kappa called the weighted kappa statistic (k,,.), to measure the proportion of weighted agreement corrected for chance. Either degree of disagreement or degree of agreement is weighted, depending on what seems natural in a given context.

The statistic $\hat{k}_w$ provides for the incorporation of ratio-scaled degrees of disagreement (or agreement) to each of the cells of the *m* x *m* table of joint assignments so that disagreements of varying gravity (or agreements of varying degree) are weighted accordingly. The non-negative weights are set prior to the collection of the data. Since the cells are scaled for degrees of disagreement (or agreement), some of them are not given full disagreement credit. However, $\hat{k}_w$ like the unweighted $\hat{k}$ is fully chance-corrected.

Assuming that w_{ij} represents the weight for agreement assigned to the i,j_{th} cell (i, j = 1,..*m*), the weighted kappa statistic is given by:

$$\hat{k}_w = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}p_{ij} - \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}p_i p_j}{1 - \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}p_i p_j}$$

Eq. 5-16

Note that the unweighted kappa is a special case of $\hat{k}_w$, with w_{ij}= 1 for $i = j$ and w_{ij} = 0 for $i \neq j$.

Translating Eq. 5-13 and Eq. 5-16 introducing our terms in the confusion matrix reported in Table 5-9 for a two-class problem, we obtain:

$$\hat{k} = \frac{TotalPopulation \cdot (TP + TN) - (CP \cdot TOP) - (CN \cdot TON)}{TotalPopulation^2 - (CP \cdot TOP) - (CN \cdot TON)}$$

Eq. 5-17

$$\hat{k}_w = \frac{TPR \cdot (ACC + TPR - 2 \cdot PPV \cdot TPR + PPV - 1)}{TPR + PPV \cdot (ACC - 2 \cdot TPR)}$$

Eq. 5-18

In summary, we will see in the next paragraphs that the evaluation of the results of the Bayesian network is provided in tables where not only the number of true positives (called 'a') is reported but also both Cohen's Kappa and the normalised Kappa are reported.

5.5.2 Data fusion by Bayesian Network

The data fusion approach based on Bayesian networks was carried out on a step by step basis, considering different levels of inputs, coming from the different modules.

The first easiest approach is based on a "*Rule Fusion*" approach (see Par. 4.2.2) where we simply used the probabilities of damage as evaluated by the relevant single modules.

Therefore, in Figure 5-9 each node represents the probability that the optical (popt), SAR (psar), structural (pstruct) and geotechnical (pgeotec) modules have individually identified. The fusion module of Figure 3-1, in this example, takes as input the evaluation already made by the other modules.

In this simplified topology of our network, we omit the relations among the parents. This means that we assume that the results of each module are fully independent from one another and that there is no causal relationships between the nodes. Anyway, this step is

interesting to start understanding the behaviours when considering a binary classification for the damage assessment and when the parent node is described in terms of probability of damage (i.e., rule fusion).

The relevant network is reported in the following figure. The resulting net, though very simple, already presents some important issues that we need to solve to manage the fusion module better, when coming to more complex network architectures.
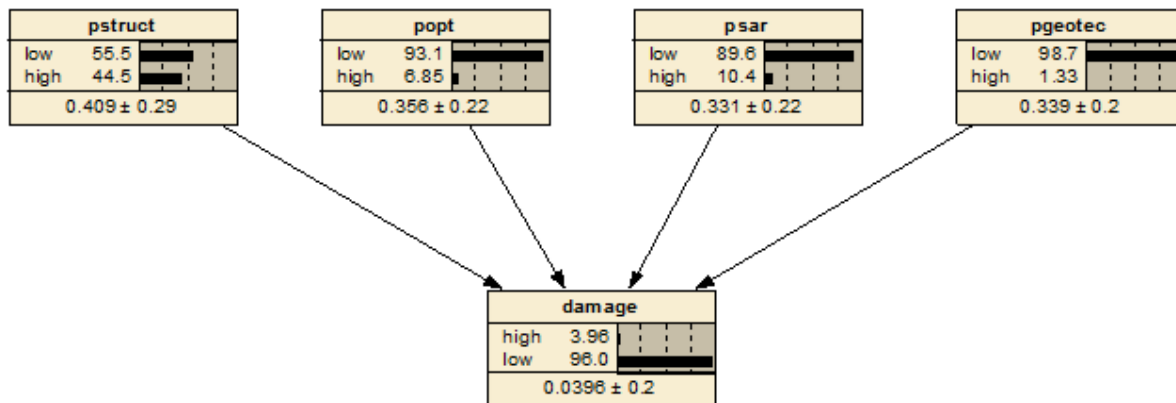
Figure 5-9: Probabilities Fusion Network

In fact, the probability nodes represent a continuous distribution, as evaluated by each input module. It is worth noticing that it is common to have a continuous variable that you want to break up into intervals so that you can treat it as a discrete variable, which is known as discretizing the variable. If the variable is truly continuous, it is usually best to make it a continuous node, and then discretize it with an interval list, rather than just making it a discrete node. This provides better documentation of the node, and makes it easier if at a later time you want to discretize it another way (i.e. with a different number of states, or different cutoff points – thresholds - for the state intervals). In addition, that way the network is able to accept continuous values when learning from cases, or generate them when simulating cases. Such cutoffs (or thresholds) are needed to discretize continuous variables.

Alternately, the final assessment on the damage node is represented as a discrete node that may have numeric quantities attached to each state, so that each state can represent a number, but the variable is incapable of representing numbers between those of each state. In our case, label "*high*" for damage is associated to value 1, while label "*low*" is associated to 0.

In summary, when dealing with continuous variables, we have an additional step that consists in the identification of the optimal threshold for each node. However, our structural input module, having been evaluated using a simplified mechanical approach (see Par. 5.2), in L'Aquila case only identifies four ranges of values for the collapsed probability, namely:

- 0 – 0,0025
- 0,0025 – 0,0582
- 0,0582 – 0,3725
- >0,3725

Consequently, in order to compensate for the highly unbalanced collapsed dataset, the threshold for the structural module was set to the highest evaluation coming from this module (>0,3725).

Considering a continuous distribution of the probabilities coming as inputs of the different classifiers of the modules, the operation to carry out is an optimisation of the thresholds, evaluated with the goal of maximising the k performance. The evaluated thresholds are the following:

Table 5-10: Optimised Cutoff for probabilities node

popt	psar	pgeotec	pstruct
0,6721	0,5180	0,6654	0,3725

The results that can be obtained using the probabilities with the above-mentioned thresholds are reported in the following, where the individual modules did not present any capability to detect collapsed buildings, while a joint combination of the modules improves the results of the single ones. As from now on, all the results related to the performances of the Bayesian Network are reported in the form shown in Table 5-11. These tables show all the components of the confusion matrix, as presented in Par. 5.5 and Cohen's Kappa and Normalised Kappa. Our optimisation is mainly based on K, however it is also important to take note of the True Positives identified (a).

All the results shown as from now on, are also validated in terms that they are representative of a k-fold cross validation where the data set was divided into ten (10) sub-sets. One of the main reasons for using cross-validation instead of using the conventional validation (e.g. partitioning the data set into two sets of 70% for training and 30% for test) is that there is not enough data available in our L'Aquila Test case. In fact, having a small number of collapses (see Par. 5.1) the error on the test data set does not properly represent the assessment of model performance. In order to properly estimate model prediction performance, we used cross-validation as a powerful general technique [142]. In our case we used a k-fold cross-validation (with k=10), where the original sample was randomly partitioned into k equal sized subsamples. Of the k subsamples, a single subsample is retained as the validation data for testing the model, and the remaining k – 1 subsamples are used as training data. The cross-validation process is then repeated k times (the folds), with each of the k subsamples used exactly once as the validation data. The k results from the folds were then summed up to produce a single estimation. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and validation, and each observation is used for validation exactly once [143].

The same training and testing sets were used for all Bayesian Networks identified in this study.

Table 5-11: Performances of the Probabilities Network

	confusion matrix	SAR+GEOTEC	SAR+GEOTEC+STRUCT	opt+sar	opt+sar+struct	opt+sar+geotec	opt+sar+struct+geotec
a	detected collapsed	2	1	21	15	23	21
b	false positive collapsed	0	0	9	8	9	9
c	misdetected collapsed	73	73	54	60	52	54
d	detected not collapsed	1602	1602	1593	1594	1593	1593
performance	Cohen's Kappa	5,0	2,6	38,4	29,1	41,4	38,4
	Cohen's Kappa normalised	2,7	1,4	27,4	19,5	30,1	27,4

As we can see from the results above, the performances of our network are quite low.

In fact, the individual module do not seem to work correctly on their own, even though an improvement in the performances can be established when combined together. Although, the fully combined module (including optical, SAR, geotechnical and structural information) reaches a good level of Cohen's Kappa, we can see that the normalised coefficient is not as good, namely because of the relatively level of True Positives identified.

As said above, the use of such network is useful to understand the behaviour of continuous node networks but the simplified one reported above does not fully exploit the inference capabilities of a Bayesian network. This is why we identified more complex networks that establish relations among the features coming from the different modules. In this way we mix optical, SAR and structural features, aiming (Feature Fusion) at obtaining a more significant network able to be extended to additional scenarios. The use of Feature based networks is also more representative from a physical point of view.

Here in the following we examine the single network blocks representing the features of each module. We start with the Optical Module, followed by the SAR, structural and geotechnical modules.

5.5.2.1 Optical Module

Referring to the problem faced in AphoRISM, we can start from building the Bayesian network for the optical module, as in the following picture, where each node represents an available set of features and each arc connects a parent (influencing) node to a child (influenced) node.

As presented in Par. 5.3, for the optical module in L'Aquila Test case, we have thirteen (13) optical features available for all individual buildings.

Having carried out an optimisation of the combinations of the features in terms of performances (with respect to Cohen's Kappa), the resulting net shows that no significant improvement in the performances is obtained if all features are used together. The following correlation matrix is also coherent with this result. The yellow cells represent a medium level of correlation (assumed in the ranges between 0,3÷0,7 or -0,7÷-0,3) while the green ones represent a high level of correlation (assumed in the ranges between 0,7÷1 or -1÷-0,7).

Table 5-12 shows Pearson's correlation coefficient, commonly called simply "the correlation coefficient". This is obtained by dividing the covariance of the two variables by the product of their standard deviations.

The population correlation coefficient $\rho_{X,Y}$ between two random variables X and Y with expected values μ_X and μ_Y and standard deviations σ_X and σ_Y is defined as:

$$\rho_{X,Y} = corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y]}{\sigma_X \sigma_Y}$$

Eq. 5-19

where E is the expected value operator [144].

Table 5-12: Correlation Matrix for the optical features

		INFORMATION THEORY FEATURES				TEXTURE FEATURES					COLOUR FEATURES			
		MIsb	KLDsb	MImb	KLDmb	dcon	dcor	dene	dhom	dent	dhue	dsat	dlum	diff
INFORMATION THEORY FEATURES	MIsb	1,00	-0,11	0,71	-0,16	-0,10	-0,04	0,05	0,16	-0,11	0,03	0,10	-0,13	0,15
	KLDsb	-0,11	1,00	-0,09	0,58	0,00	-0,02	0,15	0,06	-0,09	0,00	0,05	-0,01	0,04
	MImb	0,71	-0,09	1,00	-0,14	-0,10	0,02	0,09	0,16	-0,14	0,02	0,04	-0,10	0,12
	KLDmb	-0,16	0,58	-0,14	1,00	0,01	0,02	0,05	-0,04	0,03	-0,12	0,08	0,17	-0,13
TEXTURE FEATURES	dcon	-0,10	0,00	-0,10	0,01	1,00	-0,28	-0,06	-0,55	0,31	0,04	-0,15	0,36	-0,37
	dcor	-0,04	-0,02	0,02	0,02	-0,28	1,00	-0,01	0,23	-0,02	-0,01	0,15	-0,06	0,07
	dene	0,05	0,15	0,09	0,05	-0,06	-0,01	1,00	0,47	-0,65	-0,06	-0,04	0,06	-0,06
	dhom	0,16	0,06	0,16	-0,04	-0,55	0,23	0,47	1,00	-0,69	-0,06	0,15	-0,32	0,33
	dent	-0,11	-0,09	-0,14	0,03	0,31	-0,02	-0,65	-0,69	1,00	0,01	-0,13	0,24	-0,24
COLOUR FEATURES	dhue	0,03	0,00	0,02	-0,12	0,04	-0,01	-0,06	-0,06	0,01	1,00	-0,33	-0,21	0,19
	dsat	0,10	0,05	0,04	0,08	-0,15	0,15	-0,04	0,15	-0,13	-0,33	1,00	-0,20	0,28
	dlum	-0,13	-0,01	-0,10	0,17	0,36	-0,06	0,06	-0,32	0,24	-0,21	-0,20	1,00	-0,98
	diff	0,15	0,04	0,12	-0,13	-0,37	0,07	-0,06	0,33	-0,24	0,19	0,28	-0,98	1,00

However, it is worth highlighting that, being the correlation coefficient a measure of linear relationship, a value of ρ = 0 does not imply that there is no relationship between the variables. Also at the same time, the existence of a strong correlation does not imply a causal link between the variables. This analysis is instead useful for our purposes to understand which variables, added to our network, do not provide significant additional information.

Among the thirteen optical features, an optimisation was carried out to identify the optimal combination in order to identify those with best performances, in terms of Cohen's Kappa. From this optimisation process, a set of four optical features was identified as optimal. This includes the Kullback-Leibler Divergence, contrast, correlation and energy.

The identified continuous features and their relevant thresholds are hereinafter reported:

$KLD_{mb} >$ -0.225

dcon <   0.2433

dcor >   0.0351

dene <   -0.006
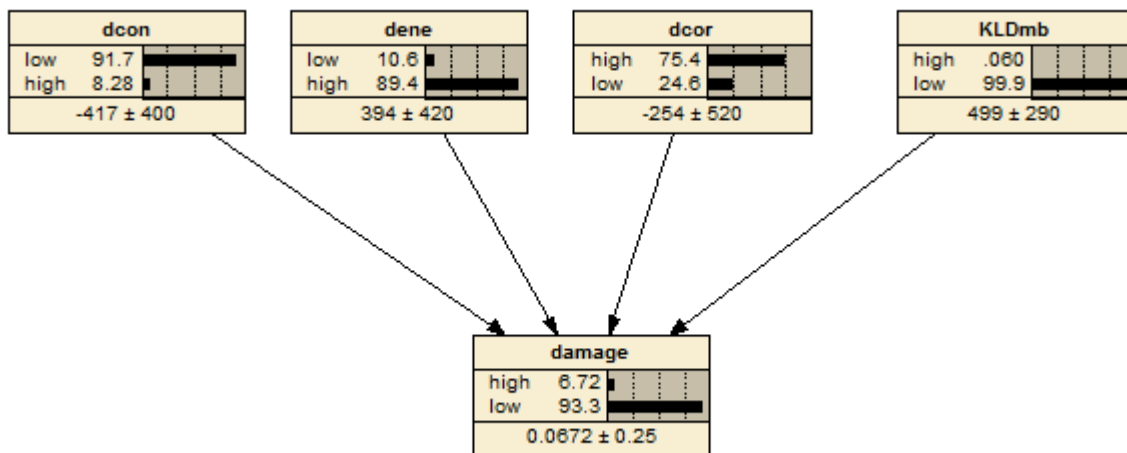
Figure 5-10: Optical Feature module

The threshold for the damage node is set to 4,5%. The reason for this choice is because we need to compensate for the low number of collapses in our unbalanced dataset. By analysing the combinations of features values that are mostly associated to collapsed buildings, we can identify the following combinations where 0 represents "*low*" damage probability and 1 a "*high*" damage probability.

Table 5-13: Combinations of optical features associated to Collapsed Buildings

dcor	dene	dcon	KLDmb
Low	High	High	Low
High	Low	High	Low
High	High	High	Low

The results of the net above are reported in the following table. In this table we compare both the results of the non validated test (i.e. the complete data set was used for both training and testing) which may, of course, lead to overfitting problems [145]. Therefore, as also done for the previous network, we present the results for the k-fold cross-validation (using k=10 subsets).

Table 5-14: Optical Feature Module Performances

confusion matrix		OPT_no validation	OPT_validation
a	detected collapsed	36	36
b	false positive collapsed	96	103
c	misdetected collapsed	39	39
d	detected not collapsed	1506	1499
	Cohen's Kappa	30,8	29,6
	Cohen's Kappa normalised	42,0	41,6

The optical module, the richest one in terms of features, produces a generally good classification at the expense of some false positives. As usual, the Cohen's Kappa measures the performance of the classifier, with and without normalisation with respect to the number of samples within each class (cardinality of the test set). In this case, the optical module score is $k \cong 29,6\%$.

An extension of this net is built in the following to better address the combination of all the modules. Two nets were tested, one based on an optical continuous probability node (p_{opt}), the other with a new discrete classification node (c_{opt}). The introduction of this additional node allows using the graphical structure as an easy way to specify the conditional independencies, and hence to provide a compact parameterization of the model. We carry out this step in view of the extension of this architecture to the other modules and to their integration.

These nets and their relevant results are reported hereinafter where we can see that their performances are substantially the same. This is not surprising since the two newly introduced "hidden nodes" do not provide additional information with respect to the previous ones.
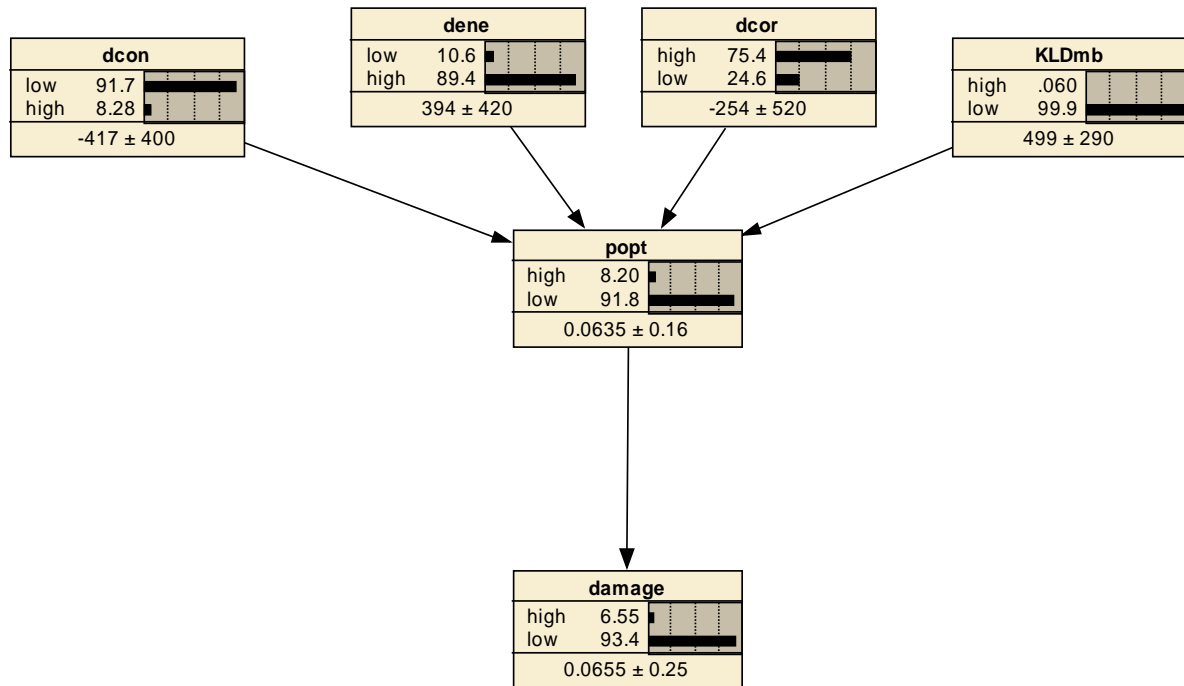
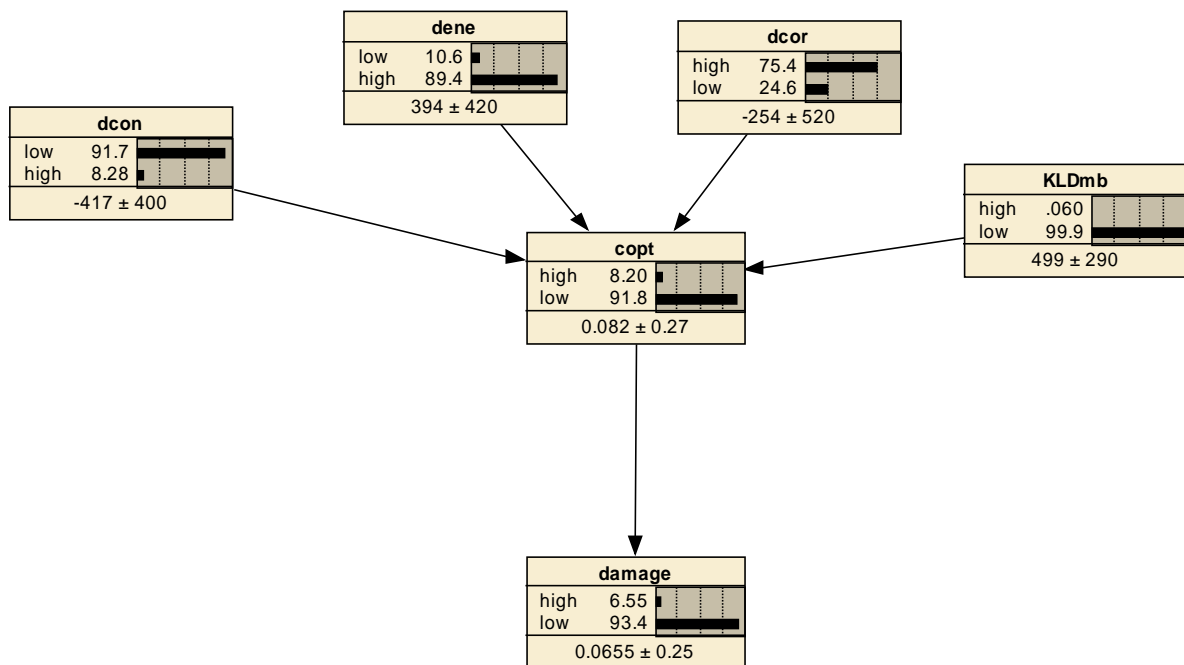**Figure 5-11: Optical Feature module with a continuous probability node**

Figure 5-12: Optical Feature module with a discrete classification node

The results of the above mentioned nets are reported in the table below. This table also clearly shows that we do not lose information in the net when adding a node, such a p_{opt} or c_{opt}, that is useful for further extension of the net when additional modules are added bringing in information at the level of features that can be clustered into different groups.

This approach does not only simplify the network-building phase but also enables a clearer inference when some variables are missing.

Following the same methodology as for the previous network, we provide hereinafter the results for both the non validated and k-fold cross-validated tests, where it is possible to highlight that the same results could be reached as for the previous network in Figure 5-10.

Table 5-15: Optical Module Performances with probability and classification nodes

confusion matrix	OPT+popt_no validation	OPT+popt_validated	OPT+copt_no validation	OPT+copt_validation
a	detected collapsed	36	36	36	36
b	false positive collapsed	96	103	96	103
c	misdetected collapsed	39	39	39	39
d	detected not collapsed	1506	1499	1506	1499
Cohen's Kappa		30,8	29,6	30,8	29,6
Cohen's Kappa normalised		42,0	41,6	42,0	41,6

5.5.2.1 SAR Module

We operate in the same way as done for the optical module with the SAR module. Even though theoretically we have eight (8) features for SAR, we have to highlight that all the inc (increased in the backscattering objects) – as reported in Table 5-6 – are mostly NaN (Not a Number). Therefore, we have to consider them as missing data. As already done for the optical features in Table 5-12, we carry out the same correlation for the SAR DBO features. The following table shows the correlation among the SAR features. Again, the existence of a strong correlation does not imply a causal link between the variables, but gives us information on which variables could provide additional information if added to the network.

The module provides as output, for each building, either the change features of DB/LO and INC objects (NaN when no object is associated to a building).

Table 5-16: Correlation Matrix for the SAR features

	SAR							
	dbo cbkld	inc cbkld	dbo coh	dbo corr	dbo i ratio	inc coh	inc corr	inc i ratio
dbo cbkld	1,00	-0,09	-0,21	0,45	0,29	0,08	0,15	0,10
inc cbkld	-0,09	1,00	0,38	-0,20	-0,22	0,14	0,05	0,22
dbo coh	-0,21	0,38	1,00	-0,24	-0,14	0,03	0,05	0,11
dbo corr	0,45	-0,20	-0,24	1,00	0,18	-0,14	-0,11	-0,12
dbo i ratio	0,29	-0,22	-0,14	0,18	1,00	0,07	0,26	-0,41
inc coh	0,08	0,14	0,03	-0,14	0,07	1,00	0,68	0,12
inc corr	0,15	0,05	0,05	-0,11	0,26	0,68	1,00	-0,09
inc i ratio	0,10	0,22	0,11	-0,12	-0,41	0,12	-0,09	1,00

The optimised selection carried out to identify the optimal combination uses all the four DBO features. The identified continuous features and their relevant thresholds are reported in the following. Since most of the INC data are NaN, the selection of features was oriented only to DBO features, namely coherence, correlation, the Double Bounce ratio between the pre- and post-event backscatter intensity images and the Kullback-Leibler Divergence.

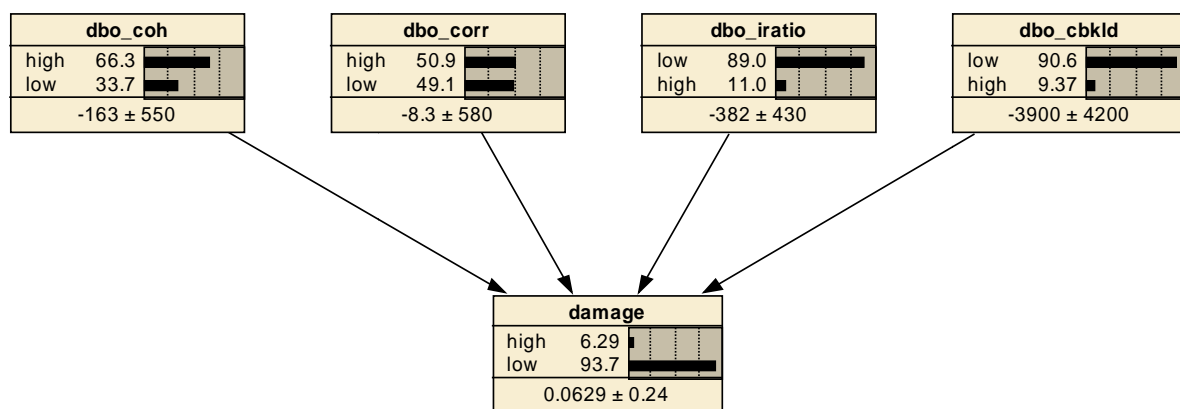|	
dbo_coh <	0.5131
dbo_corr <	0.6619
dbo_iratio >	14.9633
dbo_cbKLD >	339.931

Figure 5-13: SAR Feature module

The threshold for the damage node is set to 4,5% since also here we want to compensate for the unbalanced collapsed dataset. By analysing the combinations of features values that are mostly associated to collapsed buildings, we can identify the following combinations of features that are associated to collapsed buildings. The table shows if the feature is labelled as "*low*" damage probability or "*high*" damage probability.

Table 5-17: Combinations of SAR features associated to Collapsed Buildings

dbo_coh	dbo_cbKLD	dbo_iratio	dbo_corr
Low	High	High	Low
High	Low	High	High
High	High	High	High

The results of the net above in Figure 5-13 are reported in the following table, both for the non validated run and for the k-fold cross validation (k=10):

Table 5-18: SAR Feature Module Performances

	confusion matrix	SAR_no validation	SAR_validation
a	detected collapsed	24	24
b	false positive collapsed	63	182
c	misdetected collapsed	51	51
d	detected not collapsed	1539	1420
	Cohen's Kappa	26,1	11,3
	Cohen's Kappa normalised	28,1	20,6

In general, we can state that the SAR module has worst performances than the optical one, since it is able to detect a lower number of collapses at the expense of a higher number of false positives (k≅11,3%).

Similarly to what was done for the optical module, also in this case, an extension of this net was tested to better address the combination of all the modules. The first net is based on SAR continuous probability damage node, the other one with a discrete SAR classification node. These nets and their relevant results are reported hereinafter where we can see that their performances are substantially the same, as expected since the two additional nodes are substantially summarising the modules features and can be interpreted as latent variables. These variables have the capability of capturing complex relationships among the manifest variables.
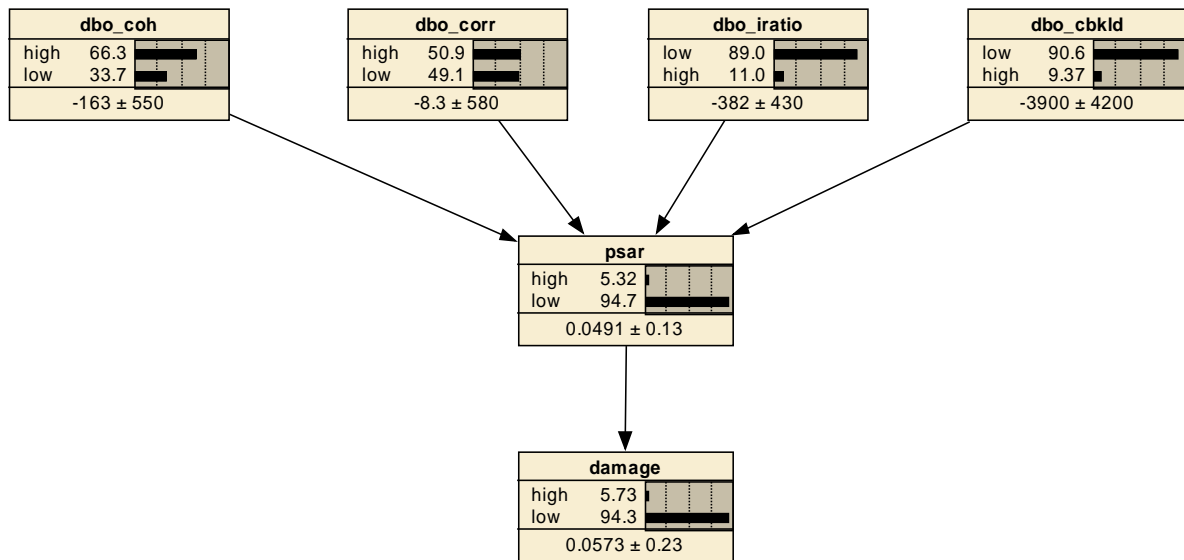
Figure 5-14: SAR Feature module with a continuous probability node

Figure 5-15: SAR Feature module with a discrete classification node

The results of the above-mentioned nets are reported in the table below. This table also again clearly shows that we do not lose information in the net when adding a node, such a p_{sar} or c_{sar}.

The results of the above mentioned nets are reported in the table below, in both non validated and validated (k-fold) cases, as usual:

Table 5-19: SAR Module Performances with probability and classification nodes

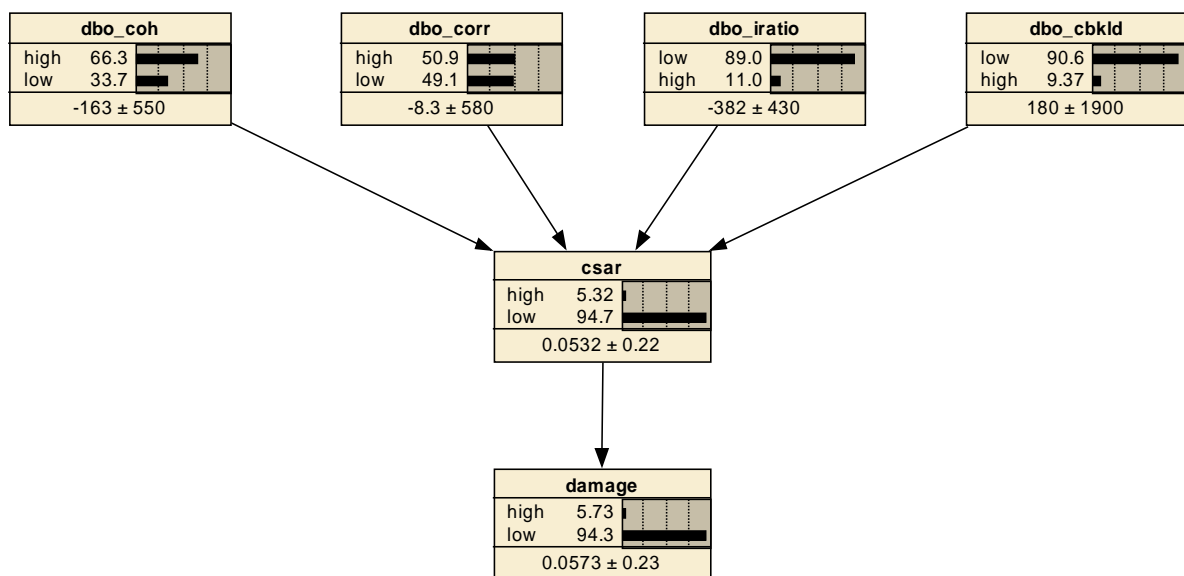|	confusion matrix	SAR+psar_no validation	SAR+psar_validation	SAR+csar_no validation	SAR+csar_validation
a	detected collapsed	24	23	24	23
b	false positive collapsed	62	179	62	179
c	misdetected collapsed	51	52	51	52
d	detected not collapsed	1540	1423	1540	1423
	Cohen's Kappa	26,3	10,8	26,3	10,8
	Cohen's Kappa normalised	28,1	19,5	28,1	19,5

Similarly, to the performances described before for the network in Figure 5-13, also, the SAR networks in Figure 5-14 and Figure 5-15, the performances are quite low and a high number of false positives are identified.

5.5.2.1   STRUCT Module

For the sake of completeness, we also built a full structural module that includes most of the features available. The extended version of the Structural module can be created depending on the relations between the different building features, as shown in the following picture. The macroseismic intensity is given as a set value when a post-event simulation is carried out. In our case, this value is 8,5 for L'Aquila event.

In this network, we included the relationship between the regularity in plant (planReg node) of the building and the Earthquake Resistance Design (ERD) to the type of building (namely masonry of reinforced concrete – codCIEMS98). Also the node InfillReg node is related to the type of building since this feature addresses the position of the building with respect to an aggregate. The following figure better explains the meaning of the two nodes "PlanReg" and "InfillReg" which might not always be available.
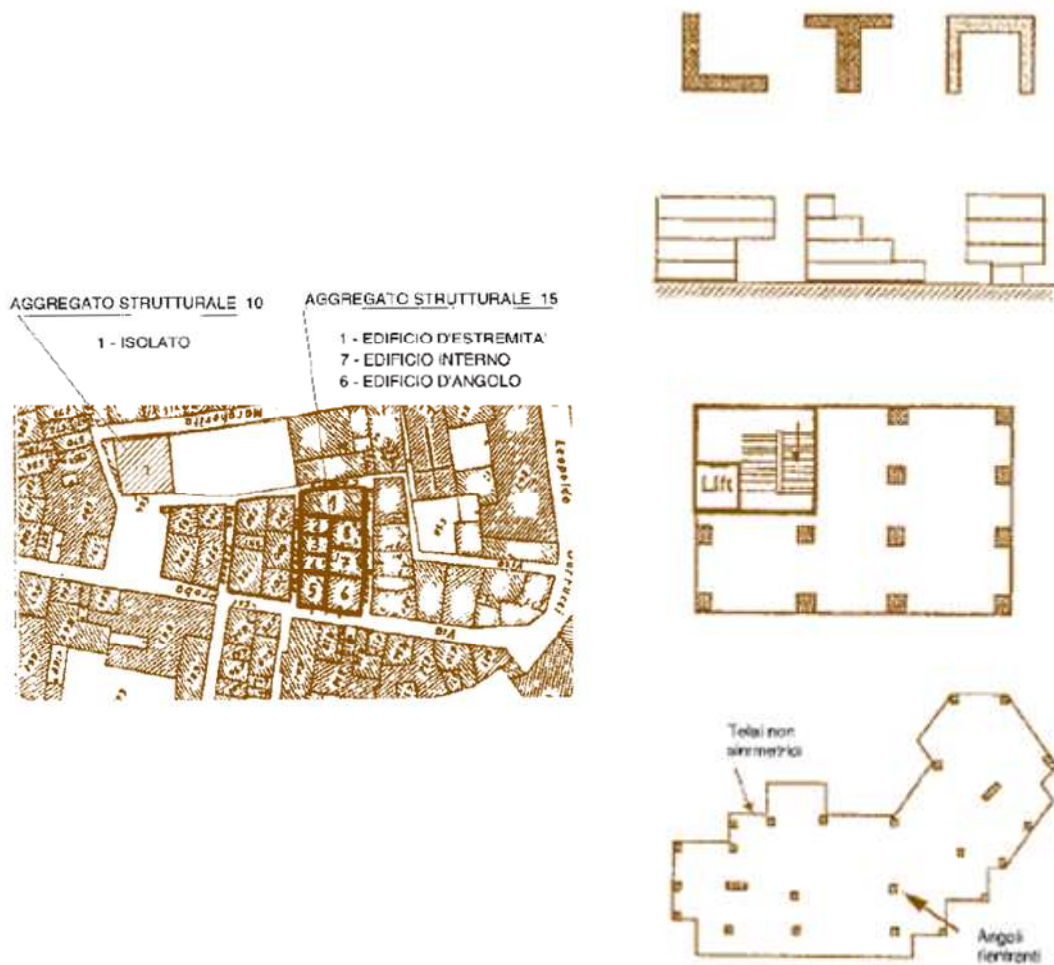
a) The building may be in isolated, header, internal, or corner position

b) Regularity in Plant

Figure 5-16: Description of the a) InfillReg and b) PlanReg nodes

The complete structural net also includes information on the building age (codYear) and the building height (codHeight).

However, all the nodes mentioned lead in the end to the μ_D node that makes use of the vulnerability concept and calculates the data using the model already described in Par. 5.2.

The following figure shows the above described network, without being trained:
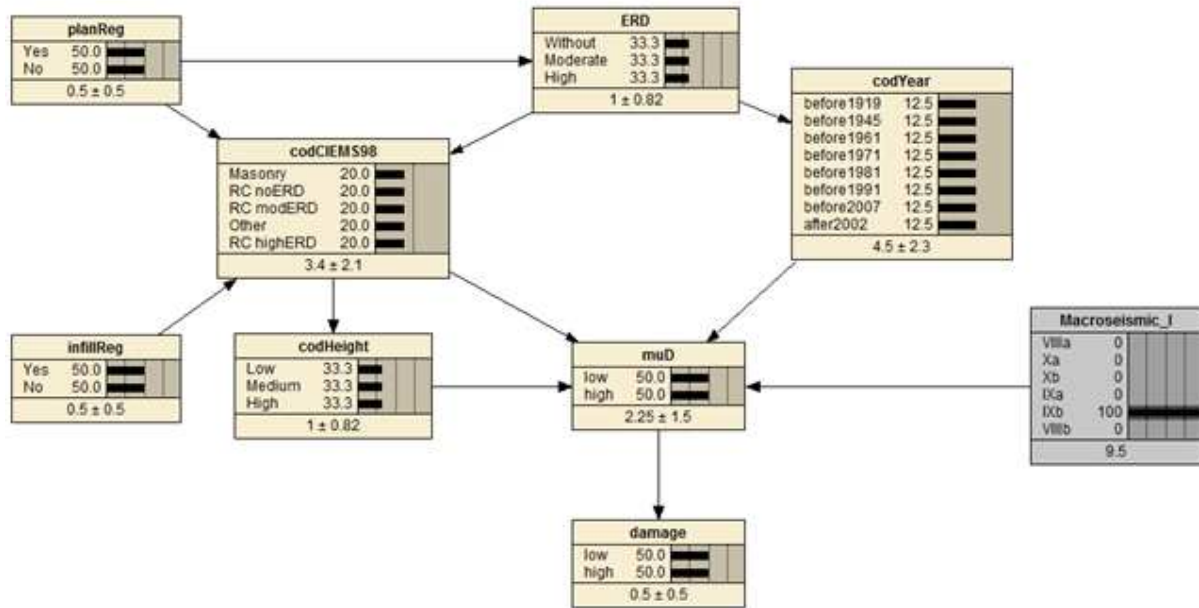
Figure 5-17: Complete Structural Module

However, since the only feature that comes into the damage assessment is summarised in μ_D, we can focus on a simplified version of the net since in our case we already have this value. The exercise of building the whole new is useful in extending the net for future cases where we might not have all features available. Our simplified structural network can be then summarised as follows:
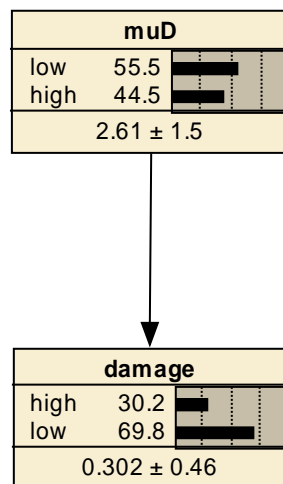
Figure 5-18: Simplified Structural Module

where μ_D only has few values, included between 0 (no damage) and 5 (Collapsed). Therefore, in the net above it is assumed that the μ_D node is continuous (see Eq. 5-3) and a threshold was selected to identify when the building should be considered as collapsed:

$$\mu_D > \quad 3$$

The results of the above mentioned simplified net are reported in the following table:

Table 5-20: Structural Module Performances

|	confusion matrix	STRUCT_no_valid ation	STRUCT_validation
a	detected collapsed	25	25
b	false positive collapsed	580	580
c	misdetected collapsed	50	50
d	detected not collapsed	1022	1022
	Cohen's Kappa	-0,7	-0,7
	Cohen's Kappa normalised	-2,9	-2,9

As expected, the use of only the structural module, that uses just one feature (the building vulnerability class), is not able to detect collapsed buildings. The reason for this is due to the fact that, using the simplified model explained in Par. 5.2.1, only a small number of collapses is identified jointly with a high number of false positives.

5.5.2.1 GEOTEC Module

Similarly, to the structural module the geotechnical module can be represented by one node. In our case, the fusion module makes use of the only available data, which is represented by the evaluation of the geotechnical probability.
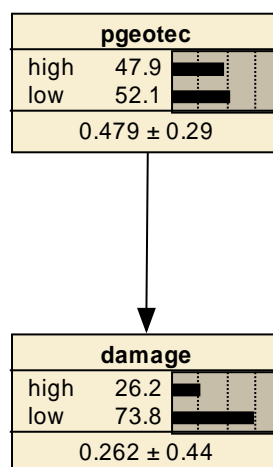

pgeotec	
high	47.9
low	52.1
0.479 ± 0.29	

damage	
high	26.2
low	73.8
0.262 ± 0.44	

Figure 5-19: Geotechnical Module

The optimised cutoff threshold is set to

$p_{geo} > 0.4793$

The results for this net are the following:

Table 5-21: Geotechnical Module Performances

confusion matrix		GEO_no validation	GEO_validation
a	detected collapsed	47	46
b	false positive collapsed	741	730
c	misdetected collapsed	28	29
d	detected not collapsed	861	872
	Cohen's Kappa	3,0	2,9
	Cohen's Kappa normalised	16,4	15,8

The geotechnical module, in L'Aquila test case, has just one feature related to the soil resonant period. It is then not surprising that the performance of this module is poor as reported in the table above (k ≅ 3%).

5.5.3   Data Fusion Module

Also in this case, as already explained for the structural module, we can think of an extended network that could potentially make use of different features, including the structural, optical, SAR and geotechnical features, as shown in the following picture in a non-trained example:
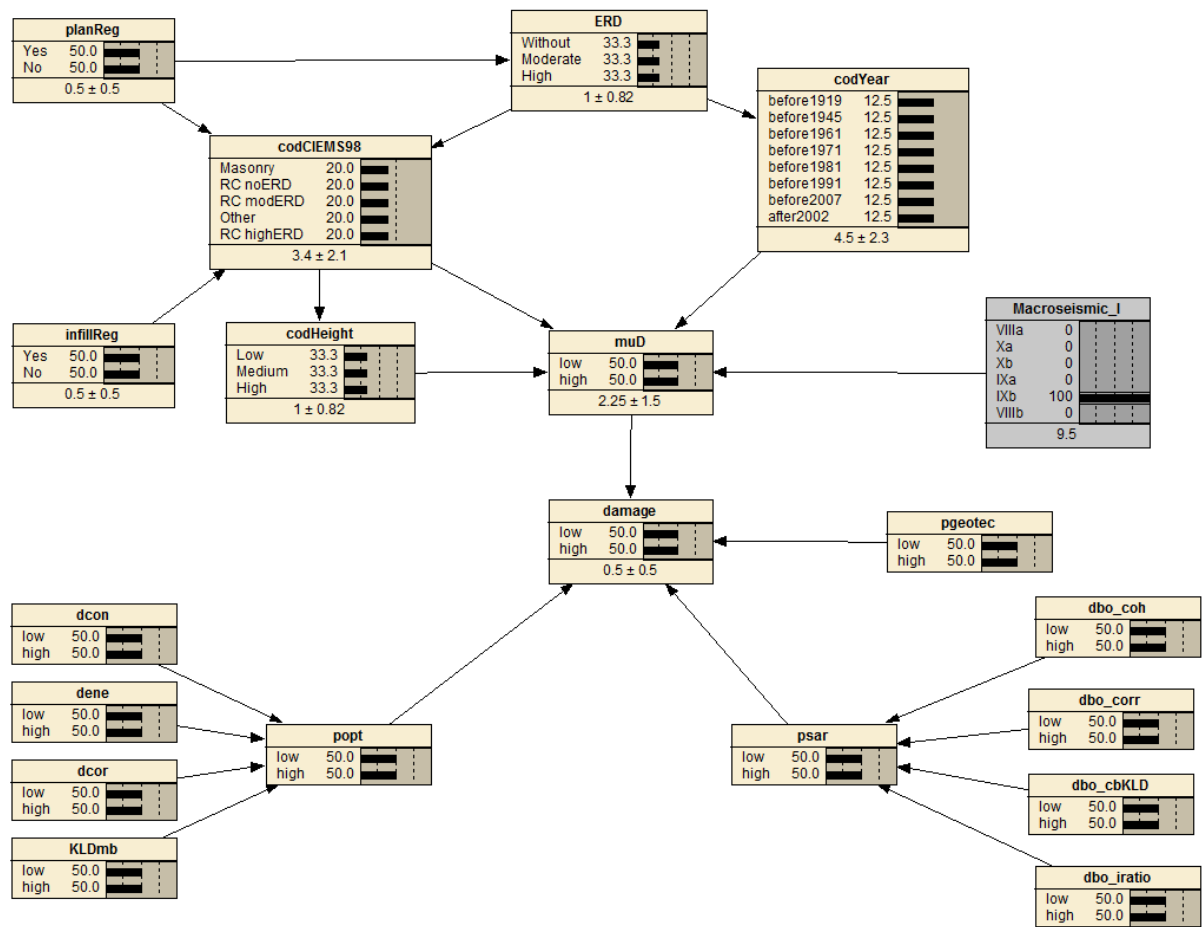
Figure 5-20: Complete Bayesian Network

However, we can focus on a simplified version of the net, since the only features that come into the damage node are reported in the Bayesian network hereinafter:
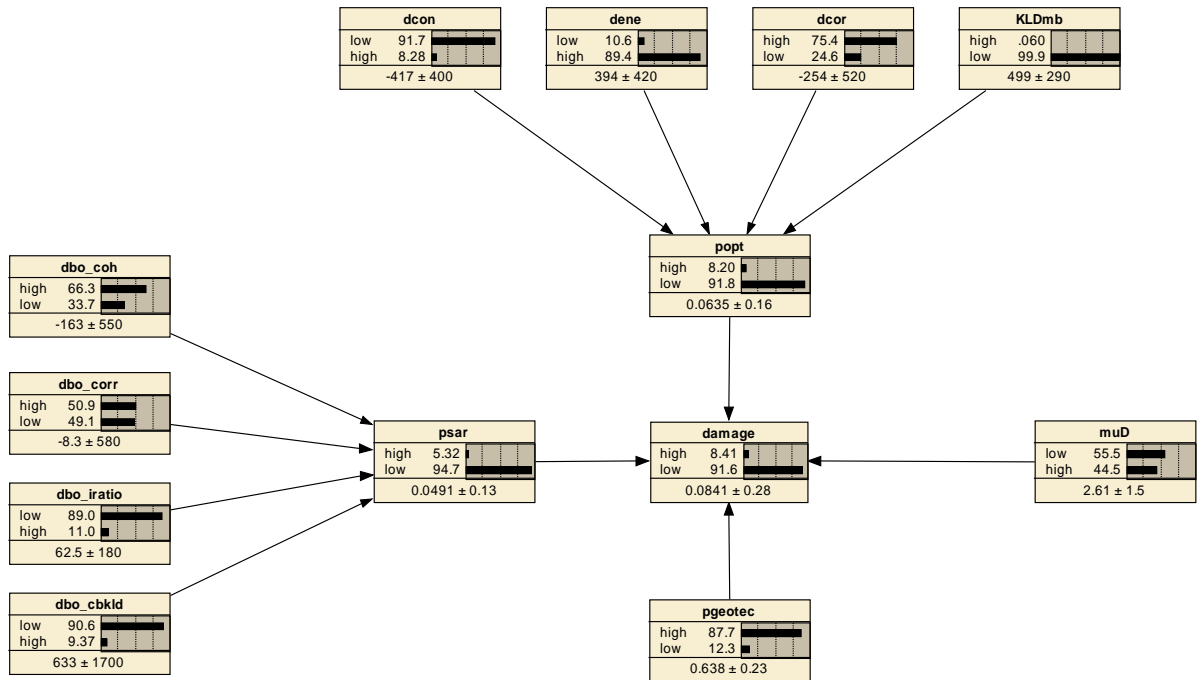
Figure 5-21: Bayesian Network for Data Fusion model in APhoRISM

The advantage of using this type of configuration is that we can mainly work on "building blocks" that are already trained and we can focus on their combinations.

If we the results obtained by the single individual modules above with the ones that can be obtained by adding *a-priori* information we obtain some combinations that are reported in the following table.

Table 5-22: Data Fusion Results

	confusion matrix	OPT_SAR_no validation _prob	OPT_SAR_vali dation_prob	OPT_SAR_STRUCT _no validation	OPT_SAR_STRUCT _validation	OPT_SAR_STRUCT _GEO_no validation	OPT_SAR_STRUC T_GEO_validatio n
a	detected collapsed	36	36	44	43	43	42
b	false positive collapsed	96	103	157	247	141	215
c	misdetected collapsed	39	39	31	32	32	33
d	detected not collapsed	1506	1499	1445	1355	1461	1387
	Cohen's Kappa	30,8	29,6	27,1	17,7	28,7	19,7
	Cohen's Kappa normalised	42,0	41,6	48,9	41,9	48,5	42,6

The combination of only the Optical and SAR modules does not significantly increase the performances while the overall effect is quite negligible. This happens most likely because the Optical dataset holds a similar and, in the average, richer information with respect to the SAR dataset.

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 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) implies a very low probability that the building is collapsed or not, regardless of what is observed from remote sensing (Optical or SAR). We can state that while the structural module has a tendency in identifying the highest level of collapsed buildings, however this is also reflected in the increased number of false positives. This tendency can be mitigated by the remote sensing performance, therefore when the structural module identified a collapsed building; jointly with any of the remote sensing modules (either Optical or SAR) the assessment is that the building is highly likely to be collapsed.

Considering also the geotechnical probability evaluation, the results, compared to the ones obtained by combining only the Structural, Optical and SAR modules are confirmed but do not significantly improve, as shown in Table 5-22. However, a small contribution to the reduction of false positives seems to be realised by the addition of this module.

The contribution of the Geotechnical module is not highly significant, even though it contributes in reducing the false alarms, not changing the general overall classification. This is also because the correlation observed in Figure 5-8b is obtained after averaging within each damage class but the data at individual building scale are very much scattered.

The addition of the structural module is however very important in this scheme to catch seven (7) collapsed buildings; this means a 12% increase in the identification of true positives. Such result essentially confirms the objective pursued in the APhoRISM project, i.e., *a-priori* information, not coming from remote sensing, notwithstanding its poor classification performances, can significantly increase the performances of any classifier based on remote sensing datasets only.

The advantage of using the network also relies in the fact that, if any data is missing we can infere the result starting from a trained network.

6 CONCLUSIONS & RECOMMENDATIONS

In this work, we presented a possible approach to data fusion of different types of data available from different observations to post-earthquake damage. The study was carried out in the framework of the APhoRISM FP7 funded project, which, among others, aims at improving the post-seismic event damage evaluation by combining remote sensing and a-priori in-situ data. The goal is to combine heterogeneous data for a final assessment of the damage at individual building level. In particular, the discrimination of fully collapsed buildings was here considered as the main objective.

Our test case is based on L'Aquila main seismic event that took place on April 2009, since this was the event on which it was possible to gather a sufficiently significant set of data for both remote sensing and *a-priori* datasets, although not all those potentially valuable for this objective. Actually, we have features from optical and SAR change detection analysis for each building in the central part of the town, together with so-called "*a-priori*" data related to structural vulnerability and soil frequency. In this test area the interseismic velocity from InSAR, as well as other soil instability factors and ground acceleration were not available or not exhibiting a suitable variability to assume a positive contribution to the final assessment. Therefore, their evaluation, though included in the general approach proposed for the complete APE flowchart (see Figure 3-1) is not herein presented. However, considering that we aim at developing a generally applicable methodology, the additional data can be considered as a more detailed elaboration of the geotechnical module, therefore already included in the overall architecture. In future validation test sets, if this data is available a dedicated analysis can be extended on the geotechnical module.

Following our approach which, as mentioned, aims at providing a methodology that could easily be extended to future events, regardless of their geographical location and magnitude, we designed an architecture based on Bayesian networks. The reasoning behind this choice is, in fact, that (as also experienced in L'Aquila test case) the possible situation of missing data is quite likely to occur also in other cases. The use of a Bayesian approach allows exploiting the hypothesis of conditional independence between remote sensing observations and other data sources. The approach based on networks also enables the gradual upgrade of the nodes, once new data becomes available, through the intrinsic Bayesian inference of the network.

The main conclusions of our work demonstrates that the use of *a-priori* structural data with a simplified model, notwithstanding their poor capability to predict the damage of buildings when used on their own, are able to improve the final damage assessment product. In particular, the main contribution of the structural data is in the reduction of false alarms of the image change detection that makes use of remote sensing data available (either SAR or optical or both).

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, though observable, was very small. As mentioned, in our L'Aquila test case, the maps of soil shaking, which are supposed to improve the overall damage assessment, were completely uniform through the whole test area. Therefore, this aspect could not be further exploited in our work.

Having summarised our main results, we can identify some actions that could be implemented in future works to improve our tool and extend its applicability. We can summarise the next actions as in the following, where we clustered the actions into two main groups. One is related to algorithm and computational improvements; the second group is instead more related to the methodological and practical implementation of this kind of tool in real post seismic event situations.

Computational Improvements

- Paragraph 5.5 presents a heuristically augmented optimisation algorithm to identify the best thresholds for continuous variables that need to be discretised in the Bayesian network node. This optimisation is fully based on the identification of the maximum Cohen's Kappa. The construction of an optimisation module already integrated in NETICA for the automatic identification of cut off thresholds for more efficient BN formulations should be developed to increase efficiency allowing consideration of larger systems.

- In our work, we have classified the output in either "*Collapsed*" or "*Not Collapsed*". However, a more sophisticated approach could make us of at least three levels classifications for the building damage (for example, low, medium, high damage depending on the percentage of building damage that can be observed). Such extension can be implemented in our model, given that the available data set includes enough information to provide this kind of difference.

- Seismic events provide information about seismic hazard and system performance. This information should be used to update the models, and the system characterizations that are used in BNs. The methodology described here should be extended to facilitate this updating.

- Paragraph 5.1 presents the difficulty in building the ground truth. The highly fragmented information and the different sources that provide the information are not always easy to interpret. Ideally, an automated GIS based interface to harmonise the different layers information should be developed to easily identify the ground truth reference.

Methodological Improvements

- In a practical scenario, where decision need to be taken to implement the required actions (e.g. post-seismic rescuers after an earthquake or allocation of resources for repair and reconstruction), we need to further extend our model to provide, after the evidence on the building damage, a clear action in collaboration with actual decision-makers.

- We have built a highly modular architecture that, for the moment, includes modules specifically dedicated to earthquakes. However, the same methodology could be extended to other types of natural disasters, such as volcanic eruptions, floods, etc. The so-called structural module would need to be replaced with other kinds of a-priori

data, while remote sensing data could add new or make use of similar features, if needed. The development of models for additional disasters will increase the harmonisation process that we are aiming at in the APhoRISM framework.

- For real post-seismic event scenarios, the proposed methodology described in this study must be integrated directly with tools that process data from external sources, (e.g. structural health monitoring sensors, ground motion recording instruments, remote sensing data). Ideally, the data should be fed automatically into the BN without external actions. This would require a well developed interface that processes both input and output data.

REFERENCES

CHAPTER 1

[1] United Nations Resolution adopted by the General Assembly 63/217. Natural disasters and vulnerability - on the report of the Second Committee (A/63/414/Add.3) - http://www.unisdr.org/files/resolutions/N0848381.pdf - A/RES/63/217 – December 2008

[2] UN Resolution A/RES/44/236 - International Decade for Natural Disaster Reduction – http://www.un.org/documents/ga/res/44/a44r236.htm - December 1989

[3] CEOS – Committees on Earth Observation Satellites - Satellite Earth Observations in Support of Disaster Risk Reduction - 3rd United Nations World Conference on Disaster Risk Reduction (WCDRR) – March 2015

[4] http://www.copernicus.eu/

[5] http://www.aphorism-project.eu/

CHAPTER 2

[6] Solano E., Remote Sensing to Assessment Damages Post Earthquakes, focusing in Urban Structures, March 2013

[7] http://www.disasterscharter.org/home

[8] http://www.ceos.org/

[9] AO7665 – EOEP-EOMD-EOPS-SW-13-0002 - Statement of Work - Support to the Agency concerning Disaster Risk Management activities under CEOS and the International Charter 'Space & Major Disasters'-September 2013

[10] http://www.protezionecivile.gov.it/jcms/it/rischio_sismico.wp

[11] http://www.protezionecivile.gov.it/jcms/it/attivita_sismico.wp%3Bjsessionid=02A60F 31380F483F18849198BCC706E1?pagtab=3

[12] http://www.protezionecivile.gov.it/jcms/it/glossario.wp?contentId=GLO13255

[13] http://www.datapiano.it/prodotti/sige.php

[14] C. Baggio, A. Bernardini, R. Colozza, L. Corazza, M. Della Bella, G. Di Pasquale, M. Dolce, A. Goretti, A. Martinelli, G. Orsini, F. Papa, G. Zuccaro - Field Manual for post-earthquake damage and safety assessment and short term countermeasures (AeDES) – European Commission - JRC Scientific & Technical Reports - 2007

[15] UNISDR – http://www.unisdr.org/we/inform/terminology

[16] ftp://ftp.ingv.it/pro/gndt/

[17] MIT – Infrastructure Science and Technology Group – Seismic risk assessment and loss estimation - http://web.mit.edu/istgroup/ist/documents/earthquake/Part1.pdf

[18] Giardini et al - The GSHAP Global seismic hazard map – Annali di Geofisica Vol. 42 n°6 – December 1999

[19] D' Andrea, Condorelli – Metodologie di valutazione del rischio sismico sulle infrastrutture viarie – World Road Association

[20] http://www.mi.ingv.it/pericolosita-sismica/

[21] http://www.ingv.it/it/

[22] http://www.protezionecivile.gov.it/jcms/it/classificazione.wp

[23] http://earthquake.usgs.gov/earthquakes/shakemap/

[24] http://rts.crs.inogs.it/it/project/15_tipologia-di-mappe.html

[25] Allen, T.I., Wald, D.J., Hotovec, A.J., Lin, K., Earle, P.S., and Marano, K.D., 2008, An Atlas of ShakeMaps for selected global earthquakes: U.S. Geological Survey Open-File Report, 2008-1236, 34 p.

[26] Advanced National Seismic System - ShakeMap® Manual Technical Manual , Users Guide, and Software Guide – Version 1.0 June 2006

[27] http://www.protezionecivile.gov.it/jcms/en/microzonazione.wp

[28] Grunthal – European Macroseismic Scale – 1998

[29] Gem – open earthuake oq risk manual

[30] Zuccaro – vulnerabilità presentazione

[31] Calvi, Pinho, Magenes, Bommer, Resrepo-Velez, Crowley – Development of seismic vulnerability assessment methodologies over the past 30 years – Journal of Earthquake Technology , Paper No 472, Vol. 43, No 3, pp. 75-104

[32] Barbara Borzi, Helen Crowley, Rui Pinho - Simplified Pushover-Based Earthquake Loss Assessment (SP-BELA) Method for Masonry Buildings- International Journal of Architectural Heritage: Conservation, Analysis, and Restoration Volume 2, Issue 4, 2008

[33] Crowley H, Bommer J, Pinho R. J.F. Bird- Displacement based earthquake loss assessment – DBELA –Research Report Rose 2006/01 IUSS Press

[34] Polese M., Di Ludovico M., Prota A., Manfredi G., (2013) "Damage-dependent vulnerability curves for existing buildings", Earthquake Engineering and Structural Dynamics, 42 (6), 853-870, DOI: 10.1002/eqe.2249

[35] I. Iervolino, G. Manfredi A review of ground motion record selection strategies for dynamic structural analysis - 2008

[36] Rossetto, T., Elnashai, A., 2003, Derivation of vulnerability functions for European-type RC structures based on observational data, Engineering Structures, (25) 1241-1263

[37] Freeman S.A. – Review of the development of the Capacity Spectrum Method – ISET Journal of Earthquake technology , Paper n° 43 8, Vol. 41, N°1, March 2004, pp. 1-13

[38] HAZUS® - MH 2.1 - Multi Hazard Loss Estimation Methodology – Earthquake Model Technical Manual - http://www.fema.gov/hazus

[39] Fajfar P., A non linear analysis method for performance based seismic design – Earthquake Spectra Vol. 16, n°3, pp. 573-592, Augus t 2000

[40] Di Pasquale, Dolce, Martinelli – Analisi della Vulnerabilità

[41] Attolico – Rischio sismico

[42] A. Bernardini, S. Giovinazzi, S. Lagomarsino, S. Parodi – Matrici di probabilità di danno implicite nella scala EMS-98 – 2008

[43] http://www.provincia.pu.it/protezionecivile/piano-provinciale-di-emergenza/lanalisi-dei-rischi/il-rischio-sismico/scenari-di-rischio-terremoti/

[44] D. Benedetti, V. Petrini (1984), "Sulla vulnerabilità sismica degli edifici in muratura: un metodo di valutazione", L'Industria delle Costruzioni, n. 149, Rome

[45] Liberatore, Zuccaro - A trent'anni dal Terremoto dell'80 La Prevenzione del Rischio Sismico tra Memoria e Innovazione – Novembre 2010

[46] Taubenbock, Post, Roth, Kosseder, Strunz, Dech – A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing - 2008

CHAPTER 3

[47] IASPEI Commission on Earthquake Hazard, Risk and Strong Ground Motion (IASPEI, 2004) http://www.iaspei.org/

[48] http://www.evoss-project.eu/

[49] http://www.subcoast.eu/

[50] http://www.doris-project.eu/index.php?option=com_content&view=category&layout=blog&id=39&Itemid=430

[51] http://www.pre-earthquakes.org/

[52] http://www.lampre-project.eu/

[53] http://www.ssg.group.shef.ac.uk/semep/

[54] http://www.prefer-copernicus.eu/

[55] http://www.pharos-fp7.eu/

[56] http://www.floodis.eu/

[57] http://www.rasor-project.eu/

[58] http://www.floodis.eu/

[59] http://www.rasor-project.eu/

[60] http://www.fast-space-project.eu/

[61] http://www.sensum-project.eu/home

[62] http://www.move-fp7.eu/index.php?module=main

[63] D.4.2 Handbook of Vulnerability Assessment in Europe –EC DG Environment – September 2011

[64] http://www.emergencyresponse.eu/gmes/en/ref/home.html

[65] http://www.gmes-gmosaic.eu/

[66] http://externalaction.security-copernicus.eu/

[67] http://www.eugene-fp7.eu/

[68] http://www.increo-fp7.eu/

[69] http://www.geo-pictures.eu/

[70] http://www.pangeoproject.eu/home

[71] http://vulsar.galileianplus.it/

[72] http://www.epo.org/searching/free/espacenet.html

[73] http://massive.eu-project-sites.com/Default.aspx

[74] http://ec.europa.eu/echo/files/civil_protection/civil/prote/pdfdocs/massive_final_techn ical_report_en.pdf

[75] Sedan O., Negulescu C., Terrier M., Roulle A., Winter T., Bertil D. - Armagedom – A tool for seismic risk assessment illustrated with applications – Journal of Earthquake Engineering, 17:2, 253-281, 2013

[76] Panagiota M., Jocelyn C., Erwan P. - State of the Art on Remote Sensing for vulnerability and damage assessment on urban context – March 2011

[77] Aphorism - Advanced PRocedures for volcanIc and Seismic Monitoring THEME [SPA.2013.1.1-07] [Remote sensing methods] - Grant Agreement No. 606738 http://www.aphorism-project.eu/

[78] Chini, M., Pierdicca, N. and 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, Vol. 47, No. 1, pp. 145 - 152, January 2009.

[79] Stramondo, S., Bignami, C., Chini, M., Pierdicca, N. and Tertulliani, A. "Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies". International Journal of Remote Sensing, Vol. 27, No. 20, pp. 4433 – 4447, October 2006.

CHAPTER 4

[80] Castanedo F. - "A review of data fusion techniques" – Hindawi Publishing Corporation – The Scientific World Journal, Vol. 2013, Article ID704504

[81] Foo P.H., Gee W.NG – "High Level information fusion: An Overview" – Journal of Advances in Information Fusion, Vol. 8, No. 1 – June 2013

[82] Steinberg A., Bowman C., White F. – "Revisions to the JDL data fusion model" - ERIM International, Inc. 1101 Wilson Blvd. Arlington, VA – 22209 – 1999 - www.dtic.mil/dtic/tr/fulltext/u2/a356422.pdf

[83] S. S. Blackman, "Association and fusion of multiple sensor data," in Multitarget-Multisensor: Tracking Advanced Applications, pp. 187–217, Artech House, 1990.

[84] S. Lloyd, "Least squares quantization in pcm," IEEE Transactions on Information Theory, vol. 28, no. 2, pp. 129–137, 1982.

[85] Y. Bar-Shalomand E. Tse, "Tracking in a cluttered environment with probabilistic data association," Automatica, vol. 11, no. 5, pp. 451–460, 1975.

[86] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-target tracking using joint probabilistic data association," in Proceedings of the 19th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes, vol. 19, pp. 807– 812, December 1980.

[87] D. B. Reid, "An algorithm for tracking multiple targets," IEEE Transactions on Automatic Control, vol. 24, no. 6, pp. 843–854, 1979.

[88] C. L. Morefield, "Application of 0-1 integer programming to multitarget tracking problems," IEEE Transactions on Automatic Control, vol. 22, no. 3, pp. 302–312, 1977.

[89] K. C. Chang, C. Y. Chong, and Y. Bar-Shalom, "Joint probabilistic data association in distributed sensor networks," IEEE Transactions on Automatic Control, vol. 31, no. 10, pp. 889–897, 1986.

[90] Y. Chong, S. Mori, and K. C. Chang, "Information fusion in distributed sensor networks," in Proceedings of the 4th American Control Conference, Boston, Mass, USA, June 1985.

[91] Y. Chong, S. Mori, and K. C. Chang, "Distributed multitarget multisensor tracking," in Multitarget-Multisensor Tracking: Advanced Applications, vol. 1, pp. 247–295, 1990.

[92] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Mateo, Calif, USA, 1988.

[93] Koller and N. Friedman, Probabilistic Graphical Models: Principles and Techniques, MIT press, 2009.

[94] C. Brown, H. Durrant-Whyte, J. Leonard, B. Rao, and B. Steer, "Distributed data fusion using Kalman filtering: a robotics application," in Data, Fusion in Robotics and Machine Intelligence, M. A. Abidi and R. C. Gonzalez, Eds., pp. 267–309, 1992

[95] R. E. Kalman, "A new approach to linear filtering and prediction problems," Journal of Basic Engineering, vol. 82, no. 1, pp. 35–45, 1960.

[96] R. C. Luo and M. G. Kay, "Data fusion and sensor integration: state-of-the-art 1990s," in Data Fusion in Robotics and Machine Intelligence, pp. 7–135, 1992.

[97] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practitioners," IEEE Transactions on Signal Processing, vol. 50, no. 3, pp. 736–746, 2002.

[98] J. Martinez-del Rincon, C. Orrite-Urunuela, and J. E. Herrero- Jaraba, "An efficient particle filter for color-based tracking in complex scenes," in Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 176–181, 2007.

[99] M. E. Liggins II, C.-Y. Chong, I. Kadar et al., "Distributed fusion architectures and algorithms for target tracking," Proceedings of the IEEE, vol. 85, no. 1, pp. 95–106, 1997.

[100] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," in Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03), pp. 138–149, November 2003.

[101] S. Bashi, V. P. Jilkov, X. R. Li, and H. Chen, "Distributed implementations of particle filters," in Proceedings of the 6th International Conference of Information Fusion, pp. 1164–1171, 2003.

[102] M. Coates, "Distributed particle filters for sensor networks," in Proceedings of the 3rd International symposium on Information Processing in Sensor Networks (ACM'04), pp. 99–107, NewYork, NY, USA, 2004.

[103] D. Gu, "Distributed particle filter for target tracking," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '07), pp. 3856–3861, April 2007.

[104] J. K. Uhlmann, "Covariance consistency methods for fault tolerant distributed data fusion," Information Fusion, vol. 4, no. 3, pp. 201–215, 2003.

[105] P. C. Mahalanobis, "On the generalized distance in statistics," Proceedings National Institute of ScienceIndia, vol. 2, no. 1, pp. 49–55, 1936.

[106] C. Cou´e, T. Fraichard, P. Bessì`ere, and E. Mazer, "Multi-sensor data fusion using Bayesian programming: an automotive application," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 141–146, October 2002.

[107] D. L. Hall and J. Llinas, Handbook of Multisensor Data Fusion, CRC Press, Boca Raton, Fla, USA, 2001.

[108] P. Dempster, "A Generalization of Bayesian Inference," Journal of the Royal Statistical Society B, vol. 30, no. 2, pp. 205–247, 1968.

[109] A. Shafer, Mathematical Theory of Evidence , Princeton University Press, Princeton, NJ, USA, 1976.

[110] A. M. Abdelbar, E. A. M. Andrews, and D. C. Wunsch II, "Abductive reasoning with recurrent neural networks," Neural Networks, vol. 16, no. 5-6, pp. 665–673, 2003.

[111] J. R. Aguero and A. Vargas, "Inference of operative configuration of distribution networks using fuzzy logic techniques. Part II: extended real-time model," IEEE Transactions on Power Systems, vol. 20, no. 3, pp. 1562–1569, 2005.

[112] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349–1380, 2000.

[113] Moser G., S. B. Serpico – "Unsupervised Change Detection From Multichannel SAR Data by Markovian Data Fusion" - IEEE Transactions on Geoscience and Remote Sensing, Vol. 47, No. 7, July 2009 – pp. 2114 - 2128

[114] Makarau A., Palubinskas G., Reinartz P. Remote Sensing Technology Institute German Aerospace Center (DLR)– "Multi-sensor data fusion for urban area classification" - In: Stilla U., Gamba P., Juergens C., Maktav D. (Eds) JURSE 2011 - Joint Urban Remote Sensing Event Munich, Germany, April 11-13, 2011

[115] Sportouche H., Tupin F., Denise L. – "Building Detection by Fusion of Optical and SAR Features in Metric Resolution Data" - 978-1-4244-3395-7/09 ©2009 IEEE - IGARSS 2009 – IV 769 -772

[116] Fauvel M., Chanussot J., Benediktsson J.A. – "Decision Fusion for the Classification of Urban Remote Sensing Images" - IEEE Transactions on Geoscience and Remote Sensing, Vol. 44, No. 10, October 2006 – pp. 2828 – 2838

[117] Pohl C., Van Genderen J.L. – "Review Article multisensory image fusion in remote sensing: Concepts, methods and applications" - International Journal of Remote Sensing, 1998, Vol. 19, No. 5823 -5854

[118] Sonia Giovinazzi - "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

[119] Bernardini et al.- "Matrici di probabilità di danno implicite nella scala EMS-98", XII convegno ANIDIS, 2007

[120] Bernardini et al. - "Vulnerabilità e previsione di danno a scala territoriale secondo una metodologia macrosismica coerente con la scala EMS-98", XII convegno ANIDIS, 2007

[121] McGuire R. K. (2004). Seismic Hazard and risk analysis. EERI Earthquake Engineering research Institute

[122] http://research.microsoft.com/en-us/um/redmond/groups/adapt/msbnx/msbnx/basics_of_bayesian_inference.htm

[123] O. Bousquet, U. von Luxburg, and G. RÄatsch (Eds.) - Bayesian inference: An introduction to Principles and practice in machine learning - Advanced Lectures on Machine Learning, pp. 41-62. Springer, 2004.

[124] Mihajlovic V., Petkovic M. – Dynamic Bayesian Networks: A state of the Art – University of Twente

[125] Spiegelhalter D.J., Dawid A.P., Lauritzen S., Cowell R. - Bayesian Analysis in Expert systems – Statistical Science, Vol. 8 n°3, 219-283 – 1993

[126] http://www.norsys.com/

[127] K. B. Korb, A.E. Nicholson - Bayesian Artificial Intelligence – Chapman & Hall, London, UK– 2004

[128] Netica-J Manual - Version 4.18 and Higher – Java Version of Netica API Norsys Software Corp - October 2010

[129] R., Stuart, P. Norvig (2009) Artificial Intelligence: A Modern Approach (3rd edition), Prentice Hall, Englewood Cliffs, NJ

CHAPTER 5

[130] 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 EMS98 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. DOI:10.1007/s10518-010-9188-4

[131] 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. DOI:10.1785/0120110205

[132] Kramer, S. L. (1996). Geotechnical Earthquake Engineering. Engineering (Vol. 6). http://doi.org/10.1007/978-3-540-35783-4

[133] S. Visa, B. Ramsay, A. Ralescu, E. Van Der Knaap - Confusion Matrix-based Feature Selection – 1998

[134] M. Banerjee, M. Capozzoli, L. McSweeney, D. Slnha - Beyond kappa: A review of interrater agreement measures - The Canadien Journal of Statistics Vol. 27, No. 1, 1999, Pages 3-23 - La revue canadienne de statistique

[135] Cohen, J. (1960). A coefficient of agreement for nominal scales. Edu. ond Psych. Meas., 20, 3 7 4 6

[136] Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psych. Bull., 70, 213-220

[137] Donalek C. - Supervised and Unsupervised Learning Ay/Bi 199 – April 2011 - http://www.astro.caltech.edu/

[138] P.N. Tan, M. Steinbach, V. Kumar - Introduction to Data Mining - Pearson Addison-Wesley – 2006

[139] Freund Y, Schapire RE. Experiments with a new boosting algorithm. Proceedings 13th International Conference on Machine Learning 1996; 148–156

[140] Breiman L. Bagging predictors. Machine Learning J 1996; 24(2): 123–140

[141] Ho TK. The Random subspace method for constructing decision forests. IEEE Trans Pattern Analysis and Machine Intelligence 1998; 20(8): 832–844

[142] Grossman,, Robert; Seni, Giovanni; Elder, John; Agarwal, Nitin; Liu, Huan (2010). Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions. Morgan & Claypool. doi:10.2200/S00240ED1V01Y200912DMK002

[143] McLachlan, Geoffrey J.; Do, Kim-Anh; Ambroise, Christophe (2004). Analyzing microarray gene expression data. Wiley

[144] Rodgers, J. L.; Nicewander, W. A. (1988). "Thirteen ways to look at the correlation coefficient". The American Statistician 42 (1): 59–66. doi:10.1080/00031305.1988.10475524. JSTOR 2685263

[145] Luke ZeBlemoyer - CSE546: Linear Regression Bias/ Variance Tradeoff http://courses.cs.washington.edu/courses/cse546/12wi/slides/cse546wi12LinearRegression.pdf